Kant's Pedagogy of Hope:

A Reading of the 'Doctrine of Method' in the *Critique of Practical Reason*

Adam Blazej

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

Abstract

Kant's Pedagogy of Hope:

A Reading of the 'Doctrine of Method' in the *Critique of Practical Reason*

Adam Blazej

Why and, if so, how should educators cultivate hope in hopeless times? I defend a novel interpretation of Immanuel Kant's theory of moral education - specifically, what I call his "pedagogy of hope," a pedagogical method Kant prescribes to moral educators in the 'Doctrine of Method' of the Critique of Practical Reason for the purpose of cultivating virtuous character. According to Kant, moral educators should cultivate students' hope for moral progress in order to sustain their moral motivation in the face of uncertainty, failures, and suffering. Kant's two-step pedagogical method amounts to an aesthetic education, in the sense that it mirrors his account of the relationship between feeling and judgment in experiences of the beautiful and the sublime. Drawing on that account, I describe how, for Kant, moral educators can cultivate hope by developing students' judgment through deliberation of examples of moral conduct and of moral exemplars.

Table of Contents

List of Abbreviations

Kant

References to works by Kant are to the volume and page of *Kants gesammelte Schriften* (except citations to KrV, which are to the A and B pagination) and accord with the following abbreviation scheme:

ApH: *Anthropologie in pragmatischer Hinsicht / Anthropology from a Pragmatic Point of View* (vol. 7)

B: *Briefwechsel / Correspondence* (vol. 12)

Blom-L: *Logik Blomberg / Blomberg Logic* (vol. 24)

GMS: *Grundlegung zur Metaphysik der Sitten / Groundwork of the Metaphysics of Morals* (vol. 4)

IaG: "Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht" / "Idea for a Universal History with a Cosmopolitan Aim" (vol. 8)

KpV: *Kritik der praktischen Vernunft / Critique of Practical Reason* (vol. 5)

KrV: *Kritik der reinen Vernunft / Critique of Pure Reason*

KU: *Kritik der Urteilskraft / Critique of the Power of Judgment* (vol. 5)

MS: *Metaphysik der Sitten / Metaphysics of Morals* (vol. 6)

RGV: *Die Religion innerhalb der Grenzen der bloßen Vernunft / Religion Within the Bounds of Reason Alone* (vol. 6)

Rousseau

References to works by Rousseau accord with the following abbreviation scheme:

DI: *Discours sur l'origine et les fondements de l'inégalité parmi les hommes / Discourse on the Origin and Foundations of Inequality among Men*, in *The 'Discourses' and Other Early Political Writings*, trans. Victor Gourevitch (Cambridge: Cambridge University Press, 1997), 111-222.

E: *Émile, ou De l'éducation / Emile, or on Education*, trans. Allan Bloom (New York: Basic, 1979).

Other sources

All other sources are cited according to author name and initial date of publication (or composition). The full citations can be found in the References, where the initial date of publication or composition is followed by the publication of the cited edition in brackets.

Acknowledgments

During those times when I felt a loneliness in standing at the feet of Philosophy that seemed to paralyze me, I often drew strength from my gratitude to the many who supported me in this intellectual endeavor.

I feel fortunate to have been able to study under my adviser, Pat Kitcher. Pat was unwavering in her support, starting from the moment almost ten years ago that I first approached her with my interest in writing about Kant's *Critique of Judgment*, despite insisting up to this day that she is not a "third Critiquer." The intimate knowledge of Kantian scholarship that Pat generously shared in the many hours of conversation we had together in her office or on Zoom deepened my thinking, as well as made me feel there was a space for someone like me in philosophy. Particularly in the final years, Pat dutifully navigated me through deadlines and bureaucracy, while, at the same time, and with a sincere warmth, recognizing that I needed to care for my young sons. This kind of support is a blessing for any parent trying to complete a dissertation, let alone a parent trying to do so during a global pandemic.

I am also thankful to the other members of my committee – Francey Russell, Fred Neuhouser, Megan Laverty, and Michele Moody-Adams – for their commitment to seeing me through the completion of my defense. Additionally, I feel fortunate to have benefitted, even if only in passing, from the camaraderie and stimulating environment Megan has had a hand in cultivating at the Philosophy and Education program at Teachers College. I also feel fortunate to have been able to participate in Fred's dissertation colloquia, which made the research process less isolating, and his German translation group, which bolstered my confidence in my first year

at Columbia. Michele's seminars sparked my initial interest in Kant's social and political philosophy that led me to the topic of this dissertation.

From comments on drafts to general camaraderie, the community of graduate students in Columbia's Philosophy Department was indispensable. I am especially thankful to my cohort – Isabel Kaeslin, Jake McNulty, Laura Martin, Matt Heeney, Sébastien Rivat, Simon Brown, and Thimo Heisenberg – as well as others with whom I formed bonds or crossed paths, whether at MAP meetings, on the picket line, or in the grad lounge – Amogh Sahu, Brittany Koffer, César Cabezas, Chuyu Tian, Helen Zhao, Iain Atkinson, Jonathan Fine, Katie McIntyre, Lin Zhang, Phil Yaure, Qian Cao, Usha Nathan, and Yarran Hominh. Special thanks is also owed to Helen and Simon for taking walks with Penny, as well as Jeremy Forster, whose late night conversations were sorely missed at the onset of the pandemic and in the later years of my dissertation.

I am thankful, as well, for the work of all the staff in the Philosophy Department over the years, most of all: Asha Panduranga, D.D. Meakin, Maia Bernstein, and Stacey Quartaro. Without their work, the bureaucracy would be insurmountable. Additionally, they were so kind in allowing my sons to explore the department offices (even playing along themselves!), which helped to make the Department (and, by extension, philosophy) feel more like a home.

Among some of the most valued relationships I formed in graduate school were those related to my involvement with Rethink. Through Rethink, I met John Fantuzzo, Max Hayward, and Robbie Kubala, each of whom proved, in their own way, to be a valuable mentor in my first years at Columbia. I also made friends with whom I learned a great deal: Abram de Bruyn, Alex Rigas, Andy Jampol-Petzinger, Borhane Blili-Hamelin, Conor Cullen, Ignacio Ojea Quintana, Sara Hardman, Veronica Padilla, and Vita Emery. And I formed relationships with Mia Ruyter

from the Heyman Center for the Humanities, and staff at multiple community organizations, including the Osborne Association and Sanctuary for Families. These relationships, along with my work with other public philosophers in New York City like Ian Olasov, provided moral support and intellectual nourishment that sustained, as well as informed, my research when I had begun to lose touch with other institutions in academic philosophy. With regards to keeping in touch, Conor and Veronica deserve special thanks for their love and support in maintaining our friendships and pushing me to finish.

If not for Janice Daurio, my first philosophy teacher, it is likely I would not have started out on the path to graduate school in philosophy. And if not for the encouragement that she and Paul Ford afforded me, or the kindness they both showed towards me and my family at their home, it is likely I would not have continued on the path for as long as I did. Janice's classes at Moorpark College were, with no exaggeration, transformative for me, and continue to orient me in my research and teaching. I feel so fortunate to have had a position teaching at Moorpark the past few years that has helped sustain me and my family as I have completed this dissertation.

Some of the ideas of this dissertation began to percolate while I was at the University of Wisconsin–Milwaukee, where I met David Shope, Jon Grandits, Mathew Snow, and Nick Nash. Their love and friendship over the last ten years has meant more than I can put into words. Needless to say, this dissertation benefitted deeply from our conversations, and I would be hard pressed to turn down wasting time together with them. Much of the same goes for my friendships dating back even further with Tyler DeGroot and Matt Gillooly, except I would add that Tyler owes me another meal at Don Cucos in the hopefully not too distant future.

Finally, I am deeply grateful to my family who have continuously encouraged me in my philosophical endeavors despite, perhaps, not fully understanding what I do or why I do it. My

parents, Cathy and John, have been there since day one in support of my academic pursuits and drive for knowledge. While my brothers, Christopher, Matthew, and Joshua, have not all been there since day one, I have only felt their love and support to grow over the years. My sons, Callum and Holden, who were born at different stages of me writing this dissertation, have sustained (but also exhausted!) me through our many hours of playing together. And, most of all, I am so grateful for my wife, Christina, to whom this dissertation is dedicated. Her continued support and unwavering care, especially during the most demanding parts of the writing process, mean the world to me, as does her love and companionship, in general.

All errors are, of course, mine.

For Christina.

Introduction

In the autumn of 1791, a young disciple wrote to Immanuel Kant with a plea. The disciple was Maria von Herbert, an Austrian noblewoman who had participated in an intellectual salon centered around discussion of Kant's critical philosophy. She explained to the "Great Kant" that she had made a confession to a man she loved, which appeared to have led him to end his love for her. Utterly distraught, Maria tells Kant that she had contemplated suicide until realizing that doing so would violate the moral prohibition against suicide for which he argued: "your theory stops me," she wrote. Since, however, this had not put a stop to her torment, she called upon Kant "for solace, or for counsel preparing me for death," neither of which she had yet procured from his philosophy. The solace for which Maria longed was sought by her in being shown that this life has a value that "could replace the good I have lost."[1] What she seemed to be searching for, in other words, were grounds to hope in a life otherwise full of suffering, and she was searching for those grounds in, of all places, Kant's critical philosophy, if not in the philosopher himself.

Kant wrote a lengthy reply to Maria von Herbert's letter. There, he specifically avoids directly addressing her interest in finding hope. Instead, he focuses on those matters in her letter which he believes "spoke of truth and trust" and, accordingly, delivers an analysis of friendship and secrecy, much of which he would repeat word for word in *The Doctrine of Virtue*, published some six years later.[2] Speaking as her "moral physician," Kant advises Maria to ask herself whether her tormenting self-reproach stems from regret at having done her duty in disclosing the

[1] B 11:273-4.
[2] B 11:273n1.

secret to her friend, or from remorse over having deceived him in the first place.[3] The former, Kant says, is no moral reason for self-reproach, and the latter is justified only within limits. To that point of moderation, Kant says that "to make one's whole life useless by continuous self-reproach on account of something that happened once upon a time and cannot be anymore…would be a fantastic notion of deserved self-torture."[4] He continues: "It would be like many so-called religious remedies that are supposed to consist in seeking the favor of higher powers without one's even having to become a better human being."[5] This is what Kant calls *Schwärmerei*, a term roughly equivalent to "enthusiasm," which he deems morally objectionable.

At the end of his reply, Kant offers what is, for my purposes, a key to interpreting his seemingly roundabout response to Maria's calls for hope in a life otherwise full of suffering. What he says is that his letter followed "the customary divisions of a sermon: doctrine, discipline, and solace, of which I beg you to pay attention somewhat more to the first two, since the last, and your lost contentment with life, will surely be recovered by itself when once these others have had their effect."[6] Thus, according to Kant, matters of hope should not be of primary concern; rather, what should be of primary concern are the specific matters of theory and practice to which he devotes the analysis in his letter, since those provide the "doctrine" and "discipline" from which "solace," or hope, "will surely be recovered by itself." Without adhering to doctrine and discipline, Maria's calls for hope in finding a new purpose in life will amount to mere fantasizing that undermines her moral agency, manifesting, for example, in her tormenting self-reproach. By contrast, Kant claims, as if making a psychological prediction, that the right

[3] B 11:333.
[4] B 11:333-4.
[5] B 11:334.
[6] Ibid.

kind of hope, one necessary for moral agency, can only result from Maria attending to more fundamental matters of theory and practice.

But, for Maria, hope in this life was not to be recovered. About the same time that she wrote a second and third time to clarify the hope she was seeking and why it called for a different response than the one Kant had given, Kant received a letter from their mutual acquaintance, J.B. Erhard. Erhard told Kant that Maria was not in a state in which she would be receptive to moral instruction, for she had lost all "delicacy": "her moral feeling is totally severed from prudence and is therefore coupled with fantasy, a more subtle sensibility."[7] Trusting Erhard's judgment, Kant decided not to respond to Maria's second letter or a third she subsequently sent. Instead, he passed her two letters and Erhard's on to the young daughter of a friend as an "example of warning" of "sublimated fantasy."[8] Kant thus abandoned hope in a most devoted disciple of his critical philosophy who had been seeking grounds in that same philosophy for hope in the face of hopelessness. Maria von Herbert committed suicide some ten years later in 1803.

Kant's poignant correspondence with Maria motivates the central question of this dissertation: why (and how) should we cultivate hope through moral education? Consider, for instance, Kant's initial willingness to give some hearing to Maria's belief that his moral philosophy ought to provide hope. This willingness suggests that, for Kant, one possible function of moral philosophy is to be a source of hope – to not only articulate principles for how one ought to act, but also to explain how, through the recognition and application of those principles, we can cultivate hope, perhaps even in seemingly hopeless conditions. There is also Kant's refusal to answer Maria's subsequent calls for hope because he apparently came to believe she would not be receptive to the necessary moral education in her current state. What this indicates

[7] B 11:203-4.
[8] B 11:204.

is that Kant understood it to be a goal of moral education in particular to cultivate hope and that he believed he possessed some pedagogy that could achieve that goal under the right conditions, conditions apparently lacking in the case of Maria von Herbert. My chief aim in this dissertation is to defend a novel interpretation of Kant's account of moral education that articulates both why hope ought to be an aim of moral education and how that aim can be achieved. More simply, my aim is to reconstruct a Kantian pedagogy of hope.

That hope would be of some significance to Kant should be no surprise. He famously proposed three questions that answer "all my reason's interest": "What can I know?" "What ought I to do?" and "What may I hope?"[9] Admittedly, the first two of these are frequently discussed and recognizably central philosophical questions, whereas the third can seem unusual: Why is hope of central interest to reason? What does philosophy have to offer concerning hope, specifically (as it sounds) permission to hope? Moreover, Kant's own understanding of the question about hope, and his answers to it, are not as easy to identify and isolate as his treatment of the other two questions: his first two Critiques (along with the *Groundwork of the Metaphysics of Morals* and other works) respectively aim to systematically answer the questions concerning human knowledge and moral obligations. By contrast, Kant mentions hope infrequently and never provides an explicit analysis of what it is or what its philosophical significance might be. However, with the rise of interest in the topic of hope among philosophers over the past couple decades, there has also been increasing attention afforded to Kant's hope question, including, even if only marginally, its relevance in his correspondence with Maria von Herbert.

Much of the recent literature on Kant's account of hope takes the central target of analysis to be what Kant argues is the object of rational hope, viz. the 'highest good.' Generally speaking, Kant's highest good is that state of affairs, or that "world," in which the greatest happiness

[9] KrV A805-806/B833-834.

corresponds to the greatest morality. The highest good is, therefore. Kant's answer to – and interpretation of – the problem of justice in the world or, in Leibniz' language, to the problem of theodicy in a special (moral) restriction of its meaning. What cannot be accepted in Kant's eyes, from the moral point of view, is the impossibility for morally good human beings to hope for happiness as the just retribution of their morality. Hence, Kant's answer to his own question, 'What may I hope?' is the highest good, and what permits us to hope for the highest good (albeit, under certain conditions) is the authority of the moral law.

This brief explanation of Kant's notion of the highest good as the object of hope gives me reason to specify the more precise object of hope as it pertains to Kant's moral education. As I argue in later chapters where I discuss the highest good in more detail, the object of hope in Kant's moral pedagogical method is that respect for the moral law will constitute one's character, and so continue to influence one's future actions. In other words, students are taught to hope that they will continue to have the right motivation to do the right thing for the right reasons in the future, even against certain errors or obstacles. Here, the object of hope is not identical to the highest good, or a just world in which happiness is proportioned to virtue; rather, the object of students' hope is more akin to their own moral development, though that may arguably be a part of a hope for a better world. I will return in subsequent chapters to say more about the precise object of the hope at issue in Kant's moral pedagogy, as well as its relationship to his notion of the highest good.

Kant's account of moral education had been neglected up until recently for reasons similar to those that had led to lack of attention to his account of hope. For one thing, there appears to be no single writing of Kant's that explicitly targets the topic of education. Furthermore, his moral theory can make it seem as though moral education is unnecessary, since,

at times, he appears to argue that moral knowledge and decision-making can be self-taught. Moral education has even appeared impossible in light of Kant's theory of freedom: in short, the idea of education, as it is usually understood, implies the possibility of influencing another person's processes of learning, but, according to Kant, we can be free only if our will is not influenced by empirical causes, that is, if it stands outside the causal chains of the empirical world. Yet, scholars have recently explained how, in fact, Kant was deeply concerned with moral pedagogy and how such a pedagogy is a consistent, even essential, part of his moral theory. Some have gone so far as to argue that there is a pedagogical purpose at the core of Kant's entire critical philosophy. Again, the significance of moral development is evident in his correspondence with Herbert both in his sole reply and in his use of her example for encouraging the moral development of the daughter of a friend.

My interpretation of Kant's account of moral education is novel in at least two respects. The first is that it treats the cultivation of hope figures centrally in that account. While other commentators have articulated how, for Kant, the aim of moral education is the formation of a fully virtuous character, few have paid any attention to the role of hope in the formation of such a character.[10] But even those who do mention hope either do so in passing or lump hope together with seemingly related items. For example, Robert Louden has examined hope as one among many other facets in the role of moral exemplars in Kant's theory of moral education.[11] By contrast, in my interpretation, the problem for which Kant seeks to provide moral education as a solution, as well as the pedagogical method he develops, cannot be understood without recognizing that hope is a central concern in the formation of a fully virtuous character. Kant's

[10] Moran (2009 & 2012).
[11] Louden (1992).

correspondence with Maria von Herbert underscores this point that moral agency must be grounded on the right kind of hope, which can result from the right kind of moral education.

The second respect in which my interpretation of Kant's account of moral education is novel is that it provides a close reading of a text to which little attention has been previously given. This is the 'Doctrine of Method of Pure Practical Reason,' the penultimate chapter of Kant's *Critique of Practical Reason*. The 'Doctrine of Method' has often been ignored in discussions of Kant's moral philosophy, perhaps because it does not appear to contribute to the project of the second Critique. By contrast, in my interpretation, the Doctrine of Method is a crucial component of that project. Moreover, while some scholars have provided close analyses of the Doctrine of Method as it relates to Kant's second Critique or entire critical philosophy, there has not been a rigorous reconstruction of the moral pedagogical methodology he lays out in that chapter. For example, G. Felicitas Munzel has written extensively and insightfully about the Doctrine of Method, both in terms of its historical context and its theoretical significance in Kant's critical philosophy.[12] However, her analysis does not illuminate the specific model of pedagogy that I argue Kant sees it as important to deliver to his readers. Providing a reconstruction of that pedagogy – what I am calling a Kantian pedagogy of hope – is one of my main tasks in this dissertation.

Admittedly, the preceding discussion of Kant's correspondence with Maria von Herbert may overstate the promise and understate the limitations of a Kantian pedagogy of hope. For many readers, the problems of greatest concern in that correspondence have to do with interpersonal forgiveness and structural injustices, such as the oppression of women living in a "sexual marketplace" wherein they, as persons, have a "price" that is contingent upon their sexual history. With these problems at the forefront, some readers conclude that what Maria von

[12] Munzel (1999, 2002, & 2012). Also, Bacin (2010).

Herbert's letters represent are a critique of Kant's moral and political philosophy. As Elvira Basevich has argued, "Immanuel Kant does not countenance, much less redress, that the issues of gender and race *empirically* define modern republics and *normatively* constitute central obstacles to progress."[13] Ultimately, I do agree with this view that there are important limitations to any Kantian moral or political theory, specifically surrounding race and gender, and that some of those limitations are, indeed, highlighted by Kant's correspondence with Maria von Herbert. However, I also agree with those who argue that, while Maria does present a challenge to Kant's moral theory, their correspondence also invites us to rethink that theory in a way that overcomes that challenge. Most influential among these readings is Rae Langton's essay "Duty and Desolation," in which she articulates a model of Kantian ethics under oppressive conditions. As Langton points out, Kant's moral theory provides the conceptual resources for persons to understand and undo the domination contract, even if that theory, as stated, cannot completely redress such injustices.[14] Likewise, I believe that a Kantian pedagogy of hope invites us to rethink Kant's moral theory in terms of its promises and limitations in addressing certain injustices – specifically, the moral-psychological foundations for overcoming injustices.

While hope and education have not heretofore been jointly investigated as themes in Kant's philosophical writings, these themes have been amply explored by other practitioners and researchers in the field of education. "Education itself seems to be a hopeful endeavor," writes philosopher of education Sarah Stitzlein, "insofar as schools are focused on preparing for the future and aim to make that future better than the present."[15] Moreover, the vast popular and academic literature on teaching is peppered with proclamations "'Hope is the essence of

[13] Basevich (2022: 541).
[14] Langton (1992: 501-505).
[15] Stitzlein (2019: 81).

teaching,' 'To teach is to be full of hope,' and 'Teaching is . . . in every respect a profession of hope.'"[16]

Of course, in an age of test-based accountability that often overshadows educational aims beyond test scores, one might not expect school administrators and parents to view education as an essentially hopeful endeavor. But 83% of superintendents and 77% of parents apparently believe that getting children to have hope in the future is an indicator of school effectiveness. This particularly comes in the wake of Barack Obama's 2008 presidential campaign slogan of hope, which ushered in some schools being labeled "schools of hope" and others being celebrated for the hope they purported to offer in films like *Waiting for Superman* and *The Lottery*.

It is beyond the scope of this dissertation to meaningfully interject the resulting analysis of Kant's pedagogy of hope into contemporary scholarship on hope in education. But gesturing here at the possible fruits of the analysis helps to highlight what are my aims here. In particular, what Kant's pedagogy of hope potentially contributes to contemporary education scholarship is both the idea that hope is a fundamental part of moral agency and that not just any kind of hope will be adequate for developing, and sustaining, such agency, particularly in the face of hopelessness. These two ideas, which are focal points of this dissertation, can help to challenge contemporary views in which hope is valued merely for its apparent benefits in students' success as students and not necessarily as developing moral agents. In this way, my articulation and defense of Kant's pedagogy of hope can be understood as seeking to insert his educational thought into the tradition of progressive education that holds a more holistic view of the aims of education and of the value of hope in achieving those aims.

[16] Edgoose (2010).

The first two chapters of this dissertation provide some framing to guide and support my reconstruction of Kant's pedagogy of hope. In Chapter One, 'Rousseau, Kant, and the Problem of Education,' I critically examine aspects of Jean-Jacques Rousseau's influence on Kant's philosophical thought. Rousseau's own educational philosophy is developed in response to his diagnosis of the source of ills in human society, which he identifies as being amour propre, a kind of self-love that is traditionally interpreted to be as nothing but the source of vindictiveness, cruelty, and conflict. I draw on a recent, and influential, revisionist interpretation of Rousseau's theory of *amour propre* on which that desire is not in itself bad, but is bad only in its corrupted, or inflamed, varieties. On that revisionist interpretation, the goal of education, for Rousseau, then becomes to forestall the awakening of amour propre and cultivate its health, uncorrupted, development. This framing of education as a remedy developed in response to a diagnosis of the source of human evil will allow me to draw similarities and contrasts with Kant's own theory of moral education that will ultimately help to sharpen our understanding of the goals and methods of his pedagogy of hope. In particular, I argue that the Kantian analogue to *amour propre*, self-conceit, is best understood as posing a distinct kind of threat to moral development that ultimately creates divergences in Rousseau's and Kant's theories of moral education.

In Chapter Two, 'What is hope?' I try to make some progress towards articulating Kant's account of hope. Despite the centrality of hope in Kant's philosophical writings, reconstructing his conception of hope turns out to be challenging, given that he never explicitly defines or systematically develops it. My aim in this chapter is not to provide a comprehensive and systematic analysis of Kant's account of hope. Rather, my aim is to explicate some features of the kind of hope that is central to my interpretation of that part of Kant's theory of moral education which I am calling his 'pedagogy of hope.' To do so, I draw on the two main

approaches taken by those who have sought to reconstruct Kant's account of hope. One of these is to begin from the so-called 'standard account' of hope, according to which hope is a compound attitude consisting of a desire for an outcome and a belief that that outcome is possible but not certain. Another of the strategies I examine is one I mentioned earlier that seeks to analyze hope in terms of its object. As I briefly explained above, for Kant, the object of reasonable hope is the highest good. So, I dedicate a portion of this chapter to his remarks and arguments concerning the highest good, specifically in the second Critique, since that is the text to which I will turn my attention in reconstructing Kant's pedagogy of hope. Drawing on contemporary analyses of the standard account of hope and of Kant's notion of hope will help me to specify the object of hope as it pertains to Kant's moral pedagogical method, namely, the hope on the part of students that respect for the moral law will constitute one's character, and so continue to influence one's future actions.

In Chapter Three, 'Hope and the Doctrine of the Method of Pure Practical Reason,' I finally arrive at my close reading of the Doctrine of the Method in the *Critique of Practical Reason*. This is where I give my reconstruction of Kant's pedagogy of hope, that is, of the pedagogical method he prescribes to moral educators in the Doctrine of Method. I provide a novel interpretation of Kant's aim in that text, according to which what he seeks to accomplish in that chapter is to show how the feeling of respect for the moral law can have a lasting influence on our moral strivings. To, once again, guide and support my efforts, I draw on one other resource – namely, Kant's aesthetic theory. I argue that Kant's aesthetics – particularly, his accounts of the beautiful and the sublime – illuminates key features of the pedagogical method Kant develops in the Doctrine of Method. My core claim is that that method is designed to elicit

not just students' respect for the moral law, but also their hope to sustain that respect in the face of certain hardships.

In the concluding chapter, I return to Kant's correspondence with Maria von Herbert. With the preceding analysis of his pedagogy of hope in hand, we can better make sense of the rationale, and possible limitations, of Kant's attempt to indirectly address Maria's calls for hope in a life otherwise full of suffering. In particular, I show how, for Kant, such hope is not something that can be given by him in his capacity as a philosophical interlocutor, but rather that hope can only be expected from commitment to the doctrine he articulates throughout his writings on moral philosophy and through a discipline, or form of practice, that he outlines in his writings on moral pedagogy, one of which centrally being the Doctrine of Method in the *Critique of Practical Reason*.

Chapter 1: Rousseau, Kant, and the Problem of Education

My aim in this chapter is to provide one part in the framing for my reconstruction of Kant's pedagogy of hope. This framing concerns the influence of certain aspects of eighteenth-century French philosopher Jean-Jacques Rousseau's writings on Kant's account of moral education. I draw on recent interpretations of Rousseau's notion of *amour propre* to illuminate analogous aspects of Kant's moral psychology that inform his account of moral education. In short, just as Rousseau's notion of *amour propre* can be understood as the source of the problem for which education is a (partial) solution, Kant's analogous notion of self-conceit explains the central problem for which his pedagogy of hope is to be a solution.

In section one, I lay out a revisionist interpretation of Rousseau's notion of *amour propre*, on which the educational program laid out in *Emile* is to forestall the development of a corrupted *amour propre* and, instead, cultivate its healthy development. In section two, I articulate some of Rousseau's concrete pedagogical measures to achieve that end in *Emile*. In section three, I compare the revisionist interpretation of *amour propre* to Kant's notion of self-conceit, both of which figure centrally in the respective philosophers' diagnoses and pedagogical prescriptions to the sources of corruption in individuals and society.

1. Rousseau's diagnosis of *amour propre*

To understand the general problem for which Rousseau's *Emile* is a partial solution, it is useful to turn to his *Discourse on the Origin and Foundations of Inequality among Men* (hereafter, Second Discourse), where he provides the clearest diagnosis of that problem –

namely, the problem of *amour propre*. I begin my explanation of Rousseau's diagnosis of that problem in the Second Discourse with his distinct conception of the so-called 'state of nature.'

Rousseau understands the state of nature for the purposes of his Second Discourse somewhat differently from how Hobbes and Locke before him had construed it. For them, the state of nature was an analytic construct, rather than a true characterisation of humanity in an observable, empirical state.[17] The idea was to simply eliminate the state from our reflection on human relations and consider what human life would be like if we did not have the state and the law to govern us. They postulated that, if we simply subtract the state from the picture, we would be left with a state of nature that would be rife with inconveniences and insecurity. From there, they argue that individuals who found themselves in that situation would find it intolerable and would, in a variety of ways, agree to establish a state on certain terms, an argument they regard as defending and legitimating the distinctive powers of the modern sovereign state.

Rousseau regards that approach as naive. He thinks that Hobbes and Locke have not articulated the true state of nature, which he construes in more anthropological terms.[18] Rousseau invites us to imagine a more radical state of nature, one that represents the point of origin, as it were, in the development of the human species before any culture, any civilization, has taken root. He further invites us to imagine individuals who interact even before the establishment of any linguistic capacities or any linguistically mediated consciousness about who they are, what their lives mean, or anything of that kind. It is necessary to go back further in this way, he thinks, in order to reflect in a critically detached way on the gains and losses that the human species has

[17] For an in-depth analysis of the Hobbesian state as an analytical construct, as opposed to an empirical reality, see McClure (2014: 114–125). McClure and others argue that Hobbes' rationale for using this construct was connected to indoctrination into the authority of the state as much as explanation of that authority. On this view, the *Leviathan* is not just a theoretical text, but a pedagogical one, aimed at socializing the reader into the commonwealth. See Bejan (2010: 607–626).

[18] Part of Rousseau's criticism is that theorists like Hobbes and Locke mistakenly assume aspects of individuals, such as ideas of self-ownership, that would not exist in a state of nature independently of any processes of socialization.

incurred in the course of the civilizing process. The state of nature as Rousseau describes it is an almost idyllic natural state in which individuals live together in a state of innocence, a lack of self-consciousness, and in harmony with each other and with nature.[19] Individuals in that situation would have enjoyed, according to Rousseau, a kind of independence from each other. They are indeed activated by a rudimentary kind of self-love – *amour de soi*, about which I will say more later – centered around relatively satiable natural drives and instincts. That self-love is qualified only by a sense of natural sympathy, or pity, in the face of suffering experienced by fellow sentient beings, which Rousseau thinks prevents self-love from acquiring any kind of acquisitive or vindictive quality.

However, as the civilizing process grinds on, Rousseau postulates that something gets introduced, such that we learn dispositions that lead us, against our natural sentiment of pity, to be indifferent to the suffering of others and to contrive various forms of vindictiveness and even cruelty towards others. Rousseau's central question, then, is how do violence, predation, acquisitiveness, and related phenomena gain ground as the human species develops? How do we learn, or acquire, the various forms of vindictiveness and cruelty towards one another that we observe? By pursuing this question, Rousseau thinks we might acquire some critical distance on that tendency, recognize it for what it is, and become conscious of a tension between the world that we have created and the more innocent situation that he thinks is more proper to, and reflective of, our natural dispositions towards each other. So, whereas Hobbes and Locke conceived of the state of nature as an analytical construct in their efforts to defend and legitimize the distinctive powers of the modern sovereign state, Rousseau conceives of the state of nature as

[19] To be sure, Rousseau was not so naive as to think that the state of nature would be without any misfortune. Rather, the state of nature, as he conceives it, would be peaceful just in the sense that those misfortunes would not be seen as ones in which there is some intention behind the will. The reason for this will be clearer when I analyze Rousseau's notion of *amour propre*.

a diagnostic tool to critically reflect on the perils (but also, as we will see, the promise) of human existence.

Given this description of his project, many are inclined to compare Rousseau's depiction of the seemingly idyllic, primal state of nature of the so-called 'noble savage' to the traditional theistic story of the 'Fall of Man.'[20] In many ways, there is considerable overlap between the Christian account of that Fall, in which human existence gets poisoned by primal sin in the otherwise idyllic garden of Eden and that then condemns humans to divided and violent forms of life. In this traditional narrative, which also influenced Hobbes, there is a strong emphasis on the problem of vanity. This tendency has us living in the opinions of others, constantly worrying about controlling how we appear to others in order to achieve advancement in our social lives.[21]

There are also some important differences from the traditional Christian Fall of Man narrative and Rousseau's anthropological account in the Second Discourse. One difference is that there is no primal sin in Rousseau's account of the state of nature. His account of the origins of inequality does not condemn everybody to division because the noble savage or anyone else committed an original sin as it is described in traditional Christianity. Rather, for Rousseau, the sins, as it were, come later. The inequalities attributable to vanity arrive later in the course of human development, and the problem with these inequalities is that they represent, not disobedience before the will of God, but a kind of disobedience towards ourselves in relation to others. So, part of what Rousseau is trying to get across is the sense that we have betrayed ourselves in the way in which we have constructed systems of power and domination in the course of civilizing processes. Another difference is that Rousseau's account in the Second Discourse has a stadial theory of human development under which human civilization goes

[20] See, for example, Neuhouser (2008), who explicitly characterizes Rousseau's project as a theodicy.
[21] This emphasis on vanity and vainglory goes back at least into medieval christianity. For example, writing in the 13th century, Saint Thomas Aquinas listed envy, discord, and contention as the "daughters of vainglory.".

through various phases or stages. Though common currency among 18th century commentators, it is not obvious that a stadial theory corresponds to anything in the Christian theistic narrative of the Fall of Man. For my purposes, I will just mention that it is in third stage outlined by Rousseau, where humans start living together in villages, that you get the first stirrings of the kind of self-love – or *amour propre*, about which I will say more soon – that Rousseau thinks are going to get us into trouble down the line. Nonetheless, he thinks of this period as a kind of golden age. So, Rousseau's account in Second Discourse is not a straightforward story of decline; instead, there is improvement on the way towards our state of unfreedom, and that is something that he acknowledges in this third stage during which we were able to take advantage of some forms of productive mutual cooperation. As we will see later, these views of the delirious, though convoluted, effects of civilizing processes will shape the pedagogical measures Rousseau's develops as a (partial) solution.

So, Rousseau's central claim in his secularized version of the Fall of Man – that is, his diagnosis of the source of ills in human society – is that the the natural balance that we observe, or at least imagine, in the initial phase of human development between the self-interest of *amour de soi* and sympathy for the suffering of fellow sentient beings gets disrupted. We have become confused, Rousseau thinks, about the relationships between our concern for self and our concern for others, because *amour de soi* is replaced by *amour propre*. The distinction between these two kinds of self-love is key to Rousseau's diagnosis of, and remedies, to the ills of human society, both of which illuminate Kant's theory of moral education. So, I will now elaborate on Rousseau's crucial distinction between *amour de soi* and *amour propre*.

Let us begin with perhaps Rousseau's clearest and most explicit formulation of the two kinds of self-love:

> *Amour propre* and *Amour de soi-même*, two very different passions in their nature and their effects, should not be confused. Self-love [*amour de soi-même*] is a natural sentiment which inclines every animal to attend to its self-preservation and which, guided in man by reason and modified by pity, produces humanity and virtue. Amour-propre is only a relative sentiment, artificial, and born in society, which inclines every individual to set greater store by himself than by anyone else, inspires men with all the evils they do one another, and is the genuine source of honor.[22]

While Rousseau's theory of self-love is certainly subject to controversy, this passage relatively neatly lays out the basic contours. Most importantly, Rousseau defines *amour propre* in contrast with *amour de soi* by reference to three core features: *amour propre* is (1) "a relative sentiment" that is (2) "artificial, and born in society," and (3) "inclines every individual to set greater store by himself than by anyone else." I will briefly expand on what had until recently been a standard interpretation of these features before turning to an influential revisionist interpretation that has a Kantian twist.

First, while *amour de soi* is an "absolute sentiment" which "regards only ourselves," *amour propre* is a form of self-love that manifests in an individual's relation to other people.[23] For Rousseau, this relation to others occasions a self-relation, since to regard and to be regarded are psychologically intertwined. Put another way, in judging other individuals, I become aware of their ability to judge me, and this awareness stirs my interest in how I fare in their eyes. In the Second Discourse, Rousseau writes that, when humans initially emerged from their naturally solitary existence, "Everyone began to look at everyone else and to wish to be looked at himself, and public esteem acquired a price."[24] In *Emile*, he adds that "as soon as *amour propre* has developed, the relative *I* is constantly in play, and the young man never observes others without returning to himself and comparing himself with them."[25] So, as a relative sentiment, the

[22] DI 218.
[23] E 215.
[24] DI 166.
[25] E 243.

satisfaction of *amour propre* depends on one being aware of having the recognition of other people.

Second, what does Rousseau mean by describing *amour de soi* as natural and *amour propre* as "artificial, and born in society"? *Amour de soi*, as a drive for self-preservation, is manifested in responsiveness to an instinctual concern for one's physical well-being, expressed in behaviors like foraging for food in nature, seeking shelter from the element, defending against attacks and, on the level of the species, mating.[26] The goods sought by *amour de soi* change over time with the advancement of civilization, but the basic, pre-social needs remain the same. *Amour de soi* moves us to put our own interests above those of others in times of scarcity, but it never desires anything *in order to* deprive someone else of it. By contrast, *amour propre* is artificial in that it has little, if anything, to do with an instinct for self-preservation. The chief concern of *amour propre* is status through comparisons with other people, a concern that Rousseau considers a direct result of processes of socialization. This concern drives us to take more than we need in order to have more than others. So, for Rousseau, *amour de soi* is a kind of self-love on which natural goodness is founded, while *amour propre* is a kind of self-love that is the source of vindictiveness, cruelty, and conflict.

Third, as just alluded to, the relativity of *amour propre* is often regarded as inseparable from a competitiveness that, as Rousseau puts it in the quote from earlier, "inclines every individual to set greater store by himself than by anyone else." For Rousseau, once we find ourselves in a relation to others, we are driven to compare ourselves with them and, as a result, to desire to surpass them. As he writes of the imagined youth who is the namesake of *Emile*, "the first glance he casts on his fellows leads him to compare himself with them. And the first

[26] As I explain below, this description of *amour de soi* is somewhat standard, though alternative, revisionist readings have been advanced more recently. Without taking a stance on the merits of either reading, the revisionist one will be of greater use for my purposes.

sentiment aroused in him by this comparison is the desire to be in the first position."[27] This competitive impetus "to be in the first position" makes into prized possessions what had previously been superfluous advantages. For Rousseau's noble savage, being the best singer or dancer were not meaningful achievements. But when we began to judge one another, and so become aware of being judged by others, "the one who sang or danced best; the handsomest, the strongest, the most skillful, or the most eloquent came to be the most highly regarded, and this was the first step at once toward inequality and vice."[28] As soon as *amour propre* is activated, to be better than others becomes a goal in itself, overriding any other benefits to be sought after in singing or dancing well. Superiority itself becomes an object of desire and *amour propre* emerges, in the words of Judith Sklar, as a "desire for inequality."[29]

The above description of Rousseau's distinction between *amour de soi* and *amour propre* relays what had until recently been a standard interpretation. To summarize, on that interpretation, the two kinds of self-love are mutually exclusive in character: *amour de soi* is a pre-social concern for one's physical well-being, whereas *amour propre* is an excessive craving for superiority, motivating one to pursue a higher status over others on the social ladder, sometimes even at the expense of one's physical well-being. However, a revisionist interpretation of Rousseau's theory of *amour propre* has been increasingly adopted, particularly in Anglophone scholarship on Rousseau. This interpretation critically departs from the standard one in at least two key ways, which, as I will explain later, serve my purposes in comparing Rousseau's notion of *amour propre* with Kant's notion of self-conceit.

One key departure is that, on the revisionist interpretation, *amour de soi* and *amour propre* are not mutually exclusive. As Nicholas Dent argues, *amour de soi* is not a desire for just

[27] E 235.
[28] DI 166.
[29] Shklar (1969: 35).

physical well-being, but instead "signifies a concern, a care, to look to, guard, preserve and foster one's own personal wellbeing."[30] It identifies a basic psychological drive to pursue "the preservation, growth, amplitude, of secure, rich, being and life for a creature."[31] Put another way, *amour de soi* is a general concern for one's overall individual well-being, understood more broadly than mere physical preservation. While the specific goods sought by *amour de soi* may concern physical well-being, that concern could also extend to an individual's moral well-being as a "right-possessing member of society."[32] If the conception of the self that is the object of self-love extends beyond a purely, or perhaps narrowly defined, physical one, then *amour de soi* will extend beyond the cares of one's physical condition. Indefinitely many goods may be pursued under the guise of different self-conceptions. So long as the desired object is not sought on the basis of some comparison in order that one may gather esteem for oneself, it may yet be understood as a possible object of *amour de soi* on the revisition line.

Another key departure concerns the claim that *amour propre* is an altogether bad and unhealthy passion. Recall how, on the standard interpretation, *amour propre* is itself, in all of its manifestations, an excessive craving for superiority that leads to ills in human society. Proponents of the revisionist interpretation challenge this claim, arguing in various ways that *amour propre* is a benign desire to "have a position, to be a part, to count for something."[33] For example, John Rawls regards a "drive toward equality" as the appropriate manifestation of *amour propre*.[34] Similarly, Frederick Neuhouser argues that the desire for equal standing in relation to others is an expression of the more general good of having relative standing with

[30] Dent (1988: 20).
[31] Dent (1988: 74-5).
[32] Dent (1992: 33).
[33] E 160.
[34] Rawls (2000: 222).

respect to others.[35] And Joshua Cohen maintains that *amour propre* corresponds to two sorts of self-worth, one benign and one bad:

> A person with the first form – call it the egalitarian form – regards his own worth as equal to that of others …The second, inegalitarian form of self-regard is to think oneself more worthy of regard than others, and find it insulting if they reject this elevated conception – if they reject the view that one is more worthy of consideration and respect than they are themselves.[36]

Hence, proponents of this line of interpretation distinguish an original, "egalitarian" version of *amour propre* from its corrupted, or inflamed, versions, which are the source of ills in human society.[37] This raises the question of how it is that a good, or at least benign, sort of *amour propre* becomes inflamed, as well as what, if anything, can be done to avoid that inflammation or mitigate its harms.

To recap this section, Rousseau's Second Discourse presents the clearest articulation of his secularized version of the Fall of Man – that is, his diagnosis of the source of ills in human society. On his diagnosis, the principal source of those ills is *amour propre*, a kind of self-love that manifests in a desire for recognition from others. However, on a revisionist interpretation of *amour propre*, that desire is not bad in all of its forms. Indeed, on that interpretation, the goal of the educational program in Rousseau's *Emile* is to forestall the degeneration of *amour de soi* into an inflamed *amour propre* and, instead, develop the latter in such a way that its redemptive features flourish.

2. Rousseau's (partial) remedy in education

[35] Neuhouser (2008: 32-3).
[36] Cohen (2010: 102-3).
[37] The idea of an inflamed *amour propre* comes from Rousseau (see E 247). But the concept, as well as the general revisionist approach, was popularized following Nicholas Dent's examination of it in Dent (1989, 56-58). For a detailed description of the variety of inflamed *amour propre*, see Neuhouser (2008).

Rousseau tells us that his principal works – his first two discourses, the *Social Contract*, and *Emile* – should be evaluated as a coherent system.[38] Following this suggestion, some commentators view the first two discourses as presenting the problems of a society that enshrines the inflammation of *amour propre*, while the *Social Contract* presents the social solution, and *Emile* its domestic solution.[39] While Book V of *Emile* incorporates a political education at the beginnings of adulthood, the bulk of that work focuses on an education that would preserve and cultivate the youth's original good nature until the child becomes a free, morally responsible adult. To that end, Rousseau describes how the tutor, named Jean-Jacques, forestalls the development of an inflamed *amour propre* in the youth after whom the novel is named. In this section, I will articulate some of the concrete pedagogical measures outlined in Rousseau's *Emile* that make up his partial solution to the problems (and promise) of *amour propre*.

To contextualize Rousseau's pedagogical measures, it is useful to note that his views on the malleability of human psychology were not unique. For example, John Locke's widely read *Some Thoughts Concerning Education*, which is commonly understood to have had a great influence on Rousseau's writing of *Emile*, expresses a view on which a stable moral character comes from habituation:

> Every man must sometime or other be trusted to himself and his own conduct; and he that is a good, a virtuous, and able man, must be made so within. And therefore what he is to receive from education, what is to sway and influence his life, must be something put into him betimes; habits woven into the very principles of his

[38] See, for example, Rousseau's *Dialogues*, pp. 209-212. However, some prominent commentators (e.g. Shklar 1969: 5) do not approach Rousseau's writings as forming a coherent system, either because they maintain that no such coherence can be found or because they assume that was not Rousseau's intention. For my purposes, I am setting aside debates about how Rousseau's writings relate to one another and following those who treat the *Social Contract* and *Emile* as offering remedies for the problems Rousseau diagnosis in the Second Discourse.

[39] This distinction between Rousseau's "social" and "domestic" remedies for the problem of *amour propre* is taken from Neuhouser (2008). Since my focus is on the latter insofar as it serves to illuminate Kant's theory of moral education, I will have little to nothing to say about the social remedy Rousseau offers in the *Social Contract*, which focuses on the social and political institutions necessary to foster virtue in citizens.

> nature; not a counterfeit carriage, and dissembled outside, put on by fear, only to avoid the present anger of a father, who perhaps may disinherit him.[40]

Education, in Locke's view, could work upon the soul of every child similar to how God could work upon the soul of man wholly prepared for grace. Our one innate tendency – namely, the tendency to avoid pain and seek pleasure – makes it possible for educators to inculcate habits by means of a system of rewards and punishments. This view was more well received in France than in England. Étienne Bonnot de Condillac and Claude Adrien Helvétius, for instance, developed a more extreme version of Locke's view. In the *Traite des Sensations*, Condillac proposes that all powers of the mind are habits. And in *De I'Homme*, Helvetius argues that the quality of the mind is entirely a result of education, in direct opposition to the view that it was instead the product of physical or psychological constitution. Hence, Helvetius' celebrated line, "Education can do all."[41]

Rousseau's belief in the possibility, and limits, of education to shape our psychology – particularly, *amour propre* – should be examined at this juncture. Although not entirely identical, his ideas remained closer to Locke's than to the more extreme views of Condillac and Helvetius. Unlike those Frenchmen, Rousseau always believed that there were at least some potentialities natural to man: the ability to reason and make judgments; the ability to feel pity at the sight of suffering; the capacity for self-perfection; and a sense of obligation to realize these potentialities. But he thinks that these natural capacities were, to some degree, subject to change. Indeed, as I mentioned above, the stated goal of the educational program Rousseau lays out in *Emile* is to forestall the degeneration of *amour de soi* into *amour propre*, in order that one's natural goodness can flourish. From the Second Discourse, we know that the moment that saw the birth of *amour propre* was when opinion became a dominant psychological force in human life. One

[40] Locke, *Some Thoughts Concerning Education*, §42
[41] Helvétius, Claude Adrien. 1773. *De L'Homme, de ses Facultés Intellectuelles et de son Education*, Londres: 223.

key element of Rousseau's educational program, then, is that human beings become independent, so that, when they enter society, they do not have to depend on others' opinions in ways that stifle their natural goodness.

What are the particular pedagogical measures Rousseau prescribes for cultivating independent judgment in order to tame *amour propre*? His description of Emile's educational program is winding and seemingly without any clear shape. However, Neuhouser helpfully divides that program into three phases.[42] The first of these encompasses Books I-III of *Emile*, the second Book VI, and the third Book V. For my purposes, I will focus on just the first two phases.

The first phase of young Emile's education is to take place before the emergence of *amour propre* in his adolescence, with the aim of prolonging its emergence. To appreciate this phase, it will be helpful to return to the question of how *amour propre* becomes inflamed, since its early stirrings in Emile, shape the measures Rousseau believes ought to be taken by his tutor. According to Rousseau, its occurrence arises as follows: "the first condition of man is want and weakness," a child's tears are an expression of felt needs which he cannot satisfy independently. It is "from these tears that we might think so little worthy of attention" that "is born man's first relation to all that surrounds him; here is formed the first link in that long chain of which the social order is formed."[43] Rousseau goes on to say:

> The first tears of children are prayers. If one is not careful, they soon become orders. Children begin by getting themselves assisted, they end by getting themselves served. Thus, from their own weakness, which is in the first place the source of the feeling of their dependence, is subsequently born the idea of empire and domination. But since this idea is excited less by their needs than by our services, at this point moral effects whose immediate cause is not in nature begin to make their appearance, and one sees already why it is important from the earliest age to disentangle the secret intention which dictates the gesture or the scream.[44]

[42] Neuhouser (2008: 171-183).
[43] E 65.
[44] E 66.

One important point here is that a desire for domination, or an inflamed *amour propre*, emerges from how those tasked with raising children react to the latter's experiences of dependency. Rousseau tells of how an infant, in his first attempts to grasp an object beyond his reach, does so without tears, since he does not yet anticipate the frustration of his coming failure. The infant does not yet judge the distance as something he cannot overcome on his own, though we know it to be inaccessible without assistance. Once the infant, however inchoately, becomes aware of this, his tears follow, representing that he is commanding either the object to come to him or those surrounding him to bring it to him.

According to Rousseau, if those around the child do not refuse the command that his tears represent, this will, over time, establish in him the view that other persons are but instruments to satisfy his desires, and and it is from this view that children "become difficult, tyrannical, imperious, wicked, unmanageable – a development which does not come from a natural spirit of domination but which rather gives one to them." The experience of "how pleasant it is to act with the hands of others, and to need only to stir one's tongue to make the universe move," transforms the child's otherwise benign desire for recognition into a desire for domination over others.[45] Moreover, that desire is at risk of becoming further inflamed, since, at some point, the caregiver will have to decline some demand issued by the child. The child will then see another will as not merely an instrument, but as a source of resistance to the satisfaction of his desires. For, Rousseau writes, "if he is obeyed, as soon as something resists him, he sees in it a rebellion, an intention to resist him."[46] The harm is no longer experienced by the child like the harm experienced by Rousseau's "noble savage," for whom such harm is merely mechanical, without

[45] E 68.
[46] E 213.

any intention behind the will that can be a source of blame or contempt.[47] Rather, the awareness of an intention of harm, even if just an act of denial, is what incites us to the vindictive passions that take root and manifest in corrupted versions of *amour propre*. This is at least one reason why, once *amour propre* arises, it is almost natural for it to take a turn towards cruelty and conflict.

The first phase of Emile's educational program, then, consists of carefully orchestrated episodes so as to encourage that he learn, play, and work, while keeping his *amour propre* at bay. In one of these episodes, Emile and his tutor encounter a magician at a local fair.[48] In attempting to learn and reveal how he does his tricks, the pair embarrasses the magician in front of others, making him lose face, while Emile accidentally tastes public applause. After that encounter, Emile begins to recognize that learning is no longer pertinent to only his personal well being, or *amour soi*, and so "the first movement of vanity" appears to threaten to transform his curiosity into an inflamed *amour propre*: "If you know thus how to make humiliation and disgrace arise from it, be sure that a second movement will not come for a long time."[49] Later, however, the magician comes to their door to scold them, saying that in their quest to gain knowledge they have humiliated him and endangered his livelihood. The scene ends with Emile and his tutor returning to the fair and watching the trick performed once again, though this time with a strong sense of shame. Where before Emile may have been eager to show off his knowledge, that drive for vanity is now quieted, and he watches in silence. He ends up learning something about the magnets that made the trick possible, but, more importantly, he learns about the value in preserving others' dignity through proper sensitivity and respect. At the same time, Emile is taught that he can maintain his own dignity through independence from others' judging gaze.

[47] See footnote 3.
[48] E 173-5.
[49] E 175.

Consider another episode in this first phase of Emile's education. Though Emile comes from a background of privilege, his tutor instructs him to learn a trade – specifically, carpentry. The goal of this is that no matter what he will be able to take care of himself and will not be dependent upon the beneficence of others. Rousseau does not want Emile to be treated like a gentleman who is dabbling in a hobby, but like a true apprentice carpenter. His honor is not one of aristocracy but of usefulness and self-sufficiency. Locke also encouraged pupils to become trained in some trade, but Rousseau disagrees about his choice of trade. Rousseau writes, "I do not want him to be an embroiderer, a guilder, or a varnisher, like Locke's gentlemen."[50] While the wealthy would be aghast at the idea of their sons learning any trade, considering such work dishonorable, Rousseau says "I want to give him a rank which he cannot lose, a rank which does him honor at all times."[51] He will not be trapped in a role where he must pay constant attention to the opinions of others. He is invited to take up a job that relies, instead, on his solid understanding and which can be useful anywhere. With it, Emile will be in a better position to free himself for corrupting influences of wealth and society. Rousseau differs profoundly with Locke for whom the ever-expanding fruits of civilization are always preferable to a limited life of solitude. While Rousseau praises the life of one man who is able to subsist on his own efforts and who looks with suspicion upon the division of labor under which "a hundred men working in harmony will earn enough to give subsistence to two hundred."[52] Locke thought that such a solitary life was practically unlivable and certainly undesirable and declared the division of labor essential to human fulfillment.[53]

[50] E 197.
[51] E 196.
[52] E 185.
[53] Doubt about the practicality of the sort of solitary life Rousseau describes will reemerge when I turn to Kant's own pedagogical method.

In each of these two episodes drawn from the first phase of Emile's educational program, we see the development of Emile's capacity to judge independently of others' opinions. That training follows a trajectory in which Emile's faculties of reason and of sentiment are to work on one another. On the one hand, the tutor claims that "In developing [Emile's] nature, we have sidetracked its nascent sensibility; we have regulated it by cultivating reason."[54] He also stresses the importance of letting reason, or the "correctness of ideas" "to illuminate" or "order" all the affections of the soul. On the other hand, the tutor says that he wants to "perfect reason through sentiment," and that "by reason alone...no natural law can be established." Neuhouser puts this chief task of the first phase as follows:

> The most important result…is that before the passions of adolescence force him to give weight to the opinions of others, Emile will have developed a substantial reservoir of self-esteem as well as the capacity to evaluate himself according to transparent, non-fluctuating standards and independently of others' judgments.[55]

We may conclude, then, that cultivating Emile's independence, keeping his *amour propre* dormant, requires that he develop a capacity for autonomous judgment (including judgment about what is good) through a balance between his faculties of reason and sentiment. Among those sentiments is a sort of confidence in his own ability to judge in a way that is, if not certain, at least more stable than the opinions of others.

The second phase of Emile's education encompasses Book VI, with the full awakening of *amour propre* in his adolescence. As with the first phase, it is useful here to consider how *amour propre* is at risk of becoming inflamed at this stage of Emile's development, since this, once again, shapes what pedagogical measures are advised. That can be traced to some time after the first appearance of domination in the child's tears when *amour propre* inevitably stirs with the maturation of sexuality in adolescence:

---

[54] E 314
[55] Neuhouser (2008: 174).

> [L]ove must be reciprocal. To be loved, one has to make oneself lovable. To be preferred, one has to make oneself more lovable than another…at least in the eyes of the beloved object. This is the source of the first glances at one's fellows…the first comparisons with them…the source of emulation, rivalries, and jealousy.[56]

Due to the natural desire to be loved, the adolescent will feel a need to tailor himself based on those criteria set by social convention. Emile will eventually have to pursue esteem from others to be successful in courtship, as "he who senses how sweet it is to be loved would want to be loved by everyone; and all could not want preference without there being many malcontents." Once preference emerges, the desire for love soon becomes the competition for preference between competitors. Because sexuality is something that "whatever we may do…will be born in spite of us," the genesis of *amour propre* is inevitable.[57] However, unlike the process in the Second Discourse in which *amour propre* arises from emerging economic inequalities, the situation in *Emile* focuses on how that passion is not eliminated, but rather prevented from finding expression in social dimensions that inflame it.

While the details of the second phase of Emile's education are complex, the basic strategy is to shape his *amour propre* through the development of his imagination and the sentiment of pity. For Rousseau, pity is the only moral sentiment that has characteristics of both *amour de soi* and *amour propre*, since it involves the act of putting oneself in the place of another. Though pity shares some features of *amour propre*, pity differs in that it depends on the imaginary transportation of the self to a less fortunate being, the result of which is a sweet sentiment. Pity is sweet, because, when we put ourselves in the place of someone who suffers, we are aware of the pleasure of not suffering. Pity seems to exempt us from the pains he suffers.

> In fact how do we let ourselves be moved by pity if not by transporting ourselves outside of ourselves and identifying with the suffering animal by leaving, as it were, our own being to take on his being? We suffer only so much as we judge

[56] E 214.
[57] E 215.

> that it suffers. It is not in ourselves, it is in him that we suffer. Thus, no one becomes sensitive until his imagination is animated and begins to transport him out of himself.[58]

The imagination that stirs pity also restrains that sentiment by holding the suffering of the other – indeed, the other themselves – at a certain distance. Due to the nature of pity and its cultivation, Emile is able to learn to relate to others in a way that does not inflame his *amour propre*.

The cultivation of pity can forestall the degeneration of *amour de soi* into *amour propre* because, in doing so, the tutor is shaping Emile's faculty of imagination, which is responsible for the emergence of *amour propre*. In the Second Discourse, Rousseau claims that pity in civilized men is developed but weak, whereas it is obscure but lively in the case of the so-called "noble savage." For the former possesses imagination, but the emergence of *amour propre* weakens the sentiment of pity. And in the savage man pity is lively, since it arises immediately from a physical sensation, namely, the simple observation of the suffering of another living creature. Rousseau himself points to this difficulty in *Emile* when he describes the first sensations of pity: "If he had remained stupid and barbaric, he would not have them; if he were more learned, he would know their source. He has already compared too many ideas to feel nothing and not enough to have a conception of what he feels."[59] So, part of the challenge the tutor faces is how to make sure that the pity Emile feels is both developed and lively, instead of obscure and weak, without at the same time stimulating the passions which force him to make comparisons that inflame *amour propre*.

In order that Emile will not feel those selfish passions and thereby appropriately cultivate his imagination, he needs to exercise the right kind of judgment. We can see how this is so from Rousseau's most substantive pedagogical measure in this second phase of Emile's education:

---

[58] E 223.
[59] E 222.

reflection on examples from reading historical works, such as those of Thucydides and Plutarch. Rousseau believes that Emile can see the heart of men – the knowledge of which is important for Emile to judge himself as well as other human beings – without any judgment on the part of the historians.[60] Reflecting on these historical works shows Emile persons in times and places other than the one in which he finds himself. This is important so that Emile remains a simple spectator – disinterested and without passion – a judge instead of an ally or accuser. Even though he is making a comparison to judge that the historical figures are like him, he will neither be envious of nor despise them. Hence, in shaping Emile's sentiment of pity, the tutor is habituating him to a mode of judgment that does not involve any harmful passions.

The tutor proceeds through stages in developing Emile's sentiment of pity. First, the tutor waits until Emile's basic faculties of reason and judgment are first developed in the first phase of the educational program. Then, the tutor directs Emile's gaze to the poor, the sick, the oppressed, and the unfortunate, in order to help him recognize the similarity and vulnerability of all persons. After that, the tutor uses examples drawn from history to show how even the more fortunate can fall victim to corruption under their own ambitions. And finally, through fables, he wants Emile to know that he himself is a potential sufferer. These moments move progressively, such that the ideas and knowledge that Emile acquires will help him finally develop the concept of equality in humanity. The increasing complexity of ideas at work suggests that, as Emile's pity becomes more and more cultivated, his reason and judgment, which allow him the ability to grasp those ideas, also become more advanced. This progress is necessary, not only so Emile can develop the idea of equality, but also because he will encounter more complicated problems and thus greater risk of inflaming *amour propre* as he matures. In those circumstances, he needs yet a greater ability to exercise sound judgment.

[60] E 239.

So, in this second phase of Emile's education, the tutor is trying, through the presentation of examples, to let sentiments perform their task in sharpening his judgment. In letting Emile's sentiments perfect his intellectual faculties, the tutor perhaps finds a more effective means to counter the stirrings of *amour propre* than by providing step-by-step instructions. Taking into consideration that the overall goal of Emile is to help youth attain independence, part of the significance of Book IV lies in how it makes Emile's independence not just a matter of accomplishment in his powers of reasoning, as it is in Books I-III, but also of achievement in his faculty of sentiment. Emile's independence is not complete, that is, if his reason is not perfected by his sentiments. So, not only did Rousseau believe that the nature of given sentiments or passions can change, but that sentiments such as pity could shape the intellectual faculties, which subsequently shapes *amour propre* so that he can judge autonomously.

To recap, in this section I explained some of the pedagogical measures of Rousseau's educational program designed to forestall the degeneration of *amour de soi* into an inflamed *amour propre* and, instead, develop *amour propre* in such as way that its redemptive features flourish. The pedagogical measures are introduced in two phases: the first consists of carefully orchestrated episodes designed to encourage Emile's confidence in his own ability to judge independently of others' opinions; the second consists exercising Emile's imagination through the use of examples drawn from history that stir his sentiment of pity and shape his *amour propre*. While this is not the end of Emile's education, these two phases will be important to keep in mind when I turn to Kant's own pedagogical method in the next chapter. I will end this chapter with an examination of the source of the problem for which Kant offers his pedagogical method as a solution. As I will show, the source of that problem is similar to, but also different from, Rousseau's notion of *amour propre*. Seeing how central that notion is to the shape of Rousseau's

educational program, it seems reasonable to expect what Kant views as the source of evil will likewise be an important factor in his moral pedagogical method.

3. *Amour propre* and Kantian self-conceit

In the last two sections, I explained the role of *amour propre* in Rousseau's diagnosis of, and remedies for, the ills of society. For Rousseau, inflamed *amour propre*, or a corrupted desire for recognition, is the source of those ills; and his "domestic remedy" found in the educational program of *Emile* chiefly aims to forestall the emergence of *amour propre* and cultivate its healthy development. In this section, I will turn to Kant's notion of self-conceit, which many commentators have regarded as similar to Rousseau's notion of *amour propre*. Examining Kant's notion of self-conceit in light of that comparison and the foregoing analysis of *amour propre* will help illuminate the problem for which Kant's designs his pedagogical method as a solution.

Let's begin with why some commentators are drawn to compare *amour propre* and self-conceit. Consider Kant's discussion in *Religion within the Boundaries of Mere Reason* in which he distinguishes three predispositions to the good. First, there is the "predisposition to animality," which involves "physical or *merely* mechanical self-love," a kind of self-love that does not involve reason.[61] This predisposition consists, roughly, in our basic drives tending toward not only self-preservation, but also species-preservation through sexual reproduction. Second is the "predisposition to humanity," which Kant says "can be brought under the general title of self-love which is physical and yet *involves comparison* (for which reason is required); that is, only in comparison with others does one judge oneself happy or unhappy." This second kind of self-love is a desire "to gain worth in the opinion of others, originally, of course, merely *equal worth*: not allowing anyone superiority over oneself, bound up with the constant anxiety

[61] R 6:26

that others might be striving for ascendancy." However, Kant goes on to say that this comparative desire that tends toward equality becomes corrupted, turning to a desire to gain superiority for ourselves, which leads to our vices: "e.g. in *envy*, *ingratitude*, *joy in others' misfortunes*, etc., they are called *diabolical vices*." The third predisposition is one to personality, which is the "susceptibility to respect for the moral law *as of itself a sufficient incentive to the power of choice*."[62]

The two variants of self-love Kant attributes to the predispositions to animality and humanity are reminiscent of Rousseau's descriptions of *amour de soi* and *amour propre*. Like *amour de soi*, the "merely mechanical" self-love in the predisposition to animality prompts human beings to care for the satisfaction of their basic needs. And like *amour propre*, a self-love "that involves comparison" in the predisposition to humanity implies a mediated and calculative mode of relation in which an agent's sense of worth depends upon the opinion of others. Conflict is inevitable here, for what starts as a drive for "*equal worth*, not allowing anyone superiority over oneself," is soon "bound up with the constant anxiety that others might be striving for ascendency," and hence triggers the "unjust desire to acquire superiority for oneself over others."[63] These are the comparative tendencies Kant identifies in *Idea for a Universal History* with "the *unsociable sociability* of human beings," that is, "their propensity to enter into society, which, however, is combined with a thoroughgoing resistance that constantly threatens to break up this society."[64] As he explains in that text:

> The human being has an inclination to become *socialized*, since in such a condition he feels himself as more. But he also has a great propensity to *individualize* (isolate) himself, because he simultaneously encounters in himself the unsociable property of willing to direct everything so as to get his own way.[65]

[62] R 6:27.
[63] Ibid.
[64] IaG 8:20
[65] IaG 8:20-21.

Like Rousseau, Kant connects these competitive tendencies with tendencies that develop in processes of socialization, thinking that on them "can be grafted all sorts of vices," though they "do not of themselves issue from this predisposition at its root."[66] Instead, these vices are observable manifestations of the propensity to evil ("a *radical* innate *evil*") to which all human beings are subjected, according to Kant.[67]

These remarks by Kant on self-love bring us closer to his notion of self-conceit. The "vices that are grafted on" the comparative form of our self-love are not necessarily violations of the moral law. After all, for Kant, though jealousy, envy, cruelty, and the like may be vices, that does not yet make them motives for our actions. I can be jealous of some person who possesses something I desire without acting on that jealousy; the moral law can be an incentive for my acting from duty, contrary to my jealousy. So, for Kant, it is only when we are led to think that self-love overrides the moral law as an incentive for action that it becomes self-conceit. To explain further, then, what self-conceit is, I need to briefly explain what it means for the moral law or self-love to be an incentive before turning to how one can be overridden by the other.

An incentive is something which moves a person to act. It is, as Kant defines it, "the subjective determining ground of the will of a being whose reason does not by its nature necessarily conform with the objective law."[68] In other words, it is something which moves a person to act. Many of the incentives we experience come from what Kant calls our "sensibility." We are sensible in that we are physical beings who can sense or be affected by the world. The desires we form on the basis of how our sensibility is affected are what Kant calls our "inclinations."[69] I am inclined to act in a certain way to the extent that my desire to act that way

[66] R 6:26.
[67] R 6:32.
[68] KpV 5:72.
[69] GMS 4:414n.

is based on my expectation that it will bring me pleasure or the avoidance of pain.[70] The thought that a certain action would satisfy an inclination is often an incentive for human action.

For Kant, the things that we morally ought to do are determined not by our sensibility, but by our nature as rational beings. Nonetheless, these two aspects of ourselves – our sensibility and our rationality – are both essential for the moral law to be an incentive for us. To understand why, it is helpful to consider Kant's remarks on what he calls the "dear self," which is opposed to what he elsewhere, in the *Groundwork of the Metaphysics of Morals*, calls the "proper self."[71] Beatrice Longuesse helpfully delineates these terms as follows: the "proper self" refers to the subject of the proposition 'I ought to X' when, and only when, 'ought' is determined under the categorical imperative; whereas the "dear self" refers to the subject of the proposition 'I ought to X' when 'ought' is determined, not under the categorical imperative, but under prudential or technical rules dictated by our sensible impulses or inclinations.[72] Now, the subject of any maxim must, on Kant's view, refer to what, in the *Critique of Practical Reason*, he calls the "entire self," which is both sensible and intellectual.[73] For, as Longuenesse explains, "the ends we set for our actions depend both on the ways we seek or avoid the objects affecting us with pleasure or pain, and on the command of duty."[74] Hence, the categorical imperative is legitimate – that is, the moral law can be an incentive for us – only insofar as "we regard ourselves as belonging to the world of sense and yet *at the same time* to the world of understanding."[75]

But a problem, as Kant sees it, is that the dear self, "even though by its maxims it is entirely unsuitable for universal legislation, nonetheless endeavors – just as if it amounts to our

[70] For a careful analysis of Kant's hedonism, see Herman 2008. Herman argues that Kant accepted a sophisticated form of hedonism that is worth taking seriously and that his larger aim is to show that without objective, moral values grounding our incentives, "hedonism is the true theory of motivation and choice" (Herman 2008: 177).
[71] GMS 5:407, 5:457.
[72] Longuenesse (2013: 93-94).
[73] KpV 5:74.
[74] Longuenesse (2013: 94).
[75] GMS 4:453, emphasis added.

entire self – to validate its claims beforehand and as primary and original."[76] Against this obstacle to our awareness of our proper self, or moral being, one of Kant's chief aims in the second Critique is to argue that the moral law can be an incentive. The moral law becomes an incentive, he explains, by calling up a moral feeling he calls of respect. This finally brings us to the topic of self-conceit. For it is through the moral law "striking down" the self-conceit of the dear self that we become aware of the proper self *as* our proper self and to feel respect for the moral law.

Kant's most explicit and comprehensive discussion of self-conceit comes in the Third Chapter of the *Critique of Practical Reason*. In the following passage, Kant explains that self-conceit is one of two types of "self-regard" that are made up by our inclinations:

> All inclinations together (which presumably can also be brought into a tolerable system, and the satisfaction of which is then called one's own happiness) amount to *selfishness* (*solipsismus*). This is either that of *self-love*, a paramount *benevolence* toward oneself (*philautia*), or that of liking for oneself (*arrogantia*). Specifically, the former is called *love for oneself*, the latter *conceit for oneself*. Pure practical reason merely *impairs* love for oneself, inasmuch as it only restricts such love—as natural and as astir in us even prior to the moral law—to the condition of agreement with this law, and this love is then called *rational self-love*. Self-conceit, however, pure practical reason *strikes down* altogether, inasmuch as all claims of self-esteem that precede agreement with the moral law are null and without any authority, since precisely the certainty of an attitude agreeing with this law is the primary condition of all worth of a person (we shall soon make this more distinct), and all presumption prior to this certainty is false and unlawful. … Hence inasmuch as—in contrast to its subjective opposite, viz., the inclinations in us—it *weakens* self-conceit, the moral law is at the same time an object of *respect*; and inasmuch as it even *strikes down* self-conceit, i.e., humbles it, the moral law is an object of the greatest respect and thus also the basis of a positive feeling that is not of empirical origin and is cognized a priori. Therefore respect for the moral law is a feeling that is brought about by an intellectual basis, and this feeling is the only one that we cognize completely a priori and the necessity of which we can have insight into.[77]

[76] KpV 5:74
[77] KpV 5:73.

In other words, self-conceit is a kind of self-love that, though a hindrance to morality, is central to our moral experience, insofar as we come to feel respect for the moral law as an incentive only when it "strikes down" self-conceit. Now, I want to set aside difficulties surrounding Kant's claim that the moral law strikes down self-conceit.[78] My interest here is in whether self-conceit is the same thing as *amour propre*, as often claimed. So, I will examine one common interpretation of Kant's notion of self-conceit on which it appears strikingly similar to *amour propre* before considering an alternative interpretation.

Many commentators interpret self-conceit as a kind of interpersonal arrogance, a tendency to erroneously understand ourselves as having some rank above all others, leading us to think that only our own needs and desires matter.[79] Self-conceit, on this line, is an unconditional satisfaction with oneself – that is, seeing oneself as superior to others – out of the implicit belief that one's own self-love has the authority of a universal law and that one must, in their own eyes and those of others, be seen as having a higher value in comparison. The agent's orientation towards others is precisely what is morally objectionable, and erroneous, in this view. Their error is thinking that others owe them too much respect or esteem and that they owe others too little of either.[80]

---

[78] Some of the philosophical and interpretative challenges surrounding Kant's claim that respect for the moral law is issued from its "striking down" self-conceit are discussed in Reath (1989).

[79] Stephen Engstrom, Allen Wood, and Andrews Reath are three examples, whom I will discuss below. But even some who do not think that self-conceit is directly rooted in an inegalitarian claim of one's worth in relation to others nonetheless agree that self-conceit must involve a desire for some form of recognition, such as esteem for oneself that implicitly ranks oneself above others. For example, Frederick Neuhouser (2008) affirms this sort of interpretation when he writes that people often misinterpret Rousseau's notion of *amour propre* because they take it to be the same as Kant's notion of self-conceit. As Neuhouser understands it, self-conceit "can be defined as taking one's own happiness—rather than the moral law, which also takes the happiness of others into account—as the supreme criterion for action [and] since this is in effect to regard oneself as worthier of happiness than others, it can plausibly be viewed as embodying an inflated sense of self-worth)" (Neuhouser 2008: 61). The problem in projecting this conception onto Rousseau, according to Neuhouser, is that it "overlooks the distinction made sometimes between respect and esteem" (ibid.).

[80] Again, this may overlook a distinction between respect and esteem, as argued for by Neuhouser. For more analysis on such a distinction, see Darwall (1977: 36-49).

One proponent of this line of interpretation is Stephen Engstrom, according to whom self-conceit arises when we take the mere fact that we have a conception of happiness as confirmation that we, or our moral claims, are superior to others. As he explains it, in self-conceit, we come to think of ourselves "as worthy of esteem on account of [our] cognitive power… to determine an end" and form conceptions of the good.[81] But being esteemed requires a point of comparison against which that which is esteemed appears favorable. But the fact that one has the power to set ends does not readily suggest some such point of comparison. So, Engstrom concludes, "[s]elf-conceit's esteem for self can therefore only be indirect, through depreciatory judgments concerning others."[82] That is to say, the self-satisfaction in self-conceit is based on a belief that the exercise, or mere existence, of our capacities proves others' capacities to be inferior. In this way, self-conceit must be a kind of interpersonal arrogance that assumes superiority over others.

Engstrom is not alone in thinking that self-conceit must have something to do with an arrogance in which we judge that we are better, or more deserving of happiness, than other people. Allen Wood, for instance, argues that, for Kant, human beings first learn to esteem themselves on the basis of their superiority over non-rational animals. This leads them to want to use similar grounds to claim superiority over other human beings, and this illicit claim is what Kant means by self-conceit.[83] Andrews Reath is another commentator who sees self-conceit as a kind of interpersonal arrogance, describing it as "a disposition to assign oneself a standing to treat oneself and one's subjective concerns as objective reasons that one does not and cannot

[81] Engstrom (2010: 108).
[82] Engstrom (2010: 109).
[83] Wood (1999: 241).

acknowledge in others…to accord oneself a special standing to make claims on one's own behalf in virtue of one's superior worth."[84]

There are at least two reasons why, on its face, this reading of self-conceit is attractive. One is that it sounds similar to what we usually mean when we call someone self-conceited. Another reason, however, is that it frames Kant's moral theory in a way that is more amenable to certain trends in contemporary moral thought. Kant's remarks on the immoral person as someone who mistakenly prioritizes sensibility over rationality seems to miss the apparent fact that morality importantly has to do with how we relate to others. Understanding self-conceit as a sort of interpersonal arrogance helps with this problem by framing the overvaluing of sensibility as just a way of valuing oneself – or one's "dear self" – over others. Since it is much easier for us to see how the person who demands that others bend to their will is immoral, this interpretation seems to bring Kant's theory closer to one seemingly common conception of morality.

But there are some challenges to this way of understanding Kant's notion of self-conceit. One has to do with how, as I mentioned earlier, Kant takes the error of self-conceit itself to be the root of immoral action. Because self-conceit is involved in the stirrings of the feeling of respect, and the moral law should be at least capable of calling up respect in the face of any temptation to act contrary to duty, self-conceit – or, more precisely, the error behind it – must be at issue in any case in which immoral actions seem attractive to human beings. But not all immoral actions involve arrogantly thinking oneself to be superior to others. To the contrary, immorality can sometimes be a matter of failing to stand up for others because of a lack of self-esteem or self-importance. Consider, for example, a bystander witnessing, and failing to intervene in, some wrongdoing. There might be numerous reasons to be a mere bystander, but just one possible

[84] Reath (2006: 25).

source of motivation could come in the form of struggling to convince oneself that one has the status, knowledge, or capacity to say that something is wrong, let alone take action to intervene.

With that challenge in mind, consider another interpretation of self-conceit that is not centered around interpersonal arrogance. According to Kate Moran, self-conceit fundamentally involves distorting moral principles in such a way that we can reassure ourselves that we are moral. In other words, self-conceit is a kind of misapprehension of the moral law that serves to protect our "dear self" from harm and to elevate its status. Moran supports her reading, in part, by pointing out that Kant describes the self-conceited person as solipsistic. Referring to Kant's lectures on ethics, Moran asserts, "self-love becomes blameworthy when it becomes solipsistic, or when it 'excludes others from our love or inclination toward them.'"[85] She also supports her reading by drawing a comparison between moral and logical self-conceit, the latter of which Kant discusses in his lectures on logic.[86] In both cases, the agent isolates himself by refusing to accept the disagreeing judgments of others, perceiving himself to be correct (morally good or logically sound), despite their lack of real evidence. More precisely, the self-conceited person tends to interpret evidence to his own liking, "see[ing] no possible need for intersubjective verification as a touchstone against error."[87]

Moran's interpretation of self-conceit as a misinterpretation of the authority of moral law is also supported by some of Kant's remarks on the feeling of respect. After he has introduces that feeling in the second Critique, he emphasizes that the feeling of respect for the moral law is distinct from a "voluntary liking" of the moral law:

> The moral level on which a human being… stands is respect for the moral law. The disposition incumbent upon him to have in observing it is to do so from duty, not from voluntary liking or [as something he] undertakes unbidden, gladly and of

[85] Moran (2014: 433).
[86] Moran (2014: 434).
[87] Moran (2014: 427).

> his own accord; and his proper moral condition, in which he can always be, is virtue, that is, moral disposition in conflict, and not holiness in the supposed possession of a complete purity of dispositions of the will.[88]

Rather than respect for the moral law, self-conceit is a type of disposition toward morality not appropriate for us: namely, "voluntary liking." The self-conceited person does not view the authority of the moral law as one of necessity (let alone, an authority which "strikes down" their "dear self"), but rather as something to be distorted or manipulated to suit their conception of happiness. By identifying what the moral law commands of us with what we want, we act as though we have the sort of nature that naturally aligns perfectly with the moral law. Kant uses self-conceit in this way, for example, in his discussion of "moral enthusiasm," which involves the insistence that morally good actions are, or should be, done from a sort of liking, or feeling of gratification:

> By exhortation to actions as noble, sublime, and magnanimous, minds are attuned to nothing but moral enthusiasm and exaggerated self-conceit; by such exhortations they are led into the delusion that it is not duty—that is, respect for the law whose yoke… they must bear, even if reluctantly—which constitutes the determining ground of their actions… but that it is as if those actions are expected from them, not from duty but as base merit.[89]

Self-conceit, Kant goes on to say, leads people to a "fantastic cast of mind, flattering themselves with a spontaneous goodness of heart that needs neither spur nor bridle."[90] According to Moran, the problem with thinking that we are naturally morally good in this way is not that it would be unfortunate if it were true. Nor is the problem that it involves arrogantly asserting one's superiority over others. The problem is that it is not true.

Now, it so happens that the sort of moral misinterpretation that Moran thinks is at the heart of self-conceit does tend to lead to a kind of interpersonal arrogance. By distorting moral

[88] KpV 5:84.
[89] KpV 5:84-5.
[90] KpV 5:85.

principles in such a way that a self-conceited agent can reassure themselves that they are moral, they tend to demand too much respect from others, since the self-conceited agent believes they are perfect and hence morally (or logically) superior to (imperfect) others. Nonetheless, Moran assets that self-conceit is not fundamentally a kind of interpersonal arrogance:

> Self-conceit is thus not the decision to *ignore* the constraint of impartiality and impose one's own conclusions upon others. Rather, it is the tendency to construct a fiction that one is abiding by this constraint when one has actually failed to heed it. The failure of self-conceit is thus more intractable and insidious than the imposition of one's will, or the claim to superiority, even if it often ultimately takes this form.[91]

This in turn leads self-conceited agents to demand too much respect of others (if they happen to be around), since self-conceited agents believe they are perfect and hence morally superior to (imperfect) others. Nonetheless, self-conceit is not fundamentally the arrogant assertion of oneself over others, though that may be a consequence. Rather, self-conceit is primarily a preoccupation with the value of the self. As Moran puts it, "the fundamental failure of self-conceit is a failure in the way an agent assesses herself."[92]

Moran's interpretation of self-conceit as just one instance of a more basic error concerning how an agent views herself helps to meet the challenge from earlier. Recall how the view that self-conceit is a kind of interpersonal arrogance faces a difficulty in explaining how the error behind self-conceit can be active in any instance in which immoral action compels someone. For some such actions are based in a kind of self-doubt, rather than any claim of superiority. Moran's interpretation, however, can meet this challenge by explaining how self-conceit is just one way that the basic error at the root of immorality shows up. Another version of that same error is moral despondency. According to Moran, "Despondency occurs when, upon realizing his moral shortcomings, an agent decides that none of his actions could

[91] Moran (2014: 422).
[92] Moran (2014: 438).

ever conform with the moral law, a conclusion that leads to a type of resigned indolence."[93] Moran points out that Kant contrasts despondency with self-conceit: whereas the self-conceited agent distorts moral principles in such a way that he cannot help but be moral, the despondent agent distorts those principles in such a way that he can never be moral. For Kant, arrogant self-conceit and despondency are the Scylla and Charybdis between which we must steer towards becoming moral.

Though opposites, self-conceit and despondency stem from a common error that explains immoral action. The error common to both, as Moran described it, is "the failure to recognize that virtue consists of a continuous struggle":

> The conceited agent recasts his failures, via fanaticism and rationalization, as moral successes. The despondent agent, on the other hand, takes his shortcomings to be evidence that he will never be adequate to the moral law. Notably, however, both errors stem from the failure to recognize that virtue consists in a continuous struggle.[94]

Immoral action has its source, then, in a more general error of self-knowledge. Self-conceited and despondent agents fail to know themselves as sensible and rational beings, for whom morality requires constant striving. Moreover, as in the case of self-conceit, despondency can lead to the imposition of one's erroneous self-evaluation on others, in that the despondent agent may ultimately impose his judgment about his failures on others to maintain the delusion that those failures are insurmountable. Still, such an imposition of one's will on another is not, Moran argues, the heart of either self-conceit or despondence. At the heart of each is, instead, a failure concerning a self-relation.

This interpretation of self-conceit, and the related moral failure despondency, will be helpful when I turn, in the next chapter, to reconstructing Kant's pedagogical method in the

[93] Moran (2014: 426).
[94] Ibid.

Doctrine of Method of the *Critique of Practical Reason*. For, as I will argue, that pedagogical method is designed as a remedy to the problems of self-conceit and despondency, to help us steer between the Scylla and Charybdis of failing to recognize that virtue requires continuous striving. In this section, my goal was to explore Kant's notion of self-conceit in light of comparisons with Rousseau's notion of *amour propre*. On some interpretations of self-conceit, the two notions are quite similar, both being grounded in claims of superiority based on comparative judgments of our value with that of others. Whereas, on interpretations such as Moran's, in which self-conceit is, at bottom, a failure of self-knowledge that does not fundamentally involve comparative judgment, the differences with *amour propre* appear greater.

In this chapter, I have drawn on a revisionist interpretation of Rousseau's notion of amour propre, in order to explain both his diagnosis of the problem of *amour propre* and the (partial) solution he outlines in the educational program of *Emile*. As I explained, on that interpretation, the problem lies in *amour propre* becoming, in various ways, inflamed, leading to harmful effects on individuals and society as a whole. When inflamed, *amour propre* manifests in excessive cravings for superiority, which can be satisfied only at the cost of others' well-being and, sometimes, even one's own. The task of the educational program in *Emile*, then, is to forestall the emergence of *amour propre*, as well as to provide grounds for its proper cultivation. Rousseau's framing of the problem of education – particularly, when viewed through the revisionist interpretation – importantly influences Kant's own conception of the problem for which moral education is to provide a solution. However, Kant's analogue to *amour propre*,

self-conceit, is importantly distinct. Whereas *amour propre* depends on judging our worth in comparison to, and in the eyes of, others; self-conceit is not fundamentally based on any kind of comparative judgment, but rather on a misapprehension of the authority of the moral law. These differences in conceptions of the problem of education ultimately bear out certain differences in Rousseau's and Kant's respective theories of education.

Chapter 2: What is hope?

In the previous chapter, I drew comparisons between Rousseau's notion of *amour propre* and Kant's notion of self-conceit as one piece in the framing of my reconstruction of Kant's pedagogy of hope. In this chapter, I provide another important piece of framing. The guiding question here is, 'What is hope?' I will neither provide an exhaustive summary of the many efforts to define hope, nor attempt to give any comprehensive definition. Instead, what I aim to do in this chapter is examine two strategies recent commentators have taken to articulate Kant's account of hope, putting the strategies in conversation with Kant's writings on education. My purpose is not to assess the merits or demerits of either strategy, but rather to use them to bring to the fore features of Kant's account of hope that are most relevant for my purposes in understanding its role in his theory of moral education. So, the more precise question of this chapter may be stated as, 'What is hope, specifically in the context of Kant's moral pedagogical method?'

In section one, I present what is often referred to in the scholarly literature as the 'standard account of hope,' which is a starting point for many recent attempts by analytic philosophers to analyze hope in terms of its underlying mental state. In section two, I focus on one take on this approach – namely, Adrienne Martin's 'incorporation analysis' of hope, which takes inspiration from an influential (not to say uncontroversial) interpretation of Kant's theory of moral psychology. For my purposes, one important lesson of the standard account of hope and Martin's analysis is that hope is a mental state distinguished by its sustaining power to overcome the possibility of despair. In section three, I turn to a second strategy towards giving an account of hope, which is to analyze hope in terms of its object. Since, for Kant, the object of reasonable

hope is the so-called 'highest good' – a world in which happiness and moral goodness are in perfect proportion – this strategy usually involves honing in on those contexts in which he invokes that notion. As I will explain, those efforts often revolve around what are often understood as two competing interpretations of the highest good – one theological, the other secular. Building on the upshot of sections one and two, I explain how the highest good, whether viewed through a theological or secular lens, helps to explain how, for Kant, the highest good explains the role of hope in sustaining our moral motivation in the face of error and evil. In section four, I analyze Kant's remarks on the highest good in the *Critique of Practical Reason* in order to expand on the motivational role of the highest good as an object of hope. This sets the stage for the next chapter in which I reconstruct the pedagogical method Kant presents at the end of the second Critique, since, as I show, the rationale for that method is to explain how we can educate youth to sustain their moral motivation, something for which hope is necessary.

1. The standard account of hope

Many analytic philosophers, including commentators on Kant's philosophy, begin their analysis of hope with what is sometimes referred to as the "standard account."[95] While there are some proponents of that account, there is widespread agreement in the recent philosophical literature that it is inadequate. In this section, I summarize the standard account of hope and some of the challenges that have been raised against it. My main aim is not to assess the merits of either the standard account or the challenges against it, but rather to use both of these to

[95] I will mention some examples of contemporary philosophers who take this approach later in this section. For one early analytic investigation of hope along these lines, see Day (1969). Versions of the standard account of hope are also attributed to some modern philosophers, like Hobbes, who defines hope as an "appetite [i.e., desire] with an opinion of attaining" some future good. Thus, perhaps it should not be surprising to discover that Kant accepted the orthodox definition of hope, if only some other version, as well.

elucidate certain features of Kant's account of hope, the main one being that hope works to sustain our moral motivation in the face of certain failure, injustice, and evil.

On the standard account of hope, there are two jointly necessary and sufficient conditions for some person to count as having hope for an outcome. To hope for some outcome is to both desire it (call this the 'desire condition') and believe it is possible but not certain to obtain ('belief condition'). For example, my hope that the Los Angeles Dodgers win the World Series is, on this account, constituted by my desire that that be the outcome of this season of professional baseball, on the one hand, and my belief that that outcome for the Dodgers is realizable but not a sure thing, on the other.

If we take for granted that hope can be analyzed as a combination of belief and desire, or in terms of necessary and sufficient conditions, then the standard account may seem intuitive.[96] For one thing, the belief condition distinguishes hope from less modally constrained attitudes, like wishing, and more confident ones, like optimism. After all, while it may make sense for me to wish that the Dodgers won the World Series last year, it does not make sense (at least, not on the face of it) to hope for an outcome different than the one that happened, since that would be to hope for the impossible. Likewise, while it may not be nonsensical to hope for something we are optimistic about, a *ceteris paribus* 'assert-the-stronger' norm usually leads us to assert the strongest justified attitude that we have towards a state of affairs in order to give others a better sense of our information state.[97]

With regards to Kant's account of hope, some version of the standard account seems to be present in his writings. In various places, he appears to assume the desire condition in speaking

[96] For a criticism of analyzing hope as a combination of a certain belief and desire, see Segal, Gabriel, and Mark Textor (2015). Also, Blöser (2019).

[97] On this point, see Chignell (2013: 200). Chignell claims that we assert the stronger attitude (e.g., we assert that we know *p*, even if we also hope that *p*), but we might question whether that supposed norm of assertion does, in fact, support a conceptual claim that it is irrationality or impossible for agents to hold both the stronger and weaker attitude.

of hope as a kind of desire, or inclination. For example, in some of his lectures, he says that people hope something will be true "because they would like to have it," that "all hope concerns happiness" and, elsewhere, that "[h]appiness is the satisfaction of all our inclinations."[98] Kant likewise appears to assume the belief condition of the standard account. For example, on his view, the moral person hopes for, but is not certain to attain, a degree of happiness that is in proportion to their virtue: "Logical certainty of a future life is very hard to attain … if one considers that here on earth happiness is not always a consequence of good behavior, hence another world is to be hoped for in which this will occur."[99] In addition to these examples, one might find a version of the standard account of hope in Kant's political writings – specifically, his remarks on the hope for the collective progress of the human species towards a more just and peaceful world, conceived as a distinctly shared and historical task. Take, for instance, these remarks in his essay *Toward Perpetual Peace*:

> [R]eason is not enlightened enough to survey the entire series of predetermining causes that foretell with certainty the happy or unhappy consequences of humankind's activities in accordance with the mechanism of nature (although it does let us hope that these will be in accord with our wishes).[100]

In other words, historical progress is uncertain – something which "reason is not enlightened enough to survey" – yet reason does permit that we hope that that state of affairs obtains "in accord with our wishes."[101] So, it appears as though a version of the standard account of hope is at least consistent with some of Kant's scattered remarks that most closely approach something like a definition of hope.

While there are some proponents of the standard account of hope, there is widespread agreement in recent philosophical literature that its two conditions concerning belief and desire

[98] Blom-L 24:93; KrV A805/B833.
[99] Blom-L 24:200.
[100] PP 8:370
[101] Ibid.

are insufficient to constitute an adequate analysis of hope. Although my main aim is not to assess arguments for or against the standard account of hope, I will consider three challenges raised against that account, since I believe these will help to illuminate certain facets of Kant's own remarks, particularly with respect to what he takes to be the role of hope in motivating us to bring about its object.

One challenge raised against the standard account of hope is that its two stated conditions are insufficient for distinguishing hope from despair. Ariel Meirav illustrates this challenge with a thought experiment (hereafter, 'Shawshank') based on Stephen King's novella, "Rita Hayworth and the Shawshank Redemption," and the film based on it.

> Shawshank: Imagine two men, Andy and Red, who are both serving life sentences in a state prison. They both desperately want to be free, and they both believe it is possible, though not likely, that they might succeed in escaping. Yet Andy holds out hope for an escape and a life of freedom, while Red only despairs.

Insofar as Andy hopes for freedom and Red does not, in spite of their similar desire for freedom and expectation of attaining it, Shawshank seems to present a case in which the standard account of hope fails to distinguish hope from despair. Hence, assuming that an adequate analysis of hope must be able to explain how it differs from despair in every case, that account must either be jettisoned or supplemented by some further conditions. As I will show in the next chapter, Kant's description of what I call his pedagogy of hope shows that he is interested in not only, or even primarily, in distinguishing the concepts of hope and despair, but also in showing how these attitudes function together.

A second challenge to the standard account of hope is that it does not capture what is distinctive about those hopes that seem to matter to us. Philip Pettit describes the standard account as seeking "a lowest common denominator of usage" for the concept 'hope.' He argues that, while this approach might capture "superficial hopes," it fails to illuminate a more narrow

range of hopes that are substantial rather than superficial.[102] So, for example, my hope that the Dodgers win the World Series, though sincere, is less substantial, as it happens, than my hope that my sons know that they are loved and that they are, in fact, loved. After all, I would be more devastated if the latter turned out not to be the case than I would be if the Dodgers were to lose. And both of those hopes are yet more substantial for me than my hope for clear, blue skies this weekend. According to philosophers like Pettit, hope should not be defined in terms of its lowest common denominator, but rather in terms of substantial hopes, in order to put us in a position to explain why some hopes matter to us. This is a point that Kant appears to appreciate. For, as I will discuss in a subsequent section of this chapter, his attention seems to be on what he takes to be the object of not just any sort of hope but of *rational* hope, in particular.

The third challenge to the standard account of hope that I will consider concerns the motivational role of hope. On the one hand, it is argued that the standard account cannot accommodate the intuition that hope has the power to sustain us through difficult trials. Adrienne Martin summarizes this point as follows: "Desiring, even desperately so, to overcome such situations doesn't have any special kind of motivational power. Moreover, recognizing extremely slim odds seems likely only to hold one back or make one's efforts more timid."[103] On the other hand, it is also argued that the standard account is unable to explain how hope can, instead, undermine our efforts, particularly those part of time-extended pursuits aimed at realizing ends to which we attach relatively high significance. Along these lines, Martin argues that hope can be dangerous, because it can easily lead us to become lost in outcome-oriented fantasies. That is to say, some hopes might lead us to acting, or fantasizing, without any regard for the means required to fulfill the hopeful fantasy, or to fixate on fantasies that orient us towards the

[102] Petit (2004: 153-4).
[103] Martin (2013: 5).

unlikelihood of the hoped-for outcome. While such fantasizing can contribute to our efforts, Martin points to empirical findings that "[t]hey can also propose reasons to give up a struggle, doubt the effectiveness of one's means or plans, and lose confidence in one's effectiveness as an agent."[104] Thus, the two-fold challenge here is that the standard account appears unable to explain either the motivational benefits or dangers we might reasonably associate with hope.

This final challenge concerning the motivational role of hope seems in accordance with Kant's moral psychology. Recall, in particular, what I said in the previous chapter regarding his remarks on moral despondency. For Kant, moral despondency is opposite self-conceit, though both are based on the same moral error: whereas the self-conceited agent distorts moral principles in such a way that she cannot help but be moral, the despondent agent distorts those principles in such a way that she can never be moral. Writing in the context of Kant's account of hope, Andrew Chignell explains, "Kant was acutely aware of our psychological vulnerability to perceived inefficacy in the face of large-scale evil and injustice."[105] Perhaps the most striking example of this is where, in the *Critique of Judgment*, Kant considers the psychology of an atheistic but "righteous man (like Spinoza)" who is still liable to be left despondent by the perception that his own moral strivings, as well as those of others, are inefficacious against large-scale structural evil and injustice:

> But his effort is limited… Deceit, violence, and envy will always surround him, even though he is himself honest, peaceable, and benevolent; and the righteous ones besides himself that he will still encounter will, in spite of all their worthiness to be happy, nevertheless be subject by nature, which pays no attention to that, to all the evils of poverty, illnesses, and untimely death, just like all the other animals on earth, and will always remain thus until one wide grave engulfs them all together (whether honest or dishonest, it makes no difference here) and flings them, who were capable of having believed themselves to be the final end

[104] Martin (2013: 92).
[105] Chignell (2020: 224).

> of creation, back into the abyss of the purposeless chaos of matter from which they were drawn.[106]

Ultimately, when faced with such conditions, the righteous Spinoza must either "give up [his end of being righteous] as impossible" or "assume the existence of a moral author of the world, i.e., of God, from a practical point of view, i.e., in order to form a concept of at least the possibility of the final end that is prescribed to him by morality."[107] What this amounts to, according to Chignell, is that "even someone as righteous as Spinoza simply cannot maintain his resolve without some source of hope that, ultimately, justice will prevail."[108] For Kant, too, it seems that the standard account of hope is inadequate insofar as it cannot explain the power of hope to sustain our moral strivings through difficult trials.

In light of the three challenges described above, as well as others not mentioned,[109] some philosophers have proposed supplementing the two stated conditions of the standard account of hope. In addition to desiring the hoped-for outcome and believing that outcome is possible but not certain, some proposals include: Luc Bovens argues that we must engage in a kind of "mental imaging" about the projected outcome or the means for bringing it about; we must "orient our agential energies" towards its chance of occuring;[110] Philip Pettit says we must resolve to act "as if the desired prospect is going to obtain";[111] Ariel Meirav, who popularized the Shawshank example, argues that we must view the "external factors" in the possibility of the outcome that are beyond our control as good – that is, we must focus on the outcome under the aspect of its possibility.[112] While I think each of these proposals has its merits and could illuminate Kant's

---

[106] KU 5:452.
[107] KU 5:452-3.
[108] Chignell (2020: 225).
[109] Kwong (2018), for example, adds to these criticisms that the standard account of hope cannot explain the varying strengths of different hopes.
[110] Bovens (1999).
[111] Petit (2004).
[112] Meirav (2009).

varied remarks of hope, I will focus on just the account of hope developed by Adrienne Martin in her book *How We Hope*. Her account is not only comprehensive but also distinctly Kantian in ways that suit my purposes.

2. Martin's 'incorporation analysis'

Like those of other contemporary analytic philosophers, Martin's theory of hope takes its cues from the standard account (or what she refers to as the "orthodox definition") and the challenges which it is up against. However, Martin is careful to differentiate her analysis from attempts to supplement the standard account with some third condition. She proposes, instead, that we think of hope as a "'syndrome' of attitudes and feelings," rather than as being constituted by a set of necessary and sufficient conditions.[113] Begin, she suggests, by considering how certain activities are ones that hopeful agents tend to engage in: thinking, imagining, anticipating, fantasizing, planning, intending, and acting or feeling in certain ways towards a desired outcome are some examples of characteristically (not to say always) hopeful activities. To use the Shawshank thought experiment from earlier, Andy's hope for freedom might be characterized by him planning his escape, fantasizing about life beyond bars, and anticipating the feeling that fulfillment of his desire for freedom will bring. By contrast, Red's despair might be characterized by him doubting Andy's plans, fatalistically repeating ineffective courses of actions, and his gloomy feelings in response to Andy's fantasies.

Worried that such an analysis runs the risk of being ad hoc, Martin maintains there must be some element of hope that unifies the activities characteristic of it as a syndrome. Martin argues that the unifying element of hope is 'incorporation,' which she describes here:

[113] Martin (2013: 24).

> to hope for an outcome is to desire (be attracted to) it, to assign a probability somewhere between 0 and 1 to it, and to judge that there are sufficient reasons to engage in certain feelings and activities directed toward it. The element that unifies hope as a syndrome is this final element, which I argue is a way of incorporating hope's other elements into one's rational scheme of ends.[114]

Martin's 'incorporation analysis' of hope is a direct reference to Henry Allison's influential interpretation of Kant's moral psychology in *Kant's Theory of Freedom*. One of Allison's central (and much disputed) interpretative claims about Kant's moral psychology is that our capacity to act freely relies on our ability to endorse or reject some sensible inclination as a reason to act. Endorsing an inclination, or desire, as a reason to undertake some action amounts to 'incorporating' that desire into one's maxims and being disposed to justify one's actions by citing that desire as a reason.[115] Likewise, Martin's incorporation analysis of hope relies on an opposition between, on the one hand, desires conceived merely as causal forces capable of moving us to act but that are not sensitive to norms and, on the other, desires conceived as reasons that are sensitive to norms and that allow us to determine our actions as rational agents. For Martin, the hopeful person not only desires the relevant outcome, but also endorses, or incorporates, her desire as a reason to pursue valuable activities such as planning and fantasizing.

Within this framework, Martin conceives of hope as requiring two acts of judgment. In the first act, the hopeful person judges that her belief about the possibility or uncertainty of the outcome's occurring to 'licenses' activities characteristic of hoping, such as planning, fantasizing, and entertaining feelings of anticipation. The possibility of the outcome is treated not yet as a reason to hope but as an 'enabling condition' for other reasons. This licensing act

[114] Martin (2013: 26).
[115] Allison (1990: 40). Allison offers the label "Incorporation Thesis" to refer to the following quote from Kant's *Religion within the Boundaries of Mere Reason*: "an incentive can determine the will to an action only insofar as the individual has *incorporated* it into his maxim" (R 6:24, emphasis added).

requires, if you will, that the hopeful person take a second-order stance on the probability she assigns to the hoped-for outcome, as exhibited in the following syllogistic reasoning:

> I'm facing incredibly bad odds, but if I focus on this fact, I'm going to be paralyzed by despair. … So I'd best not think about how poor the odds are, and instead focus on the fact that they are in the realm of possibility; … "miracles" happen, and so it is possible that I will get what I so desperately desire.[116]

The avoidance of despair is the given end, and the licensing stance is the chosen means. Martin claims, it is "psychologically possible to adopt the licensing stance on the basis of deliberation about how useful it would be to see things as that stance dictates," and "whether it is successful or defective as the kind of state that it is depends on its responsiveness to considerations of rational-ends-promotion."[117] Hence, taking up the licensing stance is justified whenever it coheres with and promotes a rational scheme of ends, perhaps most prominently by overcoming the possibility of despair that threatens the realization of that same scheme of ends.

To illustrate this initial licensing act of judgment, recall Shawshank, the case challenging the standard account of hope on the grounds that it is unable to distinguish that attitude from despair. Meirav's presentation of that case is useful here:

> Andy lives in the hope of escape, whereas Red despairs of this. Indeed, Red thinks that hope should be resisted, suppressed, for hoping in this virtually hopeless situation would threaten his sanity … Red will say, "I grant you it is possible, but the chance is only one in a thousand!" whereas Andy will say, "I grant you the chance is only one in a thousand, but it is possible![118]

In terms of Martin's account, what distinguishes Andy and Red are their second-order judgments about the probability of escaping from prison: although they assign similar, if not identical, probabilities to that outcome, they take different activities to be licensed by that outcome. Whereas Andy focuses on the possibility of that outcome ("I grant you the chance is only one in

[116] Martin (2013: 96-7).
[117] Martin (2013: 101).
[118] Meirav (2009: 221).

a thousand, but it is possible!"), Red focuses on its significant improbability ("I grant you it is possible, but the chance is only one in a thousand!"). Consequently, Andy judges that the probability of escape licenses activities characteristic of hope (planning his escape, fantasizing life beyond bars, etc.), whereas Red judges that probability of escape does not license such activities, to the point of resisting even entertaining the idea of such activities and advising Andy to do the same.

With this license in hand, the hopeful person is free to make what Martin's conceives of as the second act of judgment on the part of a hopeful person – namely, to judge her desire for the hoped-for outcome to be a practical reason to undertake activities characteristic of hope towards that outcome. Whereas the licensing act of judgment amounts to a person permitting herself to see certain considerations as practical reasons for someone, this second act of judgment amounts to her acting on those considerations as sufficient practical reasons for her. It is possible, then, that two persons similarly judge the probability of some outcome to license treating certain considerations as practical reasons and yet diverge in judging whether those considerations provide sufficient practical reason. Two persons may even agree in judging that those other considerations are sufficient for acting and yet decide to act in different ways. However this second, incorporating-act of judgment unfolds, the hopeful person who acts from this judgment "stands ready to offer justification for her action, even though she may be mistaken about whether her action is actually justified."[119]

What makes Martin's incorporation analysis useful for my purposes is how she conceives of hope as an act of judgment in a distinctly Kantian register. From a Kantian perspective, the standard account of hope can seem misguided insofar as it does not treat hope as a normative activity. That hope is normative, for Kant, is nowhere more apparent than in what he says is one

[119] Martin (2013: 75).

of the core questions of his philosophical thought: "What may I hope?" Setting aside for the moment what Martin takes to be the nature of the normativity of hope, the fact is that, for Kant, we must view ourselves as subject to certain norms insofar as we are to hope rationally. Martin's framing hope as an act of judgment, as opposed to a belief-desire mental state, draws on what, in a different context, Robert Brandom calls "perhaps Kant's deepest and most original idea":

> What distinguishes judging and intentional doing from the activities of non-sapient creatures is not that they involve some special sort of mental processes, but that they are things knowers and agents are in a distinctive way responsible for. Judging and acting involve commitments. They are endorsements, exercises of authority. Judgments and actions make knowers and agents liable to characteristic kinds of normative assessment.[120]

One way of construing Martin's insight is to extend this idea articulated by Brandom: we are subject to normative assessment not only as knowers and doers, but also as hopers. As hopeful agents, we are liable to certain norms because, as Martin emphasizes, hope is fundamentally an activity of judgment in the sense articulated by Brandom. Later in this chapter, I show how, in his pedagogical writings, Kant appears to likewise treat hope as a normative judgment.

In a paper entitled "Fundamental Hope and Practical Identity," Claudia Blöser and Titus Stahl raise an objection to Martin's analysis that is further useful for my purposes towards describing the pedagogical dimension of Kant's account of hope. They object to what they call Martin's "purely instrumentalist" view of the normativity of hope. Blöser and Stahl take Martin's analysis to imply that "it is rational to treat one's attraction to an outcome as a reason to engage in hopeful activities if the *consequences* or *effects* of one's engaging in these activities or simply engaging in the activities themselves coheres with my rational scheme of ends."[121] For instance, what makes it rational for, say, Andy to treat his desire for freedom as a reason for planning his escape, fantasizing about life beyond bars, and so on, is that engaging in these activities

[120] Brandom (2009: 32).
[121] Blöser and Stahl (2017: 353).

contributes to the realization of his rational scheme of ends. Blöser and Stahl object to this purely instrumentalist view of the rationality of hope on the grounds that it cannot accommodate cases such as ones where it might be rational for a person to hope, despite the lack of any clear benefits. For example, it might be appropriate for a political organizer to hope for the end of racial, gender, or class inequality, "[e]ven if her hoping for political change does not generate any recognizable benefits for her or promote any ends."[122] A purely instrumentalist view of the rationality of hope such as Martin's seems to not capture the appropriateness of hope in such cases.

In light of that limitation, Blöser and Stahl propose supplementing Martin's incorporation analysis with a non-instrumentalist view of the rationality of hope. Their proposal draws on the notion of a "practical identity," which is a characterization of who or what an individual is that entails she has certain reasons for acting.[123] For example, to say someone "is a political organizer" is to say something about who she is, or what her socially-constructed role is, while at the same time giving a broad characterization of how she ought to act. More specifically, to say that someone is a political organizer is to say she has reason to, for example, build a base of people whom she can mobilize to win an election or carry out some direct action, or to work with some community to create a power map visualizing who to identify the best individuals to target to promote the desired social change. So, were we to ask a person why she routinely takes time to participate in these activities, "I'm a political organizer" would be an intelligible response, and were she to deny that she had reason to do any of these things, we'd take this as a denial that she's a political organizer at all.

[122] Blöser and Stahl (2017: 354).
[123] See, especially, Blöser and Stahl (2017: 355ff). The conception of practical identity they draw on is one developed by Christine Korsgaard (1996 & 2009).

Utilizing this notion of practical identity, Blöser and Stahl argue that it can be non-instrumentally rational for some person to hope if it is part of her practical identity, or if "*these hopeful activities are essential to the hopeful person's being the person*."[124] The rationality of such hopefulness is not grounded in "that of end promotion but that of upholding one's personal integrity," such that "the relevant hope is a constituent part of a scheme of ends that one has reason to uphold."[125] For example, a political organizer might have reason, in virtue of that practical identity, to hope for the end of global inequality in full knowledge of its unlikeliness. Generally speaking,

> the crucial condition is that the identity of the agent must be partly constituted by a certain *perspective on the world* in which certain considerations count as reason-giving, and that giving up hope would entail the unavailability of those lines of reasoning. ... [T]he fact that the hopeful activities form part of a person's practical identity provides a reason to continue to hope that may outweigh the reasons given by cost-benefit analysis.[126]

To say that the rationality of hope is, in part, grounded in an agent's practical identity, Blöser and Stahl are not denying that it can sometimes be rational to give up hope. After all, it might make sense for someone to shed the practical identity in virtue of which they have reason to hope. Nonetheless, by supplementing Martin's incorporation analysis of hope with a normativity grounded in practical identity, that analysis can now not only accommodate the problematic cases described by Blöser and Stahl, but also explain what might make it rational (or irrational, as the case may be) to continue to hope under difficult circumstances.

Before wrapping up this section, I want to anticipate how the preceding discussion on hope will be brought to bear on Kant's theory of moral education. Briefly: There is, arguably, one practical identity we cannot shed, that is, our identity as a human being. For Kant, our humanity

[124] Blöser and Stahl (2017: 354).
[125] Blöser and Stahl (2017: 355).
[126] Blöser and Stahl (2017: 357).

commits us to certain moral duties that we cannot waive. Among these is a duty to self-improvement. For Kant, this includes both my moral development and what we might call my personal development, or the cultivation of my talents and capacities. With respect to the former, Kant basically thinks that we are morally obligated to work on our moral characters. It is something that we owe to ourselves as a matter of self-respect. Moral improvement is a way of valuing ourselves as rational, finite beings who are ends in themselves. When we make ourselves morally better, we make ourselves more capable of acting in accordance with rational principles and so more capable of acting autonomously. In this sense, improving our moral character is a way of making ourselves free.

Now, the goal of moral education, for Kant, is the formation and perfection of moral character. While he certainly thinks we should prioritize the cultivation of our rational abilities, he does also invite us to think about our natural perfections more broadly, including emotional, spiritual, and physical capacities as these relate to moral and non-moral aspects of our characters. As I will show in greater detail in the next chapter, among those capacities he thinks moral educators ought to cultivate is hope for moral progress. As I will argue, someone can successfully adopt the end of self-improvement only if she hopes in the sense, following Martin, that she incorporates her desire for self-improvement into her agency to give justification for participating in activities characteristic of hope. However, the hope for self-improvement is not only instrumentally valuable towards, say, overcoming despair; rather, following Blöser and Stahl, that hope is valuable in itself also as a matter of respect for the integrity of our humanity.

3. Two conceptions of Kant's 'highest good'

In the previous two sections, I examined two accounts – the so-called standard account of hope and Adrienne Martin's incorporation analysis – that take a similar approach to defining hope. That approach seeks to define hope in terms of some mental state, or perhaps mental act, of the hopeful agent: in the case of the standard account, hope is defined as a certain combination of belief and desire; in the Martin's incorporation analysis, hope is defined in terms of the mental acts of licensing and incorporating. I showed some of the ways in which each of these illuminates features of Kant's account of hope. One key takeaway is that, on this approach to defining hope, it is a mental state distinguished by its sustaining power to overcome the possibility of despair. In this section, to build on that lesson, and further illuminate Kant's account of hope, I examine a different approach to defining hope, one that takes the central target of analysis to be not hope itself but its object. As I will explain, what Kant takes to be the object of rational hope also importantly works to sustain our moral motivation in the face of error and evil.

An object-oriented approach to analyzing hope is encouraged in no small part due to Kant's formulation of the question of hope in the *Critique of Pure Reason*: "What may I hope?"[127] Since questions of hope, according to Kant, are always related to happiness, a more precise formulation of his question is: What can we hope for if we act as we should and become worthy of happiness? Kant's answer in the first Critique is that we can hope to attain the 'highest good,' which consists of "happiness distributed in exact proportion to morality (as the worth of a person and his worthiness to be happy)."[128] To put it in simple terms: a morally good person is rationally permitted to hope to obtain the happiness he deserves. However, Kant's conception of the highest good is widely debated. The crux of that debate concerns the relative merits of what

[127] KrV A804/B832.
[128] KpV 5:110.

are usually regarded as two competing conceptions of the highest good in Kant's corpus that constitute the objects of two kinds of hope. In this section, I lay out those two conceptions of the highest good and their corresponding kinds of hope. This will put me in a position to, in the next section, identify a third conception of the highest good in some of Kant's writings on education.

Debates concerning Kant's two conceptions of the highest good are helpfully set up by an influential paper by Andrews Reath. In that paper, Reath distinguishes a theological conception of the highest good from a secular conception of the highest good. According to Reath, the theological conception of the highest good refers to a "state of affairs that comes about in another world through the activity of God," whereas the secular conception of the highest good can be described in entirely naturalistic terms as a state of affairs to be achieved in this world through our actions.[129] An important distinction that Reath points out between these two conceptions is that, in the theological one, Kant stresses the proportionality between two kinds of good, viz. virtue and happiness, while in the secular one there is not a necessary connection between the two. Instead, the secular conception of the highest good is merely a union (or combination) of virtue and happiness, where happiness is subordinated to virtue.[130] For Reath, we can imagine a state in which people establish just social institutions to ensure that agents who are virtuous are proportionately happy. Hence, the requirement that happiness be proportional to virtue in the world can be achieved through mere human agency via secular means.

What leads Kant to the theological conception of the highest good, on Reath's reading, is the assumption that all individuals will attain the highest good.[131] Reath sees the theological conception of the highest good as having two requirements: (1) all individuals are virtuous and (2) happiness is distributed in exact proportion to virtue. Since these requirements cannot be

[129] Reath (1988: 601).
[130] Reath (1988: 605).
[131] Reath (1988: 603).

satisfied for everyone in the empirical world through human agency alone, the theological conception describes an otherworldly goal. The second of these claims about the theological conception of the highest good is articulated powerfully by Frederick Beiser. According to Beiser, Kant's conception of the highest good is "fundamentally Christian, indeed Protestant" and that it is "utterly contrary to his intention to interpret it as secular and humanist."[132] Beiser argues that Kant saw his entire ethics as Christian doctrine and his theory of the highest good as being essentially a moralized version of the Augustinian notion of the city of God. The existence of God is not only posited as a practical postulate in the second Critique, about which I will say more later, but the justification of such a postulate closely resembles the Old Testament, more precisely the "ancient lament of Job: Life is unfair, and we live in an unjust world where the vicious prosper and the virtuous suffer."[133] On this interpretation of Kant's notion of the highest good, justice can only be restored through divine intervention. In light of these religious underpinnings, Beiser concludes that Kant's account of the highest good is decisively apolitical: "unlike his more idealistic successors, Kant does not think that the highest good can be a political ideal, one achieved through the state."[134]

However, Reath argues that the theological conception of the highest good lacks a clear basis in Kant's critical philosophy and does not fit in with his other commitments. On Reath's view, the theological conception of the highest good cannot be generated from Kant's philosophy internally, and thereby, it "represents a departure from Kant's basic principles."[135] That is why Reath favors the secular conception of the highest good over the theological conception. For

[132] Beiser (2006: 590).
[133] Beiser (2006: 597).
[134] Beiser (2006: 603).
[135] Reath (1988: 613).

Reath, the secular conception of the highest good represents Kant's mature view of the highest good.[136]

What is important for my purposes is not important whether Kant's notion of the highest good is viewed through either a theological or secular lens. Rather, what is important is how, on either conception, the highest good ultimately helps to explain how that object of hope, and so hope itself, works to sustain our moral motivation in the face of error and evil help. To see how that is the case, consider that, while Beiser is correct that Kant didn't envision the highest good as attainable through human efforts alone, this acknowledgment does not preclude its viability as a political ideal, or one that requires some collective struggle. Indeed, in *Religion Within the Boundaries of Mere Reason*, Kant explicitly frames the duty to bring about the highest good as a collective endeavor, for it "will not be brought about solely through the striving of one individual person for his own moral perfection but requires rather a union of such persons into a whole toward that very end."[137] Kant more precisely characterizes the resulting union as an "ethical society" or "Kingdom of Virtue," wherein all rational beings share the obligation to coexist.[138]

Kant delves into the implications of our duty to collectively advance the highest good in the third and fourth books of the *Religion*. There, he envisions a universal ethical society where rational beings, united by virtue, constitute an "absolute ethical whole."[139] Kant contends that there is a moral obligation for individuals to join forces with other rational beings, describing it as a "duty sui generis of the human race toward itself." This duty involves establishing a society dedicated to the common pursuit of the highest good, providing a formidable defense against radical evil.[140] While a juridico-civil state involves individuals coexisting under enforceable

[136] Reath (1988: 601).
[137] RGV 6:97-8.
[138] RGV 6:95.
[139] RGV 6:96.
[140] RGV 6:97.

public laws, an ethical society, or an ethico-civil state, is a community where people abide by laws of virtue voluntarily, without coercion. In his political writings, Kant asserts our moral entitlement to compel other rational beings to live in a "rightful condition," safeguarding the outer freedom of all through government protection. In contrast, in his *Religion*, an ethico-civil state is not subject to coercion; its formation depends on the free choice of individuals to depart from what Kant terms the "ethical state of nature." This state represents a society where humans have yet to unite morally to overcome their inclination towards evil.

How might a universal ethical society coexist within the framework of a political state? For Kant, the solution to bringing about an ethical society within the framework of a political state lies in the formation of a church — a community of rational beings who perceive themselves as living "under divine moral legislation." Kant goes so far as to assert that the ethical society can only find realization in the structure of a church, emphasizing the non-secular nature of his conception of the highest good in the *Religion*. This is underscored by Kant's exploration of the relationship between the visible church and the invisible church, a distinction that marks his clearest reference to a theological conception of the highest good. Within Kant's philosophy of religion, the invisible church represents the complete actualization of the highest good through a "union of all upright human beings under direct yet morally governed divine authority." Termed "invisible" because it is "not the object of a possible experience," this church aligns with the idea of the "Kingdom of God on earth."[141] Unfortunately, its manifestation in the sensory world is indefinitely deferred, existing in an unseen distance.[142]

Nonetheless, the mere noumenal potential for realizing a universal ethical society does not exempt rational beings from the duty to establish a phenomenal "union of human beings into

[141] RGV 6:101.
[142] RGV 6:151.

a whole that accords with this ideal" — namely, a visible church enabling collective promotion of moral conduct in society.[143] Despite the commendable creation of visible churches, it is crucial for members not to lose sight of the fact that any institutional church remains an imperfect realization of the highest good, merely a "schema of the invisible Kingdom of God on earth."[144] Kant's philosophy of religion rejects the notion that the highest good can be fully achieved through human agency, portraying institutional ethical societies as mere approximations. He goes as far as to argue that viewing a visible church as the actual embodiment of the Kingdom of God on earth is a form of "religious illusion."[145] In the fourth book of the *Religion*, Kant acknowledges the admirable nature of all churches as "attempts by poor mortals to give sensible representation to the Kingdom of God on earth," yet he also assigns blame when, within a visible church, they mistake the form of representation for the essence of the idea.[146] Members who believe in realizing the Kingdom of God on earth through their religious endeavors may be moral saints, but according to Kant, they are nonetheless susceptible to "madness."

In Kant's view, churches contribute to the cultivation of a moral life for their members through at least three avenues. First, they serve as a supportive network offering advice, aiding individuals in navigating moral dilemmas, and fortifying their disposition to adhere to the moral law. As noted by Reath, the public backing of one's commitment to ethical principles, coupled with the assurance of similar commitments from others, reinforces one's moral resolve.[147] Second, churches function as social environments where individuals engage with each other as collaborators rather than adversaries. In contrast to the competitive nature often inherent in the pursuit of personal interests, religious associations foster a sense of cooperation, allowing

[143] RGV 6:101.
[144] RGV 6:151.
[145] RGV 6:168.
[146] RGV 6:175.
[147] Reath (1988: 617).

members to mutually benefit from each other's progress toward virtue. Unlike other social formations where personal success may come at the expense of others, religious communities emphasize shared advancement. Thirdly, and in connection with the second point, churches actively promote solidarity and mutual assistance. Through acts of goodwill, they afford individuals the opportunity to fulfill their duty of contributing to the happiness of others, particularly by assisting those in need.

To recap, where we find the clearest discussion of the theological dimensions of Kant's notion of the highest good, we find the imperative of joining forces with others to attain the moral fortitude essential for leading a virtuous life. Kant contends that churches facilitate this collective moral strength by fostering mutual moral support among rational beings. In this way, the theological conception of the highest good explains how that object of hope helps to sustain our moral motivation. Nevertheless, as I aim to show, churches are not the exclusive associations required for fulfilling the duty to advance the highest good. In addition to establishing connections within ethical communities, rational beings must also coexist in a rightful condition with others and engage in a federative union of political states dedicated to safeguarding freedom through perpetual peace. Let us turn, then, to Kant's secular conception of the highest good, showing how on that conception, too, the highest good as an object of hope works to sustain our moral motivation.

In his political writings, Kant outlines our responsibility to advance the highest good in seemingly secular terms as "the duty to bring it about as far as we can that such a relation (a world in keeping with the moral highest ends) exists."[148] Entering what Kant describes as "a rightful condition" – achieving, that is, perpetual peace through the formation of a federative union of states "having as its only purpose the avoidance of war" – is a necessary (albeit not

[148] PP 8:280.

sufficient) condition of the realization of the highest good.[149] Until rational beings coexist in a federative union that respects sovereignty peacefully, the world will lack perfect virtue and happiness, and the highest good will remain elusive. Moreover, as in the case of its theological dimensions, Kant's doctrine of the highest good can motivate political agents in fulfilling their duties by reinforcing their commitment to acting justly.

Let's delve into how, for Kant, forming a lawful association — that is, a political state where the outer freedom of all is ensured — is a prerequisite for the highest good. If this is the case, then rational beings can only advance the highest good in a political state where citizens enjoy freedom. Kant, in the Metaphysics of Morals, asserts our duty to depart from the state of nature. As he writes, "a civil constitution, though its realization is subjectively contingent, is still objectively necessary, that is, necessary as a duty."[150] Refusal to enter a rightful condition implies neglect of a rational imperative, for, Kant asserts, "You ought to enter this condition."[151] Virtue, being integral to the highest good, necessitates life in a free political state. Deprived of outer freedom, fulfilling duties of virtue becomes nearly impossible. Kant argues that rational moral agents must wish to possess outer freedom, and only a rightful condition can secure this freedom. Hence, living in a free political state becomes indispensable for pursuing the highest good.

Moreover, Kant views rightful conditions not only as necessary for the highest good but also as efficiently promoting the happiness and virtue of those governed by the rule of law. In the Doctrine of Virtue, Kant delineates the obligatory ends that rational beings should pursue. The highest good, as he defines it, encompasses a world in which both these moral objectives and the prudential goals individuals establish for themselves in their pursuit of happiness find fulfillment. The establishment of a political state is geared towards ensuring that each rational

[149] PP 8:285.
[150] PP 8:264.
[151] PP 8:306.

being possesses the necessary means to attain both their moral and prudential objectives. Rightful conditions secure the outer freedom vital for fulfilling moral and prudential ends. A human lacking outer freedom would struggle to meet moral obligations, showcasing the importance of political states in providing the means for moral and prudential pursuits. It is evident that a person lacking outer freedom would face significant challenges in fulfilling their duties of virtue. For instance, a slave confined to the house would find it notably difficult to meet their moral obligation to, not least of all, cultivate their natural talents. Considering the normativity of instrumental reason, it becomes apparent that it would be irrational for an individual to embrace the duty to promote the highest good without simultaneously desiring the means necessary for realizing both moral and prudential ends. Kant asserts this perspective, stating that "whoever wills the end, also wills (insofar as reason has decisive influence on his actions) the indispensably necessary means to it that are within his power."[152] Consequently, a rational moral agent aspiring to fulfill their moral and prudential aims must implicitly wish to possess the outer freedom essential for such pursuits. As Kant firmly maintains that only a rightful condition can secure the outer freedom of individuals, it follows that a rational moral agent cannot reject living in a free political state; citizenship is a necessary enabling factor for the pursuit of the highest good.

The political implications stemming from the duty to promote the highest good go beyond the mere obligation to enter a rightful condition. Rational beings not only bear a moral obligation to inhabit a political state that ensures a certain level of outer freedom, but also, and crucially, the political states to which they belong are duty-bound to form a federative union with the citizens of other states. This union is envisioned to lead to a continuous approximation of the

[152] GMS 4:417.

highest political good—perpetual peace.[153] Only within a federative union which is "given a priori and is necessary by principles of right" is "the harmony of politics with morals" possible.[154] To be clear, Kant does not equate perpetual peace with the highest good; only the highest good, not perpetual peace, demands the proportional alignment of happiness with virtue. Rather, the point is that achieving perfect virtue, a prerequisite for the highest good, requires us to strive for peace. While ordinary citizens are not the primary bearers of this duty, but rather political leaders, the point remains that if the highest good requires virtue and virtue necessitates life in a political state and the pursuit of perpetual peace, then the highest good inherently mandates life in a political state and the endeavor for perpetual peace. In this sense, our political duties serve as necessary conditions for the highest good.

Yet, how does the awareness that living in a rightful condition and renouncing war are prerequisites for the highest good impact the practical lives of political agents? First and foremost, it is crucial to acknowledge that engaging in politics can be particularly disheartening for those dedicated to promoting justice. Kant's suggestion that politics is seldom the domain of genuinely moral politicians resonates with many. Instead, it often becomes a sphere dominated by what he terms "political moralists" — individuals who "deal in machinations" to gratify their self-serving desires while feigning adherence to the principles of right and virtue.[155] Unfortunately, politics frequently rewards immoral behavior, contrasting the struggles of peace activists and NGO workers striving for altruistic goals, sometimes at the expense of their own freedom. Corrupt and power-hungry politicians, meanwhile, ascend to power. On a global scale, nations with violent and oppressive historical records often thrive today, and leaders who disregard peace occupy influential seats in international organizations. In essence, if there is one

[153] MS 6: 355
[154] PP 8: 385
[155] PP 8:373.

realm where Kant's insight that happiness does not seem proportionate to virtue holds true, it is undoubtedly politics.

For political leaders who choose to prioritize morality over amoral or immoral politics, and who often face setbacks from the unscrupulous proponents of *Realpolitik*, Kant's notion of the highest good can serve as a source of moral solace. As stated earlier, this notion rests on the practical postulate that virtuous agents will ultimately be rewarded with happiness, even if this reward is deferred to a future life. It provides a basis for believing that the prosperity of political moralists and the adversity faced by moral politicians are transient, as the world extends beyond our current observations in nature. Although the realization of the Kingdom of God on earth is postponed to an unseen distance, faith in an eventual state of affairs where happiness aligns with virtue persists. Moreover, by positing the existence of an omnipotent being ensuring this alignment, Kant's highest good can free political leaders from the perilous illusion that governments should actively strive to proportion happiness to virtue, such as through punitive wars against states deemed unworthy. In this regard, Kant's notion of the highest good imparts a profound lesson in political humility: the role of political leaders is to secure the outer freedom of all through republican government and peace, rather than asserting self-conferred moral superiority over foreign nations. And this humble self-conception can be a source of comfort in the face of political machinations that distort or subvert morality.

However, the postulate for the possible existence of God and the requirement for proportionality are not the sole components of Kant's notion of the highest good that can serve as motivational forces in politics. Equally significant is Kant's depiction of a universal ethical society in which all rational beings abide by laws of virtue. Earlier, I noted that a considerable part of our lives is spent in competition for social distinctions. Yet, when rational beings engage

in war, their interactions go beyond treating each other as social adversaries; they become enemies prepared to take lives. To psychologically reconcile with the atrocity of war, combatants often undergo a process of moral depreciation, making it bearable for them to end others' lives. Convincing oneself that killing is morally acceptable involves portraying enemies as criminals or infidels and willfully forgetting that they are fellow rational beings. The historical narrative of politics teaches us that religious differences frequently contribute to the mental process whereby enemies gradually lose the dignity and respect owed to each other. Kant's discourse on the universal ethical society acts as a philosophical antidote to this process and to religious discrimination at large, fostering a sense of universal peerhood. It explicitly reminds us that, despite the temptation to exclude members of other religions from equal moral consideration, the highest good necessitates the formation of a universal union with rational beings of all faiths. As Kant elucidates in Perpetual Peace, "there can indeed be historically different creeds... and just as many different religious books... but there can be only one single religion for all human beings and in all times."[156] While we may harbor diverse sectarian affiliations, we are nonetheless bound by the same universal moral requirements and the duty to unite with others in our collective struggle against radical evil.

These observations might appear idealistic, and the political demands of Kant's practical philosophy can seem utopian. However, a closer examination of everyday political thinking reveals its foundation on a broad spectrum of regulative ideas. The mere possibility of these ideas is enough to instill in us a sense of duty to promote them. For instance, acknowledging the unlikelihood of completely eliminating gun violence in our country does not deter us from making reasonable efforts toward that goal. Similarly, envisioning a democracy without a gap between the will of the people and decisions of its representatives can inspire political reforms,

[156] PP 8:368.

even if skepticism lingers about closing such a gap. Kant's entire philosophy is grounded in the notion that representing ideals beyond human cognition or agency can still be practically effective. Just as the representation of a nature governed by empirical laws motivates the pursuit of theoretical knowledge, the idea of a world where moral subjects are virtuous and happy can guide our political conduct. Thus, Kant's practical philosophy aligns with our ordinary ways of thinking, marked by the hope that diligent efforts bring us closer to realizing our goals. Like our perspective, his concept of the highest good relies on both moral faith and the necessity of human action. Kant's notion of the highest good calls for the creation of a federative union of free political states that empowers citizens to pursue virtue and happiness by safeguarding their outer freedom.

In this section, I have examined two conceptions of Kant's notion of the highest, that which he argues is the object of rational hope. Those two conceptions of the highest good result in two kinds of hope: one is religious hope, which refers to the hope for divine assistance in performing a "revolution of the will" through which we overcome our natural propensity to evil and adopt the moral law as our fundamental maxim; the other is historical hope, which refers to the hope for the collective progress of the human species towards a more just and peaceful world, conceived as a distinctly shared and historical task. But whether viewed through a theological or secular lens, the highest good rests on human reason's hope that the steadfast pursuit of moral perfection will bring about happiness. So, analyzing the object of hope brings us to the same conclusion as when I surveyed attempts to analyze hope in terms of its underlying mental state. That conclusion is that hope, whether through its distinctive underlying mental activity or its object, importantly serves to sustain our moral motivation.

4. Hope and the highest good in the 'Dialectic of Pure Practical Reason'

Up to this point in the chapter, I have examined two different strategies towards giving an account of hope. One of these analyzes hope in terms of its underlying mental state, the other focuses on the object of hope, which, for Kant, is the highest good. What I hope to have shown is how both of these strategies emphasize that one of the functions of hope is to sustain our motivation against threats of despair or other obstacles. In this section, I both build on the previous one in which I discussed two conceptions of Kant's notion of the highest good and move closer towards my reconstruction of Kant's pedagogy of hope. More specifically, I will closely examine Kant's remarks on the highest good in the *Critique of Practical Reason*. My aim is to show how those remarks support the claim that the highest good, as the object of rational hope, serves to sustain our moral motivation, which sets up the rationale for the pedagogical method Kant presents at the end of the second Critique.

The Preface to the second Critique is where Kant first mentions the highest good. After explaining why his book does not amount to a critique of *pure* practical reason, but rather a critique of practical reason. He asserts that this critique, together with the existence of pure practical reason, proves transcendental freedom, a concept that, in his opinion, is "the *keystone* of the whole edifice of a system of pure reason, even of speculative reason."[157] In fact, Kant argues that freedom is proved by what he calls "fact of reason," that is, our awareness of the authority of the moral law.[158] The necessity of concepts of God and immortality is established through practical reason, as well, although as mere postulates, a qualification which I will explain in more detail later on. Among these concepts – freedom, God, and immortality of the soul – Kant claims that the first (freedom) is "the only one whose possibility we *know* a priori," since it is the

[157] KpV 5:4.
[158] The fact of reason, which is the practical basis for the belief that we are free, is first mentioned at KpV 5:6. Kant discusses it in greater detail at KpV 5:31.

condition of the moral law to the extent that the moral law would be illegitimate without the existence of freedom.[159]

Whereas freedom is the condition of the moral law, the concepts of God and immortality are not. They are, instead, "conditions only of the necessary object of a will determined by this law, i.e., conditions of the merely practical use of our reason[,] … conditions for the application of the morally determined will to its object that is given to it a priori (the highest good)."[160] Here, Kant initially presents the highest good as both a necessary object of the will determined by the moral law and an object given a priori by reason. But, for Kant, just as freedom is the condition of the moral law, the concepts of God and immortality are necessary conditions for the highest good. Hence, while we cannot obtain theoretical knowledge of these concepts, morality leads us, nonetheless, to assume their possible existence. In other words, for Kant, it would be unreasonable to not assume, or postulate, the possible existence of God and immortality, since these are conditions of that object – the highest good – which "one ought to set unfailingly as the aim of one's conduct."[161]

Moving past its Preface, the *Critique of Practical Reason* is, like the first Critique, divided into two parts, a 'Doctrine of the Elements of Pure Practical Reason' and a 'Doctrine of the Method of Pure Practical Reason.' I will examine the latter of these in the next chapter. In the remainder of this section of this chapter, I will focus on Kant's remarks on the highest good in the 'Doctrine of Elements,' which is yet further divided into two parts, an "Analytic of Pure Practical Reason" and a "Dialectic of Pure Practical Reason." I will first briefly comment on the former (that is, the Analytic), though it is the latter (the Dialectic) that is most pertinent for my purposes in showing how Kant's remarks on the highest good bear on his account of hope.

[159] KpV 5:4.
[160] Ibid..
[161] KpV 5:5.

In the "Analytic" of the second Critique, Kant mentions the highest good no fewer than two times. Kant tells us how "the moral law transfers us, in [our] idea, into a nature in which pure reason, if it were accompanied by the physical power adequate to it, would produce the highest good."[162] The basic idea here is that the highest good, which follows from the moral law, is ultimately possible under the existence of pure reason and some "physical power adequate to it," one beyond our own, as was argued in the first Critique and as will be reiterated in the 'Dialectic of Pure Practical Reason.'[163] Kant also mentions the highest good in relation to the ancients philosophers, who he thinks mistakenly regard the highest good as the determining ground of the will, without having first established the moral law as its unconditioned condition. This, Kant argues, leads to heteronomy, which is precisely what he tries to avoid by establishing that it is the moral law that determines what is good, and not the other way around. While he will go on to assert the importance of the concept of the highest good for morality, it is important, from his perspective, that this concept comes later. In his words, the highest good is an object which can be represented as the object of the will, only after the moral law has been justified as the immediate determining ground of the will. Precisely how the highest good is to be represented as an object of the will is the goal of the 'Dialectic of Pure Practical Reason.'[164]

It is in the Dialectic of the second Critique that Kant presents his most developed, though not to say unproblematic, account of the highest good.[165] Before turning to the Dialectic proper, Kant makes a key, though seemingly surprising claim: the highest good can be considered the determining ground of the will without bringing heteronomy into morality. How can that be the

[162] KpV 5:43.
[163] KrV A813/B841.
[164] KpV 5:64.
[165] As I mentioned earlier, dialectic is a common facet of the architecture of Kant's critical philosophy. Each of his three Critiques includes its own dialectic, though the extent the purposes and arguments of these are the same is open to debate. Eric Watkins (2010) draws on parallels with the Dialectic of the first Critique to reconstruct Kant's argument in the 'Dialectic of Pure Practical Reason.'

case when, in the Analytic, Kant had chided the ancients for making the highest good the determining ground of the will? To be sure, he still maintains that the highest good, as object of a pure practical reason, cannot be considered the determining ground of the will, since no empirical object can determine the will without bringing heteronomy with it. Nonetheless, he asserts that there is a sense in which the highest good can be the determining ground of the will. This is possible, Kant says, since the moral law is part of the highest good; indeed, it is its "supreme condition."

> It goes without saying, however, that if in the concept of the highest good the moral law as supreme condition is already likewise included, then not only is the highest good [the] *object* of the pure will, but the concept of this good and the presentation of its existence as possible through our practical reason are also at the same time the *determining basis* of the pure will; for then it is in fact the moral law, already included and likewise thought in this concept, and not any other object, which determines the will in accordance with the principle of autonomy.[166]

In this way, the highest good is both the determining ground of the will and the object of pure practical reason, without undermining the autonomy of the will, according to Kant. This move will be important when it comes to explaining how hope can play a sustaining role in our moral strivings for the highest good without thereby making those strivings heteronomous.

Kant goes on to clarify the distinction between the "complete good" and the "supreme good," recognizing that both could be identified as the highest good and hence lead to confusion. The supreme good is virtue, since it is the "condition which is itself unconditioned." However, virtue cannot be regarded as the complete good, that is, "as the object of the power of desire of rational finite beings."[167] Instead, the complete good must also include, as one of its components, happiness. To establish this, Kant refers, as he did in his discussion of the highest good in the

[166] KpV 5:109.
[167] KpV 5:110.

Critique of Pure Reason, to the idea of an "impartial judge."[168] Any such judge with sufficient power to proportion virtue in accordance to happiness will find it improper not to do it, according to Kant. This is how he arrives at his first substantial definition of the highest good in the second Critique:

> virtue and happiness together amount to possession of the highest good in a person, and thereby happiness distributed [to persons] quite exactly in proportion to [their] morality (as a person's worth and his worthiness to be happy) amounts also to the highest good of a possible world, the highest good means the whole, [i.e., it means] the complete good.[169]

So, the highest good is composed of two components – virtue and happiness – with virtue as the "supreme good," as we said earlier. Happiness is also good, but not unconditionally so. Note that Kant also asserts that the highest good can be viewed from two mutual perspectives: that of a person as a member of an ideal world, and that of an ideal world as a whole, all of its members included.

With this basic definition of the highest good in hand, the next question Kant addresses is how the highest good is practically possible. According to Kant, the unity of virtue and happiness in the highest good must be either analytic or synthetic:

> Two determinations necessarily linked in one concept must be connected as basis and consequence, and so connected, moreover, that this unity is regarded either as analytic (logical connection) or as synthetic (real connection), the former according to the law of identity, the latter according to the law of causality.[170]

The Stoics and the Epicureans, according to Kant, considered that the unity of virtue and happiness was analytic, i.e. that the pursuit of happiness and of virtue were one and the same. Those two schools of thought only differed when it came to deciding which of the two elements was more fundamental and grounded the other. For the Epicureans, happiness was fundamental,

[168] KrV A813/B841.
[169] KpV 5:110.
[170] KpV 5:111.

and so they argued that to seek happiness is to be virtuous; whereas, for the Stoics, virtue was fundamental, so they proposed that striving to be virtuous leads to happiness. In relation to this, Kant explains,

> The Stoic asserted that virtue is the *whole highest good* and happiness is only the consciousness of the possession of this virtue as belonging to the subject's state. The Epicurean asserted that happiness is the *whole highest good* and virtue is only the form of the maxim [whereby one is] to pursue this happiness, [consisting,] viz., in the rational use of means to it.[171]

Nevertheless, both the Epicureans and the Stoics were mistaken, according to Kant, to the extent that they sought an analytic unity in two heterogeneous forces. Virtue and happiness are, in fact, so heterogeneous, he observes, that they often collide in a person's ordinary life. That is why the moral law requires that we restrain our inclinations when the satisfaction of them implies, for instance, treating others or ourselves as mere means, and not as ends in themselves. In this way, virtuous action makes us worthy of happiness, but often at the cost of renouncing happiness. Now, with this being the case, Kant says that he still has not arrived at an answer to the question, "how is the highest good practically possible?"[172]

If the analytic union of virtue and happiness is not possible, then that unity must be synthetic. That leaves us with two possibilities: either the desire for happiness is the basis of morality, or virtue causes happiness. Kant rejects the first option, on the grounds that our inclinations are the basis of heteronomy, which is opposed to the autonomy that morality requires. And he says the second option is also impossible, since virtue does not necessarily produce happiness. Happiness is a consequence of rules of prudence, not of the moral law. Now, if the highest good happens to be impossible, Kant argues, the moral law must be regarded as false, as "fantastic":

---

[171] KpV 5:112.
[172] Ibid.

> since the promotion of the highest good, which contains this connection in its concept, is an a priori necessary object of our will and inseparably bound up with the moral law, the impossibility of the first must also prove the falsity of the second. If, therefore, the highest good is impossible in accordance with practical rules, then the moral law, which commands us to promote it, must be fantastic and directed to empty imaginary ends and must therefore in itself be false.[173]

This point clearly reiterates what Kant had previously asserted in the first Critique.[174] What is at stake is nothing less than the moral law, consciousness of which establishes the idea of freedom, the "*keystone* of the whole edifice of a system of pure reason, even of speculative reason."[175] So, it is not only the authority of the moral law that is at risk, but the whole system of Kant's critical philosophy. This explains the urgency of a proper solution of the antinomy, i.e. of this apparent conflict of practical reason with itself.

Kant's solution is developed later in chapter two of the Dialectic. He explains that, to say that virtue is the cause of happiness, is only conditionally false for humans, since we are rational, finite creatures who are members of not only the sensible world but the intelligible world, as well. That we also partake of the intelligible world is shown by the fact that we can think of ourselves as free in the world of understanding, mainly thanks to the moral law, which is a "purely intellectual determining basis of my causality (in the world of sense)."[176] Hence:

> it is not impossible that the morality of [one's] attitude should have a connection, and moreover a necessary one, as cause with happiness as effect in the world of sense, if not a direct connection then still an indirect one (by means of an intelligible originator of nature), a linkage which in a nature that is merely an object of the senses can never take place except contingently and cannot be sufficient for the highest good.[177]

From this it follows that the highest good, i.e. the synthetic union of virtue and happiness, in which the former produces the latter, is practically possible. The validity of the moral law is thus

[173] KpV 5:114.
[174] KrV A811/839.
[175] KpV 5:4.
[176] KpV 5:115.
[177] Ibid.

reasserted, this time in relation to its highest end. In this passage, we also see that the connection between the two elements is a mediated one, since, as was stated earlier, virtue does not necessarily grant happiness in our ordinary lives. What mediates that connection are the postulates of pure practical reason concerning God and the immortality of the soul.

At this point, Kant introduces a concept for a distinctive "liking for one's existence," namely "*self-satisfaction*," presumably as an example of how virtue can be the immediate source of some enjoyment distinct from that which is associated with happiness.[178] Contentment is the (negative) satisfaction with one's existence that accompanies the consciousness of striving to become virtuous. This feeling of contentment is important for my purposes, since I will later associate that feeling with a kind of hope for moral progress felt on the part of students subjected to Kant's moral pedagogy. For now, note that, despite the fact that the feeling of contentment is the result of putting the moral law above our inclinations, this cannot be a proof by example of virtue as source of happiness, because happiness is the satisfaction of our inclinations. In other words, happiness is a positive feeling, whereas contentment is a negative one, in that it is the result of freedom from inclinations, not their satisfaction. However, Kant also wants to clarify that contentment is not yet beatitude, which requires complete independence from the inclinations. Nonetheless, according to Kant, contentment is closer to beatitude than to happiness, since both beatitude and contentment have the same origin – independence from inclinations – whereas contentment and happiness have different sources and hence different natures.

To briefly take stock, Kant's stated purpose in the Dialectic is to resolve an antinomy of practical reason. That antinomy emerges from the recognition that there are not one but two ways in which morality and happiness can be linked in the concept of the highest good: the synthetic

[178] KpV 5:117-8.

connection can go from happiness to morality or, alternatively, from morality to happiness. These represent the thesis and antithesis that make up the antinomy of practical reason. Kant's resolution is to argue that while the thesis is "*false absolutely*," the antithesis is "false only *conditionally*."[179] The antithesis – that virtue produces happiness – is false only insofar as existence in the sensible world is taken to be the only kind of existence. Put another way, while virtue does not necessarily produce happiness in this world, it can make us worthy of happiness, a worthiness that is felt in the negative satisfaction of contentment with oneself as exercising freedom from one's sensible inclinations. After claiming to have solved the antinomy of pure practical reason by showing that the highest good can be thought of as possible, Kant wants to "to exhibit the bases of that possibility," both in relation to what is in our hands, namely virtue, and what is in greater measure beyond our power, namely happiness as a reward for that virtue.[180]

Kant's postulates of pure practical reason – of immortality and God – are meant to provide the grounds for the possibility of the highest good. In other words, those postulates are to show how we may hope for happiness as a result of our virtue. The practical necessity of the postulate of the immortality of the soul is defended as follows. As was already stated, the supreme condition for the highest good is virtue. But virtue to what degree? For Kant, what is required is "[c]omplete adequacy of the will to the moral law," or holiness.[181] This creates a problem, since human beings, who are rational but finite beings, are not capable of holiness at any time of their lives. On the contrary, holiness can only be thought of as possible if we assume endless progress, which in turn is only possible in beings possessing immortal souls. Therefore, Kant asserts, "the highest good is practically possible only on the presupposition of the

[179] KpV 5:114.
[180] KpV 5:119.
[181] KpV 5:122.

immortality of the soul, and hence this immortality, as linked inseparably with the moral law, is a **postulate** of pure practical reason."[182]

This idea of holiness as attainable only through an endless progress is useful for Kant not only because it extends some degree of legitimacy to an idea (namely, the immortality of the soul) that was not possible to secure from the speculative perspective, but also for religious purposes. Without this postulate (and hence without this conception of our moral vocation) one could either become indulgent with respect to the moral law, or else get the false illusion that it is possible to attain holiness in this life. Both options, according to Kant, hinder moral progress:

> [B]oth of which [consequences] only prevent the unceasing *striving* toward meticulous and thoroughgoing compliance with a strict and unforbearing but nonetheless true rather than ideal command of reason. For a rational but finite being only the progression *ad infinitum* from lower to the higher levels of moral perfection is possible.[183]

This justification of the postulate of immortality makes it reasonable for rational, finite beings to hope for that immortality, since that will allow us to think of ourselves as approaching holiness, which in turn is commanded by the moral law and is contained in the highest good as its supreme condition. In Kant's words, humans "hope for a further uninterrupted continuation of this progress, however far his existence may extend, even beyond this life."[184]

Whereas the postulate of immortality grants what is in greater degree within our power in relation to the highest good – i.e., virtue – the postulate of God is to show how the second element – i.e., happiness in proportion to virtue – could be possible. The argument begins with a definition of happiness: "Happiness is the state of a rational being in the world for whom in the whole of his existence everything proceeds according to his wish and will; it therefore rests on the harmony of nature with his whole purpose as well as with the essential determining basis of

[182] Ibid.
[183] KpV 5:123.
[184] Ibid.

his will."[185] But the moral law, as established in the Analytic of the second Critique, commands independently from, and sometimes contrary to, happiness. So, there is no obvious way to think of virtue as a sufficient cause of happiness in the natural world. In light of that, Kant denies that humans can bring about that connection completely on their own. In his words, a being that is rational but that belongs to the sensible world "cannot through his will be the cause of this nature and, as far as his happiness is concerned, cannot by his own powers make it harmonize throughout with his practical principles."[186] Nevertheless, the highest good must be possible, since the moral law commands that we ought to promote it (and we know that, for Kant, ought implies can).[187] At this point, Kant argues that the only plausible possibility is to refer to the concept of an author of nature, i.e. God, considered as "a cause of nature as a whole, distinct from nature, which contains the basis of this connection, namely the basis of the exact harmony of [one's] happiness with [one's] morality."[188] And that is how "the postulate of the possibility of the *highest derivative good* (the best world) is simultaneously the postulate of the actuality of a *highest original good*, viz., [the postulate] of the existence of God."[189]

It is important to recognize that, for Kant, there is no obligation to believe in God, despite the above deduction. What he concludes is that it is morally necessary to assume the existence of God, but that "this moral necessity is subjective, i.e., a need, and not objective, i.e., itself a duty; for there can be no duty whatever to assume the existence of a thing (because doing so concerns only the theoretical use of reason)."[190] God is not the ground of morality; that position is held

[185] KpV 5:124.
[186] Ibid.
[187] For one expression of 'ought implies can,' see Kant's remarks on the so-called 'Gallows Man': "He judges, therefore, that he can do something because he is conscious that he ought to do it, and he cognizes freedom within himself—the freedom with which otherwise, without the moral law, he would have remained unacquainted" (KpV 5:30).
[188] KpV 5:125.
[189] Ibid.
[190] Ibid.

solely by the principle of autonomy. From the standpoint of speculative philosophy, the postulate of God is just a hypothesis. However, because that postulate shows how the highest good is possible, it is elevated to the condition of a "pure *rational faith*."[191] This subjective character of the postulate of God (and immortality) will be important later when I explain how Kant's true purpose in showing that the highest good is practically possible has less to do with the objective basis for morality and more to do with staving off the threat of despair that stands in the way of us becoming virtuous.

This explanation for the possibility of the highest good also allows Kant to explain why the Stoics and the Epicureans failed in their attempt to explain that possibility in relation to morality. The Epicureans failed from the beginning, because they chose the wrong principle (viz. happiness) as the ground for morality. The Stoics, on the other hand, were right in the principle they identified (viz. virtue), but committed two mistakes. First, they thought that perfect virtue was possible in this life, and second, they took contentment as equivalent to happiness; that is to say, the Stoics left happiness out of the highest good. From Kant's perspective, Christianity is, in at least these respects, a doctrine closer to his ethics.

The postulate of God, like the postulate of immortality, makes it reasonable for rational, finite beings such as ourselves to hope for happiness in proportion to virtue, i.e. the highest good. According to Kant, "we can hope to reach the highest good, which the moral law makes it our duty to posit as the object of our endeavor, only through a will that is morally perfect (holy and benign) and simultaneously also all-powerful, and thus through harmony with this will."[192] This does not lead to heteronomy, Kant reiterates, since the highest good, as a necessary object of a morally determined will, follows from the moral law. The moral law is not a doctrine of how to

[191] KpV 5:126.
[192] KpV 5:129.

become happy, but of how to become worthy of happiness. Nevertheless, there is a sense in which morals can be defined as a doctrine of happiness. As Kant explains,

> when morality (which merely imposes duties and does not provide us with guidelines for self-interested wishes), has been set forth completely, then—after the moral wish, based on a law, to further the highest good (to bring the kingdom of God to us), which could not previously have sprung up in any self-interested soul, has been awakened and for the sake of this wish the step to religion has been taken—then for the first time can this doctrine of morals also be called a doctrine of happiness, because only with religion does the *hope* for happiness first arise.[193]

From this it also follows, according to Kant, that the final end in God's creation of the universe is the realization of the highest good, i.e. happiness not as a result of mere beneficence but as a reward for virtue. In fact, if the final end were happiness resulting from mere beneficence, then God would be an object of love, but not of respect. His plan, then, consists in the gradual sanctification of human beings, together with the final "crowning [of] such a beautiful order with corresponding happiness."[194]

Kant dedicates comparatively much less space in the Dialectic to the postulate of freedom. This is so perhaps because freedom received nearly all of Kant's attention in the Analytic, and because it is not only a necessary presupposition for the possibility of the highest good but, indeed, the keystone of Kant's whole philosophical system. Without freedom, any other considerations about the highest good become meaningless. Hence, the postulate of freedom must have been established before the deduction of the highest good. As was mentioned earlier, Kant proves the possibility of freedom through that fact of reason, or our consciousness of the moral law as authoritative for us. So, Kant reasserts that the postulate of freedom follows "the necessary presupposition of independence from the world of sense and of the ability to determine one's will according to the law of an intelligible world, i.e., the law of freedom."[195]

---

[193] KpV 5:130.
[194] KpV 5:131.
[195] KpV 5:132.

Kant also reaffirms that postulate "displays through the moral law, and with this law also the law of an intelligible world, [a world] to which speculative reason could only point but the concept of which it could not determine)."[196]

The ideas of God, immortality, and freedom are only hypotheses from the speculative perspective. For, from that perspective, they answer a need of pure reason to attain the unconditioned condition in the cognitive sphere, that is a natural need for answers to the most fundamental human questions. From the perspective of reason in its practical use, there is also a need, but one that follows from a moral duty, namely that of promoting the highest good. Now, as I have explained, if it is a duty to promote the highest good, then it must be possible, and hence we have to figure out the conditions of that possibility, which Kant argues happen to be the same ideas of God, immortality, and freedom. He does also say that the duty to promote the highest good is objective, since we are compelled to obey it, without additional conditions, as rational beings, Nonetheless, Kant adds,

> the subjective effect of this law, viz., the *attitude*, adequate to it and also necessary through it, to further the practically possible highest good, nonetheless presupposes at least that this good is *possible*; otherwise striving for the object—of a concept that basically would be empty and without an object—would be impossible practically.[197]

The need of pure speculative reason led us to the ideas of God, immortality, and freedom as hypotheses that help us to satiate our unlimited desire for knowledge, albeit without theoretical certainty. The need of pure practical reason leads us to those same ideas, but as postulates, i.e. as necessary assumptions for the possibility of the highest good, whose promotion is commanded by the moral law. Hence, the postulates are not, as in the speculative realm, the consequence of curiosity seeking to satisfy some arbitrary investigations, but are instead regarded as necessary

[196] KpV 5:133.
[197] KpV 5:143.

according to the moral law. The postulates, hence, appear "for the sake not of a discretionary speculative aim but of a practically necessary purpose of the pure rational will. Here this will does not *choose*, but rather *obeys* an unremitting command of reason."[198]

Worried that the above explanation may lead to misinterpretation, Kant devotes the remainder of the Dialectic to clarifying how hope for the highest good is not a duty. For Kant, "practical reason must admit [the highest good] without solicitation," but from this doesn't follow a command to believe in the possibility of the highest good.[199] Speculative reason must concede it, he thinks, because "no one can wish to maintain that a worthiness—commensurate to the moral law—of rational beings in the world to be happy, as combined with a possession of happiness proportionate to this worthiness, is in itself *impossible*."[200] To deny the supreme condition of the highest good, i.e. virtue, is senseless since that amounts to the denial of the authority of the moral law, but things are different in the case of the second element, i.e. happiness in proportion to virtue. In the latter case, there is no need of a command for its possibility, but,

> only *the manner in which* we are to think such a harmony of the laws of nature with those of freedom has something about it in regard to which a choice belongs to us, because concerning it theoretical reason decides nothing with apodeictic certainty, and with regard to this manner there can be a moral interest that turns the scale.[201]

Reason finds it impossible to conceive of a necessary connection between virtue and happiness in the world, but this impossibility is only subjective. It is an impossibility of our reason. So, while we cannot conceive such a connection as a necessary one, we also cannot prove by scientific methods that there is no such connection. According to Kant, we are justified to resort,

---

[198] Ibid.
[199] KpV 5:144.
[200] Ibid.
[201] KpV 5:144-5.

at this point, to the postulate of God, though only as the consequence of "*subjective* conditions of our reason," namely that this is one way in which our reason can conceive of a perfect connection between virtue and happiness.[202] The manner in which this connection is possible, therefore, rests on a choice, "in which however a free interest of pure practical reason decides for the assumption of a wise originator of the world."[203] Nonetheless, since the promotion of the highest good and the assumption of its possibility are objectively necessary, "the principle which determines our judgment in this is the basis—*subjectively* indeed as a need, but simultaneously also as a means of furthering what is *objectively* (practically) necessary— of a *maxim* of assent for a moral aim, i.e., a *pure practical rational faith*."[204] In consequence, the postulate of God is not a command but a pure practical rational belief, a belief that can vacillate but never be denied as impossible.

This point about the postulate of God being subjectively chosen clarifies Kant's rationale for establishing the highest good by means of the postulates. Recall that, according to Kant, the impossibility of the highest good would make the moral law must be false, or "fantastic": "If, therefore, the highest good is impossible in accordance with practical rules, then the moral law, which commands us to promote it, must be fantastic and directed to empty imaginary ends and must therefore in itself be false."[205] On the face of it, the threat may appear to be that the unrealizability of the highest good undermines the authority of the moral law. In that case, the completion of the deduction of the highest good by means of the postulates appears to rescue the moral law from the threat of "fantasy." But if that were the threat at issue in the deduction of the highest good, then that deduction would either be circular or show the threat to have been empty.

[202] KpV 5:145.
[203] KpV 5:146.
[204] Ibid.
[205] KpV 5:114.

After all, as mentioned earlier, the fact of reason was supposed to have already established the authority of the moral law, and it is only because of that authority that we recognize that the highest good is an end that we are morally obligated to promote and, so, that must be possible. So, the morally necessary highest good creates the conditions of its own possibility, i.e. the postulates of God and immortality. If this is illegitimately circular, then it proves nothing, let alone the postulates, and morality succumbs to the threat of fantasy. Alternatively, if morality creates its own conditions of possibility, then the threat of fantasy undermining its authority was empty; the moral law never needed saving in the first place. So, the threat of fantasy for which the deduction of the highest good is a solution must be interpreted as something other than a threat to the legitimacy of the moral law.

Drawing on the discussion from earlier about the role of hope in motivation offers an alternative interpretation of the threat of fantasy motivating the deduction of the highest good. One main takeaway from examining the two strategies analyzing hope – one in terms of its underlying mental state, the other in terms of its object – was that a function of hope is to sustain our motivation against threats of despair or other obstacles. This suggests the deduction of the highest good is responding, not to the threat of the legitimacy of the moral law, but to the threat of demoralization. Given our limitations and propensities as rational, sensible beings, and in light of apparent injustices, we are susceptible to becoming demoralized in our efforts to sustain the moral motivation necessary for promoting the highest good. As Andrew Chignell writes:

> It's not that overwhelming injustice and our perceived inefficacy make us think (as a consequentialist might) that working for justice is no longer required (since doing so won't make a difference). Rather, it's that the rank injustice of the massive systems of which we're inexorably a part, combined with the awareness of our acute impotence to change them, threaten to sap our resolve to keep trying.[206]

[206] Chignell (2023: 53).

In other words, the threat of fantasy motivating the deduction of highest good does not concern the a priori status of the moral law, but rather an empirical-psychological claim about what we need in order to sustain our moral resolve and avoid demoralization. The postulates of immortality, God, and freedom are not meant to secure the authority of the moral law, but to give us some means of overcoming the threat of demoralization that undermines our ability to satisfy the demands of the moral law.

This reading of Kant's deduction of the highest good in the Dialectic is more plausible when we help ourselves to that passage from Kant's *Critique of Judgment* that I referred to earlier in this chapter about the psychology of an atheistic but "righteous man (like Spinoza)." Recall how, according to Kant, such an individual, despite being righteous, is still liable to be left despondent by the perception that his own moral strivings, as well as those of others, are inefficacious against large-scale structural evil and injustice:

> But his effort is limited… Deceit, violence, and envy will always surround him, even though he is himself honest, peaceable, and benevolent; and the righteous ones besides himself that he will still encounter will, in spite of all their worthiness to be happy, nevertheless be subject by nature, which pays no attention to that, to all the evils of poverty, illnesses, and untimely death, just like all the other animals on earth, and will always remain thus until one wide grave engulfs them all together (whether honest or dishonest, it makes no difference here) and flings them, who were capable of having believed themselves to be the final end of creation, back into the abyss of the purposeless chaos of matter from which they were drawn.[207]

As I explained in relation to this passage before, What Kant claims is that even a righteous individual cannot maintain the resolve needed to sustain their moral striving without some hope that, ultimately, those efforts will be efficacious. The reason why such an individual cannot maintain moral resolve has to do with Kant's idea, expressed elsewhere in the third Critique, that "[t]he attainment of every aim is combined with the feeling of pleasure."[208] Falling short of some

[207] KU 5:452.
[208] KU 5:187.

goal – such as becoming virtuous, or making oneself worthy of happiness in order to achieve the highest good – will result in displeasure or dissatisfaction. Typically, Kant associates the experience of pleasure with the happiness derived from fulfilling our desires. However, in works like *Groundwork*, Kant acknowledges that fulfilling our moral duties can also bring about "a feeling of pleasure or of delight"; in fact, he goes so far as to argue that "[i]n order for a sensibly affected rational being to will" what the moral law demands, capacity for this feeling is "required."[209] So, maintaining a healthy dose of moral pleasure or satisfaction is crucial for our perseverance in the pursuit of morality, especially when it clashes with our prudent pursuit of happiness in the current world. With the example of the Spinoza-like individual, Kant is emphasizing how interpreting our moral life as an endless progression toward the ultimately unattainable goal of holiness threatens to eliminate the possibility of moral satisfaction, posing a practical threat. What the postulate of the existence of God satisfies is a subjective need, one endemic to the human conditions – namely, the need to overcome despair or certain failure in our efforts towards becoming moral, or making ourselves worthy of happiness. Kant's discussion of the Spinoza-like individual is entangled with his teleological arguments which may or may not be relevant for my purposes. At any rate, the basic upshot of that discussion from the third Critique is the same as what I argued to be the point of the deduction of the highest good in the second Critique: the impossibility of the highest good does not directly threaten the legitimacy of the moral law, but rather our ability to sustain the motivation it demands of us and avoid demoralization.

This is not to say that the deduction of the highest good is intended by Kant to produce hope for that outcome. For Kant's question concerning hope, "What may I hope?" may be

[209] GMS 4:460.

understood to presume that some kind of hope is a default attitude of humans. As Rachel Zuckert observes:

> The phrasing of this question suggests that human beings already hope, or have tendencies to do so. Reason, then, is to establish which objects are proper or permissible for that preexisting attitude – the "what" that is its focus, and the "may" that governs or restricts it – not to produce it.[210]

Kant's arguments about immortality, God, or the possibility of the highest good in general clarify for us the possible, as well as necessary, objects of our moral strivings. They "just provide us with a 'way of thinking' about the world, with tentative (subjective, practical) grounds for belief." They are not meant to prove definite conclusions that persuade someone who is hopeless, but rather "to ward off "crafty hope" and "wild despair," to use Kant's language; that is, not to produce hope, but to prevent the corruption or loss of preexisting hope.[211] Drawing on the analysis from the previous chapter, we can understand crafty hope as a sort of self-conceit: it is the corrupted hope that one will receive salvation from God, or obtain the highest good, without engaging in moral striving. And, on the flip side, wild despair can be understood as a kind of despondence, in which one distorts moral principles in such a way that he can never be moral, particularly in light of inevitable mistakes or failures. So, Kant's point in establishing the possibility of the highest good is to prevent these corrupted forms of the hope for moral progress that is our default attitude, in order to sustain our moral resolve and avoid demoralization.

In this way, the argument of the Dialectic in the second Critique grounds our commitment to realizing the highest good in the psychological effects that can be induced in us by striving for that end. Indeed, this line of argument is made more explicit in Kant's subsequent writings.[212] I

[210] Zuckert (2018: 256).
[211] Ibid. The phrases Zuckert refers to, "crafty hope" and "wild despair," come from Kant's RGV 6:160 and 6:71, respectively.
[212] On my reading, there is more similarity between Kant's treatment of the highest good in the second Critique and his subsequent texts than other commentators have argued to be the case. For commentators like Sussman (2015) and Chignell (2023), Kant is not preoccupied with the threat of demoralization until the *Critique of Judgment* and on. However, if the reading I present over the course of this chapter and the next is accurate, then Kant did become

have already discussed the passage about the Spinoza-like individual from the *Critique of Judgment*, according to which it is not the authority of the moral law that is at stake in the highest good, but rather our ability to acknowledge that authority in a way that sustains our motivation and avoids demoralization. Kant is even clearer in *Religion within the Boundaries of Mere Reason* that:

> All human beings could sufficiently partake of this [moral] incentive too if they just adhered (as they should) to the rule of pure reason in the law: What need have they to know of the outcome of their doings and non-doings that the world's course will bring about? It suffices for them that they do their duty, even if everything were to end with life in this world, and in this life too happiness and desert perhaps never converge.[213]

In pointing us toward the highest good, "practical reason reaches beyond the law," so as to adapt the law to "one of the inescapable limitations of human beings and of their practical faculty of reason."[214] Similarly, in the third Critique Kant writes that the highest good is an end "irresistibly imposed" on human beings not by reason alone, but by "their nature (as finite beings)" in light of the demands of reason.[215] As he explains:

> Now, in this end human beings seek something that they can love, even though it is being proposed to them through reason alone. Hence the law that only inspires respect in them, though it does not recognize this sought-after something as [its own] need, nonetheless extends itself on its behalf to include the moral ultimate end of reason…[216]

So, it is not enough that we intellectually affirm the pursuit of virtue towards achieving the highest good as our end. Instead, we must orient ourselves towards that end in a way that can engage our passions, sustaining our hope that our moral strivings will be efficacious.

---

concerned about the threat of demoralization in the final stages of the second Critique, even if his thoughts on the topic were not yet fully fleshed out.

[213] RGV 6:7n.

[214] Ibid.

[215] KU 5:541.

[216] RGV 6:7n.

Of course, grounding our commitment to the highest good in our passions, even if only for sustaining our moral resolve, can seem inconsistent with the judgment Kant makes in some places about passions. For example, in *Anthropology from a Pragmatic Point of View*, he plainly states that they are "cancerous sores for pure practical reason" and "without exception evil."[217] And, in the *Metaphysics of Morals*, he insists we should try to rid ourselves of them all as much as we can, making no allowance even for those passions that might be morally salutary. For, he argues, true virtue involves a kind of "apathy" marked by "a tranquil mind with a considered and firm resolution to put the law of virtue into practice."[218]

But other remarks suggest that Kant is more amenable in his assessment of the role of feelings in moral motivation. In *Religion*, he claims that the "aesthetic constitution" of the virtuous must be joyous.[219] If our general frame of mind is such that we experience the demands of the moral law as an imposition or burden, then we are at risk of fostering a "hidden hatred" of that law. In "The End of All Things," Kant makes a similar point but in a way more pertinent to the role of hope in sustaining our moral motivation: "[W]hat one does not do with liking he does in such a niggardly fashion—also probably with sophistical evasions from the command of duty—that the latter as an incentive, without the contribution of the former, is not very much to be counted on."[220] In other words, we need a feeling of satisfaction to make our moral motivation sufficiently reliable. Such satisfaction is not just a motivational counter-weight against our non-moral incentives. Rather, this feeling of satisfaction enables us to act from duty glady, and so prevent us from deceiving ourselves about the true nature of morality's demands. Kant never denies that we may, in certain instances, experience the demands of morality as constraints, and

[217] ApH 7:266-7.
[218] MS 6:409.
[219] RGV 6:23-5n.
[220] EAD 8:338.

so fulfill them joylessly; indeed, in some cases, it may be better that we perform our duties without a feeling of satisfaction. Problems arise, however, when there is insufficient feeling of satisfaction to sustain our moral resolve against the threat of demoralization.

Ultimately, I think Kant found the arguments of the Dialectic to be inadequate for this purpose of achieving moral resolve. The postulates are supposed to satisfy a subjective need for some feeling of satisfaction that enables us to overcome despair and sustain the hope for the highest good. Kant's postulate of immortality, for example, is meant to explain how we can be said to at least appear before a divine judge to have attained this complete conformity to the moral law. Yet, endless progress resulting in only appearing to have attained that final end without ever truly attaining it, would rule out the possibility of true satisfaction in our moral life. Construing our moral life as endless progress toward the ultimately unreachable goal of moral perfection threatens to wipe out the prospect of moral satisfaction, which is disastrous from the practical standpoint where we need to sustain our motivation in light of certain failure. So, the deduction of the highest good by means of the postulates may present some intellectual tools for preventing the corruption of our default hope, but the threat of demoralization still looms at the end of the Dialectic.

This limitation of the arguments of the Dialectic sets the stage for the section that follows it, the 'Doctrine of Method for Pure Practical Reason.' The Doctrine of Method will be the central focus of my next chapter, but it will be useful to make some anticipatory remarks here that help to bring to a close this discussion of hope and the highest good. In short, the Doctrine of Method is where Kant outlines a moral pedagogical method for cultivating hope in young students in order to sustain their moral resolve and avoid demoralization. The hope resulting from pedagogical exercises is not identical to hope for the highest good in either its theological

or historical conception. As we will see in the next chapter, Kant's pedagogical method is crafted in such a way as to preserve, or perhaps nourish, a default hope similar to what Zuckert says is the function of the postulates. But, at least in the case of moral education, the default attitude is not hope for achieving the highest good, but rather our hope to be able to sustain respect for the moral law, as the source of moral motivation, in the face of hardships that threaten to demoralize us. That hope to sustain our moral resolve is what I will associate with a pleasure we take in a feeling of the harmony of our faculties that is brought about by the moral law. Unlike religious hope for divine assistance, or historical hope for collective progress of the human species, this pedagogical hope is for something in our control, which is our ability to continue to strive to become moral. The goal of Kant's pedagogical exercises is to cultivate that feeling of the harmony of our faculties so that young students can hope to avoid demoralization and continue to act from respect for the moral law. In this way, as I will argue in greater detail in the next chapter, the Doctrine of Method completes the aim of the 'Dialectic of Pure Practical Reason' of showing how we can overcome the threat of demoralization and sustain moral resolve necessary to make the highest good possible.

In this chapter, I examined two strategies taken by recent commentators to articulate Kant's account of hope. One of these strategies is to begin with a so-called standard account of hope, according to which hope is a composite attitude of a desire for some outcome and a belief that that outcome is possible, though not certain. I showed how some philosophers, including ones writing in a Kantian vein, seek to address some inadequacies in the standard account of hope. A

desideratum for such accounts of hope is to explain the apparent motivational function of hope to sustain our efforts in the face of despair. Another of the strategies I examined analyzes hope in terms of its object, which, for Kant, is the highest good. I explained how there are interpretative disputes concerning theological and secular conceptions of that notion, but that Kant viewed the highest good, however conceived, as an object of rational hope that sustains the moral motivation against the threat of demoralization. I showed how this function of the highest good, and of hope, is apparent in Kant's remarks and arguments concerning that notion in the Dialectic of the *Critique of Practical Reason*. This sets the stage for the next chapter, in which I reconstruct the pedagogical method Kant outlines in the subsequent chapter of the second Critique, the 'Doctrine of Method for Pure Practical Reason.' For, as I show, the purpose of that pedagogical method is to educate youth so as to sustain their respect for the moral law, as the source of moral motivation, in the face of hardships. There, Kant introduces a pedagogical dimension of hope, where the object of that hope is not the highest good, per se, but rather that component of the highest good that is in our control, namely, moral virtue. Thus, as we will see in the next chapter, Kant's pedagogical exercises seek to cultivate students' hope that respect for the moral law will constitute their character, and so continue to influence their future actions, which is a necessary condition for achieving the highest good.

Chapter 3: Hope in Kant's 'Doctrine of the Method of Pure Practical Reason'

My aim in this chapter is to, at last, share a close reading of the 'Doctrine of the Method of Pure Practical Reason' (hereafter, Doctrine of Method) in the *Critique of Practical Reason*. The centerpiece is a reconstruction of Kant's pedagogy of hope, that is, of the pedagogical method he prescribes to moral educators in the Doctrine of Method. One key passage for my purposes is where Kant most succinctly states the result of this method said to directly elicit a feeling of respect as an incentive for the moral law: "this incentive and the initially slighter attempts at using it, give hope of being brought about [more frequently], so as to produce in us little by little the greatest but pure moral interest."[221] My central claim is that the pedagogical method of the Doctrine of Method is designed to elicit not just students' respect for the moral law, but also their hope to sustain that respect and avoid the threat of demoralization.

In section one, I explain Kant's overarching aim in the Doctrine of Method. Briefly stated, Kant's goal is to show how the feeling of respect for the moral law can influence our maxims over time and in the face of demoralizing obstacles. That goal both advances the arguments from preceding sections of the second Critique and anticipates issues subsequently taken up in a more programmatic fashion in the *Critique of Judgment*. In section two, I describe the first of two pedagogical exercises that comprise the method Kant prescribes to moral educators in the Doctrine of Method. The task of that initial exercise is to use examples of moral conduct to sharpen students' moral judgment so that they acquire a sort of confidence in their capacities for said judgment. As I explain, this pedagogical exercise draws explicitly on the

[221] KpV 5:159.

theory of judgment from Kant goes on to develop in the third Critique – specifically, his account of judgments of beauty. In section three, I describe the second of the two moral pedagogical exercises Kant outlines in the Doctrine of Method. This exercise builds on students' newly acquired confidence to present them with carefully chosen examples of moral conduct to elicit not just a feeling of respect for the moral law, but also a hope that that respect will continue to have a lasting influence. This pedagogical exercise, too, draws on resources to be more fully flushed out in the third Critique, particularly his account of the sublime. To elucidate both of Kant's pedagogical exercises, I draw extensively on his theories of the beautiful and sublime. My purpose is not to defend any particular interpretation of those theories from the third Critique, but rather to use those theories to expound on Kant's relatively brief description of the pedagogical exercises from the second Critique that are my main focus.

1. Aim of the Doctrine of Method

The Doctrine of Method constitutes Part II of the *Critique of Practical Reason*, the overarching aim of which is "to establish merely that there is pure practical reason."[222] In other words, Kant sets out in the second Critique to show that reason itself contains an objective principle of conduct – viz., the moral law – which applies independently of individuals' preferences and contingent ends. Kant's stated aim in the Doctrine of Method is to present how "one can impart to the laws of pure practical reason admittance to the human mind and influence on that mind's maxims."[223] Perhaps more in speaking to the overarching aim of the second Critique, Kant restates this aim as being to show "the way in which one can make objectively practical reason subjectively practical as well."[224] To better understand this aim of the Doctrine

[222] KpV 5:3.
[223] KpV 5:151
[224] Ibid.

of Method, it is helpful to consider it in relation to Part I of the second Critique, 'The Doctrine of the Elements of Pure Practical Reason.' Ultimately, what I argue in this section is that the task of the Doctrine of Method is to provide moral educators a pedagogical method by which to orient students to hope for moral progress so that those students are aware not only that there is pure practical reason in them but also that that reason in them can overcome certain obstacles to their moral strivings.

The Doctrine of Elements, which immediately precedes the Doctrine of Method in the second Critique, is composed of two parts: an "Analytic of Pure Practical Reason" and a "Dialectic of Pure Practical Reason." In the Analytic, Kant is preoccupied with showing that the possibility of a pure moral incentive – viz., respect for the moral law – must be presupposed as a condition for the possibility of moral action. Particularly important for my purposes is Chapter III of the Analytic, 'On the Incentives of Pure Practical Reason.' There, Kant is concerned with what moral experience must be like for every rational being that has the capacity to feel pleasure and pain, regardless of their particular sensibility. The conclusion he arrives at is that, when confronted with the moral law, we feel a certain humiliation of the pretensions of our sensuous nature, which is felt at the same time as an elevation of our rational nature.[225] That is to say, in criticizing our tendency to overvalue the pursuit of happiness, we must also be aware of the capacity of our will to be determined independently of pathological incentives. That awareness elicits a kind of pleasure identified with the feeling of respect for the moral law. Since that feeling is brought about by our self-criticism and awareness of our autonomy, it can serve as an incentive for moral action.

In a previous chapter, I explained how the moral law becomes an incentive in terms of Kant's notion of self-conceit. Recall that, for Kant, self-conceit is a kind of self-love that, though

[225] KpV 5:74-5.

a hindrance to morality, is central to our moral experience, insofar as we come to feel respect for the moral law as an incentive only when it "strikes down" self-conceit. Recall, too, that there are at least two ways of construing Kant's notion of self-conceit. One is to construe self-conceit as a kind of interpersonal arrogance akin to Rousseau's *amour propre*, i.e., a tendency to erroneously understand ourselves as having some rank above all others, leading us to think that only our own needs and desires matter. Alternatively, self-conceit can be construed as one manifestation, opposite of moral despondency, of a more general failure to know oneself as a sensible and rational being, for whom morality requires constant striving: whereas the self-conceited agent distorts moral principles in such a way that he cannot help but be moral, the despondent agent distorts those principles in such a way that he can never be moral. On this alternative interpretation, self-conceit may lead to, but is not fundamentally based on any kind of comparative judgment of our worth, let alone, an interpersonal arrogance. I previously argued in favor of this latter construal of self-conceit. Later in this section, I will show how that construal illuminates issues of self-conceit with which Kant is preoccupied in the Doctrine of Method.

Now, if Kant is successful in the 'Incentives' chapter, through the notion of self-conceit, in showing how the moral law can become an incentive, then the Doctrine of Method can appear redundant. Why, after Kant's explanation of how the moral law elicits an incentive of respect upon striking down self-conceit, do we then need to be shown how "one can impart to the laws of pure practical reason admittance to the human mind and influence on that mind's maxims"?[226] Or how is the feeling of respect as a moral motive not an adequate explanation of "the way in which one can make objectively practical reason subjectively practical as well"?[227] On its face,

[226] KpV 5:151.
[227] Ibid.

the Doctrine of Method can seem to be included for no apparent reason, except maybe for the sake of Kant's commitment to architectonic symmetry across his critical philosophy.

I propose that the Doctrine of Method is not redundant but, on contrary, advances the overarching goal of the *Critique of Practical Reason*. In particular, the aim of the Doctrine of Method is to show how the moral law can have not just any kind of "admittance to the human mind and influence on that mind's maxims," but a *lasting* admittance and influence. What needs to be established, in other words, is how our awareness of the moral law can have an effect on sensibility not just *at* some time, but *over* time, as well. For a genuine moral disposition involves a refusal to reconsider the value of the moral law relative to that of our happiness, as well as a refusal to demand a search for reasons for compliance beyond those contained in that law itself. That this particular disposition is not yet established is indicated by this striking remark towards the end of the 'Incentives' chapter:

> *Duty!*—you sublime, grand name which encompasses nothing that is favored yet involves ingratiation, but which demands submission, yet also does not seek to move the will by threatening anything that would arouse natural aversion in the mind and terrify, but merely puts forth a law that on its own finds entry into the mind and yet gains grudging veneration (even if not always compliance), a law before which all inclinations fall silent even if they secretly work against it[.][228]

Note that, while Kant admits that the moral law "finds entry into the mind" through the feeling of respect elicited by the sublimity of duty, that law "yet gains grudging veneration." This is so, because the inclinations the moral law restricts still "secretly work against it." Because we are sensuous beings, vulnerable to various forms of self-deception, we are liable to be led to resist the moral law and to become suspicious of its adequacy as an incentive. The aim of the Doctrine

[228] KpV 5:86.

of Method, then, is to outline "a method of founding and cultivating genuine moral attitudes," such that the moral law can have a lasting effect on our sensibility.[229]

This aim is more apparent when we consider that the Doctrine of Method is to provide a "method of founding and cultivating genuine moral attitudes," that is, some theory of moral education.[230] And, for Kant, the aim of moral education is the formation of fully virtuous persons, which goes beyond having what Kant calls a 'good will.' In Kant's conception, virtue signifies a strength of will to act from duty despite any opposing inclinations: "For a human being's moral capacity would not be virtue were it not produced by the strength of his resolution in conflict with powerful opposing inclinations."[231] According to Anne Marget Baxley, the moral strength constitutive of virtue is the strength of the executive power of the will, which should not be equated with the legislative power. She writes that the legislative power is "a prior condition" for acquiring the "executive strength of will," meaning that the moral strength that is central to virtue, is not needed for self-legislation, or the adoption of maxims. In her view, the role of moral strength comes into play only after we have adopted moral maxims. As she explains, moral strength is needed "to enforce the morally good choices we legislate to ourselves as autonomous rational agents."[232] Moral strength of will comes into play once the task of self-legislation has been properly fulfilled – it enables firm stability in the observation of already-adopted maxims. It is this moral strength, understood as a sort of stability in one's intentions over time, for which Kant seeks some pedagogical measures in the Doctrine of Method.

---

[229] KpV 5:151. Stefano Bacin (2010) argues that "the goal of the method is to make the individual subject aware of his own capacity both to evaluate the moral worth of actions and to determine himself to act morally" (202). Bacin helpfully contextualizes the Doctrine of Method within Kant's architectonic system, particularly in relation to the general function of the doctrines of method present in each of the three Critiques. However, Bacin does not adequately show how the 'Incentive' chapter had not already shown how we become aware of our capacity for moral appraise and self-determination.
[230] Ibid.
[231] MS 6:477.
[232] Baxley (2010: 57).

Consider, too, the Doctrine of Method in relation to the 'Dialectic of Pure Practical Reason' that is situated between the 'Incentives' chapters and the Doctrine of Method. Recall how, in a previous chapter, I argued that the point of the Dialectic is to establish the possibility of the highest good by means of the postulates of immortality, God, and freedom as a response to a subjective need, one endemic to the human conditions. Given our limitations and propensities as rational, sensible beings, and in light of apparent injustices, we are susceptible to becoming demoralized in our efforts to sustain the moral motivation necessary for promoting the highest good. So, in the Dialectic, Kant tries to provide discursive arguments, not to secure the authority of the moral law, but to give us some means to sustain hope and avoid the demoralization that undermines our ability to satisfy the demands of the moral law. I argued, however, that those discursive arguments, the postulates of God and immortality, are ultimately inadequate for that purpose. To show how we can sustain hope and avoid demoralization in our commitment to the highest good, Kant turns to moral education. Hence, Kant's Doctrine of Method is committed to the same task as the Dialectic of addressing obstacles to our strivings – albeit through a pedagogical method that importantly, as we will see, does not rely on discursive arguments to accomplish that task.

Kant is particularly preoccupied, in the Doctrine of Method, with a certain set of tendencies that threaten the stability of our efforts to become virtuous or to promote the highest good. He introduces these tendencies by first noting how easily people from many walks of life fall into discussion about morality:

> If one attends to the course of conversations in mixed companies that consist not merely of scholars and subtle reasoners but also of business people or women, one notices that besides recounting and jesting, another entertainment has its place in them, namely arguing; … Among all [kinds of] arguing, however, there is none that more arouses the participation of persons who are otherwise soon bored with all subtle reasoning, and brings a certain liveliness into the company, than that

> about the moral worth of this or that action, by which the character of some person is to be established.[233]

For Kant, even those who find theoretical questions to be "dry and irksome," will tend to find themselves drawn into discussions where the apparent focus is on appraising the moral worth or character of some other individual. However, this moralizing tendency is prone to a kind of corruption that theoretical investigation is not: "they are then so precise, so meditative, so subtle in excogitating everything that could diminish or even just make suspect the purity of the intention—and hence the degree of virtue in it—as one does not otherwise expect of them in the case of any object of speculation." For this reason, Kant says that, "[i]n these judgments the character of the person himself who is judging others can often be seen shining forth."[234]

Kant goes on to describe two possible reactions – one positive, the other negative – to examples of moral conduct, which inform the pedagogical method he subsequently delivers. The positive reaction is where one is uplifted or inspired by some presumed righteous action, so much so that they defend the action against unfair or excessive moral criticism.

> Some seem especially inclined, since they exercise their judicial office above all upon deceased persons, to defend the good that is recounted concerning this or that deed of the deceased persons against any mortifying objections of impurity, and ultimately to defend the entire moral worth of the person against the reproach of dissimulation and secret malice.[235]

Kant appears to praise the character-type he associates with this kind of reaction, though it seems possible to imagine that this tendency to defend another against moral scrutiny can go too far. For example, in defending the moral worth of another's actions or character, one might too hastily discount or discredit some worthwhile criticism, whether by rationalizing the motives of

[233] KpV 5:153.
[234] Ibid.
[235] Ibid.

the person whose character they are appraising or by questioning the motives of the person casting criticisms.

At any rate, it is another possible reaction to examples of moral conduct that is foremost on his mind as a reflection of a character that threatens the status of the moral law as an incentive. This is the reaction of the moralistic fault-finder. Rather than be inclined to defend some example of moral conduct against any scrutiny, the characteristic reaction in this case is "to thinking up charges and accusations to challenge that worth."[236] This individual tries "to subtly reason virtue away entirely from all examples36 of human beings in order thereby to turn virtue into an empty name," while those who give a proper weight to morality seek to preserve the force of examples of good conduct, and so where there is

> the presumption of righteousness…, they would like to wipe off it even the slightest blemish—from the motive lest, if the truthfulness of all examples were disputed and the purity of all human virtue denied this virtue be in the end considered a mere chimera and thus all endeavor [directed] toward it be disdained as idle affectation and deceptive self-conceit.[237]

What may be going on here is that finding fault with apparent good examples stems from "deceptive self-conceit" and an urge to see oneself as superior, and this moralistic reaction, in turn, weakens the feeling of respect brought upon by the example. Consequently, an opportunity to make the force of the moral law vivid, to give it lasting effect that can sustain itself in the face of certain trials, is lost.[238]

---

[236] KpV 5:154.
[237] Ibid.
[238] Richard Dean (2012) describes another tendency that undermines the force of examples to elicit the feeling of respect for the moral law: "This is the tendency, when comparing oneself to others, to seek out others whose behavior provides standards that are too easy to meet" (591). However, I don't find this tendency described in the Doctrine of Method, and it is unclear whether Dean locates it there either. Without doubt, there are a multitude of ways in which we can deceive ourselves in our appraisal of examples of moral conduct. For my purposes, I focus just on those that Kant appears to explicitly have in mind in the Doctrine of Method.

The reading I just gave for the moralistic fault-finding reaction assumes that the self-conceit involved therein is one based on a kind of interpersonal arrogance. Contrary to respect for the exemplar of good conduct, what is felt is resentment, envy, hate, or another attitude that leads to the desire to claim superiority. But, as I mentioned above and discussed in detail in a previous chapter, self-conceit need not be understood as fundamentally based on a kind of interpersonal arrogance. Maybe what is driving one to find faults in another is such arrogance, or claim to superiority. But it seems possible that the subtle reasoning used to discredit any given example of virtue can also be motivated by a claim to inferiority or, more precisely, a moral despondency that is opposite to self-conceit.[239] If Kant's right that something about an individual's character can be revealed by how they appraise examples of moral conduct, then it seems the despondent person would react in just the same way Kant associates with self-conceit. The despondent person scrutinizes any given example of moral conduct to the point that "this virtue be in the end considered a mere chimera," because the despondent person has distorted moral principles in such a way that he can never be moral.[240] However, the effect on the force of the example is the same: the feeling of respect produced by the exemplar is diminished.

If the moralistic fault-finding need not be grounded on any kind of interpersonal arrogance, then there might be some other, or perhaps additional, source of the tendency with which Kant is preoccupied. Recall, then, Kate Moran's (2014) interpretation of Kant's notion of self-conceit, according to which the relevant moral error is not associated with a comparative judgment, but rather a misapprehension of the authority of the moral law, such that we can reassure ourselves that we are moral. The self-conceited person does not view the authority of

[239] Recall that a similar point is raised with respect to Rousseau's theory of *amour propre*, viz. Whether that form of self-love necessarily involves a claim of superiority, or whether it may not, instead, manifest in certain claims to inferiority.
[240] KpV 5:154.

that law as a kind of necessity liable to strike down their self-conceit, but rather as something to be distorted or manipulated to suit their conception of happiness. Looked at through that lens, the moralistic fault-finder can be understood to find such fault in others, to the point of making virtue seem impossible, as a way of preserving their own happiness. Since, if the moral law is a "mere chimera," then it cannot constrain the satisfaction of one's desires.

So, Kant's aim in the Doctrine of Method is to provide a pedagogical method for cultivating students' characters so as to avoid the sort of moral error from which self-conceit and despondency stem. That error is a failure to recognize that virtue requires continuous struggle, which requires, in turn, that the moral law can have an effect on sensibility not just *at* some time, but *over* time, as well. This error is one that threatens to demoralize us in our commitment towards becoming virtuous. To avoid such demoralization, however, Kant cannot appeal to some empirical incentive; otherwise, the incentive for moral action will not be pure. This means that what makes the feeling of respect for the moral law capable of yielding a lasting effect cannot be external to that feeling itself. Rather, the feeling of respect must be self-sustaining. But what could make us resilient to obstacles that threaten to demoralize us and undermine the feeling of respect, if not some additional incentive? For Kant, the answer is, I think, a certain kind of hope that is internal to the feeling of respect for the moral law. This hope is different from the hope for the highest good; rather, it is hope to be able to sustain respect for the moral law, as the source of moral motivation, in the face of hardships that threaten to demoralize us. And in the next two sections of this chapter, I analyze the pedagogical methods Kant prescribes to moral educators to cultivate this hope.

Before turning to Kant's pedagogical method, there is one more point of context for appreciating the aim of the Doctrine of Method. That point is that the Doctrine of Method not

only serves to advance the overall project of the *Critique of Practical Reason*, but also anticipates Kant's deeper preoccupation with hope in the *Critique of Judgment*. While 'hope' itself is not explicitly mentioned, let alone thematized, in the third Critique , the very problem that gives rise to the need for hope is announced in its Introduction as the guiding thread of the inquiry. We find it most strikingly in Kant's articulation of the "incalculable gulf" between freedom and nature that must be bridged in order to complete the system of philosophy:

> [T]here is an incalculable gulf fixed between the domain of the concept of nature, as the sensible, and the domain of the concept of freedom, as the supersensible ... the latter should have an influence on the former, namely the concept of freedom should make the end that is imposed by its laws real in the sensible world, and nature must consequently also be able to be conceived in such a way that the lawfulness of its form is at least in agreement with the possibility of ends that are to be realized in it in accordance with the laws of freedom.[241]

As Kristi Sweet explains, "a need for hope is born of our concern for freedom's efficaciousness in the natural order and has as its object a nature that is reconceived in the context of this concern."[242] Since we need our freedom to be made real and sensible through our actions, we also need to think about nature's susceptibility to our efforts to satisfy the demands of freedom. Hope is about this relation between freedom and nature. We hope for the possibility that nature, despite its seeming indifference or hostility to the ends we set ourselves, is inherently hospitable to the conditions of our happiness. We hope that there exists a profound order underlying nature, surpassing the laws we observe in our understanding of the world, and we hope for this underlying law to enable a comprehensive understanding of nature's own laws. According to Kant, what allows nature and freedom to be related to one another as parts of a larger, harmonious system, and so gestures towards a world in which we may hope, are judgments of reflection – the explicit subject matter of the *Critique of Judgment*.

---

[241] KU 5:176.
[242] Sweet (2023: 1)

I will say more about Kant's account of judgments of reflection as it pertains to his moral pedagogy in subsequent sections. For now, what I want to end with is emphasizing that Kant's aim in the Doctrine of Method in the *Critique of Practical Reason* should be viewed as a part of his more programmatic preoccupation with the question of hope in the *Critique of Judgment*. Whereas the third Critique is about how we can hope that there is nothing in the world itself, or nature outside of us, that makes it inhospitable to our efforts to promote human happiness, the Doctrine of Method in the second Critique is about how we can hope that there is nothing in us, or the nature inside of us, that makes us impotent, threatening to sap our resolve to keep trying. I will now turn to Kant's concrete pedagogical method for cultivating this hope in our ability to sustain our moral resolve and avoid demoralization.

2. First exercise: Sharpening judgment

The pedagogical method that Kant prescribes to educators by which moral educators may, "after laying a solely moral catechism at the basis," cultivate young students' virtuous disposition.[243] That method consists of two "exercises."[244] In brief, both involve presenting students with examples of moral actions: the first is designed to improve students' capacity to judge such actions; and the second is designed to produce within them certain affective responses. Awareness on the part of students of said improvement and affective responses by illustrative examples is said to elicit a feeling of respect for the moral law that is accompanied by a hope that that respect will be sustained so as to overcome threats of self-conceit and despondence. To understand how that is supposed to be brought about, I will consider each of the two exercises in more detail over the next two sections.

[243] KpV 5:154.
[244] KpV 5:159.

In the first exercise of Kant's pedagogical method, students are presented with examples of moral conduct – drawn from historical biographies rather than, say, fairy tales – in order to improve their ability to judge the actions and characters of individuals according to duties derived from the categorical imperative. Kant describes the goal of this exercise being to "to make it, as it were, a habit" for students to judge actions according to the moral law, in order "to sharpen" their moral judgment. This exercise requires students to engage one another in a "game of the power of judgment," in which they playfully deliberate with one another about how to morally judge the examples at issue.[245] You might imagine, for example, a group of students arguing about how best to interpret the case of a shopkeeper who passes up the chance to overcharge an inexperienced customer.[246] Some students in the group judge that the shopkeeper is acting from the duty not to lie, while others judge he is merely trying to be prudent in keeping from offending other customers. The students can thus improve their ability to apply the a priori moral concepts which had been elucidated by the moral catechism – concepts concerning, for example, the formulations of the Categorical Imperative, the duties entailed by the moral law, and the difference between prudence and goodness, regret and remorse, approbation and esteem, and so forth.

In addition to appraising the "legality" of the deeds with which they are presented, i.e. whether those actions conformed to duty, Kant proposes that educators direct students to appraise its "spirit," that is, to consider "whether the action was also done (subjectively) for the sake of the moral law."[247] This is puzzling, however, because it is impossible, on Kant's view, for someone to ascertain the motives of even their own actions, let alone those of others. So, the

[245] Ibid. One difference worth noting here between the pedagogical method Kant describes in the Doctrine of Method and the educational program in Rousseau's *Emile* is that Kant does not seek to isolate pupils from another. While he does appear to appreciate how comparative judgments can corrupt the pedagogical exercise, he also appears to think that those comparative judgments are essential to the task of sharpening pupils' judgment.
[246] GMS 4:398.
[247] KpV 5:159.

expectation of students cannot be that they make the correct judgment about the action. Rather, the expectation must be that this exercise of appraisal will be "carried only as a *game of the power of judgment* in which children can compete with one another" to offer the most compelling interpretation of a given action or character. In coming up with, sharing, deliberating together over these interpretations, students will be able to distinguish more clearly the action at issue in the presented examples from its underlying motivations. Allowing students to judge the same action according to different duties "teaches [them] how to distinguish different duties that come together in the same action."[248]

Kant believes that a child so guided will "soon become very acute and not a little interested, since he would feel the progress of his faculty of judgment."[249] Moreover, this felt improvement of one's capacity to judge the conduct and character of others "must gradually produce a certain interest in reason's law itself and hence in morally good actions."[250] This interest of the child, however, "is not yet interest in actions and in their morality itself." Rather, the game of judgment Kant prescribes to moral educators "merely brings someone to like to entertain himself with such an appraisal and gives to virtue or the cast of mind according to moral laws a form of beauty, which is admired but not yet on that account sought." That game, in other words, "produces a satisfaction that can also be communicated to others, while nevertheless the existence of the object remains indifferent to us, inasmuch as the object is viewed only as the occasion of our becoming aware of the tendency of talents in us which are elevated above animality."[251] Since, however, this falls short of cultivating students' interest in acting from duty, Kant prescribes to moral educators a second pedagogical exercise. Despite this

[248] KpV 5:154.
[249] Ibid.
[250] KpV 5:159-60.
[251] KpV 5:160.

limitation, this first exercise does make progress towards Kant's goal in the Doctrine of Method. In particular, as I will explain later in this section, this first exercise inspires students' confidence in their capacity for moral judgment, which helps to avoid the threats of self-conceit and despondence.

Shifting gears a bit, I want to draw attention to some parallels between, on the one hand, this first exercise in Kant's prescribed pedagogical method and, on the other, the aesthetic theory he develops in the *Critique of Judgment*. On its face, any such parallel can seem like a stretch. How can an aesthetic theory, which Kant had not yet penned, help to illuminate this particular aspect of his theory of moral education? But there are, upon closer inspection, some unmistakable parallels between this pedagogical method and the mature aesthetic theory he was, in fact, developing at the time of putting the finishing touches on the second Critique.[252] As one clue, Kant tells us that the sharpening of students' capacity for moral judgment through deliberation about examples "gives to virtue or the cast of mind according to moral laws a form of beauty."[253] So, while my aim here is not to defend any particular interpretation of Kant's theory of the beautiful, I will draw on that theory in order to help make sense of his brief remarks on pedagogy in the Doctrine of Method.

There are at least three features of Kant's account of the beautiful in the third Critique that bear on the analogy with this initial moral pedagogical exercise: (1) the subjective universality of judgments of beauty, (2) the reference of those judgments to a community of taste, i.e., *sensus communis*, and (3) the nature and possibility of disagreement concerning judgments of beauty in such a community. In the remainder of this section, I present a brief

[252] For a detailed historical account of Kant's writing of the *Critique of Judgment* during the time he was finishing the *Critique of Practical Reason*, see Zammito (1992: 46ff).
[253] KpV 5:160.

overview of Kant's account of judgements of beauty, focusing on these three features, before returning to his moral pedagogy to further elucidate its first exercise.

a. Subjective universality

For Kant, judgments of beauty are one species of aesthetic judgment of reflection, a genus distinct from both logical judgments and aesthetic judgments of sense. Judgment in general is defined as "the faculty for thinking of the particular as contained under the universal."[254] Logical and aesthetic judgments are two ways of exercising this faculty. Logical judgments, for the most part, subsume particular representations (whether intuitions or concepts) under concepts, in order to make cognitive claims. To judge, for example, 'This swan is white' or 'All men are mortal' is to bring a particular representation (some swan, in one case, and all humans, in the other) under a more general representation. Aesthetic judgments, by contrast, do not determine representations under concepts; instead, they relate given representations "to the subject and its feeling of pleasure or displeasure,"[255] that is, aesthetic judgments use "as a predicate something that is not even cognition at all, namely the feeling of pleasure (or displeasure)."[256] So, aesthetic judgments are expressed in the register of feeling (pleasure or pain), which Kant distinguishes from sensation (e.g., of color, sound, and so forth).[257] While sensation, for Kant, plays a legitimate role in empirical theoretical judgment – that is, in the determination of how, in fact, things are – feeling does not. Aesthetic judgments are, accordingly, non-cognitive.[258]

[254] KU 5:179
[255] KU 5:203.
[256] KU 5:288.
[257] KU 20:224; KU 5:206.
[258] KU 20:224; KU 5:203

Depending on the basis of the feeling of pleasure or pain at issue, an aesthetic judgment is either one of sense or reflection. Aesthetic judgments of sense express immediate liking for what is agreeable, or disliking for what is disagreeable, owing to contingent facts about the judging subject's physical constitution (e.g., 'Chocolate tastes better than vanilla'). Aesthetic judgments of reflection express immediate liking as well, but one that is not based upon contingent facts about the judging subject's physical constitution; rather, such judgments make potentially legitimate claims to "universal validity and necessity."[259] For example, in judging a sunset to be beautiful (or the idea of duty to be sublime), I am claiming that everyone else ought to judge, or feel, the same way that I do in relation to that sunset (or the idea of duty). Yet, this claim to universality does not depend on my recognition feelings aroused in me by the sunset or the idea of duty as falling under any determinate concept. For Kant, then, aesthetic judgments of reflection occupy a middle ground between, on the one hand, ordinary cognitive judgments, i.e., logical judgments, which claim universal validity on the basis of some determinate concept, and, on the other hand, aesthetic judgments of sense, which are non-cognitive but lack any claim to universal validity. That is to say, judgments of beauty make a claim to a subjective university.

The mental state definitive of judgments of the beautiful that underlies the subjective universality of such judgments is "the free play of the imagination and the understanding," which are the two faculties operative in ordinary cognitive judgment.[260] In ordinary empirical cognition, paradigmatically the perceptual recognition of an object as having certain features (e.g. that this brown table is oval), imagination and understanding work harmoniously together, resulting , but in such a way that the former is governed by concepts of the latter ('brown,' 'table,' 'oval,' etc.)

[259] KU 20:225.
[260] KU 5:218. Interpretation of Kant's account of the free play of the faculties is subject to much dispute. The interpretation that I find most compelling and which informs my analysis is developed in a series of essays by Hannah Ginsborg (2015).

which function as rules for the imagination in its synthesis of the manifold of intuitions. In a judgment of the beautiful, the faculties harmonize as they do in ordinary cognition; however, the imagination is free rather than governed by any determinate concept of the understanding. This is not to say that the understanding is inactive in judging that, e.g., 'This rose is beautiful.' After all, the object being judged is a rose, a flower, red, possessed of petals, etc. The understanding can successfully apply the concepts 'rose,' 'flower,' 'red,' and 'possessed of petals,' as well as many others, to the contents of the imagination. But in no case will the application of concepts by the understanding compel the compliance of the imagination, since the imagination will always have more to offer. Put another way, it is not that we cannot say anything about the objects we judge to be beautiful, but that, in principle, we cannot say everything about them. There is always more to say about the ceiling of the Sistine Chapel in that viewing it stimulates our imagination and understanding in a way that does not come to any end in a determinate statement about the beauty of its frescoes. The 'play' of our faculties in a judgment of beauty, then, consists in this continuous effort on the part of the understanding to conceptualize the rich contents of the imagination.

The subjective universality of judgments of beauty helps to illuminate three aspects of the first exercise of the pedagogical method Kant prescribes to moral educators in the Doctrine of Method. First, just as the legitimacy of judgments of beauty depends not on some objective element of that which is judged beautiful, but rather on a playful, yet purposive, activity of the imagination; so, too, the legitimacy of the moral judgments by students in the "game of the power of judgment" depends not on any objective quality of the examples they are judging, but rather on a feeling of the exercise and strengthening of their capacity to make moral judgments.

As Kant writes, “this occupation of the power of judgment, which lets us feel our own cognitive powers…merely brings about [the fact] that one gladly entertains oneself with such judging.”[261]

A second aspect of the first moral pedagogical exercise illuminated by the analogy with subjective universality of judgments of beauty concerns the disinterested pleasure associated with each. Judgments of beauty involve a disinterested pleasure in that they are not based on an expectation of satisfaction in the existence of an object of the pertinent kind. In other words, judging something beautiful is independent of desiring the (continued) existence of that which is judged beautiful. Likewise, as stated in the quotes in the preceding paragraph, students are intended to feel some pleasure in deliberating together about the examples with which they are presented. However, that pleasure does not result in an interest in acting from duty: “produces a satisfaction that can also be communicated to others, while nevertheless the existence of the object remains indifferent to us.”[262]

The third, and most important, aspect of the exercise to sharpen students' judgment illuminated by the analogy with judgments of beauty concerns the distinctive sort of universality of those judgments. Recall that, according to Kant, in judging, ‘This sunset is beautiful,’ I am aware that I am judging as I ought to with respect to that object, and I judge that everyone else ought to judge the way I do. Similarly, in the case of moral judgment, in judging that some character in a given example is doing as he ought to, a student is aware that he himself is judging as he ought, and he judges that everyone else ought to judge the same way as her. In both cases, the awareness of the appropriateness of the response in question is identical with the disinterested feeling of pleasure mentioned above. This is the awareness Kant is referring to when he says “we become aware of the predisposition of the talents in us, which is elevated

[261] KpV 5:160.
[262] Ibid.

above animality" through our appraisal of the actions or characters in given examples.[263] Note, however, that this sense of elevation is different from the sense at stake in our feeling of respect for the moral law. While both senses of elevation carry with them a certain feeling of pleasure and normative awareness, only the elevation involved in respect for the moral law reveals to us "a life independent of animality and *even of the entire world of sense*."[264] In other words, the student at this stage is aware that his appraisal of the examples with which he is presented is appropriate, without yet being aware that he ought to do as he judges and so do as the moral law commands.

b. Sensus communis

Shifting back, the subjective universality of judgments of beauty is importantly related to another key notion in Kant's aesthetic theory – namely, that of a *sensus communis*. As previously stated, the free play of imagination and understanding provide the basis for the universal validity of judgments of beauty. Every judgment of beauty assumes that others would likewise experience the free play of imagination and understanding when making judgments about a beautiful phenomenon, i.e., that everyone would judge such an object as beautiful. Thus, judgments of beauty always implicitly refer to a virtual community of taste, which Kant calls a *sensus communis*.

Kant distinguishes this notion of a *sensus communis* from two other notions. First off, a sensus communis is distinguished from any notion of common sense in which we might identify the common with the vulgar, or the average, i.e., the lowest common denominator upon which people would agree. As Kant puts it, the latter "is regarded as the least that can be expected from

[263] Ibid.
[264] KU 5:162, italics mine.

anyone who lays claim to the name of a human being." He further points out that "to possess [that which is encountered everywhere] is certainly not an advantage or an honor."[265] The vulgar or the merely popular is precisely what lacks the fine precision of an authentic notion of common sense. Such an authentic notion of common sense, and so whatever advantage or honor that brings, belongs to the notion of a *sensus communis*:

> By '*sensus communis*,' however, must be understood the idea of a communal sense, i.e., a faculty for judging that in its reflection takes account (a priori) of everyone else's way of representing in thought, in order as it were to hold its judgment up to human reason as a whole and thereby avoid the illusion which, from subjective private conditions that could easily be held to be objective, would have a detrimental influence on the judgment.[266]

Put another way, *sensus communis* is a way of judging that reflectively accounts for the disposition of everyone else's faculties, comparing its own feeling to the expected feelings of everyone else and discerning when its judgment is subjectively universally valid. This is in contrast with a vulgar notion, according to which common sense merely rests on prejudice and hasty generalizations that seemingly receive authentication from their being strongly felt rather than thought through.

Kant argues that we have good reason to presuppose a *sensus communis* on the grounds that it is a necessary presupposition of the communicability of the feeling of the free play of our cognitive faculties and thus of the universal communicability of cognition in general.[267] Moreover, Kant says that *sensus communis* is "the effect of the free play of our cognitive powers," rather than an additional sense.[268] He thereby avoids having to revise his transcendental psychology; instead, a new relationship between the faculties already introduced in the first Critique, specifically the imagination and understanding, suffices to explain the necessity of

[265] KU 5:293.
[266] Ibid.
[267] KU 5:238-9.
[268] KU 5:238.

judgments of beauty. One belongs to a community of taste, or *sensus communis*, to the extent that one communicates that one is judging an object to be beautiful. Potentially everyone who can experience the free harmonious play of imagination and understanding in judgments of beauty belongs to the *sensus communis*.

I take the notion of *sensus communis* to be reflected in at least two aspects of Kant's description of the first exercise in his prescribed pedagogical method. One of these is how that exercise is to take the form of deliberation among students in which they do not simply make claims about their beliefs. Instead, they are like judgments of beauty in which normative claims are made on other independent, judging subjects. Otherwise, the students could not participate in "a game of the power of judgment in which [they] can compete with one another."[269] Second, the assumption of something like a *sensus communis* seems to be in operation with Kant's observation that in "the course of conversations in mixed companies…another entertainment has its place in them, namely arguing" about the moral worth of characters or actions presented in examples.[270] Kant suggests that here lies a neglected means by which educators may promote a consciousness of one's moral being. However, he also notes that people's self-love can turn this propensity for moral conversation into a mechanism for defending themselves from criticism. Broadly speaking, it is possible that people may either be inclined to defend the virtuosity of the persons involved to the point of whitewashing their deeds, or be keen on unmasking the suspicious intentions of those same persons. Grounded in feelings of self-love, these apologetic and skeptical tendencies are based on prejudice or otherwise heteronomous forces. The cultivation of a *sensus communis*, specifically a shared moral feeling in which participants share in their pleasure of joint moral progress, can curb these twin tendencies.

[269] KpV 5:154.
[270] KpV 5:153.

c. Aesthetic disagreement

So far I have focused on the subjective universality of judgments of beauty, along with the notion of *sensus communis* that explains how such universality is possible. Before continuing onto the second exercise in Kant's moral pedagogical method, I want to introduce one more feature of his account of judgments of beauty – namely, the nature and possibility of aesthetic disagreement. For starters, I want to introduce a distinction found in Kant's *Critique of Judgment*, specifically in his presentation to the 'Antinomy of Taste.' Kant observes that, although we cannot "dispute" (disputiren) about matters of taste, we nonetheless "argue" (streiten) about them. This observation turns on what Kant takes to be two "peculiarities" of judgments of taste that he had earlier elucidated in the 'Analytic of the Aesthetic Power of Judgment.'

One of two "peculiarities" of judgments of taste, Kant argues, is that such judgments outstrip our concept-driven claim-making abilities. "There is no disputing about taste," for otherwise people could be brought to agree with our judgment "according to determinate concepts, by basing a proof on them."[271] You may be an art historian with genuine knowledge of your subject; you may try to convince me that a particular painting is beautiful by means of your superior knowledge of color, form, and perspective. But if I do not find it beautiful, I will simply "stop my ears, shall refuse to listen to reasons and arguments."[272] The distinguishing feature of a judgment of taste is that it "is always a singular judgment about the object."[273] It is I who must find a singular object (e.g., this painting, this rose, etc.) beautiful, and it is I who must judge.

[271] KU 5:338.
[272] KU 5:284.
[273] KU 5:215.

However, the fact that we nevertheless "argue" about matters of taste indicates another "peculiarity" of judgments of taste: "A judgment of taste…makes a claim to everyone's assent, as if it were an objective judgment."[274] By contrast with judgments about the "agreeable" ("I like canary wine"), judgments of taste are not merely subjective but exhibit subjective universality, as explained earlier. It would be "ridiculous" to say "this object (the building we are looking at, the garment this man is wearing, the concert we are listening to, the poem put up to be judged) is beautiful for me."[275] The claim to beauty is not in the form of an empirical prediction with which everyone will agree. It is a normative claim that everyone ought to agree. Not being grounded on any (determinate) concept, the judgment exhibits a necessity that "can only be called exemplary, i.e., a necessity of the assent of everyone to a judgment that is regarded as an example of a universal rule that we are unable to state."[276]

So, both disputing and arguing about some matter are distinct from idle chatter, since they properly concern matters of judgment, where some claim to universal agreement is at stake. As such, disputing and arguing, in the case of the beautiful, are not distinguished from one another in the ways we talk about matters of taste but in how we judge on those matters, specifically the grounds for those distinctive kinds of judgments.

> For to argue and to dispute are certainly alike in this, that they try to bring about unanimity in judgments through their mutual opposition, but they differ in that the latter hopes to accomplish this in accordance with determinate concepts as grounds of proofs, and so assumes objective concepts as grounds of the judgment.[277]

To "dispute," then, is to deliberate by means of "proofs" that are grounded on "objective concepts," or antecedently determined and discursively articulable rules for judging. In that case,

[274] KU 5:281.
[275] KU 5:212.
[276] KU 5:237.
[277] KU 5:338.

a “dispute” is a particular kind of disagreement which is recognized as taking place within an established framework where clear decision-making procedures are in place. In other words, a “dispute” is a disagreement, like a lawsuit, whose resolution is the outcome which results from following fair procedures.

By contrast, to “argue” about, say, matters of taste is to deliberate by means of something other than proofs grounded on determinate concepts. For, recall, a judgment of beauty does not consist in the recognition of an objective feature of the thing about which the judgment is made (e.g., the color of a rose judged beautiful); instead, a judgment of beauty expresses some kind of relation between the thing, or the perceptual representation of the thing, and the subject’s own feeling. But, Kant writes, “wherever it is supposed to be possible to argue, there must be hope of coming to mutual agreement; hence one must be able to count on grounds for the judgment that do not have merely private validity and thus are not merely subjective.”[278] Because judgments of beauty cannot be ultimately grounded in concepts, the “hope of coming to mutual agreement” must be based on the expectation that we can actually (somehow) change each other‘s feelings. In other words, in arguing about matters of taste, what is at stake is not what another should say about some object of art (although we do that as well) but how they should feel about it.

Moreover, the hoped-for agreement in arguing about matters of taste is different from the sort of agreement in matters of taste that Kant tries to establish elsewhere in the third Critique. The latter sort of agreement concerns whether it is possible to agree about matters of taste in any case, whereas, in arguing, what is of concern is whether it is possible to agree such matters in a particular case. If, for instance, a friend and I are arguing about whether some film is beautiful, it would not help to settle our argument to be reminded of Kant's ‘Deduction of Pure Aesthetic Judgments,’ since we are making claims upon each other’s feelings. For the same reason, the

[278] Ibid.

possibility of a *sensus communis* would also fall short of providing the agreement we are seeking. So, in arguing about matters of taste, we are trying to produce an agreement in feeling and are not merely concerned with the general possibility of such agreement.

These remarks on the nature and possibility of aesthetic disagreement help to illuminate the sense in which students are to playfully compete with one another in the first exercise of Kant's pedagogical method. Using Kant's technical sense of what it means 'to argue,' as opposed 'to dispute,' the students are to deliberate by means of something other than proofs on determinate, moral concepts. Of course, the students can apply concepts in the claims about the examples of moral conduct with which they are presented in order to express themselves in deliberations with others; just as an art curator can apply concepts to some painting judged beautiful in order to help gallery visitors to see the painting as he sees it. However, in neither case are those concepts the basis for the distinctive normative claims at issue: students and curators alike are making normative claims on others based not on any determinate concept, but rather on a subjective feeling. Put another way, the point of Kant's exercise is not for students to persuade one another, though that may happen. Rather, the point is for students to take pleasure in exercising and strengthening their capacities to make moral appraisals.

Taking stock, in this section, I presented a brief overview of Kant's account of judgments of beauty, focusing on three interrelated features: the subjective universality of judgments of beauty, the reference of those judgments to a community of taste, and the nature and possibility of disagreement in such a community. The connections drawn out in this section begin to show how hope appears in Kant's pedagogical method. For one crucial point revealed by those connections is that the student at this stage is aware that his appraisal of the examples with which he is presented is appropriate, without yet being aware that he ought to do as he judges and so do

as the moral law commands. That awareness provides the students with a confidence in his ability to make moral judgment that will be better able to withstand certain failures or mistakes and, so, provide for him a source of hope in the possibility of his continued moral progress.

3. Second exercise: Eliciting respect sustained by hope

The second exercise in the course of the method involves bringing students' attention – again, through the use of examples drawn from biographies and histories – to the purity of a good will.[279] Here, the primary goal is to draw students' attention to the fact that incentives of inclination have no influence over the will in morally good action. To clarify the proper use of examples to illustrate moral goodness, Kant suggests educators might tell "the story of an upright man whom some people want to induce to join the defamers of an innocent but otherwise powerless person."[280] Since Kant alludes to the seeming fact that the person in question is Anne Boleyn, I will refer to this as "the Boleyn example" and to its protagonist as "the Boleyn man."[281] What I show over the course of this section is how this second pedagogical exercise, and so the Boleyn example, builds on the fruits of the first. It does so by drawing further on elements of Kant's aesthetic theory to show how carefully crafted examples of virtue can produce in students a respect for the moral law that elicits hope that that respect will constitute one's character.

It is important to note that the second exercise is prompted by the fact that the first was not sufficient for Kant's purposes in the Doctrine of Method. Recall that his main purpose there is to prescribe a pedagogical method by which the moral law can be made to have a lasting influence on our sensibility. The first pedagogical exercise falls short of this insofar as it does not

[279] Kant, like Rousseau, believes that histories are a better source than, say, fairy tales for examples to shape moral character, at least in early stages of development.
[280] KpV 5:155.
[281] Jens Timmermann (2010: 81n14) suggests that the man in question is Henry Norris, a figure that Kant would have been acquainted with from David Hume's *History of England*.

cultivate students' interest in the moral law. For the game of judgment Kant prescribes to moral educators "merely brings someone to like to entertain himself with such an appraisal and gives to virtue or the cast of mind according to moral laws a form of beauty, which is admired but not yet on that account sought." This satisfaction, or what I characterized as a confidence, in one's ability to make moral judgments will prove necessary for cultivating students' character, but it nonetheless requires some additional step that directly engages students' interest in the moral law. "Now, however, the second exercise enters upon its task," Kant tells us, "namely to make the purity of the will discernible in the vivid exhibition of the moral attitude in examples, at first only as a negative perfection of the will insofar as in an action done from duty no incentives whatever of the inclinations influence the action as determining bases."[282] Whereas the examples of moral conduct in the first exercise are designed to invite playful deliberation about the actions and motives of the persons involved, the examples of the second exercise are meant to make such deliberation unnecessary. Instead, "purity of the will discernible in the vivid exhibition of the moral attitude in examples" will demonstrate that incentives of inclination have no influence over the will in morally good action.

> By this the learner is at least kept attentive to the consciousness of his freedom, and although this renunciation arouses an initial sensation of pain, yet, because it withdraws that learner from the constraint of even true needs, there is proclaimed to him at the same time a liberation from the manifold dissatisfaction in which all these needs entangle him, and the mind is made receptive to satisfaction from other sources. … And now the law of duty, through the positive worth that compliance with this law lets us feel, finds readier admittance [to the human mind] through the respect for ourselves in the consciousness of our freedom.[283]

Note that the purity of the will made vivid by the examples of the second exercise "arouses an initial sensation of pain" and "at the same time a liberation" from the weight of inclinations. By evoking "respect for ourselves in the consciousness of our freedom" that is simultaneously

[282] KpV 5:160.
[283] KpV 5:160-1.

painful and freeing, the second exercise enables the moral law to have a lasting influence. For, Kant explains, "On this respect, if it is well-founded…every good moral attitude can now be grafted, because this is the best and indeed the only guard to prevent ignoble and corrupting impulses from penetrating into the mind."[284]

Let us briefly review Kant's Boleyn example to see how this looks in practice. In Kant's recounting, the Boleyn man refuses to lie about the guilt of an innocent person in the face of increasingly severe sacrifices to his apparent happiness. According to Kant, as a result of the Boleyn man's unswerving commitment to moral principles over sensible inclinations, the student to whom the example is presented "will be elevated gradually from mere approval to admiration, from there to amazement, and finally to the greatest veneration and a lively wish to be able himself to be such a man (although not, to be sure, in his situation)."[285] Admiration, amazement, and veneration are effects that ultimately yield the feeling of respect for the moral law, "which has the greatest force directly on the spectator's mind."[286]

I will discuss in more detail later on how the Boleyn example evokes these affects that collectively yield respect for the moral law that students hope will constitute their character. Right now, I want to focus on Kant's more general characterization of the crucial, simultaneous feeling of pain and liberation that is evoked by examples in the second exercise. This characterization is an important indication that Kant is, again, anticipating his mature aesthetic theory. According to that theory, there are two kinds of aesthetic judgment of reflection: one are judgments of beauty, which I discussed in the previous section in connection with the first pedagogical exercise; the other are judgments of sublimity. Whereas judgments of beauty are based on a distinctive feeling of pleasure, Kant tells us that judgments of sublimity are based on

[284] KpV 5:161.
[285] KpV 5:156.
[286] KpV 5:157.

"a pleasure that is possible only by means of a displeasure."[287] Kant regularly associates the moral law with the sublime elsewhere in the *Critique of Practical Reason*. For example: "Duty!—you sublime, grand name which encompasses nothing that is favored yet involves ingratiation."[288] Also: "It is something very sublime in human nature to be determined to actions directly by a pure law of reason."[289] But, as was the case with the beautiful in the first pedagogical exercise, I think the allusion to the sublime in the second exercise promises to be instructive. Indeed, I will argue that the examples of moral conduct in Kant's second exercise are themselves meant to be objects of sublimity. To be sure, Kant's theory of the sublime, like his theory of the beautiful, is subject to debate, and my goal is neither to defend that theory nor any particular interpretation of it. Rather, I will look at the psychological processes in judgments of sublimity, in order to elucidate Kant's extremely brief remarks on the apparently complicated psychological processes at work in the Boleyn example and the second pedagogical exercise, in general.

a. Judgments of sublimity

As aesthetic judgments of reflection, judgments of sublimity share some of the same qualities as judgments of beauty. Both are singular and based on a disinterested feeling expressing the mental state underlying a relation of cognitive powers in the representation of an object. Judgments of sublimity also make a claim to subjective universal validity for the underlying feeling and display subjective purposiveness. But the object, the quality of the feeling, and the cognitive basis of the judgment is different from the one connected to judgements of beauty:

[287] KU 5:260.
[288] KpV 5:86.
[289] KpV 5:117.

> The beautiful in nature concerns the form of the object, which consists in limitation; the sublime, by contrast, is to be found in a formless object insofar as limitlessness is represented in it, or at its instance, and yet it is also thought as a totality: so that the beautiful seems to be taken as the presentation of an indeterminate concept of the understanding, but the sublime as that of a similar concept of reason.[290]

While the imagination relates to a capacity for concept formation in both kinds of aesthetic reflective judgements, in judgments of sublimity this capacity is not understanding, but reason. This introduces an important difference in the role of the imagination in these judgments, as well as its relationship with the other faculty involved. Concepts of the understanding are concepts of experience, i.e., concepts which give form to intuitions. By contrast, concepts of reason, or what Kant usually refers to as "ideas," have no experiential content. Rational ideas like freedom, God, and immortality do not have a corresponding object in intuition. Because concepts of the understanding give form to intuitions, the free play between that faculty and the imagination involves a limited, yet undetermined form. By contrast, ideas of reason have no form, and so the imagination can only relate to reason by encountering objects that are formless, allowing for limitlessness to be represented in those objects. There is no such thing as a formless object in intuition, but there are objects that challenge the imagination's capacity to give form to intuitions. It is these that evoke feelings of the sublime.

Kant distinguishes two kinds of sublime objects: mathematical and dynamical. The mathematically sublime encompasses that which is "absolutely great," surpassing all comparisons. Kant asserts that when we label something as "great" without specifying a comparison, we evoke the notion of greatness beyond determination. For "if I simply say that something is great, it seems that I do not have in mind any comparison at all…, since it is not thereby determined at all how great the object is."[291] Greatness, in this context, is not a concept

[290] KU 5:244.
[291] KU 5:248.

for defining an object but rather a "comparative concept." The determination of an object's magnitude necessitates a unit of measurement, requiring something else for comparison. For example, a kilometer is great in comparison with a meter, but small in comparison with a mile. The question arises: What in the contemplation of an object generates the idea of greatness independently of comparison? To this question, Kant answers:

> Nothing that can be an object of the senses is… to be called sublime. But just because there is in our imagination a striving to advance to the infinite, while in our reason there lies a claim to absolute totality, as to a real idea, the very inadequacy of our faculty for estimating the magnitude of the things of the sensible world awakens the feeling of a supersensible faculty in us; and the use that the power of judgment naturally makes in behalf of the latter (feeling), though not the object of the senses, is absolutely great, while in contrast to it any other use is small.[292]

The feeling of the sublime is evoked by an inadequacy of the imagination in relation to the demands of reason. The imagination strives to present the infinite on behalf of reason's idea of absolute totality. But this is impossible, because there is no possible presentation of infinity. Yet in its striving for the infinite, the imagination evokes the feeling of a supersensible faculty in us, namely, reason. Thus in judging an object as sublime, the power of judgment's use is absolutely great, because it serves a faculty of the mind that has no limitations.

The inadequacy between imagination and reason arises from objects challenging the imagination's ability to estimate their magnitude. Kant explains that this estimation involves apprehension and comprehension. Apprehension presents individual sensuous perceptions, while comprehension retains these presentations to form an object. The sublime is triggered by a disruption in comprehension when the imagination struggles to retain partial apprehensions,

> for when apprehension has gone so far that the partial representations of the intuition of the senses that were apprehended first already begin to fade in the imagination as the latter proceeds on to the apprehension of further ones, then it

[292] KU 5:250.

> loses on one side as much as it gains on the other, and there is in the comprehension a greatest point beyond which it cannot go.[293]

In the presence of some objects, the imagination finds itself unable to comprehend all partial apprehensions. Two of the examples Kant gives are the pyramids and Saint Peter's basilica, which suggest objects whose magnitude challenges our capacity to comprehend all its parts into one whole representation. As the imagination runs through the manifold of intuitions of these objects, it encounters difficulties retaining initial apprehensions as it proceeds to new ones. Ultimately, given that the imagination is finite, it encounters a point beyond which it cannot go. The imagination fails to give the object a form, which thus becomes "formless."

The failure of the imagination to present the object prompts the thought of infinity, which gives rise to the feeling of respect. Being a supersensible faculty, reason lies beyond the limits of the senses. Consequently, in its demand to present all the elements in a series, it "does not exempt from this requirement even the infinite (space and past time), but rather makes it unavoidable for us to think of it (in the judgment of common reason) as given entirely (in its totality)."[294] The failure of the imagination to satisfy reason's demand to comprehend the infinity of space and time makes us think of this totality, even if it cannot be represented. So, what is a failure in certain respects becomes a success from the viewpoint of thinking: while we cannot present the totality of space and time, the very attempt to do so shows that we are capable of thinking it. It is this capacity that awakens the feeling of "respect," which Kant here defines as "the feeling of the inadequacy of our capacity for the attainment of an idea that is a law for us."[295] In its effort to present an idea of reason, the imagination encounters its inadequacy, but "at the same time its vocation for adequately realizing that idea as a law," which shows that "the

[293] KU 5:252.
[294] KU 5:254.
[295] KU 5:257.

feeling of the sublime in nature is respect for our own vocation, which we show to an object in nature through a certain subreption (substitution of a respect for the object instead of for the idea of humanity in our subject)."[296] Although the failure of the imagination is felt as displeasure, it also "arouses the feeling of our supersensible vocation in us, in accordance with which it is purposive and thus a pleasure to find every standard of sensibility inadequate for the ideas of the understanding."[297] Therefore, even in their disharmony, imagination and reason find a sort of harmony. But the feeling of the sublime is not simply the feeling of this harmony. Rather, it is the "feeling of the supersensible vocation in us," that is, respect. Without this feeling of the supersensible, there would be no harmony between the faculties, because the failure of the imagination would not lead us to hear demands of reason. Consequently, Kant claims, "the pleasure in the sublime in nature… presupposes another feeling, namely that of its supersensible vocation, which, no matter how obscure it might be, has a moral foundation."[298]

Kant maintains that the mathematically sublime is a feeling of our supersensible vocation, but it is not clear why this is so when only theoretical reason is involved in this mode of the sublime. Mathematically sublime objects lead to reflection on theoretical ideas (e.g., infinity), not moral ideas (e.g., freedom or virtue). It is even less clear why the feeling at the basis of this aesthetic judgment of reflection, even if it involves moral ideas, can be a feeling of our supersensible vocation. That vocation is nothing other than to act according to the moral law and contribute to the creation of a kingdom of ends.[299] The contemplative feeling of the mathematically sublime seems far removed from this. Still, Kant clearly believes that the ability to think the infinite is of practical significance.

---

[296] Ibid.
[297] Ibid.
[298] KU 5:292.
[299] GMS 4:434f.

> Even a faculty for being able to think the infinite of supersensible intuition as given (in its intelligible substratum) surpasses any standard of sensibility, and is great beyond all comparison even with the faculty of mathematical estimation, not, of course, from a theoretical point of view, in behalf of the faculty of cognition, but still as an enlargement of the mind which feels itself empowered to overstep the limits of sensibility from another (practical) point of view.[300]

The mathematically sublime thus serves a practical function, since the feeling evoked by it indicates a supersensible side of our being, ensuring that we are not restricted by the constraints of the sensible world. Indeed, this is a function of the sublime in general. Why this is so will become clearer after we examine the dynamical sublime.

Kant delves deeper into the practical significance of the sublime in his analysis of the dynamically sublime. While the mathematically sublime is related to the faculty of cognition, the dynamically sublime is related to the faculty of desire, which is a "being's power to be, through its presentations, [the] cause of the actuality of the objects of these presentations."[301] Unsurprisingly, then, it is in his analysis of the dynamically sublime that Kant introduces human actions and affects as examples of sublimity. This does not mean, however, that that mode of the sublime is the only one of practical significance. As we saw just a moment ago, the mathematically sublime holds such significance, as well. This is further evident given that already in his analysis of the mathematically sublime Kant mentions 'The man is great' and 'The magnitude of a certain virtue' as possible instances of this kind of judgment. Still, as we will see, the dynamically sublime is where Kant draws the connection to moral ideas most explicitly.

While the feeling of the mathematically sublime is evoked by the imagination striving to represent a magnitude beyond which it is capable, the feeling of the dynamically sublime is evoked by the imagination striving to represent power – again, in a manner that extends that capacity. Kant introduces this idea as follows:

---

[300] KU 5:255.
[301] KpV 5:9n.

> Power is a capacity that is superior to great obstacles. The same thing is called dominion if it is also superior to the resistance of something that itself possesses power. Nature considered in aesthetic judgment as a power that has no dominion over us is dynamically sublime.[302]

In the dynamically sublime, nature is considered as a power, though not one having dominion over us. The dynamically sublime thus arouses admiration and awe for our own power over nature. This is the result of contemplating an object that produces fear by virtue of overcoming our capacity to resist it. Powerful objects in nature, such as

> bold, overhanging, as it were threatening cliffs, thunder clouds towering up into the heavens, bringing with them flashes of lightning and crashes of thunder, volcanoes with their all-destroying violence, hurricanes with the devastation they leave behind, the boundless ocean set into a rage, a lofty waterfall on a mighty river, etc., make our capacity to resist into an insignificant trifle in comparison with their power.[303]

Even if these objects in nature are fearful, Kant explains, "the sight of them only becomes all the more attractive the more fearful it is," given that "they elevate the strength of our soul above its usual level, and allow us to discover within ourselves a capacity for resistance of quite another kind, which gives us the courage to measure ourselves against the apparent all-powerfulness of nature."[304] In contemplating, for example, the violent eruption of a volcano, we experience a power that makes our capacity to resist insignificant. Yet, even if we are afraid of it, so long as we are at a safe distance, we can "think of the case in which we might wish to resist it."[305] According to Kant, even if resisting the power of a volcano would prove futile, the very thought of this resistance elevates our soul to become aware of a different kind of resistance of which we are capable and that is unconstrained by the bounds of physical power.

---

[302] KU 5:260.
[303] KU 5:261.
[304] Ibid.
[305] KU 5:260.

So, just as the mathematically sublime revealed a superiority of reason over nature by virtue of the incapacity of the imagination to represent ideas of reason, in the case of the dynamically sublime

> the irresistibility of its [i.e. nature's] power certainly makes us, considered as natural beings, recognize our physical powerlessness, but at the same time it reveals a capacity for judging ourselves as independent of it and a superiority over nature on which is grounded a self-preservation of quite another kind than that which can be threatened and endangered by nature outside us, whereby the humanity in our person remains undemeaned even though the human being must submit to that dominion.[306]

While our physical self-preservation can always be threatened by nature "outside of us," there is another kind of self-preservation that remains unthreatened by it. This is what Kant calls "the humanity in our person," which stems from our capacity to act on the basis of moral principles, in such a way that our principles for action can be made universally communicable and contribute to the creation of a kingdom of ends: "humanity means on the one hand the universal feeling of participation and on the other hand the capacity for being able to communicate one's inmost self universally."[307] By raising the imagination to a case in which we might wish to resist nature in all its power, we come to regard our own natural incentives for action, that is, "those things about which we are concerned (goods, health and life)," as trivial in comparison with "our highest principles," which stem from the moral law.[308] In this way, the contemplation of the power of nature "outside us" engages our imagination so as to make us aware of a power over nature inside us, because no natural incentive is so powerful as to make us disregard the moral law as an object of respect. As with the mathematically sublime, every standard for judging natural incentives for action is inferior in comparison with the moral law, which calls us to resist such incentives.

---

[306] KU 5:261-2.
[307] KU 5:355.
[308] KU 5:262.

The connection between the representation of the power of nature and the idea of resistance to it in the feeling of the dynamically sublime discloses a new possibility crucial for our capacity to act according to the moral law. Since all objects of experience, including human deeds, are ruled by the law of natural causality, all actions are determined by sensible inclinations. It would seem, then, that actions performed from moral duty simply do not exist in experience. However, if there are objects in nature which, by virtue of their resistance to the capacity of our imagination to represent their natural power, awaken the feeling of a vocation of reason that is beyond natural necessitation, then a new possibility opens up for the representation of action. While all actions are ruled by the law of natural causality, there are actions that challenge our ability to represent them as a power of nature, and consequently arouse the idea of a supersensible power. Kant mentions the warrior as "an object of the greatest admiration," someone about whom aesthetic judgment of reflection determines as deserving respect.[309] By appearing to sacrifice all inclinations – indeed, putting physical self-preservation at risk – the warrior surpasses our capacity to represent his actions as determined by nature, thus arousing the idea of a supersensible power.

There might seem something curious, however, about Kant introducing examples of human action as objects of the sublime. A morally good action is one performed from respect for the moral law, which is an idea of reason that lacks any possible representation in experience. As we saw in the previous chapters, Kant explains how the feeling of respect "is produced by a purely intellectual ground," and so is independent of sensuous sensations.[310] For Kant, there is no example of an action that embodies, or adequately represents, compliance with duty, and there is no reason to assume he had abandoned this point in the third Critique. The feeling of the

[309] Ibid.
[310] KpV 5:73.

sublime, however, is aesthetic, meaning it is connected to sensuous perceptions, but also dependent on ideas of reason; otherwise, the imagination would not feel the demands of reason outstretch its own limitations to represent the magnitude or power of the sublime objects.

So, we must regard action (like those of the warrior) that overcome obstacles of sensibility, not as examples of moral action, per se, but rather as objects that, by virtue of surpassing the capacity of the imagination to represent those actions as stemming from some power of natural necessity, awaken the idea of a supersensible power in us. In this regard, the sublimity of war is no different from that of the power of a volcano, for "sublimity is not contained in anything in nature, but only in our mind, insofar as we can become conscious of being superior to nature within us and thus also to nature outside us (insofar as it influences us)."[311] This is why Kant claims that the feeling of the sublime has its source "in the predisposition to the feeling for (practical) ideas, i.e., to that which is moral," which "is the ground for the necessity of the assent of the judgment of other people concerning the sublime to our own."[312] To be able to judge an object sublime, we need to be able to feel respect for the moral law; otherwise, we would not experience the failure of the imagination in the representation of an object as serving a purpose for our faculties, a purposiveness in the contrapurposiveness of the judgment that makes it universally communicable. This means that the representation of the morally good has both an aesthetic and an intellectual "side," as Kant explains here:

> The object of a pure and unconditioned intellectual satisfaction is the moral law in all its power, which it exercises in us over each and every incentive of the mind antecedent to it; and, since this power actually makes itself aesthetically knowable only through sacrifices…, the satisfaction on the aesthetic side (in relation to sensibility) is negative, i.e., contrary to this interest, but considered from the intellectual side it is positive, and combined with an interest. From this it follows

[311] KU 5:264.
[312] KU 5:265.

> that the intellectual, intrinsically purposive (moral) good, judged aesthetically, must not be represented so much as beautiful but rather as sublime, so that it arouses…the feeling of respect (which scorns charm).[313]

Although the moral law is the object of a purely intellectual satisfaction through the feeling of respect, its power is known to us only aesthetically through sacrifices, i.e., through the overcoming of sensible incentives for action. The aesthetic representation of the power of the moral law, however, only works by means of awakening the feeling of respect for the moral law, which is purely intellectual. So, the representation of the morally good is never purely aesthetic, since absent the feeling of respect, sacrifices would not bring to mind the idea of the moral law. In this way, the judgment of action, like action itself, has two "sides": one purely intellectual and one based in the senses.

It is this two-sided nature of judgments of sublimity that allows us to judge on the basis of ideas, despite their incommensurability with anything given in sensible experience. The imagination cannot apply ideas to intuitions, and so bridge the gap between our empirical and noumenal character, because there is no correspondence between the realms of freedom and nature. But objects in nature can nonetheless arouse ideas of reason by stretching the imagination to the limit of what it can represent. It is the awakening of the feeling of respect by virtue of this outstretching of the imagination that constitutes the feeling of the sublime. Judgments of sublimity thus relate actions and ideas to one another while preserving their heterogeneity: no action can adequately represent compliance with duty, but an action can arouse the feeling of a power in us that is beyond natural necessitation, and thus respect for the idea of duty. Superiority of reason over nature is the negative side of judgments of sublimity, since it is felt as contrary to the interests of sensibility. But that negative side, by itself, is an insufficient basis for claims to universality. The basis of such claims in judgments of sublimity lies their positive side, which is

[313] KU 5:271.

the feeling of respect for the moral law. Both sides are necessary, however, since, as we have seen, the feeling of the sublime comes from the inadequacy of the imagination in relation to the reason.

b. The sublime versus enthusiasm

There is one more point to make about Kant's account of judgments of sublimity before applying that account to his recounting of the Boleyn example as part of the second of his two moral pedagogical exercises. The point concerns Kant's remarks distinguishing the feeling of the sublime from enthusiasm. As explained above, judgments of sublimity have an "aesthetic side," which is the feeling of the superiority of a power in us over obstacles of sensible nature, and an "intellectual side," which is the feeling of respect for the moral law. Kant describes enthusiasm as "the idea of the good with affect," adding that "this state of mind seems to be sublime." He goes on to explain why this state of mind only "seems" to be sublime as follows:

> Now, however, every affect is blind, either in the choice of its end, or, even if this is given by reason, in its implementation; for it is that movement of the mind that makes it incapable of engaging in free consideration of principles, in order to determine itself in accordance with them. Thus it cannot in any way merit a satisfaction of reason. Nevertheless, enthusiasm is aesthetically sublime, because it is a stretching of the powers through ideas, which give the mind a momentum that acts far more powerfully and persistently than the impetus given by sensory representations.[314]

In other words, enthusiasm is a state of mind resulting from the feeling of the sublime, but that is crucially absent any connection to the feeling of respect. Consequently, in the case of enthusiasm, the superiority over obstacles of sensibility becomes an end valued for its own sake. The mental activity that enables one to engage in "free consideration of principles" depends on the feeling of respect for the moral law. Without this free consideration of principles, the

[314] KU 5:272.

aesthetic side of sublimity (i.e., the feeling of superiority over nature) remains "blind," in that superiority over sensible obstacles is not connected to the pure idea of morality. Kant's example is "indignation, as anger," an affect that shares with the idea of the good a resistance to sensuous self-interest, though without necessarily involving a free consideration of principles for action in accordance with the demands of reason. While indignation can predispose us to overcome our inclinations, this predisposition does not necessarily involve respect for the idea of the moral law, which is why enthusiasm "cannot merit any satisfaction of reason." Because enthusiasm is not related to a feeling of respect for the moral law, and thus to a free consideration of principles, it lacks the intellectual side necessary for judgments of sublimity. Kant claims this can be compared to the idea that sensuous representations are equivalent to that which they represent, which Kant calls a "delusion of sense."[315] So, in the case of enthusiasm, the superiority over the inclinations is not taken to be a representation of morality, but rather to be morality as such.

Indeed, in *Religion within the Boundaries of Mere Reason*, Kant identifies "delusion of sense" as belonging to the deceptive religious practices he calls "delusion of religion", which consists in "the habit of taking a mere representation (of the imagination) for the presence of the thing itself, and to value it as such."[316] The "mere representation" are those religious practices that give expression to our moral disposition, while the "thing itself" is our moral disposition as such. Kant mentions rituals of self-sacrifice as examples of delusional practices which we tend to confuse with morality, because we take its aesthetic quality (namely, the overcoming of our inclinations) as the thing itself. This remark on self-sacrifice is important for my purposes, since examples of sacrifice figure prominently in Kant's explanation of his second moral pedagogical exercise in the Doctrine of Method. His point here is that if the judgment of the morality of some

[315] KU 5:275.
[316] RGV 8:168n.

act or practice of self-sacrifice is made only in terms of its aesthetic quality, without any connection to free consideration of principles, then the feeling at the basis of that judgment will become a source of delusion. This is so, because that judgment will consist in an attachment to the appearance of virtue without a regard for the moral ideas that underlie it. Such an attempt to make the moral basis of action directly accessible to sensibility amounts to a manipulation of the faculty of judgment, insofar as it eliminates the gap between sensible representations and ideas of reason. As Kant remarks:

> Even tumultuous movements of the mind, whether they be associated with ideas of religion, under the name of edification, or, as belonging merely to culture, with ideas that contain a social interest, no matter how much they stretch the imagination, can in no way claim the honor of being a sublime presentation, if they do not leave behind a disposition of mind that, even if only indirectly, has influence on the consciousness of its strength and resolution in regard to that which brings with it intellectual purposiveness (the supersensible).[317]

No matter the degree of nobility in the idea we are presenting (besides indignation, Kant mentions courage and humility), it is only sublime if it makes us conscious of the power in us that goes beyond the greatest or mightiest object of nature. Otherwise, that seemingly noble affect remains at the level of enthusiasm and runs the risk of delusion.

To bring this discussion of the sublime to a close, I want to highlight the point above about sacrifice, since it will be instructive in turning towards the Boleyn example. Again, for Kant, sacrifice itself merits no satisfaction, let alone esteem, of reason, for it does not necessarily connect to the purely intellectual source of moral ideas. By contrast, the feeling of the sublime necessarily relates sacrifices to the feeling of respect for the moral law, thus preserving connection to the true ground for morality. In sum: whereas enthusiastic self-sacrifice consists in satisfaction in sacrifice as an end in itself, genuinely sublime acts of sacrifice consist of satisfaction in sacrifice only as a means of awakening respect for the idea of duty. Because the

[317] KU 5:273.

idea of duty is incommensurable with anything in experience, the mental state underlying the feeling of the sublime involves an inadequacy of the imagination in relation to reason, that is, between the aesthetic representation of virtue by means of the imagination and the intellectual feeling of respect for ideas of reason. Judgments of sublime do resolve that inadequacy, but rather stem from it: it is the tension between the two sides of the sublime feeling that elicits a state of mind that is purposive in its contrapurposiveness and that is universally communicable.

c. Boleyn example re-examined

I will now return to the second of the two pedagogical exercises that Kant presents to moral educators in the Doctrine of Method of the *Critique of Practical Reason*. Recall that his chief aim is to show how the moral law can be made to have a lasting influence on our sensibility. The first pedagogical exercise contributes to this by showing the "game of judgment" in appraising examples of moral conduct can lead students to take pleasure in the appropriateness of their exercise of moral judgment. That awareness provides the student with a confidence in his ability to make moral judgments that will better withstand certain failures or mistakes and, so, provide for him a source of hope in the possibility of his continued moral progress. However, that awareness of the appropriateness of his judgment is still absent any awareness that he ought to do as he judges and so do as the moral law commands. In other words, the first exercise ultimately falls short of Kant's aim in the Doctrine of Method insofar as it does not cultivate students' interest in the moral law. The second exercise is meant to complete this task, drawing once again on Kant's mature aesthetic theory to show how examples of moral conduct can contribute to the cultivation of moral character by evoke feelings of the sublime, where what we feel is the superiority of our own free will to act according to the moral law.

Here, then, is a more detailed overview of the Boleyn example, in which we find these allusions to the sublime in connection to moral development. In Kant's recounting of that example, the Boleyn man is first offered some gifts and honors in exchange for his compliance, but he turns these down, to which the student will react with "mere approbation and approval," presuming the exercise of appraisal was successful.[318] From this starting point, the story progresses through stages in which the circumstances of Boleyn man become increasingly dire as he no longer simply refuses to gain from dishonesty but refuses also to secure himself against unjust loss. First, he will lose the protection of his friends, who abandon him; next, he will be disinherited by his relatives, being left with no assets of his own; then, it will be made apparent to him that he will be unable to avoid the suffering inflicted upon him by certain powerful people; and, finally, the prince will threaten to lock him away in prison and even end his life. Meanwhile, his family pleads with him to give in and cooperate, so that his righteousness will not bring suffering upon them all. Although the Boleyn man is not insensible to their pleas, he nonetheless "remain[s] faithful, without wavering or even doubting, to his resolve of uprightness." The supposed effect of this story on the student to whom it is presented is that he "will be elevated gradually from mere approval to admiration, from there to amazement, and finally to the greatest veneration and a lively wish to be able himself to be such a man (although not, to be sure, in his situation)."[319] Admiration, amazement, and veneration are effects that ultimately yield the feeling of respect for the moral law, "which has the greatest force directly on the spectator's mind."[320] I will now briefly attend to each of these affects, explaining how they individually fall short of respect for the moral law but collectively yield that feeling, before turning to the student's wish to "be such a man" – that is, to be like the Boleyn man.

[318] Ibid.
[319] KpV 5:156.
[320] KpV 5:157.

Since Kant regards admiration, amazement, and veneration as species of affects, it will be useful to start with a definition of that genus, which we saw above bears on Kant's remarks on the sublime, particularly in relation to enthusiasm. In his *Anthropology from a Pragmatic Point of View*, Kant says that "the feeling of pleasure or displeasure in the subject's present state that does not let him rise to reflection (the representation by means of reason as to whether he as to whether he should give himself up to it or refuse it)."[321] As feelings, affects do not refer to objects as they are, as do cognitions, nor as they ought to be, as do desires; rather, affects refer merely to the subjective constitution of the subject. This means that affects belong merely to the faculty of feeling pleasure and displeasure, and are, moreover, to be distinguished from passions, which belong to the faculty of choice and "make all determinability of the faculty of choice by means of principles difficult or impossible." So, "while in the case of an affect the freedom of the mind is certainly hampered, in the case of passion it is removed."[322] However, affects like admiration, amazement, and veneration are distinct from other affects in that, although they impede reflection for the short period of time they exist, they are also the result of reflection. To show how that might be, and so how reason might allow itself to be impeded for a brief period for the sake of the cultivation of moral character, I consider these reason-produced affects in turn as they appear in the Boleyn example.

Admiration, for Kant, is a kind of astonishment, in which there is a certain excitation of feeling that exists when a train of thought is unpleasantly stopped, then followed with an influx of unexpected thoughts. Someone inexperienced in the world – someone surprised and amazed at everything he sees – feels a naïve form of this affect. Think, here, of a child bounding from one exhibit to the next at a natural history museum.[323] This is astonishment, properly speaking.

[321] Anth. 7:251.
[322] KU 5:272n.
[323] Anth. 7:261

But there is also a non-naïve form of astonishment that occurs if one follows up surprise with a certain sort of reflection. In this case, astonishment can become a genuinely reason-produced affect, and counts as admiration:

> [H]e who thoughtfully and with a scrutinizing eye pursues the order of nature in its great variety falls into astonishment at a wisdom he did not expect: an admiration from which he cannot tear himself away (he cannot be surprised enough). However, such an affect is stimulated only by reason, and is a kind of sacred awe at seeing the abyss of the supersensible opening before one's feet.[324]

In admiring some person or object, then, one's astonishment does not cease when the novelty is gone. This is, perhaps, a description of the student's response to hearing that the Boleyn man is threatened with various kinds of losses after denying gifts in exchange for his cooperation. However, this falls short of a feeling of respect, for that feeling demands certain courses of action, whereas admiration, as a mere affect, does not. In our respect for the moral law, we are aware that we can act independently of our inclinations, because we are aware that we ought to do so. By contrast, in admiration, we regard a person or object as something unusual. But, "to do one's duty," Kant says, "is no more than to do what lies in the common moral order and is not, therefore, deserving of wonder."[325]

The student's admiration for the Boleyn man gives way to amazement. While Kant does not as explicitly define this affective response as one might have liked, we might be able to work backwards from the details with which the student is presented in the Boleyn example. If the student feels admiration in response to the Boleyn man's increasingly dire situation, then amazement is presumably the student's response when told "he wishes that he had never seen the day that exposed him to such unspeakable pain."[326] What this detail effectively does is make it clear that the Boleyn man is not a superhero, someone deserving of wonder; rather, he is an

[324] Ibid.
[325] RGV 6:49.
[326] KpV 5:156.

ordinary individual, susceptible to the pull of the same inclinations the student could imagine finding herself having. Thus, whereas the student's admiration was in response to something he regarded as unordinary, his amazement is in response to something he regards as ordinary, albeit an ordinary something that causes a brief stoppage in his reflection because of an awareness of his limitations to act in the way the Boleyn man did.

In this stoppage in the student's reflection from the amazement in the Bolelyn man, we can begin to see elements of the sublime. Recall that, in Kant's analysis of the sublime, calling an object sublime is the same as claiming that the object is such that it gives rise to a particular kind of feeling in the subject experiencing the object. In particular, judgements of sublimity are based on a feeling of simultaneous attraction and repulsion evoked by the interplay of our faculties in which the imagination fails to adequately represent the sublime object under a principle of purposiveness of nature, and the object is instead judged under the principle of the purposive use of nature for our supernatural vocation. When we encounter the sublime, we experience a "resistance to the interest of the senses" that speaks to us directly and simultaneously as agents who are (in the case of the mathematically sublime) subjects trying to get to grips with something, and (in the case of the dynamically sublime) are subjects whose prospects of satisfaction out of pursuit of any ends is threatened. However, what is thereby threatened is also made vivid, namely, that we are rational beings capable of both estimating nature in light of ends we contemplate and acting to achieve ends we freely set. Ordinarily, we do not have experiences that make these features so vivid, especially not in the right balance of awareness of our limits and abilities. "The experience of the sublime," as Katerina Deligiorgi states, "requires an object that can stop us in our tracks and thereby heighten what is fundamentally a basic and rather ordinary sense of agency."[327] Here, it is the example of the Boleyn man's plain but unswerving

[327] Deligiorgi (2014: 32).

pursuit of principles that makes the student feel, in that moment of amazement, the inadequacy of any explanation of that pursuit based on natural incentives for action.

Moving on, consider now the student's feeling of veneration in response to the Boleyn example. For Kant, certain forms of veneration should be distinguished from respect for persons, which is entailed by respect for the moral law – particularly, those forms of veneration that amount to valuing an individual for his particular merits or attributes. Respect, by contrast, is a matter of valuing some person simply in virtue of him being a free, rational being. As such, respect is something to which everyone is entitled. What sort of veneration does Kant's hypothetical student feel? Given that Kant describes it as being the "greatest veneration," we might think that it is nothing other than a feeling of respect. But if that were the case, then the student's veneration would not be accompanied by a wish to be like the Boleyn man. For the feeling of respect, as the awareness of the moral law moving our will, implies an awareness that our will can be determined by the moral law. The latter awareness is not implied, however, by any wish, on Kant's definition of the term in his *Metaphysics of Morals*. Whereas in choosing to bring about some object, "it [the faculty of choice] is joined with one's consciousness of the ability to bring about its object," in wishing for some object, the faculty of desire is not joined by such a consciousness.[328] So, the student who is presented with the Boleyn example views the character of the Boleyn man as worthy of esteem and even desirable, but he is unsure about his capacity to be such a person. His esteem for the Boleyn man, then, must amount to less than respect; otherwise, the student would not merely wish to be like the Boleyn man, but rather act in such a way as to bring that about.

To appreciate why the student's veneration for the character of the Boleyn man falls short of respect for the moral law, it helps to consider what Kant takes to be conditions under which an

[328] MS 6:213

encounter with some artificial object successfully evokes the feeling of the respect for the moral law. Take this remark from the Anthropology:

> The sublime is awe-inspiring greatness (*magnitudo reverenda*) in extent or degree which invites approach (in order to measure our powers against it); but the fear that in comparison with it we will disappear in our own estimation is at the same time a deterrent (for example, thunder over our heads, or a high, rugged mountain). …
>
> The sublime is the counterweight but not the opposite of the beautiful; because the effort and attempt to raise ourselves to a grasp (*apprehensio*) of the object awakens in us a feeling of our own greatness and power; but the representation in thought of the sublime by description or presentation can and must always be beautiful. For otherwise the astonishment becomes a deterrent, which is very different from admiration, a judgment in which we do not grow weary of being astonished. … [T]he artistic presentation of the sublime in description and embellishment (in secondary works,* *parerga*) can and should be beautiful, since otherwise it is wild, coarse, and repulsive, and, consequently, contrary to taste.[329]

In short, while the feelings of the sublime and the beautiful differ, the form of the beautiful nonetheless serves as a condition for some objects to successfully evoke the former in us. By their nature, sublime objects – like the Boleyn man's unswerving pursuit of virtue – simultaneously repulse and attract through the interplay of our faculties. What the sublime invites us to do, in particular, is to compare ourselves with the power of the object. If, however, the object is presented in such a way that "is wild, coarse, and repulsive," then we will become "deterred" from, or "grow weary" of, sustaining any comparison between the power in us and the sublime object. Another way to put this point is to say that the fruits of contemplating some sublime object require sustained effort, in order to maintain focus on the purposiveness in the contrapurposiveness of the judgment of sublimity. So, there must be some explanation for why the student's veneration for the character of the Boleyn man, and initial effort to compare herself in wishing to be like him, is not met as a deterrent, but rather with sustained effort to act from respect for the moral law that that character exhibits.

[329] ApH 7:243.

To see, then, how the Boleyn example elicits a feeling of respect, we need to recall Kant's first pedagogical exercise, since it is that exercise that "gives to virtue or the way of thinking according to moral laws a form of beauty."[330] At the end of that exercise, each student was said to have become aware that his appraisal of the examples with which he was presented was appropriate, just in the sense that, in making those appraisals, he felt pleasure in improving his judgment. Moreover, on the basis of that awareness, the student was said to be entitled to judging that everyone else ought to judge the examples as he did. Yet, the student was not aware that he ought to do as he judged, for his judgment carried with it no interest in the object he was judging, except maybe an indirect interest in moral goodness. Now, having been taken through the stages of the Boleyn example from mere approval through admiration and amazement to veneration, the student is aware that he ought to do as he judges; that is, he is aware that he ought to act from duty, and cultivate character, as exemplified by the Boleyn man. For the student is aware now of not only the appropriateness of his appraisal of the Boleyn man, but of comparing herself to him, since he is like him in possessing the moral law within. And, as the story progresses, the student loses confidence that he would act in the same way. The example thus prompts, on the part of the student, reflection and comparison with the virtuous character exemplified by the Boleyn man, which brings on a feeling of humiliation. The reason why the student does not simply reject the comparison that gives rise to this disagreeable feeling is that, as just mentioned, he was brought up through the first exercise to regard his judgments as appropriate, that is, that they have the form of beauty. In other words, the student is aware of his comparison as being appropriate and, so, that he ought to feel humiliated. Since that normative awareness carries with it a feeling of pleasure, the disagreeableness of being humiliated provides the student with an incentive to act from duty, that is, a feeling of respect for the moral law. If not for the fact that he had previously

[330] KpV 5:160.

been habituated through the first exercise to feel confident in his ability to make moral judgments, then the student may grow weary in his veneration for, and comparisons with, the character of the Boleyn man. With that confidence instilled, he can sustain his judgment that the character of the Boleyn man is moral and that humiliation before such a character is fitting for a moral agent like herself, humiliation that calls for his sustained effort to do what the Boleyn did.

I have now described the two exercises that comprise the pedagogical method Kant prescribes to moral educators for the purpose of cultivating genuine moral attitudes in their students. Insofar as those exercises were shown to bring students to feel respect for the moral law, it would appear as though Kant's method was a success in showing that we can acquire an interest in the moral law. However, in a previous section, I explained that the problem for which Kant is prescribing this pedagogical method is to show how the moral law can have a lasting effect on our sensibility. What needs to be shown, more precisely, is how the feeling of respect for the moral law cultivated by moral educators can be self-sustaining. For if that feeling is sustained by anything other than itself it will no longer be a pure moral incentive. I will now try to show that the key to Kant's solution to this problem is to cultivate respect for the moral law such that that feeling gives rise to the hope that that same feeling will continue to influence one's will – the hope, that is, that respect for the moral law will constitute one's character.

Before turning to the passage that is crucial for my proposal, I want to suggest that there has already been evidence presented for thinking that, in Kant's view, moral education inspires hope on the part of students. This evidence came in the brief discussion in the previous section about the hypothetical student's wish to be like the Boleyn man. Recall that, for Kant, a wish is just a desire for an object that is not accompanied by an awareness of one's ability to bring about that object. We can further distinguish "empty" wishes from "lively" ones: an empty wish is one

that includes a belief that the wished-for object is impossible, whereas a lively wish includes a belief that the wished-for object is not impossible, though perhaps one is uncertain of the means for bringing about that object. The student's wish to be like the Boleyn man is said to be a lively one. So, the student wants to cultivate a character similar to that of the Boleyn man, and he believes that doing so is not impossible. Since, for Kant, hope is at the very least a desire for something that one believes is possible yet beyond one's powers, the student's wish is a kind of hope.

Even if, however, one grants that the student's wish does amount to a kind of hope, it is not clear how that hope is part of the student's feeling of respect for the moral law, as elicited by the Boleyn example. Why think that hope belongs to the pure incentive moral incentive, rather than being a separate, a posteriori incentive. On that point, a crucial passage to consider is the following, in which Kant summarizes the overall effect of pedagogical method he prescribes to moral educators:

> to put everything second to the holiness of duty alone and to become conscious that one can do this because our own reason acknowledges it as its command and says that one ought to do it—this is, as it were, to elevate oneself entirely above the world of sense, and is linked, as also an incentive of a power that rules over sensibility, with that same consciousness of the law inseparably, even if not always with effect; yet this effect does also, through repeated occupation with this incentive and the initially slighter attempts at using it, give hope of being brought about [more frequently], so as to produce in us little by little the greatest but pure moral interest therein.[331]

Here, Kant uses the 'ought-implies-can' principle to infer that a respect for the moral law always can be our motive, as "an incentive of a power that rules over sensibility." That is to say, in our awareness of the moral law, we are aware of that law as moving our will, or altering the structure of our practical judgments such that we subordinate our happiness to morality. Kant's hypothetical student becomes aware of this power of the moral law to some extent when he feels

[331] KpV 5:159.

admiration for the Boleyn man. I say to some extent, because, as we saw, that admiration amounts to less than a feeling of respect for the moral law: whereas in feeling respect the student is aware that certain course of action are demanded of his ('I must do such-and-such'), in feeling admiration he is simply aware that certain courses are demanded of someone ('The Boleyn man must do such-and-such'). But the student comes to be aware of the moral law as moving his will through being taken through the progressive series of affects in the manner described above: admiration of the Boleyn man's unswerving pursuit of virtue ('He is doing as he ought') becomes amazement of the character of this ordinary man ('I feel inadequate before his example'), which becomes veneration for his character ('I ought to follow his example'), which becomes respect for the moral law which he can be confident to sustain over time ('I ought to follow the moral law and hope to do so').

Crucially, however, in being aware of a power in us that rules over sensibility, we are not just aware of an alteration of our priorities, or of a temporary movement of our wills. Rather, we are also aware of what Kant elsewhere calls "reason's self-preservation" – that is, "a self-preservation of quite another kind than that which can be threatened and endangered by nature outside us, whereby the humanity in our person remains undemeaned even though the human being must submit to that dominion."[332] According to Kant, just as taking in a powerful waterfall from a safe distance elicits in us doubts about the preservation of our physical bodies and, at the same time, affirmation of the preservation of our rational capacities, so too in being presented with an example such as that of the Boleyn man we are made increasingly aware of the relative weakness and instability of our inclinations, and, at the same time, the relative strength and stability of the feeling of respect for the moral law. What makes that feeling so stable is that, in having that feeling, the student is aware of his comparison with the Boleyn man as being

[332] RGV 8:146n; KU 5:262.

appropriate. More specifically, the student recognizes the appropriateness of judging that both he and the Boleyn man are ordinary people in whom there resides a power stronger than the inclinations of sensible nature, recognition of which evokes the feeling of respect for the moral law. And since that normative awareness carries with it a feeling of pleasure, the disagreeableness of being humiliated provides the student the incentive to act from duty, that is, a feeling of respect for the moral law.

However, Kant tells us in the above passage from the Doctrine of Method that this pure moral incentive is "not always with effect." This is so, because, as sensuous beings, vulnerable to doubts and distractions, we are easily led to resist the moral law and to become suspicious of its adequacy as an incentive. This is why Kant's method for cultivating a genuine moral attitude must begin with the exercise of appraisal: for that exercise encourages students to render the legality and spirit of individuals' deeds intelligible, albeit indeterminately, so as to improve their moral judgment. The better a student's moral judgment is, the more stable the feeling of respect will be in its influence on his maxims, since pleasure in being confident in his moral judgments leads to a pleasure in the acknowledgment of the superiority of the student's reason over his inclinations. Of course, that incentive will never guarantee success; the student could, at any point, become demoralized in a way that undermines his sustained effort to cultivate moral character. That is why the student hopes, but is not certain, that respect for the moral law will continue to constitute his character.

We can now begin to make sense of how Kant thinks moral education can yield a feeling of respect for the moral law that inspires a certain hope on the part of students. The hope in question is that the feeling of respect will constitute one's character, and so continue to influence one's maxims in the future. So, the hope is not for some external state-of-affairs but, rather, the

very feeling of respect, or awareness of the moral law, that gives rise to hope in the first place through the student's pleasing awareness that that feeling is appropriate in response to the Boleyn example. Put another way, the student's hope is a refusal to despair in face of the humiliation he feels before that example; positively stated, the student's hope is a commitment to seeing his affective responses to the Boleyn man, as well as his eventual humiliation and elevation, as contributions to his cultivation of a genuine moral attitude. This is how the student's hope can sustain pure moral incentive without adding content to it: such hope does not demand further actions but requires, rather, that he views his response to the Boleyn example in a certain light. Of course, the hope inspired by respect for the moral law does not guarantee that that incentive will be effective. However hopeful the student is, he is still a sensuous being, vulnerable to doubts, distractions, and self-deception. Moreover, in hoping that the feeling of respect will constitute his character, the student is deferring to some extent to his future self. So, the hope that respect for the moral law carries with it gives the student confidence, though short of certainty, that he will cultivate a virtuous character.

The confidence inspired by students' hope that their feeling of respect will be a source of lasting influence for his moral strivings prevents the student from succumbing to either self-conceit or despondence. In the case of self-conceit, an individual isolates himself by refusing to accept the disagreeing judgments of others, perceiving himself to be correct (morally good or logically sound), despite their lack of real evidence. But Kant's pedagogical method is precisely designed so that students cannot isolate themselves. In the first exercise, they are tasked with not just appraising examples of moral conduct, but making those appraisals as part of a "game of the power of judgment," and, so, making claims on one another. And, in the second exercise, although students are not likewise tasked with deliberating with others about the example of the

Boleyn man, they are guided in such a way to compare themselves with the Boleyn man. With respect to despondency, students will recognize that, because they are like the Boleyn man in the relevant respects, that they are capable of resisting their inclinations.

In this chapter, I have provided a detailed analysis of the Doctrine of Method in the *Critique of Practical Reason*: What does Kant set out to achieve in the Doctrine of Method? Why is that aim important for him in the context of the second Critique or his broader critical philosophy? And, how does he go about trying to achieve his aim in the Doctrine of Method? In brief, what Kant set out to accomplish in the Doctrine of Method was to show how the feeling of respect for the moral law can influence our maxims over time and in the face of certain obstacles. It was important that he show this, in the context of the second Critique, in order to more fully establish that practical reason can provide us with a sufficient motivation to act from duty. Moreover, this is important in the context of his broader critical philosophy, since it is part of his aim to bridge the gulf between nature and freedom, which preoccupies him most evidently in the *Critique of Judgment*. How Kant set out to accomplish this aim in the Doctrine of Method was the centerpiece of my reading. In particular, I reconstructed the pedagogical method Kant prescribes to moral educators in the Doctrine of Method for the purpose of cultivating students' genuine moral disposition. That method is chiefly composed of a pair of pedagogical exercises. Each exercise shares illuminating structural affinities with elements of Kant's mature aesthetic theory: the first exercise builds on Kant's account of judgments of beauty; the second on his account of judgments of sublimity. Each exercise utilizes examples of moral conduct that engage

and strengthen students' moral judgment. Through first pedagogical exercise, the students' capacity for moral judgment is given the "form of beauty," in which students develop confidence in their ability to judge the motives and character of the individuals in examples of moral conduct. And through the second exercise, students are made to feel respect for the moral law in light of its power over sensible inclinations that, on the basis of successful completion of the first exercise, elicits students' hope to sustain that respect in the face of demoralizing conditions.

Conclusion

Over the course of three chapters I have reconstructed the rationale and method of Kant's pedagogy of hope, that is, the pedagogical method he prescribes to moral educators in the 'Doctrine of the Method of Pure Practical Reason' to cultivate genuine moral character. In Chapter One, I presented Jean-Jacques Rousseau's framing of the problem (and promise) of education, a framing that subsequently influenced Kant's theory of moral education. Rousseau sets out in his educational treatise, *Emile*, to provide a remedy for a problem he had diagnosed as being the source of injustice and evil in human society. As I explained, Rousseau identifies the principal source of those ills as *amour propre*, a kind of self-love that manifests in a desire for recognition from others that, when inflamed, leads to vindictiveness, cruelty, and conflict. I also explained two of the key phases in the educational program of *Emile* designed to forestall the emergence of an inflamed *amour propre* and, instead, develop that desire for recognition in such a way that its redemptive features flourish.

There are at least two pertinent respects in which Rousseau's framing of, and partial solution to, the problem of *amour propre* shapes Kant's own theory of moral education. One of these, which I examined in Chapter One, concerns Kant's analogous notion of self-conceit. For Kant, self-conceit is a kind of self-love that, though a hindrance to morality, is central to our moral experience in that the striking down of self-conceit is what engenders a feeling of respect for the moral law. I challenged the view that Kantian self-conceit is identical to, or even just a variation of, Rousseauean *amour propre.* The key difference is that the error in the latter is in understanding ourselves as having some rank above all others, whereas the error in Kantian self-conceit stems from a more general failure to know ourselves as sensible, rational beings, for

whom the pursuit of morality requires constant striving. Despite that difference, Rousseau's general framing of education as a remedy for a diagnosis of the source of evils stemming from a corrupted kind of self-love is something we can see adopted by Kant, insofar as his own theory of moral education is designed to address the threats he associates with self-conceit.

A second respect in which Rousseau's *Emile* shaped Kant's theory of moral education concerns the nature of the pedagogical he presents in the Doctrine of Method and that I reconstruct in Chapter Three. Not only do both philosophers stress the centrality of carefully crafted examples to inspire moral development, but they provide similar methods by which to utilize those examples. Just as the first phase of Emile's education consists of carefully orchestrated episodes designed to encourage Emile's confidence in his own ability to judge independently of others' opinions, so too the first exercise in Kant's pedagogical method uses a curated "game of the power judgment" in appraising examples of moral conduct in order that students acquire a sort of confidence in their capacities for moral judgment.[333] And, just as the second phase of Emile's education consists in exercising the fictional pupil's imagination through the use of examples drawn from history that stir his sentiment of pity, and thus shape his *amour propre*, so too the second exercise in Kant's pedagogical method utilizes historical examples of moral sacrifice in the face of temptation, guiding students' through certain stages of affect, so that they arrive at feeling respect for the moral law.

After discussing Rousseau's framing of the problem (and promise) of education in relation to Kant's moral theory, I turned, in Chapter Two, to addressing the question, 'What is hope?' Rather than attempt to give a comprehensive account of hope, I focused on articulating the meaning of hope in the context of Kant's theory of moral education. I examined two strategies commonly taken towards articulating Kant's account of hope. One of these takes as the

[333] KpV 5:154.

central target of analysis to be the mental state, or activity, that is presumed to define hope; the other takes that target to, instead, be the object of hope, which, for Kant, is the highest good. Rather than lay some stake in either of these strategies, I showed how each leads to the conclusion that, however hope is to be defined, its primary function is in sustaining our motivation in the face of threats of demoralization. In terms of Kant, the threat of demoralization is what motivates his arguments concerning the highest good in the *Critique of Practical Reason*, specifically the arguments for the postulates of God and immortality. Given our limitations and propensities as rational, sensible beings, and in light of apparent injustices, we are susceptible to becoming demoralized in our efforts to sustain the moral motivation necessary for promoting the highest good. The postulates are meant to address the threat of demoralization, but they are ultimately insufficient. So, Kant turns to moral education to contribute to his efforts to explain how morality can lead us to hope that respect for the moral law will continue to motivate us to act in such a way that sustains our moral strivings. The pedagogical dimension of hope, then, takes as its object the cultivation of virtue and works to sustain our efforts towards that aim, which is an essential part of the highest good. In this way, the moral pedagogy Kant outlines in the Doctrine of Method advances the general aim of the second Critique, which is to show how pure reason can be practical, and gestures at a central theme of the third Critique, which is to show how nature is not inhospitable to our efforts to achieve the highest good.

In the third chapter, I reconstructed the pedagogical method Kant prescribes to moral educators in the Doctrine of Method in the second Critique. As I explained, the aim of the Doctrine of Method is to show how respect for the moral law can not only influence, but have a lasting effect on our sensibility. The two exercises, which I briefly summarized above in relation to Rousseau's influence on Kant's theory of moral education, provide concrete measures by

which educators can elicit not just students' respect for the moral law, but also their hope to sustain that respect and avoid the threat of demoralization. In the first pedagogical exercise, students' capacity for moral judgment takes on the "form of beauty," fostering their confidence in evaluating the motives and character of individuals in moral conduct examples. The second exercise aims to instill a feeling of respect for the moral law, affirming its authority over sensible inclinations. Building on the successful completion of the first exercise, students are encouraged to harbor hope in maintaining this respect even when confronted with demoralizing conditions.

There are at least two respects in which my analysis of the role of hope in Kant's theory of moral education might bear on contemporary discourse on education. One is that my analysis invites some scrutiny of assumptions about hope in the multiple curricula being developed and sold with the promise of helping schools and teachers to increase students' hopefulness in the future. The promise of such curricula is often, though not always, based on research from the field of positive psychology. In one foundational paper in positive psychology, hope is defined as "a positive motivational state that is based on an interactively derived sense of successful (a) agency (goal-directed energy), and (b) pathways (planning to meet goals)."[334] Another defines hope as Shane Lopez, in his book, Making Hope Happen: Create the Future You Want for Yourself and Others, defines hope as "the belief that the future will be better than the present combined with the belief that you have the power to make it so!"[335] Without commenting extensively on either of these definitions, which are oft cited in curricular materials, we can at least note that these diverge from Kant's treatment of hope in that each appears to conflate hope with a sense of optimism untethered to any conception of morality. By contrast, Kant (among other philosophers) is explicit in distinguishing hope, which engages our sense of agency to

[334] Snyder, Irving, & Anderson (1991: 287)
[335] Lopez (2013: 18).

affect change, from optimism, which merely involves some expectation of change. Moreover, while the above mentioned definitions make some reference to "successful agency" and "the power" to make the future better than the present, neither of these is based in any clear or substantive account of the normativity of agency, let alone one as grounded in our moral capacities as is Kant's account of hope. My analysis of that account in the context of Kant's theory of moral education can thus provide a critical lens through which to re-examine assumptions about hope shaping contemporary educational theory and practice.

A second respect in which my analysis bears on contemporary discourse concerns what some have referred to as "educationism": the idea that poverty and inequality can be solved with education. This idea is thought to have motivated and shaped educational reform policy going back some decades. It is also an idea that has come under scrutiny, and my research shows how Kant might back that scrutiny. While, for Kant, education is a central part of our becoming virtuous agents capable of fulfilling our duty in contributing to the highest good, he also recognizes that education, whether specific pedagogical practices or entire educational institutions, are limited in the extent to which they can address either individual or social moral problems. Like Rousseau, Kant believed that education was part of a multipronged solution that also engaged general political and economic structures, as well as institutions in civil society besides educational institutions. Thus, whereas educationists place all their hope in education as a panacea to the ills of society, Kant sought to acknowledge the limits of education without sacrificing its promise.

This point about the limits of education brings me to where this dissertation began: Maria von Herbert's plea to Kant for some assurance of hope in the face of hopelessness. I used that tragic story to prompt the central question of my inquiry: why (and how) should we cultivate

hope through moral education? In his one and only response to her, Kant appears to skirt around Maria's pleas for hope, which might explain why the existing secondary literature pays little to no attention to the role of hope in their correspondence. What Kant tells Maria is that his letter followed "the customary divisions of a sermon: doctrine, discipline, and solace, of which I beg you to pay attention somewhat more to the first two, since the last, and your lost contentment with life, will surely be recovered by itself when once these others have had their effect."[336] As I mentioned in the introduction, Kant appears to be making a psychological assumption that Maria will obtain the kind of hope necessary for moral agency if she first attends to more fundamental matters of "doctrine" and "discipline." While it seems clear that the pertinent doctrine is the basis of Kant's moral theory, namely, the categorical imperative and the moral duties derived from that imperative, it is less clear on its face what Kant took to be the appropriate discipline for eliciting hope. But my detailed reconstruction of Kant's moral pedagogical method helps to elucidate the discipline needed for hope. Indeed, upon re-examining Kant's one and only response to Maria, we can find hints that he was following that pedagogical method in the matters of doctrine and discipline that shaped his letter. For example, in his letter, Kant appears to offer conceptual tools in order to strengthen Maria's capacity for thinking about friendship and deception from a moral perspective, which is similar to the task of the first of his two pedagogical exercises he outlines. At least initially, then, Kant appears willing to take on the role of Maria's "moral physician" and apply the pedagogical tools he thinks needed to instill hope in Maria.

However, as I explained before, Kant's correspondence with Maria was seemingly, and tragically, cut short. J.B. Erhard, a mutual acquaintance of Kant and Maria, wrote to the philosopher to say that she was not in a state in which she would be receptive to moral instruction, for she had lost all "delicacy": "her moral feeling is totally severed from prudence

[336] B 11:334.

and is therefore coupled with fantasy, a more subtle sensibility."[337] Trusting Erhard's judgment, Kant discontinued his correspondence with Maria, instead passing her subsequent letters, along with Erhard's, on to the young daughter of a friend as an "example of warning" of "sublimated fantasy" (11:204). Apparently, Kant saw no need in belaboring his point that the right kind of hope, or solace, would only follow from adopting the appropriate theory and practice. From his perspective, Maria suffered from a sort of enthusiasm that made his pedagogical method of no use. While Kant is right to recognize that there are limits to what moral education can do, I think he makes up those limits based on a questionable assessment of Maria's mental state. Part of what is questionable about Kant's assessment of Maria is that it seemingly disregards the oppressive societal conditions under which Maria was living and that were the subtext of both her letters and Erhard's letter to Kant. As stated, Kant's theory of moral education seems to operate under assumptions of an idealized society, but it is not obvious that need be the case. Instead, we might follow the model of Rae Langton's essay on Kant's correspondence with Maria, in which Langton articulates a model of Kantian ethics under oppressive conditions. Just as, for Langton, Kant's moral theory provides the conceptual resources for persons to understand and undo the domination contract, we might view his pedagogical method as likewise having liberatory potential, even if his theories, as stated, cannot completely redress injustices or conditions of hopelessness.[338]

[337] B 11:203-4.
[338] Langton (1992: 501-505).

References

Bacon, Stefano. 2010. “The Meaning of the *Critique of Practical Reason* for Moral Beings: the Doctrine of Method of Pure Practical Reason.” In *Kant’s Critique of Practical Reason: A Critical Guide*, ed. Andrews Reath and Jens Timmerman, pp. 197–215. Cambridge: Cambridge University Press.

Basevich, E. 2022. “The Promise and Limit of Kant’s Theory of Justice: On Race, Gender and the Structural Domination of Labourers.” *Kantian Review*, 27 (4): 541-555.

Baxley, Anne Margaret. 2010. *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge: Cambridge University Press.

Bejan, Teresa. 2010. “Teaching the *Leviathan*: Thomas Hobbes on Education.” *Oxford Review of Education*, 36 (5): 607-626.

Blöser, Claudia. 2019. “Hope as an Irreducible Concept.” *Ratio* 32(3): 205–214.

Blöser, Claudia and Titus Stahl. 2017. “Fundamental Hope and Practical Identity.” *Philosophical Papers* 46(3): 345–71.

Bovens, Luc. 1999. “The Value of Hope.” *Philosophy and Phenomenological Research* 59(3): 667–681.

Brandom, Robert. 2009. *Reason in Philosophy: Animating Ideas*. Cambridge, MA: Harvard University Press.

Chignell, Andrew. 2013. “Rational Hope, Moral Order, and the Revolution of the Will.” In *Divine Order, Human Order, and the Order of Nature: Historical Perspectives*, ed. Eric Watkins, pp. 197–218. Oxford: Oxford University Press.

—. 2020. “Hope and Despair at the Kantian Chicken Factory”. *Kant and animals*, 213.

—. 2023. “Demoralization and Hope: A Psychological Reading of Kant’s Moral Argument.” *The Monist*, 106 (1): 46-60.

Cohen, Joshua. 2010. *Rousseau: A Free Community of Equals*. New York: Oxford University Press.

Crowther, Paul. 1989. *The Kantian Sublime: From Morality to Art*. Oxford Philosophical Monographs. Oxford/New York: Clarendon Press/Oxford University Press.

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

Dean, Richard. 2012. “A plausible Kantian argument against moralism.” *Social Theory and Practice*, 38 (4): 577-597.

Deligiorgi, Katerina, 2014. “The Pleasures of Contra‐purposiveness: Kant, the Sublime, and Being Human.” *The Journal of Aesthetics and Art Criticism*, 72 (1): 25-35.

Dent, Nicholas. 1988. *Rousseau: An Introduction to His Psychological, Social and Political Theory*. New York: Blackwell.

—. 1992. *Rousseau: A Dictionary*. New York: Blackwell.

—. 2005. *Rousseau*. New York: Routledge.

Edgoose, Julian. 2010. “Hope in the unexpected: How can teachers still make a difference in the world?” *Teachers College Record*, 112 (2): 386-406.

Engstrom, S. 2010. “The Triebfeder of Pure Practical Reason.” In *Kant’s Critique of Practical Reason: A Critical Guide*, ed. A. Reath and J. Timmermann, pp. 90–118. Cambridge: Cambridge University Press.

Ginsborg, Hannah. 1991. “On the Key to Kant’s Critique of Taste.” *Pacific Philosophical Quarterly* 72(4): 290–313.

—. 2015. *The Normativity of Nature: Essays on Kant's* Critique of Judgement. Oxford: Oxford University Press.

Hobbes, Thomas. [1651] 1998. *Leviathan*, ed. J.C.A. Gaskin. Oxford: Oxford University Press.

Huber, Jakob. 2019. "Defying Democratic Despair: A Kantian Account of Hope in Politics." *European Journal of Political Theory* 20(4): 719–738.

Kant, Immanuel. [1795] 1996. "Toward Perpetual Peace." In *Practical Philosophy*, trans./ed. Mary J. Gregor, pp. 311–352. Cambridge: Cambridge University Press.

—. [1788] 1997a. *Critique of Practical Reason*, trans. Mary J. Gregor. New York: Cambridge University Press.

—. [1785] 1997b. *Groundwork for the Metaphysics of Morals*, trans. Mary J. Gregor. New York: Cambridge University Press.

—. [1797] 1997c. *Metaphysics of Morals*, trans. Mary J. Gregor. New York: Cambridge University Press.

—. [1781 / 1787] 1998a. *Critique of Pure Reason*, trans. Paul Guyer and Allen W. Wood. New York: Cambridge University Press.

—. [1792] 1998b. *Religion Within the Mere Boundaries of Reason*, trans. Allen Wood and George di Giovanni. Cambridge: Cambridge University Press.

—. 1999. *Correspondence*, Volume 11. Cambridge University Press.

—. [1790] 2000. *Critique of Judgment*, trans. Paul Guyer and Eric Matthews. New York: Cambridge University Press.

—. [1798] 2008. "Anthropology from a Pragmatic Point of View," trans. R. Louden. In *Anthropology, History, and Education*, ed. R. Louden and G. Zöller, pp. 227–429. Cambridge: Cambridge University Press.

Korsgaard, Christine. 1996. *The sources of normativity*. Cambridge University Press.

—. 2009. *Self-constitution: Agency, identity, and integrity*. Oxford: Oxford University Press.

Kwong, Jack M.C. 2019. "What Is Hope?" *European Journal of Philosophy* 27(1): 243–254.

Locke, John. [1693] 1996. Some thoughts concerning education. Hackett.

Lopez, Shane. 2013. *Making hope happen: Create the future you want for yourself and others*. Simon and Schuster.

Louden, Robert. 1992. "Go-carts of Judgement: Exemplars in Kantian Moral Education." *Archiv für Geschichte der Philosophie* 74 (3): 303-322.

Martin, Adrienne M. 2013. *How We Hope: A Moral Psychology*. Princeton: Princeton University Press.

McGeer, Victoria. 2004. "The Art of Good Hope." *Annals of the American Academy of Political and Social Sciences* 592: 100–127.

McClure, Christopher. 2014. "War, madness, and death: the paradox of honor in Hobbes's *Leviathan*." *The Journal of Politics*, 76 (1): 114-125.

Meirav, Ariel. 2009. "The Nature of Hope." *Ratio* 22(2): 216–233.

Merritt, Melissa McBay. 2011. "Kant on Enlightened Moral Pedagogy." *Southern Journal of Philosophy* 49 (3): 227–253.

Moran, Kate. 2009. "Can Kant have an account of moral education?" *Journal of Philosophy of Education*, 43 (4): 471-484.

—. 2012. *Community and progress in Kant's moral philosophy*. Catholic University of America Press.

—. 2014. "Delusions of Virtue: Kant on Self-Conceit." *Kantian Review* 19(3): 419–447.

Munzel, G.F. 2002. "'Doctrine of Method' and 'Closing'." In *Immanuel Kant: Kritik der prakitischen Vernunft*, ed. Otfried Höffe, pp. 151–163. Berlin: Akademie Verlag.

—. 1999. *Kant's conception of moral character: the "critical" link of morality, anthropology, and reflective judgment*. University of Chicago Press.

—. 2012. *Kant's conception of pedagogy: toward education for freedom*. Northwestern University Press.

Neuhouser, Frederick. 2008. *Rousseau's Theodicy of Self-Love: Evil, Rationality and the Drive for Recognition*. Oxford: Oxford University Press.

O'Neill, Onora. 1996. "Kant on Reason and Religion." *Tanner Lectures on Human Values*, delivered at Harvard University, April 1–3, Cambridge, MA.

Pettit, Philip. 2004. "Hope and Its Place in Mind." *Annals of the American Academy of Political and Social Science* 592(1): 152–165.

Rawls, John. 2007. *Lectures on the History of Political Philosophy*, ed. Samuel Freeman. Cambridge, MA: Harvard University Press.

Reath, Andrews. 2006. *Agency and Autonomy in Kant's Moral Theory: Selected Essays*. Clarendon Press.

Rousseau, Jean-Jacques. 1979. *Émile, or On Education*, trans./ed. Allan Bloom. New York: Basic Books.

—. 1990. *Rousseau, Judge of Jean-Jacques: Dialogues*, trans. Judith R. Bush and Christopher Kelly, ed. Roger D. Masters and Christopher Kelly. Hanover: University Press of New England.

—. 1997. “Discourse on the Origin and Foundations of Inequality Among Men or Second Discourse.” In *Rousseau: The Discourses and Other Early Political Writings*, trans./ed. Victor Gourevitch, pp. 111–231. New York: Cambridge University Press.

Segal, G. and Textor, M. 2015. “Hope as a primitive mental state.” *Ratio*, 28 (2): 207-222.

Shklar, Judith. 1969. *Men and Citizens: A Study of Rousseau’s Social Theory*. New York: Cambridge University Press.

Snyder, Charles, et al. 1991. “The will and the ways: development and validation of an individual-differences measure of hope.” *Journal of personality and social psychology*, 60(4): 570-585.

Stitzlein, S.M. 2019. *Learning how to hope: Reviving democracy through our schools and civil society*. Oxford University Press.

Sussman, David. 2015. “The Highest Good: Who Needs It?” In *The Highest Good in Aristotle and Kant*, eds. Joachim Aufderheide & Ralf M. Bader. Oxford University Press: 214-228.

Sweet, Kristi. 2023. *Kant on Freedom, Nature, and Judgment: The Territory of the Third Critique*. Cambridge University Press.

Timmermann, Jens. 2010. “Reversal or retreat? Kant's deductions of freedom and morality.” In *Kant's ‘*Critique of Practical Reason*’: A Critical Guide*, eds. Andrews Reath and Jens Timmermann. Cambridge Critical Guides. Cambridge: Cambridge University Press; 2010:73-89. doi:10.1017/CBO9780511770869.006

Watkins, Eric. 2014. “Nature in General as a System of Ends.” *Kant’s Theory of Biology*. De Gruyter: 117-130.

Wood, A. 1996. “Self-Love, Self-Benevolence, and Self-Conceit.” In *Aristotle, Kant, and the Stoics*, ed. S. Engstrom and J. Whiting, pp. 141–161. Cambridge: Cambridge University Press.

Zammito, John H. 1992. *The Genesis of Kant’s* Critique of Judgment. Chicago: University of Chicago Press.

Zuckert, Rachel. 2018. “Is Kantian Hope a Feeling?” In *Kant and the Faculty of Feeling*, ed. K. Sorensen and D. Williamson, pp. 242–259. Cambridge: Cambridge University Press.