NORTHWESTERN UNIVERSITY

Politics after Totalitarianism: Rethinking Evil, Action, and Judgment in Kant, Arendt, and Lyotard

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Political Science

By

Javier Darío Burdman

EVANSTON, ILLINOIS

June 2018

Abstract

Hannah Arendt claimed that "totalitarianism" emerged as a new word to designate a new form of evil in politics. However, she added, the use of a new word does not imply that we understand what this novelty consists in. This dissertation is an attempt to understand the novelty of totalitarian evil. The central argument, following Arendt, is that this understanding requires a reconsideration of the notions of action and judgment. A further argument is that a radical reconsideration of action and judgment in response to a new form of evil in modernity takes place in Immanuel Kant's practical philosophy, which influenced the political thought of two prominent post-totalitarian thinkers: Arendt and Jean-François Lyotard. The first part of the dissertation analyzes the relationship between evil and action in Kant and Arendt. For both thinkers, evil consists in the destruction of responsibility for action, which stems from the capacity to begin, transform ourselves, and establish relationships with others. The second part analyzes the relationship between evil and judgment in Kant and Lyotard. In their view, evil corrupts the power of judgment by concealing its uncertainty under absolute rules that preclude disagreement. The conclusion argues that action and judgment imply the exposure and endurance of uncertainty, while evil stems from the attempt to eliminate uncertainty.

Acknowledgments

My graduate education, which concludes with this dissertation, took place in four countries and five cities. In each place I found people and institutions that supported me. The University of Buenos Aires funded my first step in research through an "Estímulo" scholarship for senior undergraduate students. The Argentine National Council for Scientific and Technological Research (CONICET) awarded me a Scholarship for PhD studies with which I began my graduate education in Argentina. The State University of New York at Buffalo funded two years of graduate studies in Comparative Literature. Northwestern University funded my PhD studies in Political Science. As part of these studies, Northwestern's Paris Program in Critical Theory allowed me to spend a year in Paris as a visiting student at the École Normale Supérieure. The German Academic Exchange Service (DAAD) supported my research at the Goethe-Universität Frankfurt through a one-year Research Grant. I also received financial support through the Henry S. Bienen Fellowship for Holocaust Studies. I am thankful to the institutions that provided me with funding or hosted my research for their support and their trust.

I cannot possibly name each person throughout all these years whose intellectual support helped me write this dissertation, but there are some I cannot leave aside. In Buenos Aires, Emilio de Ipola, Gerardo Aboy Carlés, and Marcos Novaro helped me take my first steps in research and assisted me in every imaginable way. Ernesto Laclau generously supported my applications to start a graduate education in the United States. Ernesto's interest in the ideas of young students was and still is an inspiration for me. My two years in the Department of Comparative Literature at SUNY Buffalo were crucial for my intellectual development. I must thank especially Steven Miller, who supervised my master's thesis, David Johnson, Catherine

Malabou, Joan Copjec, and Rodolphe Gasché for their interest in my work, and their generous help in transferring to another program to further pursue my research interests. Copjec's persistent concern with the importance of psychoanalysis for our understanding of modern politics represents a crucial, even if not explicit influence in this dissertation. If Gasché had not taken a moment to tell me "I really think you have something there" after reading my seminar paper on the problem of evil in Kant and Arendt, this dissertation may have not been written. At Northwestern, I must thank the professors who discussed and expressed trust in my work: Bonnie Honig, Sara Monoson, and James Farr, in addition to the members of my dissertation committee. In Paris, Marc Crépon was always willing to discuss my research, and helped me become familiar with the French literature connected to it. In Frankfurt, I received valuable advice from Werner Hamacher and Christoph Menke.

While professors and advisors are the main reference point in the development of a dissertation, it is friends and colleagues who helped me navigate the everyday intellectual, emotional, and bureaucratic challenges involved in the process. In Buenos Aires, Belén Fernández Milmanda was an invaluable emotional and intellectual company throughout the challenging process of applying to universities in the United States. Diego Rossello's help and encouragement played an important role in succeeding. Among the many friends in Buffalo who made my adjustment to the United States smooth and enjoyable, I must thank specially Tyler Williams both for being a fun and supportive office mate, and for his help with developing my ideas through endless conversations. The secretary of the Comparative Literature Department, Mary Ann Carrick, helped me in every way she could, well beyond the scope of her duties. In Paris, Agnès Grivaux generously provided me with "inside knowledge" of the French academia,

which was crucial for making the most out of my time there. In Frankfurt, my friends Lena Krian and Nathan Taylor helped me both to adjust to a new country and to understand the German academic culture.

At Northwestern, I must thank Lucy Cane, Desiree Weber, Hye-Yun Kang, Tristan Bradshaw, Boris Litvin, and Emre Gercek for reading parts of my work or simply discussing ideas with me. Other good friends willing to hear and share experiences throughout the PhD program are Kevin Levay, Savina Balasubramanian, Rodrigo Barrenechea, and Isabel Castillo, and in the last year, Sabrina Jaromin, Fyodor Sakhnovski, Maziyar Faridi, Azadeh Safaeian, Anya Dagenshein, Brian McLoughlin, and Christine Goding. My dear friend Tomás Bril Mascarenhas was an invaluable company ever since my first steps in research in Buenos Aires. I am thankful to Tom for supporting me at each step in every way: intellectually, emotionally, and professionally. Lena Taub's company and support throughout cities and institutions was crucial for moving forward, enjoying the good moments, and feeling less lonely when things were tough. I am thankful to Lena's family, especially to her mother, Enilda, and to her grandfather, don Pepito, for welcoming me in Puerto Rico so many times.

It is not easy for parents and siblings to see a son or a brother move to a distant country. That is why I am deeply thankful to my family for its unconditional support, both before and throughout my PhD. My sister, Luciana, lent me some of her inexhaustible initiative, spiritedness, and open-mindedness when they were lacking in me, always reminding me of the importance of being happy. My mother, María Teresa, has always been there for me: listening to me, giving advice, visiting me, welcoming me back home, and cooking her delicious cakes each time I visited Buenos Aires. My mom's endless curiosity has inspired my intellectual endeavors

ever since I was a child. My father, José, enjoyed each step of my graduate education until he passed away, expressing his pride and excitement with each one of my accomplishments, either big or small. My dad's longstanding experience in medical research across countries, and his good memories of his years in the United States, awakened my interest in pursuing a graduate education in this country. My family and my good friends in Buenos Aires, Ale, Tolcha, Juan Pablo, Lío, Mena, Gonza, and Tom, made each visit back home a source of joy and renewed energy.

I learned essential lessons from each person in my dissertation committee. Samuel Weber showed me the importance of complementing critical thinking that cuts across fields with precision and intellectual rigor. Sam's passion for diverse questions, as well as his openness to different perspectives, are a reminder that good research is driven by intellectual curiosity rather than fixed disciplinary frames. I learned from Christoph Menke that the questions of contemporary political theory and the questions of the philosophical tradition are intertwined, and that there is no reason to choose between one and the other. Lars Toender taught me the importance of not losing sight of the political stakes driving one's research in political theory, mobilizing canonical texts in a way that speaks to contemporary discussions. Peter Fenves showed me that I should not take anything for granted when I read a text, and that deep knowledge is a more reliable interpretational tool than longstanding scholarly assumptions about its meaning. My advisor, Mary Dietz, taught me the importance of expressing my own viewpoint clearly and directly. Mary's generous support and open-mindedness helped me cope with the many challenges I had to face throughout the PhD program. It is not often the case that renowned academics are willing to engage with a student's own ideas and respect their intellectual drive,

which is why I am grateful to Mary for her responsiveness throughout convergences as well as divergences.

Index

Introduction

In her early essay "Understanding and Politics," commenting on the recent emergence of "the popular use of the word 'totalitarianism' for the purpose of denouncing some supreme political evil," Hannah Arendt claims:

> Yet while popular language thus recognizes a new event by accepting a new word, it invariably uses such concepts as synonyms for others signifying old and familiar evils—aggression and lust for conquest in the case of imperialism, terror and lust for power in the case of totalitarianism…. It is as though with the first step, finding a new name for the new force which will determine our political destinies, we orient ourselves toward new and specific conditions, whereas with a second step (and, as it were, on second thought) we regret our boldness and console ourselves that nothing worse or less familiar will take place than general human sinfulness.[1]

Ever since the end of World War II, we have been aware that "totalitarianism" represents a new form of evil, which is why a new word has appeared to designate it. Yet the acknowledgment of this novelty does not necessarily mean that we have come to terms with it, that we have overcome the temptation to go back to familiar ground by interpreting the new under the light of older, long-established concepts. In order to fully accept the novelty of a phenomenon, it is not enough to give it a new name—it is necessary to have the boldness to move and remain beyond our familiar conceptions about the conditions in which we live. This is what Arendt calls "understanding": the judgment of that which has ruined our standards for judgment, so that by our own initiative we are capable of finding new meaning in a world which has seemingly lost it.[2] The very appearance of the word "totalitarianism" shows that a new form of political evil emerged in the middle of the twentieth century, but the work of understanding this novelty is

[1] Hannah Arendt, "Understanding and Politics (The Difficulties of Understanding)," in *Essays in Understanding*, ed. Jerome Kohn (New York: Schocken, 1994), 312.
[2] Ibid., 321.

indefinitely open: "if we want to be at home on this earth, even at the price of being at home in this century, we must try to take part in the interminable dialogue with the essence of totalitarianism."[3]

Arendt's remarks were written in the years following the discovery of the camps, and the centrality of totalitarianism for the understanding of the contemporary world may seem less evident to our eyes. However, it would be mistaken to believe that the main questions of contemporary politics have moved beyond the problems posed by it. Together with other events of around the middle of the twentieth century, such as total war and anti-colonial struggles, totalitarianism constitutes a foundational experience for contemporary political theory. It marks the end of the Enlightenment's confidence in reason and historical progress as a ground for political action, and the transition to an intellectual context that has been repeatedly characterized as "post-metaphysical," "post-foundational" and "post-modern." The fact that terror and systematic mass murder could happen along, and not against, the forces of progress and modernization, showed to what extent the whole conceptual framework of modern political thought needed to be revised.[4] After totalitarianism, we no longer see universal cognitive foundations and totalizing narratives as a source of liberation, but rather distrust them as a potential source of domination and terror. As a consequence, most of the central figures of political theory in the second half of the twentieth century have departed from the quest for absolute foundations for politics, and stressed the irreducibility of disagreement and

[3] Ibid., 323.

[4] The interrelation between totalitarianism and the Enlightenment is of course the central concern of Horkheimer and Adorno's *Dialectic of Enlightenment*. On this point, it is worth noting Adorno's words on Auschwitz in *Negative Dialectics*: "that this could happen in the midst of the traditions of philosophy, of art, and of the enlightening sciences says more than that these traditions and their spirit lacked the power to take hold of men and work a change in them. There is untruth in those fields themselves, in the autarky that is emphatically claimed for them" (*Negative Dialectics*, trans. E. B. Ashton [New York: Routledge, 2004], 366-367).

heterogeneity to any higher, pre-political principle. Totalitarianism has been a persistent reference point among these figures, whose views emerged in part as a response to it, to the point that no comprehensive theory of politics today can leave aside an account of this phenomenon.

This dissertation is concerned with one of the central challenges that, according to Arendt, totalitarianism presented to the modern world, namely, understanding a new form of evil. The importance of this problem has been acknowledged in a number of studies and debates around *Eichmann in Jerusalem* and Arendt's polemical notion of "the banality of evil." However, scholars have rarely explored the larger implications of Arendt's reflections on evil for our understanding of modern politics. Most studies on the issue have focused on the continuities and discontinuities in Arendt's thought, illuminating the nuances of Arendt's shifting views on totalitarianism, but rarely addressing the importance of the problem of evil for questions of political agency and judgment, in Arendt as well as in other authors. If Arendt, otherwise wary of the use of moral categories for conceptualizing politics, insisted in thinking about totalitarianism in terms of "evil," it is because she believed that there is more at stake in this concept than the description of a specific phenomenon. As this study will show, the concept of evil is intertwined with two of the most central concepts in contemporary political theory, namely, action and judgment—which is why we cannot address one in isolation from the other two. In order to fully understand the novelty of totalitarian evil, it is not enough to obtain an accurate description of the psychology of the perpetrators, or of the historical causes that led to their crimes. Rather, it is necessary to rethink our most fundamental presuppositions about political action and judgment, so that we can understand not only a particular phenomenon, but also the world in which this phenomenon was possible. If contemporary political theory emerged in part as a response to a

new form of political evil, then understanding this form of evil is essential for understanding some of the essential presuppositions within the field.

The problem of evil in politics is of course not limited to the movements and regimes that we normally designate as "totalitarian." Phenomena such as mass murder, systematic violence, torture, ideological extremism, and segregation, seem to emerge again and again in the Western world and beyond. It would be evidently futile to attempt to develop a grand theory that accounts for every political phenomenon that we characterize as "evil." However, it is possible to inquire into the roots of our use of this notion, so as to make political theory more reflective about its meaning, and thus of some of the most basic presuppositions that underlie its approach to the phenomena that we subsume under it, either explicitly or implicitly. At a general level, we use the word "evil" to refer to actions that transgress universal standards determining what should and should not be done. But let us note the enormous complexity contained in this seemingly simple idea: what does is mean that an action "should" or "should not" take place? How do we determine this distinction? Can we agree on a standard for it? And even if we do, under what conditions do people comply with it or not? If we reflect upon these question, it becomes evident that whenever we address phenomena that we intuitively consider "evil," we implicitly rely on our most fundamental presuppositions about action and judgment. It is therefore worth asking: do these presuppositions adequately respond to evil as it manifests itself in the modern world? It is one contention of this dissertation that in order to properly address the perplexing recurrence of evil in politics, we need a better understanding of what we mean by this word. This understanding does not make up for the need of specific inquiries into the causes and structures of specific political processes that we associate with "evil," but it provides an orientation for how

to interrogate them and, in some cases, how to respond to them. If totalitarianism showed that political evil is not what we traditionally thought it was, then we must avoid explaining it in terms of a dated framework.

As a contribution to a new framework for thinking about evil, this dissertation presents the following interpretative argument: the concern with the emergence of a new form of evil in modernity can be traced back to Kant's critical philosophy, and this concern contributed to shape the concepts of action and judgment in Kant as well as in two post-Kantian political thinkers influenced by totalitarianism, namely, Arendt and Lyotard. Going back to Kant for an understanding of totalitarian evil may seem counter-intuitive—is not totalitarianism something new, which we can only understand through new categories? Certainly, but this does not mean, as Arendt shows in *The Origins of Totalitarianism*, that the multiple trends that would culminate in this phenomenon were not unfolding in the decades and even centuries before it. Kant knew nothing of totalitarianism as a political movement, much less a regime, but he did observe the incipient development of new forms of evil-doing that contained the seeds of the totalitarian mentality. This observation played a role in Kant's conception of action and judgment, which is why, as I will argue, we can better understand the importance of Kant's practical philosophy for contemporary political theory if we read it as a response to the problem of evil, rather than as a celebration of human autonomy. The relevance of Kant's practical philosophy for an understanding of totalitarianism becomes clear if we read it in dialogue with Arendt and Lyotard. Both thinkers are concerned with the novelty of totalitarian evil, and develop their notions of action and judgment in response to it. In so doing, they continue Kant's insight on the essence of action and judgment in modernity, while explicitly unpacking its political implications. This does

not mean that they follow every aspect of Kant's thought, nor that the three authors are homogeneous, but rather that they pursue a consistent conceptual framework stemming in part from the acknowledgement of the novelty of evil in modernity. It is one goal of this dissertation to reconstruct and show the importance of this perspective for political theory.

The Problem: A "New" Form of Evil

In what sense is totalitarianism a "new" form of evil? Arendt addresses this question in her essay "Social Science Techniques and the Study of Concentration Camps." The camps, Arendt claims, represent a "stumbling-block on the road toward the proper understanding of contemporary politics and society," which must lead social scientists "to reconsider their hitherto unquestioned fundamental preconceptions regarding the course of the world and human behavior."[5] The main perplexity that the camps present to our conception of human behavior is that they have no utility, that is, they serve no evident purpose that could be explained in terms of self-interest. Given the usual idea that evil deeds stem from some sort of benefit, the existence of the camps seems to be completely senseless: "if we assume that most of our actions are of a utilitarian nature and that our evil deeds spring from some 'exaggeration' of self-interest, then we are forced to conclude that this particular institution of totalitarianism is beyond human understanding."[6] The systematic extermination of entire populations, according to Arendt, was not only useless for the war efforts, but even detrimental to them, to the point that "it was as though Nazis were convinced that it was of greater importance to run extermination factories

[5] Hannah Arendt, "Social Science Techniques and the Study of Concentration Camps," in *Essays in Understanding*, ed. Jerome Kohn (New York: Schocken, 1994), 232.
[6] Ibid., 233.

than to win the war."[7] It is this anti-utilitarian nature of totalitarian crimes that render them "unprecedented." While mass murder for the sake of economic gains or expansion of power have been frequent throughout human history, the extermination of entire populations against any visible self-interest is entirely new: "the extraordinary difficulty which we have in attempting to understand the institution of the concentration camp and to fit it into the record of human history is precisely the absence of such utilitarian criteria."[8]

The problem, then, is how to understand the crimes of totalitarianism, and even in what sense they are "crimes," in the absence of the utilitarian motivations that we usually link to evil deeds. In *The Origins of Totalitarianism*, Arendt claims that "it is inherent in our entire philosophical tradition that we cannot conceive of a 'radical evil,'" that is, an evil beyond recognizable evil motives such as self-interest, greed, covetousness, resentment, lust for power, and cowardice.[9] Arendt does not explain why philosophers have not conceived of an evil beyond these motivations. In a later essay, however, she claims: "that evil is a mere privation, negation, or exception from the rule is the nearly unanimous opinion of all philosophers."[10] From the viewpoint of philosophy, evil is never done voluntarily, but only as the consequence of a failure to do good. If I kill an innocent person, it must be because of some sort of self-interest that prevents me from doing good, and not because killing an innocent person is an end in itself. Therefore, according to this conception, evil must always be explained in terms of a utilitarian motive that prevents that we do good, as opposed to a willful choice. In the case of

---

[7] Ibid.
[8] Ibid.
[9] Hannah Arendt, *The Origins of Totalitarianism* (New York: Harcourt Brace & Company, 1976), 459.
[10] Hannah Arendt, "Thinking and Moral Considerations," in *Responsibility and Judgment*, ed. Jerome Kohn (New York: Schocken, 2003), 179-180.

totalitarianism, however, this framework does not work, because systematic mass murder produces no evident benefit.

Arendt is not alone in viewing the camps as a rupture with the philosophical conception of evil. In *The Experience of Freedom*, commenting on Thomas Mann's words from 1939 according to which "we know once again what good and evil are," Jean-Luc Nancy claims that "the first requirement is not to understand by this the return to a 'well-known' good and evil."[11] In order to explain this "well-known" understanding of evil to which we must not return, Nancy claims that there are three lessons that we must heed:

> 1. the closure of all theodicy or logodicy, and the affirmation that evil is strictly unjustifiable;
>
> 2. the closure of every thought of evil as the defect or perversion of a particular being, and its inscription in the being of existence: evil as positive wickedness;
>
> 3. the actual incarnation of evil in the exterminating horror of the mass grave: evil is unbearable and unpardonable.[12]

These three elements constitute what Nancy calls "the modern knowledge of evil," which is "different in nature and intensity from every prior knowledge, though it still harbors certain of its traits (essentially, in sum, the evil that was 'nothing' has become 'something' that thought cannot reduce)."[13] Before modernity, most philosophers saw evil as nothing, in the sense that it had no cause or substance on its own, but was rather the deficiency of a cause or substance. We can only know what the good is, and then proceed to know evil by subtracting something from it. This means, in turn, that nothing "produces" evil. There is no force of evil, but only a force of goodness that, for reasons that must be explained, sometimes leads to evil. It is the attempt at an

[11] Jean-Luc Nancy, *The Experience of Freedom*, trans. Bridget McDonald (Stanford, CA: Stanford UP, 1998), 123.
[12] Ibid.
[13] Ibid.

explanation of the absence of complete goodness that has led to theodicy and logodicy, that is, to the idea that evil is ultimately justifiable and therefore, at least from a certain viewpoint, forgivable. According to this conception, the only reason why someone would do evil is because of an incapacity to do good, but never because of a willful choice, never as "positive wickedness."

Some prominent examples from the history of philosophy may help us clarify Arendt's and Nancy's claim that, until the camps, evil was conceived as deficiency or privation. We find a seed of this view in Plato's *Republic*, whose opening question is the distinction between justice and injustice. Faced with the challenge of explaining why justice is advantageous for its own sake, as opposed to merely when it is seen by others, Socrates explains that just actions are those that preserve and help achieve the inner harmony of the soul. Therefore, justice is advantageous independently of anything else that it may help us attain, for there is nothing more desirable than to order the parts of the soul into a harmonious unity. Yet in order to know what justice is, we need to know what produces this harmony, which is why the just person "believes that the action is just and fine that preserves this inner harmony and helps achieve it, and calls it so, and regards as wisdom the knowledge that oversees such action."[14] Wisdom is the knowledge of those actions that preserve the harmony of the soul, while its opposite, ignorance, is the belief that destroys it. No one wants to destroy the harmony of the soul, which is why only ignorance regarding what is good for it can lead to unjust actions: "in the case of just and beautiful things, many people are content with what are believed to be so, even if they aren't really so, and they

[14] Plato, *The Republic of Plato*, trans. Allan Bloom (New York: Basic Books, 1991), 120.

act, acquire, and form their own beliefs on that basis."[15] Unjust deeds do not stem from a desire for injustice, but rather from a mere belief (as opposed to knowledge) of what the good is. What is good and what is just are the same thing, yet only those who know the good will act justly. As Socrates explains, "every soul pursues the good and does whatever it does for its sake. It divines that the good is something but it is perplexed and cannot adequately grasp what it is or acquire the sort of stable beliefs it has about other things."[16] The root of unjust actions is therefore ignorance, namely, the mere belief that fails to adequately grasp what is truly good.

The view that evil is a deficiency is also present in Aristotle, whose account of it is based on error. According to Aristotle, the good is always the object of wish, but this does not mean that all men know what the good is, for "each state of character has its own ideas of the noble and the pleasant, and perhaps the good man differs from others most by seeing the truth in each class of things, being as it were the norm and the measure of them."[17] Like Plato, Aristotle describes the good man as he who knows the good and is not deceived by its mere appearance: "in most things the error seems to be due to pleasure; for it appears a good when it is not."[18] Consequently, the wicked man "is ignorant of what he ought to do and what he ought to abstain from, and it is by reason of error of this kind that men become unjust and in general bad."[19] Aristotle's conception of the good, as well as of the source of error, is fundamentally different from Plato's. However, in very general terms, both authors agree that it is always advantageous to do good, which is why

[15] Ibid., 179.
[16] Ibid.
[17] Aristotle, *Ethica Nicomachea*, in *The Basic Works of Aristotle*, ed. Richard McKeon (New York: Random House, 2001), 971.
[18] Ibid.
[19] Ibid., 966.

only error can lead a person to do evil. In this sense, evil is a privation of knowledge, for anyone who knows what is truly good would act justly.

The importance of this philosophical conception of evil for Christianity is visible in Augustine. His thought is influenced by Plotinus who, following Plato's view that the sensible world is an imperfect imitation of the intelligible world, conceives of evil as an imperfect imitation of the good.[20] Arguing against the Manicheans, for whom evil constitutes a reality independent from God (which implies that God is limited), Augustine claims that evil is mere lack of goodness: "there is no such entity in nature as 'evil'; 'evil' is merely a name for the privation of good."[21] Augustine attributes the idea of evil to our lack of understanding: if we could understand God's plan, we would see that even that which seems to be evil has a good purpose. This leaves open, however, the question of moral evil, that is, the question of why human beings do evil, even if it serves God's plan. Consistently denying the reality of evil, Augustine explains a wrong act in terms of what he calls a "deficient cause," which is a lack of sufficient cause to do good: "things have efficient causes, the higher their degree of reality, the greater their activity in good, for it is then that they are really active; but in so far as they fail, and consequently act wrongly, their activity must be futile, and they have deficient causes."[22] Activity is necessarily goodness, while failure to be active, or to have an efficient cause, is evil. The reason, Augustine clarifies, is that this lack of activity leads us to turn to things of a lower order, which have less reality than the Supreme Being. While this turning itself is a voluntary choice, which is why it can be justly punished, it does not have positive cause—it is like

[20] Plotinus, *The Enneads*, ed. John Dillon, trans. Stephan MacKenna (New York: Penguin, 1991), 10.
[21] Saint Augustine, *City of God*, ed. Henry Betteson (London: Penguin, 2003), 454.
[22] Ibid., 480.

darkness or silence: "we can only be aware of them by means of eyes and ears, but this is not by perception but by absence of perception."[23] Like silence, evil does not have a cause of its own, but stems rather from the absence of a cause.

The idea that evil stems from error or weakness plays a central role in the main philosophical systems of modernity. For both Spinoza and Leibniz, the idea of evil stems from a failure to understand the good. Influenced by the mechanistic worldview of modern science, both philosophers hold that all actions are predetermined by the universal order of things. What distinguishes a good from an evil action is the degree of understanding by the one who acts. According to Spinoza, all actions serve God, but while good actions do so willingly, evil actions do so unwillingly: "not knowing God, the wicked are but an instrument in the hands of the Maker, serving unconsciously and being used up in that service, whereas the good serve consciously, and in serving become more perfect."[24] Goodness thus depends on knowledge of God, for only those who possess this knowledge act out of love for Him. Those who lack this knowledge, by contrast, need to be constrained by rewards and punishments, which is why "the wise alone… live with a peaceful and stable purpose, not like the impious whose minds fluctuate between different passions, and therefore… possess neither peace nor calm."[25] Similarly for Leibniz, while the will strives for the good, this striving is limited by the failure to know what the perfect good is, which is why "evil is… like darkness, and not only ignorance but also error and malice consist formally in a certain kind of privation."[26]. According to Leibniz, this

[23] Ibid.
[24] Baruch Spinoza, *The Letters*, ed. Samuel Shirley (New York: Hackett, 1995), 143.
[25] Benedict de Spinoza, *Theological-Political Treatise*, ed. Jonathan Israel, trans. Michale Silverthrone and Jonathan Israel (Cambridge: Cambridge UP, 2007), 66.
[26] G. W. Leibniz and Samuel Clarke, *Correspondence*, ed. Roger Ariew (Indianapolis, IN: Hackett, 2000), 145.

limitation proceeds from the senses, which divert us from the perfect good. The limited pleasure that we obtain through our senses prevents us from fully striving for the highest form of pleasure, which is virtue and communion with God. For both Spinoza and Leibniz, to know the perfect good cannot but lead to do what is good, which is why only error can lead to evil.

How do the camps contrast with the view that evil stems from error or ignorance? Commenting on the conception of evil as privation, Arendt claims that "the most conspicuous and most dangerous fallacy in the proposition, as old as Plato, 'Nobody does evil voluntarily,' is the implied conclusion, 'Everybody wants to do good.'"[27] Evidently, the idea that "everybody wants to do good" does not mean that everybody does good, but rather that everybody aspires to do good and is deviated from it by self-interest. In other words, we can only do good voluntarily, while evil is always the indirect effect of our passions and appetites. However, as we have seen, systematic mass murder cannot be explained in terms of a motivation that leads people to do evil unwillingly, because the extermination of entire populations is an end in itself. This is why, as Arendt claims, totalitarianism represents an "unprecedented" evil, which demands that we reconsider our fundamental presuppositions about human action. As a contribution to this reconsideration, this study will pursue the following question: what sort of motivation underlies the kind of evil manifested by totalitarianism?

Antecedents and Approach

This study has four prominent antecedents, which have acknowledged the importance of the question of evil for contemporary social and political philosophy. The first one is Susan

[27] EJ 180.

Neiman's *Evil in Modern Thought*, which examines the problem of understanding evil in a series of modern thinkers from Leibniz to Freud. Neiman follows Arendt in considering the Nazi camps as the manifestation of a new form of evil, and compares this break with the one produced by the Lisbon earthquake of 1755. Although both the Lisbon earthquake and the camps presented different problems, they posited a common challenge to the longstanding philosophical attempt to explain and justify evil. Lisbon shook the faith in the goodness of Creation, which relies on the assumption that morality and nature are part of a same order: "since Lisbon, natural evils no longer have any seemly relation to moral evils; hence they no longer have meaning at all."[28] As a consequence, Neiman claims, theodicy in the sense of a totalizing account of the world that renders all events intelligible was no longer possible. Yet theodicy was not completely abandoned, but rather shifted from Providence to human practice, as in Hegel's and Marx's accounts of history. It is this second form of theodicy that had to be abandoned after Auschwitz, which "revealed a possibility in human nature that we hoped not to see"—a possibility that cannot be explained in terms of progress toward civilization and culture.[29] Neiman's goal is not present a systematic theory of evil in response to either of the two events, but rather to show that the development of modern philosophy can be interpreted as a response to both, even when addressing questions that appear to touch exclusively on matters of cognition. This is a perspective that this dissertation shares: the question of evil, far from being an isolated question, is intertwined with a series of fundamental theoretical presuppositions that affect seemingly unrelated problems. However, by contrast to Neiman's focus on questions of cognition and the

[28] Susan Neiman, *Evil in Modern Thought: An Alternative History of Philosophy* (Princeton, NJ: Princeton UP, 2004), 250.
[29] Ibid., 254.

intelligibility of the world, this study will analyze the implications of the problem of evil for questions of political agency and judgment, which are at the center of debates within political theory today.

Richard Bernstein's *Radical Evil: A Philosophical Interrogation* represents a second comprehensive attempt to analyze the problem of evil in a series of modern philosophers. One central argument of Bernstein's study constitutes an essential point of departure of this dissertation: Kant, he claims, "is the modern philosopher who initiates the inquiry into evil without explicit recourse to philosophical theodicy."[30] As we have seen, philosophical theodicy is linked to an understanding of evil as deficiency, as a mere absence of goodness. Kant's approach departs from this conception in considering evil as a choice. As Bernstein point out, the premise of this conception of evil is that human beings are inherently responsible for it: "there is one cardinal principle that Kant refuses to compromise in any way.... Human beings are morally accountable and responsible for whatever they become, for the maxims that they adopt, even for their moral disposition."[31] Building on Bernstein's approach, Chapter 2 of this dissertation analyzes the relationship between evil and responsibility in Kant, showing its connection to the phenomenon that Arendt called "the banality of evil." By reading Kant's thesis of radical evil in the larger context of his practical philosophy, the chapter shows the theoretical and historical challenges to which such thesis responds. The goal is to obtain a more systematic understanding of some of the elements identified by Bernstein in Kant's conception of evil, and to explore their reverberations in Kant's conception of action and judgment.

---

[30] Richard J. Bernstein, *Radical Evil: A Philosophical Interrogation* (Cambridge: Cambridge UP, 2008), 4.
[31] Ibid., 43.

By contrast to Neiman's and Bernstein's concern with the emergence of a new form of evil in modernity, Charles T. Mathewes's *Evil and the Augustinian Tradition* situates conceptions of contemporary evil in a line of continuity with Augustine. This approach is, in my view, misguided, and it runs counter to the one that this dissertation pursues. Adopting Augustine's conception of evil as deficiency, Mathewes situates the thought of Reinhold Niebuhr and Arendt in a line of continuity with it. While there are passages in Arendt that allow for this interpretation, a systematic analysis of her remarks on evil runs against it, as Chapter 1 of this dissertation will show. Mathewes is right in pointing out the importance of the question of evil today, but his analysis runs short of explaining how this question has been affected by the experience of totalitarianism. Seemingly unconcerned with the question of whether Augustine's understanding of evil provides an accurate framework for its manifestations in the modern world, Mathewes presents a "therapy" for (rather a "theory" of) evil—that is, a series of practices for dealing with evil, rather than an account of it.[32] This apparently transhistorical therapy does not consider historical transformations and events, relying rather on a series of general attitudes that would allow us to cope with evil in general. By contrast to this approach, this dissertation assumes that there is no such thing as a "therapy" for evil, and if there is such a thing as a response to it, it requires first a "theory" or, in Arendt's terms, an "understanding" of it.

An important contribution to this understanding, as well as to its broader implications for contemporary political theory, is presented in Simona Forti's *New Demons: Rethinking Evil and Power Today*. This study is innovational in two ways: first, it distinguishes and establishes a dialogue between pre-totalitarian and post-totalitarian views on evil; second, it pursues the

[32] Charles T. Mathewes¸ *Evil and the Augustinian Tradition* (Cambridge: Cambridge UP, 2004), 203.

implication of these views for questions that transcend the problem of evil narrowly understood. Forti claims that while modern philosophers such as Kant, Schelling, and Nietzsche conceived of evil as a strive for absolute power, post-totalitarian thinkers such as Arendt and Foucault conceive of it in terms of what she calls (changing Arendt's famous term) "the normality of evil," which unfolds through the mechanisms of biopower. Interpreting evil through the lenses of power, Forti identifies a post-totalitarian paradigm according to which evil does not stem from any subjective intention, but rather from the routinization and normalization of a series of practices that render individuals unaware of the implications of their actions. This represents a shift "from a purely subjective idea of evil—hence, aimed at grasping the actor's evil attitude and intentions—to a notion that we might call the 'bureaucratization of evil.'"[33] While Forti presents a compelling case for the need of a new approach to contemporary manifestations of evil—an approach to which I am indebted—this dissertation takes issue with her contrast between a pre-totalitarian and a post-totalitarian paradigm, and with the distinction between a subjective and a non-subjective vision on evil on which it is based. A careful reading of Kant's thesis of radical evil, as developed in Chapter 2, will show that the idea that evil is "normal" and "bureaucratic" does not stand in contrast to the idea that it is intentional and "subjective." On the contrary, it is the subjective propensity to conceal one's intentions under routinized behaviors that is at the root of modern evil.

While this dissertation is indebted to elements of Bernstein's, Forti's, and Neiman's studies, its approach to the question of evil differs from them. I do not intend to present an overview of multiple modern perspectives on evil, but rather to develop a systematic inquiry into a thought of

[33] Simona Forti, *New Demons: Rethinking Evil and Power Today* (Stanford, CA: Stanford UP, 2015), 142.

evil unfolding through the three authors involved. This approach aims at systematicity rather than extension, which means that a series of essential authors for addressing the problem of evil in modernity must be left aside (most saliently F. W. J. Schelling's *Philosophical Investigations into the Essence of Human Freedom*, whose complexity would demand an entire study on its own). The goal is to gain greater clarity on the conceptual problems involved in addressing the question of evil, and on how this problem connects to questions of agency and judgment. Scholars have often pointed out that there is no such thing as a "theory" of evil, but only a series of responses to it. This is only true if by "theory" we understand a complete, totalizing account of the notion. If, by contrast, we take "theory" in terms of what Arendt calls "understanding," that is, as the invention of new judgments that provide meaning to new phenomena, then a theory of totalitarian evil is possible. This theory does not attempt to provide a complete explanation of the essence of evil, but rather to develop new ways for thinking about it in response new experiences. It is this theory (this understanding) that Kant, Arendt, and Lyotard present us with.

Authors and Structure

The three authors that are the focus of this dissertation address the novelty of evil in modernity and its implications for the notions of action and judgment. There are certainly many other thinkers who may have been included, but there are important reasons for focusing the inquiry on these three. In responding to totalitarianism, both Arendt and Lyotard find in Kant, and especially in his later work, resources for an anti-totalitarian conception of action and judgment. Although neither of them systematically engages with Kant's conception of evil (beyond Arendt's frequent but rather brief remarks on "radical evil"), this conception is one reason why Kant's practical philosophy is nevertheless, for both, essential for understanding

political action and judgment in the post-totalitarian world. If we focus on the problem of evil in modernity, we can see a consistent line of inquiry unfolding throughout the three authors—a line of inquiry that was lost in midst of philosophies of history after Kant, and which is still ignored (as I argue in the conclusion) by those who focus on his notion of autonomy narrowly understood. One premise of this study is that a reason why Kant provides the resources for a post-totalitarian political thought is that he acknowledged the novelty of evil in modernity, which does not stem from passion and appetite. By reconstructing the importance of the question of evil in Kant's practical philosophy, and reading it in dialogue with post-totalitarian thinkers such as Arendt and Lyotard, we can better understand the implications of reflecting upon modern evil for political theory.

The ordering of the analysis will be based upon the structure of the problem, rather than the chronology of the authors. I take as a point of departure Arendt's enquiries into the novelty of evil with totalitarianism and her formulation of the concept of action as a response to it. I then proceed to connect these enquiries with Kant's practical philosophy, following Arendt's own remarks on Kant in connection to the problem of evil. Reading Kant through the lenses of the problem of totalitarian evil, I show that that his later understanding of action and judgment contributes to clarify some of the problems contained in what Arendt called "the banality of evil." Finally, in order to grasp the implications of Kant's understanding of judgment for contemporary politics, I turn to Lyotard, who found in it a model for an anti-totalitarian political judgment. Thus, rather than tracing and reconstructing the history of a concept, the dissertation develops a Kantian framework for addressing the problem of evil after totalitarianism, as well as its implications for our understanding of action and judgment. While at some points I identify

historical elements that contribute to explain the authors' views, as well as the intellectual connections between them, I proceed mainly by tracing the conceptual links connecting Kant, Arendt, and Lyotard. One goal of this procedure is to illuminate the implications of Kant's practical philosophy for contemporary political theory by reading it in dialogue with Arendt and Lyotard, as well as to clarify problematic aspects of Arendt's and Lyotard's political views by identifying some of its sources in Kant's philosophy.

Chapter 1 analyzes the relationship between evil, action, and thinking in Arendt's work. I argue that despite Arendt's shifts on the issue, she consistently links totalitarianism to a form of evil that destroys responsibility for action. This destruction proceeds by concealing the actor's spontaneity, or his capacity to begin, under fixed processes and rules that are given to him from without, in such a way that he cannot respond for them. In response to this destruction, Arendt locates responsibility not in compliance with rules, but rather in the capacity to begin by means of action, thus disclosing an actor with a unique identity that can be forgiven or punished for what he has done. Because this disclosure depends on the presence of others, it requires the establishment of relations that render the outcome as well as the meaning of the act uncertain. This means that, counterintuitively, we are not responsible for that which we can control, which would demand a complete lack of relation with others, but rather for assuming the risk of beginning something unpredictable. A similar conception of responsibility can be found in Arendt's later analysis of thinking, in response to what she characterized as Eichmann's "inability to think." Thinking is essential for responsibility because it makes us remember what we do, thus generating a stable person behind the act. Although Arendt often describes thinking as the dialogue between the self and itself performed in solitude, she also suggests that this inner

self with whom we speak is a representative of others. Therefore thinking, like action, depends on the establishment of relations with others, and it is this establishment of relations that makes us responsible. The refusal to establish relations by means of action and thinking entails a refusal of responsibility, which is an essential characteristic of totalitarian evil.

Chapter 2 connects Arendt's remarks on "the banality of evil" with Kant's notion of "radical evil." I show that Kant's "radical evil" emerges as a response to a concern similar to the one that led Arendt to the notion of "the banality of evil," namely, a form of evil-doing that seems to lack motivations and, consequently, responsibility. By focusing on the link between evil and responsibility in Kant, I show that the thesis of radical evil constitutes a response to an issue that remained unclear in Kant's practical philosophy until its development, namely, the ground of responsibility for evil-doing. While in his main writings on practical philosophy Kant situates the source of evil in our egoistic impulses, his later text *Religion within the Boundaries of Mere Reason* situates it in choice, which means that evil is not external to, but rather integrated in freedom. Yet because this choice is inscrutable, we cannot know what our intentions are, and therefore whether our actions are good or evil. For Kant, there is no rule or procedure that determines whether a certain action is good or evil, which means that the moral worth of our actions cannot be judged with certainty. To objectify the distinction between good and evil behind fixed rules or procedures is already a way to conceal our responsibility for judging our actions anew. This concealment is precisely what Kant calls "radical evil," whose radicality consists in corrupting the choice of a maxim for action, and thus responsibility for it.

Chapter 3 turns to the problem of the judgment of action in Kant. I take as a point of departure Arendt's turn to Kant's analysis of the judgment of taste in response to what she

characterized as Eichmann's "inability to judge." Scholars have often criticized Arendt's interpretation of the judgment of taste as a model for political judgment on grounds that it lacks a successful account of validity, beyond the empirical agreement of one's community. The usual solution to this problem consists in grounding judgment upon a universal moral foundation, that is, upon the moral law. I argue that this solution does not hold within the framework of Kant's thought, according to which there is an unbridgeable gap between the moral law and phenomenal actions. In order to understand how the power of judgment can establish a relationship between the two, it is necessary to analyze how it connects the faculty whose object is phenomena, which is the imagination, with the faculty whose object is the moral law, which is reason. This connection is the object of Kant's "Analysis of the Sublime" which, I argue, shows that a judgment of action is possible not by means of a rule mediating between it and the moral law, but rather by the feeling of sublimity that the act arouses in the spectator. Given that the moral law lacks any possible representation, there is no standard for determining that an action is moral. However, there are certain actions that bring the capacity to represent to its limit, thus arousing the idea of the moral law in the spectator, who experiences this arousal as a feeling of sublimity. Precisely because it stems from the idea of the law, which we share with all beings capable of action, the sublime feeling is universally communicable, beyond any empirical community or established set of norms.

Chapter 4 examines the importance of the sublime for Lyotard's understanding of the distinction between good and evil in postmodernity. Lyotard's famous diagnosis of a "postmodern condition," in which a multiplicity of discourses interact with one another without the mediation of universally shared standards, has been often interpreted as leading to political

nihilism, in the sense that political judgments should refrain from any universalist aspiration. A focus on Lyotard's interpretation of Kant shows that this view is mistaken. Reading Lyotard's main political text, *The Differend*, in dialogue with his two main studies on Kant's critical philosophy, *Enthusiasm* and *Lectures on the Analytic of the Sublime*, I show that universality and disagreement are not mutually exclusive, but rather fundamentally intertwined. The reason is that it is in the attempt to make communicable what remains incommunicable within a certain disagreement that judgments that demand universal approval emerge. This demand for approval does not stem from any rule or procedure that could be showed and controlled by a concrete subject, which is the source of what Lyotard calls "terrorism." Rather, the need to communicate what remains incommunicable produces a feeling, akin to the "sublime feeling" in Kant, which is a sign of universal ideas. The political good, according to Lyotard, consists in inventing new idioms that allow heterogeneous parties to communicate with one another, without disavowing their heterogeneity. Evil, by contrast, consists is attempting to subordinate heterogeneity under the rules of a single, totalizing language.

By reading Kant, Arendt, and Lyotard in dialogue with one another around the problem of modern evil, this dissertation challenges some common interpretations of the notions of action and judgment in the three authors. In the conclusion, I address some of the implications of this approach for debates around the nature of political action and political judgment in contemporary political theory.

Part I

Action

Chapter 1: Arendt on Evil, Disclosure, and Responsibility for Action

The problem of evil runs through Hannah Arendt's work from early on. Although it becomes a central concern in her writings only after her engagement with Eichmann and what she called "the banality of evil," references to evil appear at crucial passages in *The Origins of Totalitarianism* and *The Human Condition*. In recent years, a number of scholars have analyzed the continuities and discontinuities between Arendt's remarks on the topic. Is evil radical, as she claimed in *The Origins of Totalitarianism*, or banal, as she suggested in *Eichmann in Jerusalem*? Are the two mutually exclusive? While these questions have received broad scholarly attention, less has been said about the importance of Arendt's thoughts on evil for her conception of politics in general. Scholarship on Arendt's notion of action and on her understanding of evil have rarely entered in dialogue with one another. Margaret Canovan and Mary Dietz have argued that Arendt's political thought, as developed in *The Human Condition*, represents a response to her analysis of totalitarianism in *The Origins of Totalitarianism*, rather than a self-contained model for modern politics.[1] However, their analyses do not inquire into the connection between the concepts of evil and action. Addressing this issue, I argue in this chapter that the problem of evil plays a crucial role in Arendt's notion of action, even if she never made this connection explicit. Conversely, Arendt's notion of action illuminates some of the theoretical problems present in her remarks on evil. Totalitarianism, in Arendt's view, represented a break with the traditional understanding of *both* evil and action. On the one hand, after the camps, we can no longer explain evil as stemming from the motivations that we usually link to evil-doing, such as

[1] Margaret Canovan, *Hannah Arendt: A Reinterpretation of her Political Thought* (Cambridge: Cambridge UP, 1992), Chapter 5; Mary Dietz, "Arendt and the Holocaust," in *The Cambridge Companion to Hannah Arendt*, ed. Dana Villa (New York: Cambridge UP, 2002).

egoism, sadism, or hatred. On the other, we can no longer ground political action upon absolute moral or cognitive principles. The question that remains open is: how do these transformations (of evil and of action) connect to one another?

This chapter addresses this question by focusing on the problem of responsibility. I argue that for Arendt, despite her otherwise shifting views on the issue, totalitarianism constitutes a form of evil that destroys responsibility. In order to develop this argument, I will show that: 1) The novelty of totalitarianism consists in the destruction of the person behind the deed, thus producing a new kind of "crime by no one"; 2) Arendt's notion of action emerges in part as an attempt to find a new ground for responsibility, which she locates in the capacity to begin and the revelatory character of action; 3) Arendt's conception of action illuminates problematic aspects of her later understanding of responsibility, which she located in the activity of thinking and the realm of the inner self. Recent studies by Garrath Williams and Rosalyn Diprose have analyzed Arendt's understanding of responsibility in *The Human Condition* and in her later writings, respectively.[2] However, these studies do not address the connection between responsibility and Arendt's thought on evil in connection to totalitarianism. This is an important blind spot given that, as will be shown in this chapter, the destruction of responsibility for action is an essential element of totalitarian crimes. By understanding how this destruction of responsibility takes place, both from the viewpoint of the system and from the viewpoint of individuals, we can better understand how Arendt's notions of action and thinking respond to this issue.

[2] Garrath Williams, "Disclosure and responsibility in Arendt's *The Human Condition*," *European Journal of Political Theory* 14(1); Rosalyn Diprose, "Arendt and Nietzsche on responsibility and futurity," *Philosophy & Social Criticism* 34(6). I discuss William's and Diprose's analyses further down.

Radical Evil and the Destruction of Responsibility

Totalitarianism, according to Arendt, represents a new form of evil that disrupts our traditional understanding of this notion. This idea appears towards the end of the analysis of the concentration camps in *The Origins of Totalitarianism*, where she claims:

> When the impossible was made possible it became the unpunishable, unforgivable absolute evil which could no longer be understood and explained by the evil motives of self-interest, greed, covetousness, resentment, lust for power, and cowardice…. Just as the victims in the death factories or the holes of oblivion are no longer "human" in the eyes of their executioners, so this newest species of criminals is beyond the pale even of solidarity in human sinfulness.[3]

While traditional motivations for evil-doing entail some sort of fault with which we can at least identify, and consequently punish or forgive, totalitarian crimes stem from no recognizable motivation. The existence of the camps cannot be explained by any sort of utilitarian self-interest, for they did not serve the war efforts and were even detrimental to it at some points: "neither military, nor economic, nor political considerations were allowed to interfere with the costly and troublesome program of mass exterminations and deportations."[4] Given our traditional conception of evil as stemming from some sort of self-interest or uncontrollable impulse, it would seem that the systematic mass murder that took place in the camps lies beyond the kinds of action that we traditionally associate with it. As a consequence, Arendt claims, "we actually have nothing to fall back on in order to understand a phenomenon that nevertheless confronts us with its overpowering reality and breaks down all standards we know."[5] This lack of understanding, however, is not complete, for

[3] Hannah Arendt, *The Origins of Totalitarianism* (New York: Harcourt Bracy & Company, 1976), 459.
[4] Ibid., 411.
[5] Ibid., 459.

> there is only one thing that seems to be discernible: we may say that radical evil has emerged in connection with a system in which all men have become equally superfluous. The manipulators of this system believe in their own superfluousness as much as in that of all others, and the totalitarian murderers are all the more dangerous because they do not care if they themselves are alive or dead, if they ever lived or never were born.[6]

Even if we cannot understand the motivation behind the camps, we do recognize their goal, namely, to make everyone (the victims as well as the perpetrators) superfluous. But what does it mean to make people superfluous? And how can those in charge of the totalitarian system make themselves superfluous?

To make human beings superfluous means to eliminate their spontaneity, that is, their capacity to act in an unpredictable way. This is a necessary condition for exercising total domination over men. According to Arendt, "totalitarianism strives not toward despotic rule over men, but toward a system in which men are superfluous." While despotic rule still presupposes that men are capable of spontaneous action, and consequently that it is at least possible that they deviate from what is expected from them, "total power can be achieved and safeguarded only in a world of conditioned reflexes, of marionettes without the slightest trace of spontaneity."[7] Therefore, to be superfluous means that one is unable to act in any other way than the one determined by a series of conditioned reflexes. This entails a lack of "individuality," that is, of "anything indeed that distinguishes one man from another."[8] From the viewpoint of totalitarianism, all men are or must be nothing but the effect of a series of preexisting causes, independently of anything that can be attributed to their individual characteristics. Only this kind of men can be the subject of total domination, in the sense that every aspect of their lives is

[6] Ibid., 459.
[7] Ibid., 457.
[8] Ibid.

controllable and made entirely predictable. The superfluous man is a man whose actions are determined by larger forces and cannot do anything by his own initiative.

One of the central threads in Arendt's analysis of totalitarianism consists in explaining how this superfluousness is produced, and how spontaneity is eliminated. Of course, lack of spontaneity does not mean that people cease to act altogether, but rather that their actions can no longer be attributed to them as individual beings driven by their own unique character. Already in the initial stages of the totalitarian movement, its organization relies on what Arendt calls the "Leader principle," according to which the Leader "claims personal responsibility for every action, deed, or misdeed, committed by any member or functionary in his official capacity."[9] This organizational principle anticipates one central characteristic of the totalitarian State, namely, the fact that no one is responsible for what he or she does, because their actions are nothing but the effect of a higher law that they do not control. Given that in the totalitarian movement only the Leader is responsible for what anyone does, "every functionary is not only appointed by the Leader but is his walking embodiment, and every order is supposed to emanate from this one ever-present source."[10] Unlike even the most tyrannical regimes, the totalitarian movement lacks a hierarchical structure that distinguishes the functionaries from the leader, which means that there is no mediation between his orders and those in charge of their execution. In the totalitarian movement, "nobody ever experiences a situation in which he has to be responsible for his own actions or can explain the reasons for them," given that "since the Leader has monopolized the right and possibility of explanation, he appears to the outside world as the

[9] Ibid., 374.
[10] Ibid.

only person who knows what he is doing."[11] Even before taking full control of the State, the totalitarian movement operates according to the principle that no one knows why they do what they do, because everyone embodies a principle that they do not understand. The members and functionaries within the movement are therefore unable to respond for their actions, because they operate on the assumption that only the leader knows the reason behind them. This constitutes an incipient state of superfluousness, in which individual actions matter only as an emanation of the will of the Leader, not as an expression of the reasons and motivations of a concrete individual.

The development of totalitarianism in power intensifies this elimination of responsibility for action up to its ultimate consequences. While the Leader principle eliminates responsibility for any deed or misdeed by members and functionaries of the totalitarian movement, the Secret Police eliminates legal responsibility for any crime committed within the totalitarian State. The reason is that the Secret Police does not prosecute people on the basis of what they have done, but rather on the basis of what they are and could potentially do—it does not deal with "suspects," but rather with "objective enemies." The totalitarian State does no need to wait for specific actions to consider some people its enemies, because it "defined its enemies ideologically before it seized power, so that categories of the 'suspects' were not established through police information."[12] And given that new objective enemies may be "discovered" according to changing circumstances, no one in a totalitarian State is safe from becoming a criminal at any given moment, independently of what he or she has done. What matters is no longer the offense, either real or merely suspected, but rather the "possible crime": "while the suspect is arrested because he is thought to be capable of committing a crime that more or less

[11] Ibid.
[12] Ibid., 423.

fits his personality (or his suspected personality), the totalitarian version of the possible crime is based on the logical anticipation of objective developments."[13]

In a totalitarian State, the criminal does not deserve punishment on the basis of what he or she has done, but rather on the basis of objective developments. In its advanced stages, the totalitarian State disregards even the pretense that these developments have any connection to the characteristics of the "criminals," and those who are deemed "unfit to live" are chosen on the basis of sheer arbitrariness. This exposes the ultimate goal of destroying the link between deed and punishment, namely, to render deeds meaningless from a legal standpoint, so that not even opposition to the regime matters: "theoretically, the choice of opposition remains in totalitarian regimes too; but such freedom is almost invalidated if committing a voluntary act only assures a 'punishment' that everyone else may have to bear anyway."[14] While transgressing a law is still an act of freedom, which is why it may be punished, totalitarianism makes punishment arbitrary, with the consequence that the very distinction between a criminal and a non-criminal act disappears. People are no longer punished on the basis of their voluntary actions, but rather on the basis of developments that have nothing to do with what they have done. This instability, as Hanna Pitkin points out, is a means by which "totalitarianism undermines the very notion of agency—people's responsibility for deeds and the consequences of those deeds."[15]

The utter destruction of the legal and moral categories by which we adjudicate responsibility for action is essential to render people superfluous and dominate them completely. This is the goal of the camps, which constitute "a world which is complete with all sensual data of reality

[13] Ibid., 426-427.
[14] Ibid., 433.
[15] Hanna Fenichel Pitkin, *The Attack of the Blob: Hannah Arendt's Concept of the Social* (Chicago: The University of Chicago Press, 1998), 90.

but lacks that structure of consequence and responsibility without which reality remains for us a mass of incomprehensible data."[16] The camps develop to its ultimate implications the destruction of the distinction between guilt and innocence, because no one is in them as a consequence of what they have done: "the categories gathered in the camps—Jews, carriers of diseases, representatives of dying classes—have already lost their capacity for both normal or criminal action."[17] This is part of what Arendt calls the killing of "the juridical person in man," which consists precisely in making juridical standards for action meaningless, so that one's free choices cease to matter. For the totalitarian regime, it does not matter whether one consents with it or not, for "the arbitrary arrest which chooses among innocent people destroys the validity of free consent."[18] This arbitrariness applies to the whole population, for anyone may be subject to arrest and become an inmate of a camp at any given moment. Therefore, although only the camp inmates are subject to the most thorough destruction of spontaneity and individuality, the mere existence of the camp-system entails that the whole population is deprived of rights, given that "the aim of an arbitrary system is to destroy the civil rights of the whole population, who ultimately become just as outlawed in their own country as the stateless and homeless."[19] If anyone may become an inmate of a camp at any given moment for arbitrary reasons, the legal distinction between guilt and innocence becomes meaningless.

We can see that there is a connection between the production of superfluousness, which is the defining characteristic of the "radical evil" of totalitarianism, and the impossibility to forgive or punish it. Forgiveness and punishment presuppose that one is responsible for their actions. In

[16] Arendt, *Origins*, 446.
[17] Ibid., 447.
[18] Ibid., 451.
[19] Ibid.

a totalitarian State, however, responsibility is not only ignored, but systematically eliminated. Arendt even suggests that the desire to escape responsibility is at the root of modern ideologies, of which totalitarianism, in her view, is the ultimate outcome. Early in the analysis of antisemitism in *Origins*, she claims that "the last century has produced an abundance of ideologies that pretend to be keys to history but are actually nothing but desperate efforts to escape responsibility."[20] Ideology, according to Arendt, is the "logic of an idea," which is applied to history so as to explain all possible events in accordance with it.[21] This means that human actions, like any other natural event, stem from this logical process, and consequently "guilt and innocence become senseless notions; 'guilty' is he who stands in the way of the natural or historical process."[22] Both the victims and the executioners of totalitarian terror are merely the tools of an objective law, which is why they are all "subjectively innocent": "the murdered because they did nothing against the system, and the murderers because they do not really murder but execute a death sentence pronounced by some higher tribunal."[23] In a totalitarian State, the subjective conditions of action do not matter, because they are nothing but the effect of an objective law. As Canovan points out, "totalitarians, including the dictators themselves, had taken refuge from responsibility in ideologies that told them what must inevitably happen."[24] Therefore, if totalitarian crimes cannot be punished nor forgiven, it is because the criminals have made themselves superfluous: they do not see their actions as stemming from their own volition, or from any reason that they could explain, but merely as part

[20] Ibid., 10.
[21] Ibid., 469.
[22] Ibid., 465.
[23] Ibid.
[24] Canovan, *Hannah Arendt*, 52.

of the unfolding of a higher process. Forgiveness and punishment presuppose that action stems from the actor's own initiative and expresses his distinct identity, which is why beginning and disclosure are essential characteristics of Arendt's notion of action in *The Human Condition*, as we shall now see.

Beginning, Disclosure, Forgiveness

The Human Condition has often been read as a continuation of Nietzsche's critique of the Platonic subordination of appearances to essences. In both Dana Villa's and Bonnie Honig's readings, Arendt's emphasis on the virtuosity of the act, by contrast to its moral or cognitive foundations, seeks to question the image according to which action stems from an autonomous subject that pre-exists the act. In this regard, according to both Villa and Honig, Arendt's notion of action resembles Nietzsche's claim that "there is no 'being' behind doing… the deed is everything."[25] However, while it is clear that both Arendt and Nietzsche criticize the idea of the subject as a substance prior to the act, Arendt's emphasis on the revelatory character of action and on the disclosure of the actor in it suggests that, taken literally, the idea that "there is no 'being' behind the doing" and that "the deed is everything" does not adequately capture her conception of action. Moreover, if we focus on the problem of responsibility in connection to her analysis of superfluousness in *Origins*, this idea would seem to run counter to the concerns informing Arendt's conception of action. If we read this conception not only as a critique of the idea that there is a substantial subject that precedes the deed, but also as an attempt to develop an alternative explanation of the conditions by which we hold a person responsible for her actions,

[25] Dana Villa, "Beyond Good and Evil: Arendt, Nietzsche, and the Aestheticization of Political Action," *Political Theory* 20(2), 281; Bonnie Honig, *Political Theory and the Displacement of Politics* (Ithaca, NY: Cornell UP, 1993), 78.

we can better understand how Arendt's notion of action responds to her earlier analysis of evil and superfluousness. As we will see, Arendt's understanding of responsibility emerges in part from a reinterpretation of Kant's notion of spontaneity. While this reinterpretation challenges the idea that there is a subject that pre-exists and causes the deed, it does not lead to the conclusion that "there is no being behind the doing," but rather that the being behind the doing (or the act) is an effect of the doing (or action).

One clear connection between Arendt's notion of action in *The Human Condition* and her earlier analysis of totalitarianism is the category of "beginning." In *Origins*, as we have seen, Arendt describes totalitarianism as a system that attempts to destroy spontaneity. The use of this term is consistent with its formulation in Kant, who defined spontaneity as "the capacity to begin a state from itself," adding that this capacity is "the real ground of imputability" for action.[26] While Arendt acknowledges the importance of Kant's understanding of free action as a beginning, she finds a problem with its detachment from the realm of appearances. In her lecture notes on the "Political Theory of Kant" from 1953, Arendt claims that Kant identifies spontaneity with activity and appearance with passivity: "Being-Appearance becomes the cleavage between Ought and Is…. But this in appearance only: The true cleavage is between man as apasive [*sic*] and an active being: Spontaneity."[27] For Kant, spontaneity belongs to the noumenal realm, which means that it does not appear. The act itself, however, does appear, which introduces a duality in the actor: "man at the same time: The beginner and subject to processes: Whenever the beginning appears, it falls into the laws of appearances: Causality, or

[26] Immanuel Kant, *Critique of Pure Reason*, eds. Paul Guyer Allen W. Wood (New York: Cambridge UP, 1998), 486.
[27] Hannah Arendt, "Political Theory of Kant," in Subject File, 1949-1975, n.d., Hannah Arendt Papers, Manuscript Division, Library of Congress, Washington D.C.

non-freedom."[28] There is clearly one element of Kant's notion of spontaneity that Arendt considers worth preserving, namely, the idea that men are capable of beginning, and that this capacity is the source of freedom. However, she questions the stark dichotomy between the spontaneous, non-appearing self and the appearing self that is not spontaneous, which introduces a dichotomy between freedom and the world of appearances: "the opposite of the Givenness of the world (Affiziertheit) is the Spontaneity of Being which so to speak is not of this world."[29] If the capacity to begin is to be a worldly rather than a transcendent condition, it must be linked to the world of appearances. This complicates the distinction between freedom as activity and appearing as passivity, as we will see. In Arendt's view, Kant's stark separation between activity as freedom and passivity as necessity stems from his failure to acknowledge that there is virtue not only in what we make ourselves, but also in the endurance of what is given to us.[30]

According to Arendt, the beginning of something new through action does not entail an absolute beginning, and thus a pure activity, but rather an insertion into an existing world that is given to us. This insertion is an actualization of the condition of natality, that is, of the fact that we are a new beginning by virtue of being born into the world: "by speaking and acting we insert ourselves into the human world, which existed before we were born in it, and this insertion is like a second birth, in which we confirm the naked fact of our being-born [*Geborenseins*], and at the same time take upon ourselves the responsibility [*Verantwortung*] for it."[31] Evidently, we are not responsible for our own birth, yet we take responsibility for it from the moment we confirm

[28] Ibid.
[29] Ibid.
[30] Ibid.
[31] Hannah Arendt, *Vita activa, oder Vom tätigen Leben* (Stuttgart: R. Piper, 1960), 165. The word "responsibility" does not appear in the earlier, English version of *The Human Condition*, where Arendt claims that by word and deed "we take upon ourselves the naked fact of our original physical appearance" (pp. 176-177).

that we are a new beginning by means of action and speech. This means that beginning is neither purely active nor purely passive: the fact of birth is given to us, yet it is on the basis of this fact that we are able to actively begin. Arendt develops this idea further in "The Crisis of Education," where she claims that "insofar as the child is not yet acquainted with the world, he must be gradually introduced to it; insofar as he is new, care must be taken that this new thing comes to fruition in relation to the world as it is."[32] Beginning is therefore not opposed to the world as it is, because one can only begin by being inserted into a pre-existing world. At the same time, the existing world can only be preserved by means of its renovation, because "to preserve the world against the mortality of its creators and inhabitants it must be constantly set right anew."[33] By contrast to Kant's stark distinction between spontaneity as activity and the given as passivity, Arendt conceives of activity and passivity as interdependent: we can begin by inserting ourselves into a world that existed before us, and this insertion is essential for preserving the world as it is given.

A second, and perhaps more essential difference between Arendt's notion of beginning and Kant's notion of spontaneity concerns its connection with appearances. For Kant, spontaneity does not belong to the world of appearances, which is ruled by natural causality and necessity, but rather to the noumenal realm. For Arendt, by contrast, beginning and appearances are intertwined, because beginning is essentially appearing: "in acting and speaking, men show who they are, reveal actively their unique personal identities and thus make their appearance in the human world."[34] Just like birth represents the physical insertion of a new person into the world,

[32] Hannah Arendt, "The Crisis in Education," in *Between Past and Future* (New York: Penguin, 2006), 186.
[33] Ibid., 189.
[34] Hannah Arendt, *The Human Condition* (Chicago: The University of Chicago Press, 1998), 179.

action and speech represent a second insertion that reveals who this new person is, given that they "contain the answer to the question asked of every newcomer: 'Who are you?'"[35] This "who" is not a substance that pre-exists the act, but rather an effect of its revelatory character. By contrast to mere physical difference, which is a characteristic common to all beings, men are distinct by virtue of actively distinguishing themselves: "human distinctness is not the same as otherness—the curious quality of *alteritas* possessed by everything that is…. Only man can express this distinction and distinguish himself."[36] We are an actor with a unique identity because we reveal this identity through action and speech. Without this revelation, we remain physically different from other people and things, but we do not have a distinct identity. Thus beginning, appearing, and distinctness are intertwined: "appearance, as distinguished from mere bodily existence, rests on initiative, but it is an initiative from which no human being can refrain and still be human."[37] To refrain from beginning means therefore not to appear to others as a distinct agent, thus losing one's humanity. And if, as Arendt claims, it is by means of beginning that we take responsibility for our own birth, refusal to begin entails a refusal of responsibility altogether—just like those in charge of the camps, indifferent to their own actions, did not care about whether or not they should have been born.

Because we can only begin by appearing to others, responsibility for action presupposes a degree of uncertainty and unpredictability. Arendt claims that "this revelatory quality of speech and action comes to the fore where people are *with* others and neither for nor against them—that is, in sheer human togetherness," and therefore "although nobody knows whom he reveals when

[35] Ibid., 178.
[36] Ibid., 176.
[37] Ibid., 176.

he discloses himself in deed or word, he must be willing to risk the disclosure."[38] We do not know who we are before the act takes place, because the "who" will be determined by the way it appears to others. Arendt compares this "who" with the Greek *daimon*, which "accompanies each man throughout his life, always looking over his shoulder from behind and thus visible only to those he encounters."[39] This incapacity to know who we are before the act is seen by others is linked to its unpredictability. The meaning of the act is not something we possess before we perform it, but rather an effect of the process that is set into motion by it: "the light that illuminates processes of action, and therefore all historical processes, appears only at their end, frequently when all participants are dead."[40] This process constitutes a story, which can only be told by someone witnessing its entire development. The actor does not make the story, in the sense of having its final product in mind from the moment he acts. Rather, it is the storyteller who discloses the meaning of the actions that make up the story: "what the storyteller narrates must necessarily be hidden from the actor himself, at least as long as he is in the act or caught in its consequences, because to him the meaningfulness of his act is not in the story that follows."[41] It is in this sense that the actor must "risk his disclosure," for "one discloses one's self without ever either knowing himself or being able to calculate beforehand whom he reveals."[42] Action reveals who we are, but only insofar as it starts an unpredictable process that retroactively illuminates the meaning of the act.

---

[38] Ibid., 180.
[39] Ibid., 179-180.
[40] Ibid., 192.
[41] Ibid.
[42] Ibid.

If we do not know the meaning of the act at the moment of its performance, and if such meaning depends on the view of others, in what sense are we responsible for it? Perhaps a more fundamental question, from Arendt's viewpoint, would be: why do we assume that we are responsible only for that which we can control? This assumption underlies Kant's claim that the ground of imputability for action is its "absolute spontaneity," that is, its quality of being a beginning completely undetermined by anything external to it, and thus purely active. Arendt, as we have seen, preserves the idea that we are responsible insofar as we begin, but she questions the idea that this beginning is absolute. On the one hand, as we have seen, every beginning is an insertion into a world that pre-exists it. On the other, each beginning establishes relations with other beginnings, which means that we are never capable of controlling the consequences of what we do. Action, Arendt claims, "acts into a medium where every reaction becomes a chain reaction and where every process is the cause of new processes."[43] In part because of the unpredictability linked to the chain reaction, we do not know who will have been revealed by the act before it has taken place. Yet it is only by virtue of beginning something new that we take responsibility for being a beginning, thus becoming unique. Arendt compares the state of having a unique personal identity to the Greek *eudaimmia* which, she claims, refers to "the well-being of the *daimon* who accompanies each man throughout life, who is his distinct identity, but appears and is visible only to others."[44] To have a unique personal identity thus depends on appearing to others, thus facing the uncertainty regarding what this identity consists in: "this unchangeable identity of the person, though disclosing itself intangibly in act and speech, becomes tangible only in the story of the actor's and speaker's life; but as such it can be known… only after it has

[43] Ibid., 190.
[44] Ibid., 192.

come to its end."[45] Without appearing to others, our identity would not be uncertain, but by the same token it would cease to be, for "whatever lacks this appearance comes and passes away like a dream, intimately and exclusively our own but without reality."[46]

The implications of this conception of personal identity for our understanding of responsibility become clear in Arendt's analysis of forgiveness. Let us restate the question regarding responsibility in different terms: why do we forgive, if the actor cannot know the outcome, and consequently the meaning of his action, at the moment he acts? Arendt responds to this question by reversing the terms: it is precisely because the actor cannot know who he is at the moment of action that we have the power to forgive him. Given that action begins a process that we are unable to control, either by predicting or reversing its outcome, it places a burden upon the actor. This burden stems from the fact that the actor "never quite knows what he is doing," and "always becomes 'guilty' of consequences he never intended or even foresaw."[47] For Arendt, it is by virtue of distinguishing the actor and giving him an identity that outlasts the act that the act places a burden upon him, rendering him guilty for a process which, although set in motion by him, he cannot control. This lack of control is precisely the reason why he can be forgiven. On this point, Arendt finds a model of forgiveness in Jesus's teachings, and specifically in the idea that the reason why we must forgive is that "they know not what they do."[48] Given that we do not know the consequences of our actions, they produce what Jesus calls "trespassing" which, according to Arendt's interpretation, means "to miss" or "fail and go astray," rather than "to sin." Trespassing has nothing to do with the psychological state of the

[45] Ibid., 193.
[46] Ibid., 199.
[47] Ibid., 233.
[48] Ibid., 239.

actor, but is rather an effect of the inherently relational nature of the act: "trespassing is an everyday occurrence which is in the very nature of action's constant establishment of new relationships within the web of relations, and it needs forgiving, dismissing, in order to make it possible for life to go on by constantly releasing men from what they have done unknowingly."[49] Forgiveness thus releases the agent from his action "going astray," allowing him to begin again, and this releasement is only possible on the assumption that trespassing is an effect of the action "going astray" by virtue of its unpredictability.

Arendt's remarks on forgiveness show that, in her view, responsibility in inherently relational. On the one hand, the act discloses a unique agent who remains responsible for the process that he has begun. On the other, because this process is inherently beyond his control, his responsibility is limited by the presence of others to whom he discloses himself and with whom he establishes relationships. As Garrath Williams points out, responsibility presupposes the establishment of relationships with others, even at the expense of our capacity to control the outcome of our actions.[50] Just like the act would have no underlying agent if it did not appear to others, the agent could not be forgiven if the outcome and the meaning of this appearing were completely under his control. In the public realm, where every action entails the establishment of relationships with others, responsibility implies a certain irresponsibility. Indeed, one of the aporias of action consists in "the impossibility to make the individual responsible for what occurs," as a consequence of which "it has always been a great temptation, for men of action no less that for men of thought, to find a substitute for action in the hope that the realm of human

[49] Ibid., 240.
[50] Garrath Williams, "Disclosure and responsibility in Arendt's *The Human Condition*," *European Journal of Political Theory* 14(1), 42.

affairs may escape the haphazardness and moral irresponsibility inherent in a plurality of agents."[51] For Arendt, morality concerns the relationship between the self and itself, as opposed to its relationship with others. Alone with myself, I can control what I do, and thus remain absolutely responsible for its outcome. But as soon as my action appears to others, I lose control, and become dependent on their view and their reactions. Therefore, action in public makes the actor partly irresponsible for what he has done—but only "partly," because it is by virtue of acting in public that the actor acquires a unique identity that endures the individual act, and that can consequently be forgiven for it. Without disclosing a who to others, there would be no durable agent behind the act that could be the subject of forgiveness: "the fact that the same *who*, revealed in action and speech, remains also the subject of forgiving is the deepest reason why nobody can forgive himself; here, as in action and speech generally, we are dependent upon others, to whom we appear in a distinctness which we ourselves are unable to perceive."[52]

It is precisely because responsibility depends on the relationship to others, and thus is never absolute, that it produces the desire to escape from plurality and the realm of human affairs. Given that every action begins a process that exceeds our control, as soon as we act we are burdened with a guilt that exceeds our capacity to assume responsibility for it, and "this is reason enough to turn away with despair from the realm of human affairs and to hold in contempt the human capacity for freedom, which, by producing the web of human relationships, seems to entangle its producer to such an extent that he appears much more the victim and the sufferer than the author and doer of what he has done."[53] The only salvation from this entanglement,

[51] Arendt, *Vita activa*, 214 (my translation); *The Human Condition*, 220.
[52] Arendt, *The Human Condition*, 243.
[53] Ibid., 233-234.

Arendt claims, "seems to lie in non-acting, in abstention from the whole realm of human affairs as the only means to safeguard one's sovereignty and integrity as a person," which leads to the identification of freedom with sovereignty, "the ideal of uncompromising self-sufficiency and mastership" which "is contradictory to the very condition of plurality."[54] The identification of freedom with absolute self-sufficiency is therefore an effect of the desire to escape the burden of guilt that action imposes upon us by virtue of starting a process that we do not control. Evidently, this identification has disastrous implications for the public realm, for it implies that the existence of others is a limitation, rather than a pre-condition for freedom. If, however, we become responsible agents by means of beginning, this identification of freedom with self-sufficiency cannot but eliminate responsibility. An individual who does not appear to others remains master of his deeds, but by the same token lacks a unique identity, that is, a "who" that endures his actions and is therefore subject to forgiveness.

Arendt's account of forgiveness connects her notion of action in *The Human Condition* and her analysis of superfluousness and the elimination of responsibility in *Origins*. Superfluousness, as we have seen, is an effect of the destruction of all the standards by which we can be held responsible for our actions, in such a way that we cease to be individuals and become mere executors of an objective process. This process, unlike the processes begun by action, is not conformed by a web of relationships with multiple beginnings, but rather by a single logical development in which multiple people behave as if they were parts of one single man. Under these conditions, there is no personal responsibility, because there are no beginnings aside from that of the process itself. To make oneself superfluous is to make oneself unforgivable, because

[54] Ibid., 234.

we can only forgive those who begin, and thus accept responsibility for the uncontrollable process that they set in motion. As Williams puts it, "a person can only answer for her deeds to the extent that her conduct is not hidden but seen by others, and is acknowledged as owing to her own initiative."[55] Given that it is intertwined with disclosure, beginning reveals a distinct, enduring actor that is the *who* underlying her act. Without this *who*, there is no stable identity that can be held responsible for the act and thus released from its consequences.

Before turning to the relationship between evil and responsibility in Arendt's later writings, let us note how the focus on responsibility contributes to clarify the ethical dimension of her notion of action. Arendt's political thought has been criticized, most saliently by Geroge Kateb and Seyla Benhabib, for lacking moral foundations, that is, an account of why certain actions are considered morally valid and some are not.[56] This criticism takes for granted that people are responsible individuals who care about the moral validity of actions. As Arendt's analysis of ideology and totalitarianism shows, however, this sense of responsibility cannot be taken for granted. In order to care about the meaning and outcome of what we do, we need to disclose ourselves to others, in such a way that our actions are linked to a durable self that outlasts the fleeting instant in which they take place. Without a durable self that can respond for what he has done, moral foundations do not concern us, for the simple reason that there is no one who can be held accountable for either complying with or transgressing them. Therefore, despite Arendt's frequent opposition of politics and morality, I believe that her deeper insight lies in the political root of moral awareness, namely, in the experience of disclosing a durable identity that outlives

[55] Williams, "Disclosure and responsibility in Arendt's *The Human Condition*," 45.
[56] George Kateb, *Hannah Arendt: Politics, Conscience, Evil* (Totowa, NJ: Rowman and Allanheld, 1984), 31; Seyla Benhabib, *The Reluctant Modernism of Hannah Arendt* (Lanham, MD: Rowman & Littlefield, 2003), 194-195.

the act. It is precisely this concern with the relationship between actions and a durable self that plays a central role in Arendt's later writings on evil, as we shall now see.

Evil and Responsibility after Eichmann

The problem of responsibility becomes an explicit concern in Arendt's work after her analysis of Adolf Eichmann's trial. This analysis led Arendt to introduce a terminological shift in her characterization of totalitarian evil: while in *Origins* and *The Human Condition* she refers to totalitarian crimes as "radical evil," explicitly borrowing Kant's term, in *Eichmann in Jerusalem* she refers to them as "the banality of evil." In her letter to Gershom Scholem from 1963, Arendt claims that this shift of terms expresses a change of mind regarding the nature of evil: "you are quite right: I changed my mind and do no longer speak of 'radical evil….' It is indeed my opinion now that evil is never 'radical,' that it is only extreme, and that it possesses neither depth nor any demonic dimension."[57] The continuities and discontinuities underlying this "change of mind" have been the subject of scholarly debate in recent years. However, little has been said about one of the main and more clear continuities between the two notions, and the one that links Arendt's understanding of evil with her analysis of action in *The Human Condition*, namely, the fact that totalitarian crimes are unforgivable and unpunishable. If we focus on this element, we see that whether "radical" or "banal," totalitarian evil stems from the destruction of responsibility for action.

While in *Origins* the destruction of responsibility is analyzed from the viewpoint of the system, in *Eichmann in Jerusalem* and later writings Arendt turns to the subjective conditions of

[57] Hannah Arendt, "A Letter to Gershom Scholem," in *The Jewish Writings*, eds. Jerome Kohn and Ron H. Feldman (New York: Schocken, 2007), 470-471.

this destruction. The question is no longer "how does totalitarianism eliminate responsibility?," but rather "why are people complicit with this elimination?" Although this question is not new in Arendt's work, the trial of Eichmann brought it to the center of her inquiries, in part because personal responsibility is the fundamental presupposition of any juridical process.[58] There is in this regard a clear resemblance between the opening pages of *Eichmann in Jerusalem* and those of *Origins*, both of which warn against the tendency to narrate the events as if the actors were representative of large historical tendencies, concealing their own motivations and choices in determining their actions. In the case of Eichmann, the problem was not only that the prosecutor tried to make him representative of anti-Semitism, as opposed to focusing on his specific actions and motivations, but also, and more fundamentally, that Eichmann himself seemed to have acted with no motivations at all—as Arendt puts it in the epilogue, "except for an extraordinary diligence in looking out for his personal advancement, he had no motives at all."[59] It is precisely this seeming lack of motivations that made the adjudication of responsibility so difficult during the trial. Of course, Arendt's point was not to exculpate Eichmann, but rather to face the novelty of his crimes as well as the challenges that they presented to our usual understanding of evil and responsibility.[60]

Eichmann's seeming lack of motivations for his actions is linked to what Arendt describes as his "inability to think" and "inability to judge." The two inabilities are connected to one

---

[58] Christoph Menke points out the juridical perspective adopted by Arendt's trial report in "Auf der Grenze des Rechts: Hannah Arendts Revision des Eichmann-Prozesses," *Merkur* 67. In her essay "Some Questions of Moral Philosophy," Arendt claims that the institution of the courtroom "rests on the assumption of personal responsibility and guilt" (*Responsibility and Judgment*, 57).

[59] Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York: Penguin, 2006), 287.

[60] Pitkin.notes that some of the most polemical remarks in *Eichmann in Jerusalem* stem not from Arendt's attempt to exculpate the criminals or incriminate the victims, but rather from "the multiple incoherences of our ordinary ways of discussing agency, responsibility, and causation in human affairs" (*The Attack of the Blob*, 214).

another, but they appear in different contexts of Arendt's analysis. The idea of thoughtlessness is introduced to describe the perplexing fact that, during the trial, Eichmann described his initial years in charge of the Jewish question as if there had been a communion of interest between him and the Jews in the goal of relocation. This is an indication of an "almost total inability ever to look at anything from the other fellow's point of view."[61] This inability was supported, Arendt believes, by Eichmann's tendency to speak by means of a repetition of clichés and stock phrases, which worked as "the most reliable of all safeguards against the words and the presence of others," with the consequence that "no communication was possible with him."[62] For Eichmann, all that mattered when narrating a past event was the stock of words and phrases he had always linked to them, and which were often given to him as part of the "language rules" used by Nazi officials to conceal the true nature of what was happening (words like "extermination," "liquidation," or "killing" would be replaced by words like "final solution," "evacuation," and "especial treatment").[63] The idea of Eichmann's "inability to judge" is introduced to describe his sense that, because the respectable society around and above him had no moral problem with sending people to their death, there was no reason to judge the moral implications of what he was doing. According to Eichmann's own words, when the extermination was decided in 1942, "the most potent factor in the soothing of his own conscience was the simple fact that he could see no one, no one at all, who actually was against the Final Solution."[64] As a consequence, given that no one around and above him had any moral concerns, "*who was he to judge?*."[65] Thus while

[61] Arendt, *Eichmann in Jerusalem*, 48.
[62] Ibid., 49.
[63] Ibid., 86.
[64] Ibid., 116.
[65] Ibid., 114.

Eichmann's inability to think protected him against the viewpoint of others, his inability to judge protected him against his own viewpoint. However, both inabilities converged in soothing his moral conscience and strengthening his compliance with duty.

Lack of thinking and judgment do not amount, however, to a disregard for any standard for action, and thus to a lack of moral conscience altogether. Eichmann, like other Nazi criminals, possessed rather an inverted kind of conscience. Arendt describes it in terms of an inversion of the feeling of pity, so that it ceases to work as an indication of criminality and becomes proof of a commitment to a historic task. Nazi officials, in Arendt's account, did not look for sadists and people who could enjoy killing innocent human beings, but rather for people they could persuade of overcoming their natural aversion for murder so as to fulfill their duty.[66] The point was that murdering an innocent person had to be seen not as a crime against that person but as a weight upon the murderer, so that "instead of saying: What horrible things I did to people!, the murderers would be able to say: What horrible things I had to watch in the pursuance of my duties, how heavily the task weighed upon my shoulders!."[67] Thus the usual indications of criminality were reversed so as to become a signal of compliance with moral duty. This is one essential reason why, according to Arendt, it was so difficult to impute evil motivations to Eichmann. In complying with duty against his own private interest, he could insist that he had always been a "good citizen." A perplexing example of this logic is Eichmann's discomfort during the trial with the fact that he had made two exceptions to his duty: helping a half-Jewish cousin and a Jewish couple in Vienna to avoid the camps. The source of the discomfort is the

[66] Arendt makes a similar remark in her much earlier writing "Organized Guilt and Universal Responsibility," in *Essays in Understanding* (New York: Schocken, 1994), 129.
[67] Arendt, *Eichmann in Jerusalem*, 106.

fact that, for Eichmann, making no exceptions represented "the proof that he had always acted against his 'inclinations,' whether they were sentimental or inspired by interest, that he had always done his 'duty.'"[68] As long as one disregards one personal interests and complies with duty, Eichmann reasoned, there is no ground upon which one's actions can be judged to be criminal.

In Arendt's view, this seeming absence of reliable standards for distinguishing between right and wrong is not simply a deceptive strategy, but rather an essential aspect of the totalitarian system. Eichmann certainly did not transgress any of the laws of the Third Reich in sending innocent people to their death, because sending innocent people to their death was precisely the law.[69] It is this coincidence between criminality and compliance with the law that constitutes, Arendt claims, the "moral point of the matter," which is reached "only when we realize that this happened within the frame of a legal order and that the cornerstone of this 'new law' consisted of the command 'Thou shalt kill,' not thy enemy but innocent people who were not even potentially dangerous."[70] In this regard, Eichmann's "inverted moral conscience" seemed to be strikingly close to a normal moral conscience, which equates compliance with duty with virtue. As a consequence, even the feeling of guilt ceased to be reliable as a source of moral guidance, for "guilt-feelings can… be aroused through a conflict between old habits and new commands… but they can just as well be aroused by the opposite: once killing or whatever the 'new morality' demands has become a habit and is accepted by everyone, the same man will feel guilty if he

[68] Ibid., 137.
[69] Arendt, "Personal Responsibility under Dictatorship," in *Responsibility and Judgment*, 43.
[70] Ibid., 42.

does not conform."[71] Even temptation, one of the clearest indications of evil, ceased work as a standard of criminality, for in the Third Reich one was "tempted" not to kill innocent people, and it required an effort to overcome this temptation and fulfill one's duty.[72] The problem of holding Eichmann and other Nazi criminals responsible is that they seemed to have acted according to all the standards that we usually associate with virtue. If this is the case, on what ground can be claim that they should have been able to tell the difference between right and wrong?

This question led Arendt to inquire into the ground of moral conscience, which she identifies with the faculty of thinking. Moral propositions, according to Arendt, do not stem from any given law, but rather from the inner dialogue between the self and itself: "I cannot do certain things, because having done them I shall no longer be able to live with myself."[73] In order to have a moral conscience, it is necessary to experience this duality within ourselves, that is, the fact we are two-in-one. This is precisely what takes place in the activity of thinking, in which "I speak with myself about whatever happens to concern me."[74] In speaking with myself, I become my own company, thus developing a concern with who I am that is at the root of moral considerations. According to Arendt, the validity of the fundamental formula of morality, according to which "it is better to suffer wrong than to do wrong," stems from the insight "that it is better to be at odds with the whole world than, being one, to be at odds with myself."[75] The ultimate standard for morality is therefore whether the self is in harmony or disharmony with itself. If we can tell the difference between right and wrong even against the opinion of everyone

[71] Arendt, "Some Questions of Moral Philosophy," in *Responsibility and Judgment*, 107.
[72] Ibid., 80.
[73] Ibid., 97.
[74] Ibid., 97-98.
[75] Ibid., 100.

around us, it is because we are always in the company of ourselves and must consequently decide whether we want to be our own company: "it is better to suffer wrong than to do wrong, because you can remain the friend of the sufferer; who would want to be the friend of and have to live together with a murderer? Not even another murderer."[76] This inner dialogue, Arendt claims, does not require any standard given to it from without.[77] The only standard for distinguishing right from wrong is whether, in doing something, we can be in harmony with ourselves. Yet this standard, Arendt claims, does not crystalize in a stable system of propositions, for thinking "does not create values; it will not find out, once and for all, what 'the good' is; it does not confirm but, rather, dissolves accepted rules of conduct."[78] In this sense, thinking is only negative: it only makes us incapable of doing certain things, because we would be unable to live with ourselves if we did them, but it does not produce propositions regarding what to do.[79]

The process of thinking is not only important because it is the ground of moral propositions, but also, and more fundamentally from the viewpoint of responsibility, because it is the ground of moral personality. To talk to oneself about the meaning of what one has done is indispensable for remembering it, for "no one can remember what he has not thought through in talking about it with himself."[80] This is why thinking is the precondition for being a person, namely, a stable self that remains one throughout its individual acts. Through the process of thought, "I explicitly constitute myself as a person, and I shall remain one to the extent that I am capable of such

[76] Hannah Arendt, *The Life of the Mind* (Orlando, FL: Harcourt, 1978), 188. See also "Some Questions of Moral Philosophy," 90.
[77] Arendt, "Some Questions of Moral Philosophy," 91.
[78] Arendt, *The Life of the Mind*, 192.
[79] Arendt, "Some Questions of Moral Philosophy," 105.
[80] Ibid., 94.

constitution ever again and anew."[81] Arendt compares this constitution of the self to "root-striking," which means that one's thoughts and memories become the root of the self which will be our company wherever we go.[82] The root sets a limit to what we are capable of doing, simply because we do not want to be in the company of someone who has done certain things. "Personality" thus refers to the quality of being a stable self that remains throughout its multiple states and deeds. For Arendt, there is nothing given about being a person, because it is the result of the activity of thinking and the remembrance that stems from it. If we suspend this activity, we cease to be persons, and our actions become a series of events without an underlying identity.

It is not difficult to see why lack of thinking entails a lack of responsibility. To be responsible for one's own deeds, there must be a self behind them that can be punished or forgiven for them. If I do not remember what I do, I cannot respond for my actions, for the simple reason that I do not acknowledge them as "mine." Loss of thinking, in other words, entails "the loss of the self that constitutes the person."[83] This is precisely the reason why "rootless evil," as Arendt characterizes Eichmann's evil in the letter to Scholem, is unforgiveable, for "in granting pardon, it is the person and not the crime that is forgiven; in rootless evil there is no person left whom one could ever forgive."[84] It is in this sense that Eichmann's crimes are not "radical," but rather "rootless" or "shallow." Those who do not strike roots are capable of doing anything, because they "skid only over the surface of events," and "permit themselves to be carried away without ever penetrating into whatever depth they may be

[81] Ibid., 95.
[82] Ibid., 101.
[83] Ibid.
[84] Ibid.

capable of."[85] This superficiality enables people to engage in the most extreme forms of evil, simply because they do not care about what they do. Those who have a root in their own person cannot do certain things because they will not be able to live with themselves, which is why "extreme evil is possible only where these self-grown roots, which automatically limit the possibilities, are entirely absent."[86] As a consequence, "the greatest evildoers are those who don't remember because they have never given thought to the matter, and, without remembrance, nothing can hold them back."[87] Extreme evil and lack of responsibility are therefore not mutually exclusive, but strangely interrelated. If we strike roots by thinking and remembering what we do, we can respond for our actions. If we do not, we simply do not care, and consequently we are willing to do anything.

Evidently, this refusal to think, as well as the lack of responsibility that stems from it, is not a mere deficiency or omission, but rather a choice. As Hanna Pitkin points out, Arendt's description of Eichmann's mentality as an "inability to think" can be misleading in this regard, insofar as it seems to suggest there is no ground upon which to hold Eichmann responsible for what he had done—which is why Arendt's report of the trial was criticized for being "exculpatory."[88] However, Arendt's later writings make it clear that lack of thinking is not simply the absence of a faculty, but rather the active decision to shun it. Just like thinking "is not a prerogative of the few but an ever-present faculty in everybody," so "everybody may come to shun that intercourse with oneself."[89] If we are "unable to think," it is not because we happen to

[85] Ibid.
[86] Ibid.
[87] Ibid., 95.
[88] Pitkin, *The Attack of the Blob*, 212.
[89] Arendt, *The Life of the Mind*, 191.

lack a capacity—a deficiency for which we could not be held accountable. Rather, we are responsible for this "inability" insofar as it stems from our decision to shun our own thinking. Thus the reason why Eichmann and other Nazi criminals seem to lack responsibility, and therefore to be unforgiveable and unpunishable, is that "they renounced voluntarily all personal qualities, as if nobody were left to be either punished of forgiven."[90] It is in this sense that "the greatest evil perpetrated is the evil committed by nobodies, that is, by human beings who refuse to be persons."[91] Personality, the quality of being a durable self that can be forgiven or punished for what he has done, depends on thinking and remembering. Those who choose not to think lack a person that can be held responsible for its deeds. As Benjamin Schupmann puts it, this choice constitutes a second order act or decision, or a decision against one's own agency.[92] To shun the intercourse with oneself amounts to eliminating our concern with what we do, as if our own deeds had no meaning for us beyond the orders or the processes that determine them. This is the crucial difference between "thoughtlessness" and "stupidity," on which Arendt insisted in her later writings. While stupidity is a deficiency in our judgment, thoughtlessness stems from a decision for which we can be held responsible. If those who refuse to think seem to lack responsibility, it is precisely because they have made themselves irresponsible.

The focus on responsibility illuminates both the continuities in Arendt's conception of totalitarian evil and her shift of terminology for describing it. The shift expresses a paradox in the nature of totalitarian evil, which is perpetrated by individuals who willfully refuse to assume

---

[90] Arendt, "Some Questions of Moral Philosophy," 111.
[91] Ibid.
[92] Benjamin A. Schupmann, "Thoughtlessness and resentment: Determinism and moral responsibility in the case of Adolf Eichmann," *Philosophy and Social Criticism* 40(2), 139. According to Sophie Cloutier, "by renouncing his personal will, Eichmann believed himself to be above all responsibility" ("La question du mal chez Hannah Arendt : rupture ou continuité?," PhaenEx 3[1], 103; my translation).

responsibility for their actions. This willful refusal of responsibility can be interpreted as a "radical" motivation not to be responsible, thus rendering one's actions unpunishable and unforgivable. It can also be interpreted as a banality, as an utter lack of motivations. The two interpretations are in tension with one another, as Dana Villa claims, but there is no "contradiction" between the two "at the level of philosophical reflection on the nature of evil," as she maintains.[93] The tension is expressed by the paradox of an intention that destroys intentionality altogether, as if one's own actions had no self behind them. If we focus on the destructive intention, evil appears to be radical, that is, as stemming from an intention that is beyond all comprehensible intentions. If, by contrast, we focus on the lack of intentions, it appears to be banal, which is how Eichmann's character manifested itself, in Arendt's view. However, as long as evil, either radical or banal, presupposes responsibility, the tension contains no contradiction. Therefore, as Richard Bernstein points out, the terminological shift should be understood as a shift of focus rather than as a shift of concepts.[94]

Responsibility between Thinking and Action

Both in *The Human Condition* and in her later writings on thinking, judgment and responsibility, Arendt is concerned with the relationship between actions and their underlying self. Behind this problem lies another, longstanding concern with the problem of responsibility for action, and more specifically, with the modern tendency to escape from it—a tendency which totalitarianism carried to its ultimate conclusion. Despite this enduring thread in Arendt's

[93] Dana R. Villa, *Politics, Philosophy, Terror: Essays on the Thought of Hannah Arendt* (Princeton, NJ: Princeton UP, 1999), 56.
[94] Richard J. Bernstein, "Did Hannah Arendt Change Her Mind? From Radical Evil to the Banality of Evil," in *Hannah Arendt: Twenty Years Later*, ed. Larry May and Jerome Kohn (Cambridge: Cambridge UP, 1999), 142.

thought, there seems to be a gap between her understanding of responsibility in *The Human Condition* and in her later writings. In *The Human Condition*, there is an actor who can be forgiven and punished for his deeds as an effect of beginning and disclosing in the context of plurality. In her later writings, there is a person that can be forgiven and punished as an effect of the activity of thinking, that is, of the inner dialogue between the self and itself. This contrast is perhaps not surprising given Arendt's insistence on separating political questions, which are concerned with the relationship between the self and others, and moral questions, which are concerned with the relationship between the self and itself. Yet Arendt is not completely consistent on this point and, contrary to Diprose's claim that "Arendt always insists that thinking is individual, is distinct from and runs counter to the collective speech and action that characterizes the political," certain passages in her later writings allow for the idea that thinking and acting, while distinct, are interconnected.[95] Pursuing this line of thinking, I believe, allows us to better understand the relationship between moral conscience and political action in Arendt's thought.

In her later lectures on responsibility, Arendt considers thinking, as well as the moral propositions that stem from it, as "politically a borderline phenomenon."[96] The reason is that a moral truth, such as "it is better to suffer wrong than to do wrong," "has nothing whatsoever to do with action." Moreover, Arendt claims that "politically speaking—that is, from the viewpoint of the community or of the world we live in—it is irresponsible; its standard is the self and not the world, neither its improvement nor change."[97] Political responsibility implies a concern with

[95] Diprose, "Arendt and Nietzsche on responsibility and futurity," 625.
[96] Arendt, "Some Questions of Moral Philosophy," 105.
[97] Ibid., 79. See also *The Life of the Mind*, 192.

the world, and is therefore inherently collective because the world is never exclusively of our own making. Moral or personal responsibility, or what she designates as "guilt," implies concern with the self, and is therefore strictly individual.[98] Given that moral considerations do not tell us how to engage with the world, they remain entirely negative with regard to action, in the sense that they only tell us not to act under certain circumstances. Therefore, it is only in situations where political and personal responsibility enter in conflict with one another that the latter becomes politically relevant. When engagement in politics demands that we become someone with whom we could not live, personal responsibility demands that we withdraw from the public world, thus becoming politically irresponsible. Under conditions of totalitarianism, personal responsibility does not demand that people do anything, but rather that they cease all participation in the public world. It is only in this negative sense that moral guilt and political responsibility relate to one another.

While this contrast between moral and political responsibility seems to oppose thinking and action, other passages in Arendt's work suggest a more nuanced relationship between the two. In *Origins*, she claims that the dialogue of the two-in-one of thinking "does not lose contact with the world of my fellow-men because they are represented in the self with whom I lead the dialogue of thought," and that "this two-in-one needs the others in order to become one again: one unchangeable individual whose identity can never be mistaken for that of any other."[99] Arendt develops this idea extensively in *The Life of the Mind*, where she addresses the relationship between thinking and acting through an analysis of Socrates. The importance of Socrates stems from the fact that, unlike later "professional philosophers" entirely devoted to the

[98] Arendt, "Collective Responsibility," in *Responsibility and Judgment*, 150.
[99] Arendt, *The Origins of Totalitarianism*, 476.

activity of thinking in solitude, he "unified two apparently contradictory passions, for thinking and acting," because he was "equally at home in both spheres and able to move from one sphere to the other with the greatest apparent ease."[100] The reason why Socrates could dwell both in solitude and in public is that, as opposed to "the inclination to find solutions for riddles and then demonstrate them to others," he "felt the urge to check with his fellow-men to learn whether his perplexities were shared by them."[101] Socrates's thinking, according to Arendt, is purely critical, in the sense that it produces no belief or value as an effect of its examinations, and that it seems to say: "if the wind of thinking… has shaken you from your sleep and made you fully awake and alive, then you will see that you have nothing in your grasp but perplexities, and the best we can do with them is share them with each other."[102] Thinking does not produce any teaching that orients our actions, but rather dissolves all established doctrines. It is therefore politically dangerous, because it potentially leads to nihilism.[103] But, Arendt claims, "non-thinking, which seems so recommendable a state for political and moral affairs, also has its perils," because "by shielding people from the dangers of examination, it teaches them to hold fast to whatever the prescribed rules of conduct may be at a given time in a given society."[104] Thinking, then, allows us to examine our rules of conduct, and although this examination requires that we withdraw from the public, it also requires that we communicate our perplexities to others.

Socrates's critical kind of thinking is linked to what Arendt calls "representative thinking." In the *Lectures on Kant's Political Philosophy*, she claims that "critical thinking… exposes itself

[100] Arendt, *The Life of the Mind*, 167.
[101] Ibid., 172.
[102] Ibid., 175.
[103] Ibid., 176.
[104] Ibid., 177.

to 'the test of free and open examination,' and this means that the more people participate in it, the better."[105] Critical thinking is not merely a preparation for a more valid truth or doctrine than the predominant one, but rather the open-ended examination of one's own beliefs. As a consequence, it needs the involvement of others to whom one can communicate one's beliefs, so that they are open to examination. Even if thinking proceeds only in solitude, "unless you can somehow communicate and expose to the test of others… whatever you may have found out when you were alone, this faculty exerted in solitude will disappear."[106] Even in solitude the thinker needs to take into account others, given that "even if I shun all company or am completely isolated while forming an opinion, I am not simply together only with myself in the solitude of philosophical thought; I remain in this world of universal interdependence, where I can make myself the representative of everybody else."[107] Representative thinking takes place in solitude, where I am in the company of myself. But this self with whom I speak is a representative of everybody else, and thus of a potential public. Even in solitude I am not completely withdrawn from the world, but only distanced from it so that it becomes possible to examine my thoughts and beliefs. Without taking others into account, there would be no inspection involved, and my opinion would rely on nothing but my own individual interest. In order to make my opinion communicable to others, it is necessary to open it to their examination.

These remarks on representative thinking allow us to understand the relationship between the inner dialogue of conscience and the world of appearances. Arendt describes this inner dialogue as a process of "striking roots," in the sense that by thinking about who I am as an

[105] Hannah Arendt, *Lectures on Kant's Political Philosophy*, ed. Ronald Beiner (Chicago: The University of Chicago Press, 1992), 39.
[106] Arendt, *The Life of the Mind*, 40.
[107] Arendt, "Truth and Politics," in *Between Past and Future*, 237.

effect of my actions, I remember them, and thus become a durable person that can respond for them. But is it possible that, in order to think, I must be willing to be responsible, that is, to be accountable for my thoughts and actions to begin with? This is what Arendt suggests in the *Lectures*, where she claims that the examination of opinions characteristic of critical thinking "presupposes that everyone is willing and able to render an account of what he thinks and says."[108] In this sense, philosophy and politics share a common root: just like Plato departed from earlier "wise men" in that he gave an account of his thoughts, "to render accounts is what Athenian citizens asked of their politicians, not only in money matters but in matters of politics. They could be held responsible."[109] Responsibility, then, is at the root of all thinking, either moral or political, for without holding oneself accountable, one simply does not respond for one's own beliefs, and is therefore ready to believe anything. This is precisely one of the lessons of Eichmann's trial: because Eichmann did not respond for the ideas and the orders that he served, he did not care about what they looked like from other standpoints, or who he was as an effect of executing them. As a consequence, he was willing to hold fast to whatever system of values was given to him. As soon as someone assumes responsibility for his beliefs, he must think about them, subjecting them to inspection by others. These beliefs will therefore have a "root," not in the sense of a fixed value that is beyond examination and change, but of a stable self who can communicate to others the process by which he has arrived to them. This self can certainly change his mind, but even this change will be the effect of a process of which he can provide an account and communicate to others, as opposed to a simple adaptation to new circumstances.

---

[108] Arendt, *Lectures on Kant's Political Philosophy*, 41.
[109] Ibid.

We can see how responsibility links representative thinking to action, challenging the idea that the two activities are simply opposed to one another. To think and to act are certainly different activities and, according to Arendt, we must always stop doing whatever we are doing in order to think. But thinking is important for acting *responsibly*, that is, in such a way that we can be held accountable for what we have done. By stepping outside of the realm of appearances in order to be in the company of ourselves, we examine our beliefs and values and are therefore capable of responding for them. Although this examination is done in solitude, it presupposes the presence of others who may see things from a different viewpoint. Without this presence, representative thinking would be meaningless, for the simple reason that there would be no viewpoint other than ours. When we think in solitude, we are aware that there are other viewpoints, and that these viewpoints will play a role in determining who we are as an effect of our actions, that is, of how we appear to them. It is therefore in the back-and-forth movement between thinking and acting, solitude and appearing, that we become responsible persons capable of responding for what we believe and for what we do on the basis of our beliefs. It is this back and forth movement that puts the two activities in relation to one another—not in the sense that they are combined in a single activity, but in the sense that they inform one another, bringing to each other an element that they would lack by themselves. Without thinking, we would not care about how we appear to others, because we would simply be immersed in our uncritical beliefs. Without appearing to others through action, we would never become aware of the existence of different viewpoints that play a role in determining who we are. It is because we appear to others through action that we experience the importance of different perspectives, and it is because we can withdraw from appearances that we are able to take these perspectives into

account at the moment we appear. Without both these distinct yet interrelated elements, we would not be responsible either for our actions or for our beliefs.

Conclusion

Despite Arendt's shifts on the nature of totalitarian evil, there is a persistent concern in her work with the link between this kind of evil and the destruction of responsibility. The reason why totalitarian crimes are "unpunishable" and "unforgiveable" is that they lack a stable personality behind them, as if the deeds were committed by no one. This observation led Arendt to explore the ground of responsibility. In *The Human Condition*, she locates this ground in the actor's capacity to begin, which is intertwined with its disclosure to others. Because beginning entails an insertion into a web of relations, we are never in control of either the effect or the meaning our action. Yet by disclosing ourselves to others, we acquire an enduring identity that can respond for its deeds, and which may be forgiven or punished by them. Responsibility for action does not presuppose complete control over what we do, but rather the capacity to "risk our disclosure" and establish relations that exceed our control. After *Eichmann in Jerusalem*, Arendt's concern with responsibility shifted to the faculty of thinking which, by virtue of making us remember what we do, constitutes an enduring person. While thinking and action are different activities that take place in different realms (solitude and the public realm), Arendt's analysis of "representative thinking" shows that the two are connected. Both action and representative thinking presuppose the presence of others, and it is by taking these others into account, either in our disclosure through action or in the examination of our beliefs in thinking, that we our capable of responding for our thoughts and deeds. In both cases, responsibility presupposes that we enter into relationship with others, thus exposing ourselves to processes and viewpoints that we cannot

control. To reject this relationship by holding fast universal processes or unquestionable systems of values entails the destruction of responsibility, and thus the danger of an evil that, whether radical or rootless, is beyond forgiveness and punishment.

Chapter 2: Kant on Evil, Deception, and the Ground of Responsibility

The problem of evil represents a persistent concern in the development of Kant's thought. Although it becomes a central topic only in his later essay *Religion within the Boundaries of Mere Reason*, remarks on evil appear throughout most of his writings on practical philosophy. Recently, social and political theorists have become interested in Kant's famous notion of "radical evil," mostly due to Hannah Arendt's references to it in connection to totalitarianism and her own notion of "the banality of evil," leading to a number of studies on the relationship between the two.[1] While these studies have illuminated the connections and contrasts between Arendt's and Kant's perspectives, some of the deeper complexities of Kant's notion of radical evil in the context of his practical philosophy have received scarce attention. In particular, one of the main issues informing the thesis of radical evil, namely, the ground of responsibility for action, remains to be systematically explored—both in the context of the debates around the connection between Arendt and Kant, and within Kant scholarship more broadly. Although most studies of Kant's practical philosophy touch upon the question of responsibility one way or another, the usual perspective assumes that responsibility is contained in the notion of freedom, rather than a problem in its own right.[2] Problematizing this perspective, I argue in this chapter that identifying the ground of responsibility for action constitutes a persisting problem in Kant's

[1] Richard J. Bernstein, "Did Hannah Arendt Change Her Mind? From Radical Evil to the Banality of Evil," in *Hannah Arendt: Twenty Years Later*, eds. Larry May and Jerome Kohn (Cambridge: MIT Press, 1999); Susan Neiman, *Evil in Modern Thought: An Alternative History of Philosophy* (Princeton, NJ: Princeton UP, 2015); Adi Ophir, "Between Eichmann and Kant: Thinking on Evil after Arendt," *History and Memory* 8(2) (1996); Simona Forti, *New Daemons: Rethinking Evil and Power Today* (Stanford, CA: Stanford UP, 2015).

[2] In a volume listing bibliography on Kant in several languages between 1945 and 1990, essays or books with the term "responsibility" in the title appear twice in German, five times in English, once in French, and four times in Italian. Titles with the word "freedom," by contrast, appear nearly twenty times or more in each of these languages, and repeatedly in many others as well (Margit Ruffing [ed.], *Kant-Bibliographie 1945-1990* [Frankfurt am Main: Vittorio Klostermann, 1999], 918-920; 969).

work, and that the thesis of radical evil emerges in part as an attempt to come to terms with it. In addressing this issue, Kant reformulates important elements of his practical philosophy, moving beyond his earlier identification of freedom with autonomy. For Kant in *Religion*, evil does not stem from a *failure* to assume responsibility by becoming autonomous, but rather from an *active concealment* of responsibility.

My inquiry into the problem of responsibility for action in connection to evil in Kant's thought has three aims. First, as I shall argue, the problem of locating the ground of responsibility constitutes the main theoretical issue underlying both "radical evil" and "the banality of evil." Therefore, addressing this issue provides an opportunity to better understand the connection between the two notions. Although Arendt claimed that "the banality of evil" represents a contrast to, rather than a continuation of, "radical evil," I believe that there are important points of connection between them. Second, a focus on the problem of responsibility clarifies some of the apparent paradoxes of Kant's thesis of radical evil and its role in the development of his understanding of action. A number of political theorists have criticized Kant's practical philosophy on grounds that it disavows the contingency of politics by subordinating action to a transcendental moral law.[3] Recently, Shalini Satkunanandan has challenged this view and argued that the moral law is not an epistemological formula that we can use to know what a good action is, but rather an ontological experience of our capacity to think and transform ourselves, thus awakening a sense of responsibility that transcends moralistic

[3] William E. Connolly, *Why I Am Not a Secularist* (Minneapolis, MN: Minnesota UP, 1999), 164-177; Bonnie Honig, *Political Theory and the Displacement of Politics* (Ithaca, NY: Cornell UP, 1993), Chapter 2; Raymond Geuss, *Philosophy and Real Politics* (Princeton, NJ: Princeton UP, 2008), 7-9.

formulas.[4] Following Satkunanandan's perspective, I argue that a focus on the problem of evil and responsibility in Kant shows a concern with the ritualization of moral action. Far from grounding responsibility upon a fixed formula, Kant shows that it stems from a choice that is constitutively "inscrutable," and thus beyond any empirical manifestation that can be potentially regulated and ritualized. Evil consists precisely in the disavowal of choice by concealing it behind ritualized moral practices. Third, the place of *Religion* in the context of Kant's practical philosophy has been the subject of debate in recent years. While J. B. Schneewind and Henry Allison read the text as complementing and completing Kant's conception of moral action as based on autonomy, Susan Shell, Peter Fenves, and Ian Hunter have emphasized the ways in which it departs from such conception.[5] My reading supports this second stance by showing that Kant's understanding of responsibility and evil in *Religion* represents a departure from the one that predominates in the central texts of his practical philosophy.

Arendt, Eichmann, Kant

Kant appears at a central moment in Arendt's analysis of Eichmann's trial in *Eichmann in Jerusalem*. Arendt recounts that during the police interrogation prior to the trial, Eichmann claimed that he had lived his whole life according to Kant's moral precepts and to his definition of duty. When asked during the trial to clarify what he had meant, he answered with "an approximately correct definition of the categorical imperative," according to Arendt: "I meant by

[4] Shalini Satkunanandan, "The Extraordinary Categorical Imperative," *Political Theory* 39(2).
[5] Henry E. Allison, *Kant's Theory of Freedom* (New York: Cambridge UP, 1990); J. B. Schneewind, *The Invention of Autonomy: A History of Modern Moral Philosophy* (New York: Cambridge UP, 1998); Susan Meld Shell, *Kant and the Limits of Autonomy* (Cambridge: Harvard UP, 2009); Peter Fenves, *Late Kant: Towards Another Law of the Earth* (New York: Routledge, 2003); Ian Hunter, *Rival Enlightenments: Civil and Metaphysical Philosophy in Early Modern Germany* (Cambridge: Cambridge UP, 2003).

my remark about Kant that the principle of my will must always be such that it can become the principle of general laws."[6] Eichmann, who claimed that he had read the *Critique of Practical Reason*, added that he knew that "from the moment he was charged with carrying out the Final Solution he had ceased to live according to Kantian principles," but "he had consoled himself with the thought that he no longer 'was master of his own deeds,' that he was 'unable to change anything.'"[7] Arendt believes, however, that Eichmann did not dismiss Kant's moral philosophy altogether, but rather distorted it so that it would fit "the categorical imperative in the Third Reich": "act as if the principle of your actions were the same as that of the legislator or of the law of the land."[8] This constitutes what Eichmann himself called "the version of Kant 'for the household use of the little man.'" The one aspect of this "household use" that preserves the spirit of Kant, Arendt claims, is the demand that "a man do more than obey the law, that he go beyond the mere call of obedience and identify his own will with the principle behind the law."[9] For Eichmann, this principle was "the will of the Führer." According to Arendt, it is this complete identification with the principle behind the law of the land that explains Eichmann's thoroughness in performing his duty: "a law was a law, there could be no exceptions." This "no exception" mentality was so ingrained in Eichmann's understanding of duty that he felt ashamed about the two instances in which he had helped Jewish acquaintances to avoid the camps.[10] From Eichmann's viewpoint, this represented an exception to his compliance with duty driven by personal motives.

---

[6] Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York: Penguin, 2006), 136.
[7] Ibid.
[8] Ibid.
[9] Ibid.
[10] Ibid., 137.

Eichmann's invocation of Kant served the idea that he had simply been a replaceable cog in the totalitarian regime, and thus that he bore no responsibility for his actions beyond compliance with duty. The reasoning would seem to confirm Arendt's longstanding distrust of the philosophy of the will, and especially of its modern introduction into the public realm by Rousseau and the French Revolution. In "The Social Question," Arendt establishes a link between the development of terror and Rousseau's "general will," which she describes as "the articulation of a general interest, the interest of the people or the nation as a whole," adding that "because this interest or will is general, its very existence hinges on its being opposed to each interest or will in particular."[11] The contrast between the general and the particular will is at the core of the equation of virtue with "selflessness," which was popularized by the French Revolution, and according to which "the value of a man may be judged by the extent to which he acts against his own interest and against his own will."[12] The connection between this idea and Eichmann's self-justification is not difficult to see: no matter what a person does, the sacrifice of her personal interests and feelings is the ultimate standard of virtue. Arendt identifies a twofold problem in this idea. First, it conceals the fact that people may do evil without any egoistic motivation, and that they may even be driven to evil-doing by their attachment to self-sacrifice, as seen in the different outbursts of terror in modernity. Second, insofar as the source of duty is independent of all personal interests and feelings, it would seem like the person bears no responsibility for her actions. From Eichmann's perspective, since he had only complied with the law, only the source of the law could be held accountable for what he had done. This disavowal of one's own choice and motivations as a source of action is the reason why it was so difficult

[11] Hannah Arendt, *On Revolution* (New York: Penguin, 1965), 78.
[12] Ibid., 79.

during the trial to adjudicate responsibility to Eichmann—is not thorough compliance with the law the surest sign of virtue?

Influenced by her encounter with Eichmann, Arendt became interested in the notion of the will in her later writings, and in Kant's thoughts on it in particular. Most of her remarks criticize Kant on grounds that his conception of the will misunderstands the essence of freedom. The source of her concern is Kant's subordination of the will to practical reason, which in an early note in her *Journals of Thought* she attributes to his indebtedness to Rousseau: "practical reason as the voice of everyone (in Rousseau) or the voice of humanity in the individual (in Kant) prescribes what is to be willed."[13] In her later writings, Arendt explains that the obedience of the will to practical reason undermines its freedom, for the latter "borrows its obligatory power from the compulsion exerted on the mind by self-evident truth or logical reasoning."[14] Thus if reason commands the will, Arendt claims in another essay, then it "would no longer be free but would stand under the dictate of reason."[15] Kant sought to avoid this conclusion by defining freedom as obedience of a law of which one is the legislator (that is, as autonomy), and locating the source of such law in practical reason. But Arendt finds this solution unsuccessful, for insofar as freedom is subordinated to a general law, then it is mistaken with obedience to a fixed command, which eliminates spontaneity in the sense of beginning something new. Moreover, Eichmann showed that the demand of obedience is more firm than the authority of reason: in its "household use," Kant's notion of the will tells us that we must obey the law, whatever the law is, and that the authority of its source is unquestionable and above personal feelings and opinions. Therefore,

---

[13] Hannah Arendt, *Denktagebuch*, Vol. 1, eds. Ursula Ludz and Ingeborg Nordmann (Munich: Piper, 2002), 182; my translation.
[14] Hannah Arendt, *The Life of the Mind* (Orlando, FL: Harcourt, 1978), 63.
[15] Hannah Arendt, *Responsibility and Judgment*, ed. Jerome Kohn (New York: Schocken Books, 2002), 71.

when Eichmann claimed that not he but the Führer was responsible for what he had done, he preserved one essential aspect of Kant's subordination of the will to practical reason: the authority of the law is prior to and independent of the person's individual perspective.

Arendt criticized Kant's conception of evil, or at least its implications for our understanding of totalitarianism, in similar terms. In her view, the crimes of totalitarianism cannot be explained in terms of the traditional motivations linked to evil-doing, such as self-interest, greed, or resentment.[16] These motivations are present to some degree in all human beings, which renders the deeds that stem from them comprehensible. Totalitarian crimes, by contrast, obey no recognizable motivation, which is why Arendt defined them in *The Origins of Totalitarianism* as an "absolute" evil. Kant, Arendt claims, must have suspected of this form of evil when he coined the notion of "radical evil," but "he immediately rationalized it in the concept of a 'perverted ill will' that could be explained by comprehensible motives."[17] This distance from Kant remains after Arendt's shift to a conception of totalitarian crimes as "banal" or "rootless," in the sense of lacking motivations at all. In "Some Question of Moral Philosophy," she claims that "all transgressions are explained by Kant as exceptions that a man is tempted to make from a law which he otherwise recognizes as valid."[18] In Arendt's view, Kant remained within the traditional philosophical denial of the positivity of evil, which locates its source in limitations that prevent us from doing good. Totalitarian evil, either "absolute" or "rootless," showed that this is not the case: people may do evil for motives that do not stem from any limitation.

[16] Hannah Arendt, *The Origins of Totalitarianism* (New York: Harcourt Brace & Company, 1973), 459.
[17] Ibid.
[18] Arendt, *Responsibility and Judgment*, 62.

Although Arendt found the philosophy of the will mostly unable to cope with Eichmann's seeming lack of responsibility, she did not dismiss the possibility that it may provide certain answers. At the beginning of her more extensive engagement with this philosophy in *The Life of the Mind*, she claims that aside from its paradoxes and self-contradictions, she shall "follow a parallel development in the history of the Will," in which it "creates the *person* that can be blamed or praised and anyhow held responsible not merely for its actions but its whole 'Being,' its *character*."[19] Arendt's interest in this development is clearly influenced by her idea that Eichmann lacked personality, namely, an enduring self that can be held accountable for his actions. Given that Eichmann did not think about who he would be as an effect of his actions, he was unable to remember and to assume responsibility for them.[20] Although most of Arendt's reflections on moral personality focus on the faculty of thinking, she did not dismiss the importance of choice as the faculty that constitutes our moral character. Most of her remarks on this topic in her later writings are presented in connection to Duns Scotus. In *The Life of the Mind*, she claims that willing fashions the self "into an 'enduring I' that directs all particular acts of volition," and that "it creates the self's *character* and therefore was sometimes understood as the *principium individuationis*, the source of the person's specific identity."[21] In her essay "Some Questions of Moral Philosophy," she claims that "the will is the arbiter between reason and desire, and as such the will alone is free," adding that "while reason reveals what is common to all men, and desire what is common to all living organisms, only the will is entirely my own."[22] Arendt notes in her unpublished manuscript "Basic Moral Propositions" that the will is free "as

[19] Arendt, *The Life of the Mind*, 214-215.
[20] Arendt, *Responsibility and Judgment*, 94.
[21] Arendt, *The Life of the Mind*, 195.
[22] Arendt, *Responsibility and Judgment*, 114.

the faculty of choosing," and that "by willing I decide. And this is the faculty of freedom."[23] The idea that the will understood as choice is the faculty of freedom contradicts most of Arendt's remarks throughout her work, and brings her strikingly close to Kant's analysis of choice in his only text specifically concerned with it, namely, *Religion within the Boundaries of Mere Reason*. This is hardly surprising given that Kant's analysis in this text responds to the same concern as Arendt's, namely, how to hold people accountable for evil-doing.

Besides the theoretical connection between Kant's "radical evil" and Arendt's "the banality of evil," which I will analyze further down, there is a historical link between the two notions that is worth pointing out. It has been noted that the first suggestion that totalitarian evil, contrary to Arendt's original characterization of it as "absolute," must be confronted in its "total banality," appears in a letter by Karl Jaspers from 1946.[24] What no scholar to my knowledge has pointed out is that in 1935, two years after Hitler's coming to power, Jaspers published an essay on Kant's "radical evil" and its importance for a critical understanding of modernity. Although Kant is not explicitly mentioned in Jaspers's letter, it is evident that its main point, which is that Nazi crimes are not "demonic" but "banal," is strongly influenced by Kant's thoughts on evil. Arendt acknowledged the pertinence of Jaspers's perspective in her response letter and, years later, the term that he suggested would appear in the subtitle of her book on Eichmann.[25] Thus indirectly, and most likely inadvertently, Arendt's thoughts on totalitarian evil were influenced by Kant, via Jaspers. This intellectual trajectory connecting Kant, Jaspers, and Arendt, I shall argue, is not

[23] Hannah Arendt, "Basic Moral Propositions," unpublished manuscript from 1966, in The Hannah Arendt Papers, Courses—The University of Chicago, Series: Subject File, 1949-1975, n.d., Manuscript Division, Library of Congress, Washington D.C.
[24] Hannah Arendt and Karl Jaspers, *Correspondence: 1926-1969*, eds. Lotte Kohler and Hans Saner (New York: Harcourt Brace Javanovich, 1992), 62.
[25] Ibid., 69.

independent from a political concern running through the three thinkers, namely, the emergence of a new form of evil, linked to the destruction of responsibility, taking place with the development of the modern world. In order to understand the challenge of conceptualizing responsibility and evil in the context of modernity, let us turn to Kant.

Freedom and the Ground of Imputability

The problem of responsibility for action in Kant's thought is intertwined with the concept of transcendental freedom. Kant introduces this concept in the third antinomy of reason, which is concerned with the question of whether all events have a prior cause in time, or an event without a prior cause is possible. The solution consists in showing that the two possibilities are not mutually exclusive. As appearances, all events have a prior cause in time. But appearances do not exhaust the field of all things, which includes their source, the "things in themselves." Therefore, it is at least possible that there are things in the world capable of starting a series of events that is not determined by a prior cause in time. This capacity is precisely transcendental freedom which, Kant claims in a remark, "constitutes only that of the absolute spontaneity of an action [*absoluten Spontaneität der Handlung*], as the real ground of its imputability [*eigentlichen Grund der Imputabilität derselben*]."[26] Kant defines transcendental freedom as "the faculty of beginning a state **from itself** [*das Vermögen, einen Zustand* ***von selbst*** *anzufangen*], the causality of which does not in turn stand under another cause determining it in time in accordance with the law of nature [*Naturgesetze*]."[27] The existence of this capacity cannot be proved, because all experience is regulated by the law of natural causality. However, it cannot be disproved either,

[26] Immanuel Kant, *Critique of Pure Reason*, eds. Paul Guyer and Allen W. Wood (New York: Cambridge UP, 1998), 486; A448/B476.
[27] Ibid., 533; A533/B561.

and so it is legitimate to presuppose its existence for practical purposes. If there were no transcendental freedom, we would not be responsible for our actions, given that everything we do would follow necessarily from a cause that is external to us.

Transcendental freedom, or the capacity to begin, grounds practical freedom, or the capacity to act on the basis of imperatives. Kant defines practical freedom as "the independence of the power of choice [*Unabhängigkeit der Willkür*] from necessitation [*Nötigung*] by impulses of sensibility [*Antriebe der Sinnlichkeit*]."[28] While sensible impulses subject us to the law of natural causality, practical freedom "presupposes that although something has not happened [*etwas nicht geschehen ist*], it nevertheless **ought** to have happened [*habe geschehen **sollen***]."[29] The source of the idea that something ought to happen, or what Kant calls "imperatives," is reason. This is the only faculty capable of "complete spontaneity" [*völler Spontaneität*], because it is the only faculty that is altogether independent of affection by sensibility, and so "it makes its own order according to ideas [*nach Ideen*], to which it fits the empirical conditions and according to which it even declares actions [*Handlungen*] to be necessary that yet **have not occurred** [*nicht geschehen sind*] and perhaps will not occur [*nicht geschehen werden*]."[30] Reason, according to Kant, can produce imperatives for action based on ideas, independently of natural necessitation. These imperatives are the basis of moral judgments, for "however many natural grounds [*Naturgründe*] or sensible stimuli [*sinnliche Anreize*] there may be that impel me to **will** [*zum **Wollen** antreiben*], they cannot produce the **ought** [*nicht das **Sollen** hervorbringen*] but only a willing [*Wollen*]…, over against which the ought [*Sollen*] that reason pronounces sets a measure

[28] Ibid., 533; A534/B562.
[29] Ibid., 534; A534/B562.
[30] Ibid., 541; A548/B576.

and a goal [*Maß und Ziel*], indeed, a prohibition and authorization [*Verbot und Ansehen*]."[31] Although transcendental freedom is the ground of practical freedom, both are indispensable for imputability, insofar as it requires a standard that distinguishes between the prohibited and the approvable. Transcendental freedom shows that it is possible that we act on the basis of this standard, which for Kant amounts to the possibility that reason have causality, independently of natural impulses.

This conception of imputability opens up an important problem in Kant's practical philosophy, which is closely linked to the question of evil. Kant claims that spontaneity renders the power of choice independent of necessitation by sensible impulses, but he allows that it is nevertheless affected by them. In the only passage specifically devoted to *Willkür* in the *Critique of Pure Reason*, Kant explains that

> a power of choice [*Willkür*] is **sensible** [***sinnlich***] insofar as it is **pathologically affected** [***pathologisch affiziert***] (through moving-causes of sensibility); it is called an **animal** [***tierisch***] power of choice (*arbitrium brutum*) if it can be **pathologically necessitated** [***pathologisch necessitiert***]**.** The human power of choice [*menschliche Willkür*] is indeed an *arbitrium sensitivum*, yet not *brutum* but *liberum*, because sensibility does not render its action [*Handlung*] necessary, but in the human being there is a faculty of determining oneself from oneself [*sich… von selbst zu bestimmen*], independently of necessitation by sensible impulses [*sinnliche Antriebe*].[32]

Kant defines the human *Willkür* or *arbitrium* as both "*sensitivum*" and "*liberum*," which lays out the contrast between sensibility and freedom. The contrast preserves the distinction between choice and freedom that was present in both Wolff and Baumgarten. Following Leibniz, Wolff conceives of choice as the capacity to act on the basis of an internal ground, and freedom as the

[31] Ibid.
[32] Ibid., 533; A534/B562.

capacity to understand this ground by means of reason.[33] Baumgarten introduces an innovation that Kant takes up, namely, the contrast between a sensible and a non-sensible choice. While only non-sensible choice is free, both non-sensible and sensible choice are "mixed" in human choice.[34] For Kant, while sensibility subjects choice to the law of natural causality, freedom consists in a sort of self-determination (*sich von selbst zu bestimmen*; "to determine oneself from oneself") that is independent from it. Given that reason is the only faculty endowed with complete spontaneity, namely, with the capacity to act with complete independence of natural necessitation, freedom and reason would seem to be one and the same thing. If this is the case, however, it remains unclear what sort of agency is left for *Willkür*—why would a human being choose reason or sensibility as the source of action? And if sensibility subordinates choice to natural necessitation, in what sense is the human being imputable for a choice that is pathologically affected? Given that complete spontaneity pertains only to reason, and that it is the ground of imputability, it would seem that one is only imputable for acting on the basis of reason. If this were the case, however, we would not be imputable for actions that are conditioned by sensible stimuli and transgress reason's imperatives—that is, we would not be responsible for evil actions.

The connection between responsibility and evil represents one of the most persistent problems in Kant's practical philosophy, and has been the subject of critique ever since its

---

[33] Christian Wolff, *Rational Thoughts on God, the World and the Soul of Human Beings, also All Things in General*, in *Kant's Critique of Pure Reason: Background Source Materials*, ed. Eric Watkins (New York: Cambridge UP, 2009), 35.

[34] Alexander Baumgarten, *Metaphysics*, in *Kant's Critique of Pure Reason: Background Source Materials*, 122. An extensive analysis of the multiple influences contributing to Kant's notion of *Willkür* is developed by Katsutoshi Kawamura in *Spontaneität und Willkür: Der Freiheitsbegriff in Kants Antinomienlehre und seine historischen Wurzeln* (Stuttgart: Frommann Holzboog, 1996), Chapters 1 and 2.

contemporaries.[35] Part of the reason of this persistence is that, in his central writings on moral philosophy, Kant does not systematically analyze *Willkür*. [36] In these writings, *arbitrium liberum* becomes the will, which Kant equates with both reason and freedom. This is the basis of Kant's notion of autonomy, according to which freedom consists in complying with a law of which one is the legislator. In the *Groundowrk of the Methaphysics of Morals*, Kant claims that reason is "pure self-activity" [*reine Selbsttätigkeit*], and that it shows "a spontaneity so pure [*so reine Spontaneität*] that it thereby goes far beyond anything that sensibility can ever afford [*liefern*] to it."[37] Given that reason is the only faculty completely independent of natural causality, it is the only faculty capable of pure spontaneity, and thus of legislating what ought to be done independently of what is. If, as Kant had already suggested in the first *Critique*, freedom and spontaneity are one and the same thing, then reason is the only faculty that is completely free. Sensible impulses, by contrast, are given to us from without or "heteronomously," that is, independently of our freedom. Therefore, the question remains open: in what sense are we imputable for actions that stem from sensibility? Kant claims in the *Groundwork* that "the human being claims for himself a will which lets nothing be put to his account [*nichts auf seine Rechnung kommen läßt*] that belongs [*gehört*] merely [*bloß*] to his desires and inclinations [*Begierden und Neigungen*]," and so, by means of this will, he thinks possible "actions [*Handlungen*] that can be done only by disregarding all desires and sensible incitements [*Hintersetzung aller Begierden und sinnlichen Anreizungen*]," whose causality "lies in him as

[35] The first major critique was developed by Carl Leonhard Reinhold in "Erörterung des Begriffs von der Freiheit des Willens," in *Materialen zu Kants "Kritik der praktischen Vernunft,"* eds. Rudiger Bittner and Konrad Cramer (Frankfurt am Main: Suhrkamp, 1975).
[36] On this point, see Lewis White Beck, "Five Concepts of Freedom in Kant," in *Stephan Körner – Philosophical Analysis and Reconstruction*, ed. Jan T. J. Srzednicki (Dordrecht, the Netherlands: Martinus Nijhoff, 1987), 37.
[37] Immanuel Kant, *Groundwork of the Metaphysics of Morals*, in *Practical Philosophy*, ed. Mary J. Gregor (New York: Cambridge UP, 1999), 99; 4:452.

intelligence [*als Intelligenz*] and in the laws of effects and actions [*Gesetzen der Wirkungen und Handlungen*] in accordance with principles of the intelligible world."[38] For Kant, imputability does not stem from a person's inclinations and impulses, but rather from laws of the intelligible world, which is why

> he does not hold himself accountable for the former [*die erstere nicht verantwortet*; "does not respond for the former"] or ascribe [*zuschreibt*] them to his proper self [*eigentlichen Selbst*], that is, to his will [*seinem Willen*], though he does ascribe to it the indulgence [*Nachsicht*] he would show them [*gegen sie tragen möchte*] if he allowed [*einräumete*] them to influence his maxims to the detriment [*zum Nachteil*] of the rational laws of his will [*Vernunftgesetze des Willens*].[39]

Although only the autonomous self is accountable for his actions, this self is capable of "indulgence" to the inclinations, thus becoming heteronomous. However, if this indulgence is to be "ascribed" to the self, it would seem that it must stem from its freedom. Somewhat paradoxically, the subjection to the inclinations, and thus to natural necessitation, must be an spontaneous, free act, for otherwise there would be no ground from which the person could respond for it.

A similar idea is presented by Kant in the *Critique of Practical Reason* in terms of the will's "resistance" [*Widerstand*] to the inclinations. Given that sensible incentives for action are heteronomous to the will and, consequently, to the proper self of the human being, the will's autonomy consists in resisting them:

> a choice that is pathologically affected [*pathologisch affizierte… Willkür*] (though not thereby determined [*bestimmte*], hence still free) brings with it a wish [*einen Wunsch bei sich führt*] arising from *subjective* causes because of which it can often be opposed [*entgegen sein kann*] to the pure objective determining ground

[38] Ibid., 103; 4:457.
[39] Ibid., 104; 4:458.

> and thus needs a resistance [*Widerstand*] of practical reason which, as moral necessitation [*moralischer Nötigung*], may be called an internal but intellectual constraint [*ein innerer, aber intellektueler, Zwang*].[40]

Insofar as the power of choice is pathologically affected, which means that there are subjective causes (i.e. inclinations) opposing the pure objective determining ground of the will (the moral law), the will (which, as we have seen, is equivalent to practical reason) cannot but resist these pathological incentives. But insofar as the will is freedom itself, conceived as independence from natural causality, freedom seems to consist merely in resisting this causality. This would indicate that deeds transgressing the moral law proceed from a failure to resist, that is, from a failure by the will to "constrain" pathological incentives. Given that these incentives oppose the person's freedom, and given that only such incentives lead to actions that transgress the moral law, evil deeds cannot be freely chosen. It is only insofar as the person fails to act freely by not resisting the influence of the inclinations that evil deeds are possible.

Kant's remarks on the tension between the will and the inclinations in his main writings on practical philosophy suggest a conception of evil as deficiency, but his thought is not entirely consistent on this point. If actions that transgress the imperatives of reason are due to the will's "indulgence" or lack of "resistance," then the source of evil would lie in its incapacity to execute its own commands. Therefore, no choice or decision is involved, and freedom becomes a matter of the strength of reason to resist the assault of the inclinations, as in a battle between two opposing forces. Yet Kant suggests a different possibility in his *Lectures on Ethics*, developed briefly after the *Groundwork*, and where he explicitly addresses the problem of "imputation." Here, Kant establishes a distinction between "weakness," which "consists in its want of

[40] Immanuel Kant, *Critique of Practical Reason*, in *Practical Philosophy*, 165-166; 5:32.

sufficient moral goodness to make the action adequate to the moral law," and "frailty," which consists "also in the prevalence therein of even the strongest principles and motivations to ill-doing."[41] This is the source of what Kant defines as a "positive" or "moral" evil, which "arises from freedom, since otherwise it would not be *moral* evil, and however prone we may also be to this by nature, our evil actions still arise from freedom, on which account they are also debited to us as vices."[42] If vices can be imputed to us, as opposed to being a natural consequence of our weakness, they must stem from freedom rather than weakness. This would mean, however, that freedom lies beyond reason, and therefore beyond autonomy. This is the possibility that Kant will explore in *Religion*.

The Choice of Evil

Scholars in recent years have stressed the importance of *Religion* for a comprehensive understanding of Kant's practical philosophy. While the text is mostly known for its account of evil and its attempt to provide a rational justification of Christianity, it also plays a crucial role in clarifying problematic aspects of Kant's conception of freedom and moral action. Thus, Allen Wood reads *Religion* as continuing Kant's earlier inquiries into the necessary beliefs that are presupposed by moral agency.[43] Allison interprets it as the completion of what he calls the "incorporation thesis," according to which freedom does not stem from autonomy, that is, from acting on the basis of reason, but rather from the choice of the maxim of our actions—either

[41] Immanuel Kant, *Lectures on Ethics*, eds. Peter Heath and J. B. Schneewind (New York: Cambridge UP, 1997), 85.
[42] Ibid., 86.
[43] Allen W. Wood, *Kant's Moral Religion* (Ithaca, NY: Cornell UP, 1970), 153-155.

reason or the inclinations.[44] Schneewind presents a similar perspective in his comprehensive study of the history of moral philosophy. In his reading, Kant's notion of autonomy represents the culmination of a philosophical trend running through Leibniz, Hume, and Rousseau, which conceives morality as self-governance, as opposed to obedience to a divinely-sanctioned, humanly incomprehensible law. Schneewind is careful to point out that Kant's confidence in self-governance is nuanced by his idea of freedom as choice, yet like Allison he reads the latter as a completion, rather than as complication, of the idea of autonomy.[45] From this perspective, autonomy creates the condition for moral action by providing it with a rational incentive, while choice makes us morally imputable for either accepting or disregarding such incentive. The problem with this view, I believe, is that it neglects the tension between autonomy and choice. While autonomy grounds freedom upon a law determining action, the spontaneity of choice seems to lack any orienting principle, making its ground mysterious—or, as Kant claims, "inscrutable." As I shall argue in this section, the thesis that the moral worth of our actions is inherently inscrutable represents an innovation in Kant's understanding of responsibility and evil. My reading thus sides with Susan Shell's and Ian Hunter's interpretations of the text, according to which Kant's confidence in reason and autonomy is much more ambivalent than Wood, Allison, and Schneewind allow.[46]

Religion addresses a central issue left open in Kant's earlier writings on moral philosophy, namely, the ground of imputability for evil deeds. Retaking the link between freedom, spontaneity, and imputability that he had laid out in the first *Critique*, Kant claims that in order

[44] Allison, *Kant's theory of freedom*, 39-40.
[45] Schneewind, *The Invention of Autonomy*, 518.
[46] Hunter, *Rival Enlightenments*, 343-354; Shell, *Kant and the Limits of Autonomy*, 8-14.

for evil deeds to be imputable to the person, they need to stem from freedom, as opposed to being determined by natural causality. The problem is, as we have seen, that our natural impulses or inclinations, which subject us to the law of natural causality, are not free. If only reason is free, in the sense of spontaneous, then it is unclear on what grounds a person would be free to perform deeds that stem from the inclinations, and thus imputable for them. Kant's solution consists in shifting freedom and spontaneity from reason to choice, so that whether a person acts on the basis of reason or on the basis of the inclinations stems from his own volition. The crucial concept mediating between choice and either reason or the inclinations is "maxim": "the freedom of the power of choice [*Freiheit der Willkür*] has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive [*Triebfeder*] *except so far as the human being has incorporated it into his maxim* [Maxime]."[47] In the *Groundwork*, Kant defines a maxim as a "subjective principle of volition" or "a subjective principle of acting," and distinguishes it from an objective principle which establishes how the subject ought to act.[48] While reason and the inclinations provide the two fundamental incentives for action, choice produces a maxim that determines whether one is to comply with one or the other. This means that even when acting with the aim of satisfying our inclinations, we do so on the basis of a free choice, which is why "the ground of evil [*Grund des Bösen*] cannot lie in any object *determining* [bestimmenden *Objekte*] the power of choice [*Willkür*] through inclination [*durch Neigung*], not in any natural impulses [*Naturtriebe*], but only in a rule [*Regel*] that the power of choice itself

[47] Immanuel Kant, *Religion Within the Boundaries of Mere Reason*, in *Religion and Rational Theology*, eds. Allen W. Wood and George di Giovanni (New York: Cambridge UP, 2001), 73; 6:24.
[48] Kant, *Groundwork of the Metaphysics of Morals*, 56; 4:402; 73; 4:422.

produces for the exercise of its freedom [*Gebrauch ihrer Freiheit*], i.e., in a maxim."[49] The ground of evil is not the inclinations, but rather choice, which freely decides whether one is to subdue one's actions to the inclinations or to reason.

The location of spontaneity in choice, as opposed to reason, deprives the exercise of freedom of a ground that regulates it. This is not the case if spontaneity is located in reason, which means that to act freely and to adopt a maxim stemming from reason is one and the same thing. If, however, spontaneity is located in choice, then the question emerges of what regulates its execution. In other words: what determines whether a person chooses either a good or an evil maxim of action? In response to this question, Kant claims:

> That the first subjective ground [*erste subjective Grund*] of the adoption [*Annehmung*] of moral maxims is inscrutable [*unerforschlich*] can be seen provisionally from this: Since the adoption is free, its ground… must not be sought in any incentive of nature [*Triebfeder der Natur*], but always again in a maxim; and, since any such maxim must have its ground as well, yet apart from a maxim no *determining ground* [Bestimmungsgrund] of the free power of choice [*freien Willkür*] ought to, or can, be adduced, we are endlessly [*ins Unendliche*] referred back in the series of subjective determining grounds [*Reihe der subjektiven Bestimmungsgründe*], without ever being able to come to the first ground [*ersten Grund*].[50]

In order to explain the free choice of a maxim, we would need to rely on a prior maxim that determining choice, and thus to infinity. This is why the ultimate maxim, or the "first subjective ground," is, and even *ought to be*, "inscrutable" [*unerforschlich*]. If we were able to identify such ground, the choice of a maxim would cease to be free, which means that we would not be imputable for it. Gordon E. Michalson is therefore mistaken when he reads Kant's thesis on inscrutability as falling into an "argumentative flaw" and a "frustrating conceptual logjam," for

[49] Kant, *Religion Within the Boundaries of Mere Reason* 70-71; 6:21.
[50] Ibid., 71; 6:22.

the point is precisely that freedom and imputability presuppose a limit in our capacity to know and to explain—as both Jean-Luc Nancy and Ian Hunter point out.[51] If we could find a ground behind the adoption of an evil maxim (or a good maxim, for that matter), such maxim would cease to be freely chosen and therefore imputable—strictly speaking, it would cease to be evil at all. Therefore, Kant claims that the choice of a maxim must be represented as always already having taken place, or as "innate," even if, paradoxically, the person remains imputable for it. The first subjective ground is the inscrutable foundation of our moral character or "disposition" [*Gesinnung*], determining all particular exercises of freedom in time.

The moral worth of a maxim is determined by the ordering of the two incentives for action, the inclinations and reason. Moreover, it is this very ordering, rather than the incentives independently of one another, that constitutes the distinction between good and evil. While all actions are performed on the basis of both the inclinations and reason, only one of them can be their "supreme incentive" [*oberste Triebfeder*]. In themselves, the inclinations and reason are examples of what Kant calls the "predisposition to good [*Anlage zum Guten*] in human nature." The inclinations are part of the "predisposition to animality," which is constituted by physical self-love, and is functional to self-preservation, reproduction, and community with others. The inclinations coupled with reason are part of the "predisposition to humanity," which is constituted by physical self-love together with comparison with others, which brings about the inclination to gain worth in their opinion. These two predispositions coupled with susceptibility to respect for the moral law are part of the "predisposition to personality," which makes it

[51] Gordon E. Michalson, *Fallen Freedom: Kant on Radical Evil and Moral Regeneration* (Cambridge: Cambridge UP, 1990), 65; 67. Jean-LucNancy, *L'impératif catégorique* (Paris: Flammarion, 1983), 15; Hunter, *Rival Enlightenments*, 347.

possible for the moral law in itself to be an incentive for action. While all these predispositions demand compliance with the moral law, only the predisposition to personality makes it possible to have a "good character," namely, to make such compliance the supreme incentive for action. All maxims, according to Kant, contain more than one incentive, and so "the difference, whether the human being is good or evil, must not lie in the difference between the incentives [*Unterschied der Triebfedern*] that he incorporates into his maxim… but in their *subordination* [Unterordnung]…: *which of the two he makes the condition of the other*."[52] Evil does not consist in disregarding the moral law altogether, but rather in making compliance with it conditional on self-love, which means that "the human being (even the best) is evil only because he reverses [*umkehrt*] the moral order of his incentives [*sittliche Ordnung der Triebfedern*] in incorporating them into his maxims."[53]

Kant locates the origin of evil in what he calls a "propensity [*Hang*] to evil," which he defines as "the subjective ground of the possibility of an inclination [*Möglichkeit einer Neigung*] (habitual desire, *concupiscentia*), insofar as this possibility is contingent [*zufällig*] for humanity in general [*für die Menschheit überhaupt*]."[54] Humanity is capable of acting on the basis of the inclinations, but only contingently, that is, depending on a free choice. A propensity is the possibility of this choice which, given that it stems from freedom, is subjective. The idea of a propensity seems to clarify nothing, given that it must still be freely chosen—if we were not responsible for our own propensity to evil, the choice of an evil maxim would cease to be free.

[52] Kant, *Religion within the Boundaries of Mere Reason*, 83; 6:36.
[53] Ibid., 83; 6:36.
[54] Ibid., 76; 6:26.

Yet Kant provides an example that illuminates the relationship between a propensity (which is free) and the inclinations (which are not):

> all savages [*rohe Menschen*] have a propensity for intoxicants [*berauschenden Dingen*]; for although many of them have no acquaintance at all with intoxicants, and hence absolutely no desire [*Begierde*] for the things that produce it, let them try these things but once, and there is aroused [*hervorzubringen*] in them an almost inextinguishable desire [*kaum vertilgbare Begierde*] for them.[55]

The example suggests that a propensity is the universal ground for the choice of an evil maxim, even if no one knows that he has chosen it. The savages do not know that they have chosen the inclinations as the supreme incentive of their maxim for action, simply because they have not yet encountered the object that arouses them. From the moment they are capable of acting on the basis of the inclinations, however, they are already evil, even if their character has not yet become manifest.

The propensity to evil, Kant claims, applies to the human being considered as a species, which leads him to support the idea that "the human being is by nature evil." The claim contradicts the ideals of the Enlightenment, and generated strong criticism among Kant's contemporaries. Moreover, Kant's justification of the thesis is strange in the context of his work, basing it on "the cognition we have of the human being through experience [*durch Erfahrung… kennt*]," according to which "he cannot be judged [*beurteilt*] otherwise, in other words, we may presuppose [*voraussetzen*] evil as subjectively necessary [*subjektiv notwendig*] in every human being, even the best."[56] To ground universal assertions on the basis of empirical observation contradicts Kant's own distinction between universality and generality—making the thesis not only counter to the Enlightenment, but also unenlightened. This apparent inconsistency has led

[55] Ibid., 77-78; 6:29.
[56] Ibid., 80; 6:32.

some commentators to provide the theoretical justification for it that Kant neglected to include.[57] This justification, however, is precluded by the thesis itself, for as we have seen, any theoretical account of evil would eliminate it. Thus Kant claims that

> since this propensity must itself be considered morally evil, hence not a natural predisposition [*Naturanlage*] but something that a human being can be held accountable for [*etwas, was dem Menschen zugerechnet werden kann*]…, such maxims must be viewed as accidental [*zufällig*; "contingent"], a circumstance that would not square with the universality of evil [*Allgemeinheit dieses Bösen*] at issue unless their supreme subjective ground [*subjective oberste Grund*] were not in all cases somehow entwined [*verwebt*] with humanity itself and, as it were, rooted [*gewurtzelt*] in it.[58]

The propensity to evil is rooted in human nature in such a way that it is both universal and contingent. The idea of a universal contingency would seem like a contradiction, yet this contradiction is unavoidable when analyzing choice. Certainly there must be a propensity to evil for there to be choice, but this propensity would not be evil at all if it were not freely chosen. Choice is therefore corrupted at its very root, and the possibility of this corruption cannot be explained without eliminating its freedom. Thus the appeal to observation conceals the impossible proof for the universality of evil. However, some of Kant's examples of evil in a civilized state, such as "secret falsity [*geheimer Falschheit*] even in the most intimate friendship," a "universal maxim of prudence [*allgemeinen Maxime der Klugheit*] in social dealings," and even "many other vices yet hidden [*verborgenen*] under the appearance of virtue [*Tugendscheine*]," make it clear that the thesis works as a warning against those who believe that

[57] Allison (*Kant's Theory of Freedom*, 154-156) claims that the thesis that the propensity to evil is universal is to be taken as synthetic a priori. This is justified by the fact that human beings are sensible beings subject to temptation and, consequently, capable "a priori" (although not necessarily) of deviating from the law. Allison's argument neglects the fundamental point that, for Kant, deviation from compliance with the law cannot be caused by sensible incentives. This means that the propensity to evil has to stem exclusively from the power of choice, which is by definition unaffected by sensible incentives.

[58] Kant, *Religion*, 80; 6:32.

they are free of evil by virtue of their empirically good actions.[59] Who could deny the universality of evil, except someone who is certain that he is incapable of it, which would entail a rejection of his own *Willkür*?

Evil in civilization is closely linked to deception, Kant suggests, and this deception stems from its proximity to reason. Kant had warned in the *Groundwork* against the "coolness of a scoundrel" [*kalte Blut eines Bösewichts*; "the cold blood of the petty evil-doer"] which "makes him not only far more dangerous [*weit gefährlicher*] but also immediately more abominable [*verabscheuungswürdiger*] in our eyes than we would have taken him to be without it."[60] While a savage may simply lack reason to prevent him from a moral misuse of his inclinations, the cold blood of the petty evil-doer produces what Kant calls a "propensity to rationalize" [*Hang zu vernünfteln*], that is, to question the purity and strictness of the moral law, and "to make them better suited to our wishes and inclinations, that is, to corrupt [*verderben*] them."[61] In other words, the evil-doer with cold blood does not simply transgress the law, but uses his reason to corrupt it in a way that serves his own wishes. Kant reinforces this idea in *Religion* by showing that an evil disposition does not only stem from a propensity to rationalize our transgressions to the law, but also to *conceal* our disposition behind empirical compliance with it. Thus Kant claims that "the attitude of mind [*Denkungsart*; "kind of thinking"] that construes the absence of vice as already being conformity of the *disposition* [*Angemessenheit der* Gesinnung] to the law of duty (i.e. as *virtue*) is nonetheless itself to be named a radical perversity in the human heart

[59] Ibid., 80; 6:33.
[60] Kant, *Groundwork*, 50; 4:394.
[61] Ibid., 59-60; 4:405.

[*radicale Verkehrtkeit im menschlichen Herzen*]."[62] This kind of thinking consists in "dishonesty" [*Unlauterkeit*] regarding our incentives, and in

> a certain *perfidy* [Tücke] on the part of the human heart… in deceiving itself [*selbst zu betrügen*] as regards its own good or evil disposition and, provided that its actions do not result in evil (which they could well do because of their maxims), in not troubling itself [*sich… nicht zu beunruhigen*] on account of its disposition but rather considering itself justified [*gerechtfertigt zu halten*] before the law.[63]

Given that the maxim of our actions is inscrutable, radical evil conceals itself behind dishonesty, which consist in taking empirical compliance with duty as proof of moral virtue. As a consequence, radical evil "puts out of tune [*verstimmt*] the moral ability to judge [*moralische Urteilskraft*] what to think of a human being, and renders any imputability [*Zurechnung*] entirely uncertain [*ungewiß*], whether internal or external."[64] Because of the natural propensity to deceive oneself and others, all moral judgments are inherently uncertain. Yet the very existence of this propensity, on which the thesis of radical evil depends, allows for the presupposition that all human beings are evil.

Before proceeding to Kant's analysis of the mechanisms of moral deception, let us note one implication of Kant's shift in his understanding of evil for his views on the public role of reason and philosophy. If evil stems from a weakness in reason, as Kant suggests in the *Groundwork* and the second *Critique*, then the victory of moral virtue depends on its strengthening so that it does not yield to the inclinations. This leads to the idea that philosophers play a crucial public role, above that of theologians. The view would represent a continuation of what Ian Hunter identifies as a "philosophical theology" emerging with Leibniz, in which philosophy takes up the

[62] Kant, *Religion*, 84; 6:37.
[63] Ibid., 84; 6:38.
[64] Ibid., 85; 6:38.

role traditionally assigned to theologians, namely, the salvation of the soul.[65] If, by contrast, evil stems from a choice that is prior to reason, then the role of philosophy is much more limited—as long as philosophy is identified with rational thinking. No matter how much a person can rely on her reason, she is perfectly capable of choosing evil by virtue of her freedom of choice. Moreover, reason in the hands of an evil-doer is far more dangerous than its absence. Kant's thesis of radical evil thus represents not only a warning against the limitations of the Enlightenment, but also against some of its latent dangers. Rational evil-doers are much more pervasive than irrational ones, and so the task of the philosopher is not only to teach how to think rationally but also, and perhaps more fundamentally, to expose the conditions under which reason is misused. If this is the case, the thesis of radical evil can be read as an example of what Dieter Henrich identifies as Kant's turn from an ethics based on an abstract formula (given by reason) to another one based on the "assent" [*Zustimmung*] to the formula by the self, as part of which Kant attempts "to oppose the sophistry of reason, which stands in the service of want, and so to back the insight of the good against its dialectical tricks."[66] Philosophy, then, is as much an agent in the development of reason as in the exposure of its tricks. An essential part of this exposure consists in analyzing the mechanisms by which it is produced and sustained, as we shall now see.

---

[65] Hunter, *Rival Enlightenments*, 123-126.

[66] Dieter Henrich, "Der Begriff der sittlichen Einsicht und Kants Lehre vom Faktum der Vernunft," in *Kant: Zur Deutung seiner Theorie von Erkennen und Handeln*, ed. Gerold Prauss (Koln, Germany: Krepenheuer & Witsch, 1973), 232; my translation. Steven Lestition gives a contextual interpretation of this turn in "Kant and the End of the Enlightenment in Prussia," *The Journal of Modern History* 65(1) (1993), 91. Satkunanandan finds elements of this view on ethics in Kant's analysis of the categorical imperative in the *Groundwork* in "The Extraordinary Categorical Imperative."

Deception

There are two political developments, one immediate and the second one unfolding throughout decades, surrounding Kant's concern in *Religion* with the deceptiveness of religious practices. The first development, pointed out by a series of recent studies, is the growing public censorship in matters of theology and philosophy surrounding the writing of the text. Fredrick the Great, the "Enlightened monarch" celebrated by Kant in "What is Enlightenment?" for his support for freedom of speech, died in 1786. His son, Fredrick Wilhelm II, had more conservative views regarding religious toleration, as evidenced by the appointment as Minister of Education and Religious Affairs in 1788 of Johann Christoph Wöllner, a reactionary whose ideas were at odds with the Enlightenment. That same year, Wöllner promulgated a new edict restraining religious innovation by demanding that all teachers of theology adhere to the guidelines of their professed creed, and by reactivating censorship in matters of religion.[67] The aim and scope of Wöllner's edict are the subject of historical debate, but Kant was clearly aware that it would affect his ability to convey his thoughts honestly to the public—as the eventual censorship of *Religion*, most of which could only be published after the death of Fredrick Wilhelm II, would confirm. Susan Shell and Steven Lestition have analyzed in detail how the new political situation restrained Kant's confidence in the Enlightenment.[68] This had two implications for Kant's views on evil. First, evil is no longer seen as stemming from impulses that reason fails to overcome. If an enlightened era can be followed by an unenlightened one, the root of evil must be resistant to the progress of civilization, and perhaps may even be intertwined

[67] George Di Giovanni, Introduction to *Religion*, 42; Hunter, *Rival Enlightenments*, 338; Christopher Clark, *Iron Kingdom: The Rise and Downfall of Prussia, 1600-1947* (Cambridge, MA: Harvard UP, 2006), 267-274.
[68] Shell, *Kant and the Limits of Autonomy*, 187-211; Lestition, "Kant and the End of the Enlightenment in Prussia," 76-83.

with it. Second, the reinstatement of censorship and the demand of public declaration of faith bring about the problem of public insincerity. Is possible that, under a non-enlightened monarch, the enlightened will conceal their true opinions?

These issues are connected to Kant's broader concern with the deceptiveness of religious rituals in moral matters. Kant biographers Ernst Cassirer and Manfred Kuehn suggest that the strong discipline and control that characterized Kant's early Pietist education at the *Collegium Fridericianum* left a longstanding negative impression in him.[69] The *Fridericianum* was founded in Königsberg with the support of the two main figures of the Pietist movement, Philipp Jackob Spener and August Hermann Francke.[70] Spener, the founder of the movement, called for a transformation of the Lutheran church that shifted the emphasis from doctrinal orthodoxy to the experience of conversion and its manifestation through the performance of good deeds for the neighbor. His goal was to make religion more attuned to the needs and experiences of the laymen, by contrast to the observance of rituals and the learning of dogmas.[71] Spener's disciple Francke, a clever political strategist, played a crucial role in expanding the influence of Pietism by exploiting its affinities with the monarch's views and projects, especially in matters of education. At the time of the emergence of Pietism, the Prussian State was undergoing a process of rapid modernization under Fredrick Wilhelm I, who shared the worldly-oriented mentality of

[69] Manfred Kuehn, *Kant: A Biography* (New York: Cambridge UP, 2001), 54-55; Ernst Cassirer, *Kant's Life and Thought*, trans. James Haden (New Haven, CT: Yale UP, 1981), 16-17. See also Michel Despland, *Kant on History and Religion* (London: McGill - Queen's University Press, 1973), 106-107.
[70] Despland, *Kant on History and Religion*, 101.
[71] Spener's ideas are condensed in his main and most popular work, *Pia Desideria*. For a brief description of the book, see Lewis White Beck, *Early German Philosophy: Kant and His Predecessors* (Cambridge: Harvard UP, 1969), 159; Richard Gawthrop, *Pietism and the Making of Eighteenth-Century Prussia* (Cambridge: Cambridge UP, 1993), 107.

the new religious movement.[72] Crucial to this process was the development of the army and the bureaucracy, with their characteristic sense of discipline and obedience. The Pietist emphasis on worldly deeds over ritual, and on spiritual transformation over intellectual clarification, was functional to the development of this mentality. [73] Exploiting these affinities, Francke maneuvered to increase the number of Pietist teachers in schools, as well as the number of Pietist orphanages, which became the central institution of the movement. By Kant's time, the Pietist educational philosophy was widespread in Prussia.

Although Spener's views emerged as an attempt to reinvigorate personal religious experience against ossified symbols and rituals, Pietism under Francke led to a growing sense of discipline and control over people's actions. While both Spener and Francke emphasized the need of going through a "devotional rebirth" and overcome one's sinful self, Francke stressed the experience of repentance and uncertainty leading to it.[74] The anxiety regarding one's own moral condition translated into a growing need of purifying one's own practices, so as to prove oneself and others that one had regenerated himself.[75] Given that every proof of regeneration was uncertain, the demand for discipline and self-scrutiny was virtually inexhaustible. The only means to confirm a person's devotion was to thoroughly comply with the rules handed down by the educational and religious authorities. [76] As a consequence, some Pietist circles turned to an extreme regulation of all the spheres of life and to an ethics of blind obedience. Somewhat

---

[72] Gawthrop, *Pietism and the Making of Eighteenth-Century Prussia*, 205-206.
[73] Mary Fulbrook, *Piety and Politics. Religion and the Rise of Absolutism in England, Wüttemberg and Prussia* (Cambridge: Cambridge UP, 1983), 166-170.
[74] Martin Brecht, "August Hermann Francke und der Hallische Pietismus," in *Geschichte des Pietismus*, Vol. 1 (Göttingen, Germany: Vandenhoeck & Ruprecht, 1993), 463.
[75] Gawthrop, *Pietism and the Making of Eighteenth-Century Prussia*, 148.
[76] Klaus Schaller, "Pietismus und moderne Pädagogik," in *Pietismus und moderne Welt*, ed. Kurt Aland (Witten, Germany: Luther-Verlag, 1974), 173-174.

contrary to the original spirit of the movement, compliance with externally given rules became more important than inner regeneration—not because the latter ceased to be essential, but rather because external rules were its only available signs.[77] With time, these signs would be seen as a potential source of manipulation and deception, rather than as proofs of devotion—as indicated by Prince Wilhelm's (later Kaiser Wilhelm I) definition of a Pietist, in the context of a conversation with Otto von Bismarck in 1853, as "a man who feigns in religion in order to make a career."[78] Kant was certainly exposed to this tendency to conflate religious practices and obedience with careerism, both through his early education at the *Fridericianum* and, we may speculate, through his observation of religious development in the Königsberg of his time. If this experience made him distrustful of strict discipline and obedience in matters of religion from early on, as evidenced by the negative remarks on his childhood education, Wöllner's edict could not but intensify the fear of opportunistic insincerity in public speech and action.

Kant's concern regarding these tendencies becomes clear in the fourth part of *Religion*, titled "Concerning service and counterfeit service under the dominion of the good principle." The section analyzes deceptive forms of religious practices on the basis of the distinction between moral and statutory religion. Moral religion consists in the fulfillment of one's moral duty represented as divine commands, while statutary religion is concerned with commands given by God in history.[79] True religion, according to Kant, is moral religion, but this does not mean that

---

[77] Spener refers to deeds as "signs" [*Kennzeichen*] of inner devotion in *Der neue Mensch*, ed. Hans-Georg Feller (Stuttgart: J. F. Steinkopf, 1966), 64.

[78] Johannes Wallmann, "Was ist Pietismut?," *Pietusmus und Neuzeut* 20, 11. According to Fulbrook, "the need for Pietist testimonials to obtain positions in church and state led to superficial professions of conversion and regeneration according to the routinized general stages of Pietist experience. Pietism, conceived as a spontaneous religion of the heart, had become rationalized and mechanical as the orthodoxy of the state" (*Piety and Politics*, 170).

[79] Kant, *Religion*, 137-138; 6:102-106.

statutary religion should be disregarded altogether, given "the natural need of all human beings to demand for even the highest concepts and grounds of reason something that *the senses can hold on to* [*etwas* Sinnlichhaltbares]."[80] This demand for sensuous mediation is both indispensable and dangerous. Kant defines it as a "vehilcle" [*Leitmittel*] between the universality of the true church and the "particular validity" of a historically grounded church.[81] It is indispensable because sensuous beings cannot gain access to the rational ground of religion without the guidance of historically given commands and doctrines—just like a child, Kant explains, needs to obey seemingly arbitrary commands before he can understand the reason behind them.[82] It is dangerous, because the sensuous mediation tends to become autonomous regarding its mediating function. Ideas and practices that are meant to give a sensuous representation to moral duty are taken to be duty itself, thus transforming service into "mere *fetichism*" [*bloßes* Fetischmachen; "mere fetich-making"]. Kant calls this "delusion of religion" [*Religionswahn*].

Religious delusion is part of the self-deceptive mechanisms by which the evil disposition conceals itself. This sort of deception is self-inflicted, for it is "not merely an unpremeditated deception [*unvorsätzliche Täuschung*] but a maxim by which we attribute intrinsic value to the means rather than the end."[83] This maxim to deceive oneself is part of the "hidden inclination to deceit" [*verborgene Betrugsneigung*] that, as we have seen, is intertwined with an evil disposition. The deception has a very curious structure, for it is a deception regarding a constitutive deception. Rather than troubling ourselves with a moral improvement that is always

[80] Ibid., 142; 6:109.
[81] Ibid., 146 ; 6:115.
[82] Ibid., 151; 6:122-123.
[83] Ibid., 190; 6:170.

doubtful and problematic, we have a tendency to conceal our disposition by pretending to reveal it through ritualized moral deeds. This deception does not stem from a mistake, but rather from an intention. It is part of the corruption of choice, which conceals the choice of a maxim behind an empirically good disposition.[84] Different kinds of rituals can serve this purpose. Kant claims, however, that "sacrifices" [*Aufopferungen*] are particularly effective in this regard, for "just because they have absolutely no use in the world, and yet cost effort [*Mühe kosten*], they seem to be aimed solely at attesting [*Bezeugung*] devotion to God."[85] The deception, Kant believes, can go as far as to serve no personal interest. By sacrificing ourselves for no visible purpose whatsoever, we seem to comply with the law. The maxim to deceit thus becomes as powerful as the evil maxim to serve our inclinations, to the point that it is unclear which one comes first. At some point, the more we act against our own self-interest, the more we *seem* to be complying with the law.

The link between evil and deception represents an important innovation regarding the traditional understanding of evil. Kant's thesis of radical evil has been often criticized for remaining within the traditional conception of evil as deficiency.[86] A source of this critique is Kant's denial of the possibility of diabolical evil, that is, of an evil whose incentive is resistance to the moral law for the sake of this resistance. We can see, however, that the denial of diabolical evil does not imply a denial of the *positivity* of evil—it implies a denial of its *transparency*. Kant briefly states that an incentive to resist the law would be "too much" as a ground of moral evil,

---

[84] According to Peter Fenves's reading, "concealing from itself the source of its selfhood in the secret of freedom, the 'I' hides from itself its responsibility for the actions it undertakes, and this hiding is itself an action for which it is responsible and from which it hides itself" (*Late Kant*, 85).
[85] Kant, *Religion*, 189; 6:169.
[86] See for example Alenka Zupančič, *Ethics of the Real: Kant, Lacan* (London: Verso, 2000), 98; Jean-Luc Nancy, *The Experience of Freedom* (Stanford, CA: Stanford UP, 1993), 125.

while our sensuous nature would be "too little."[87] Given the thesis of the inscrutability of moral maxims, and the intertwinement of evil and deception, we can interpret the "too much" as "beyond inscrutability and deception." In order to be diabolical, a person's maxim would have to be transparent to her, like they would be to an innocent being, for whom the law would be the only incentive for action. The human being, by contrast, is a being exposed to corruption, which means that he is incapable of pure maxims, either good or evil. But this corruption is freely chosen, which is why it is no mere deficiency, even if it manifests itself as such. Given that evil is intertwined with deception, it appears as ignorance or weakness, that is, as an incapacity to know and love the good. The originality of Kant's thesis of radical evil consists in turning ignorance as a deficiency that explains evil into a deception that is produced by evil. Thus while we are responsible for the choice of a maxim, we are at the same time responsible for its concealment.

From Radical Evil to the Banality of Evil

Self-concealment is an essential attribute of what Arendt called "the banality of evil." As we have seen, Eichmann claimed during the trial that, since he had done nothing but thoroughly comply with the law of the land, he could not be held accountable for his actions. Given the traditional contrast between egoistic self-interest and duty, Eichmann's willingness to sacrifice himself for the latter served in his eyes as a proof of moral virtue. If there was anyone responsible for the moral implications of his deeds, it was the source of the law. For Arendt, this reasoning was not simply a lie, but rather a reflection of a longstanding conception of virtue on

[87] Kant, *Religion*, 82; 6:35.

which Eichmann relied as a means to conceal his responsibility. Arendt believed that the philosophy of the will, of which Kant was a representative, was for the most part complicit with this obliteration of responsibility. In her later writings, however, she searched for an alternative conception of the will, according to which it does not simply give a law and command its execution, but rather chooses and constitutes a character. Arendt's goal was to find the source of that which Eichmann seemed to lack, namely, a ground of responsibility. The challenge was twofold: first, to understand on what grounds Eichmann could be condemned for his actions, that is, on what grounds they could be judged to be "evil"; second, to understand the mechanism by which a person refuses responsibility—how can someone conceal or destroy her own ground of imputability?

Kant's analysis of evil and imputability in *Religion* conveys a response to both challenges. The ground of imputability lies in choice, which ought to remain inscrutable in order to be a choice at all. This is indeed what we could call the "paradox of responsibility" in Kant: in order to have a moral character that is responsible for our actions, such character must remain hidden. Every new action puts our character into question, without ever revealing it. Radical evil, according to Kant, consists in deceiving ourselves regarding our disposition, precisely by attempting to turn it into an object of experience—a fetich. Eichmann's unbreakable commitment to his duty beyond all self-interest, which made it so difficult to judge his actions as evil, can therefore be read as a mechanism by which he destroyed his own responsibility. Moreover, it is precisely because radical evil conceals itself behind empirical signs of virtue that it becomes deceitful and, consequently, utterly pervasive. Yet Kant shows that this concealment is itself freely chosen, and thus *imputable*. Therefore, the rootlessness that characterizes "the

banality of evil" and the depth of "radical evil" are not mutually exclusive: it is precisely because the person's choice can be corrupted to the core that she can appear as lacking personality, that is, as if there were nothing behind the surface of her actions.

The link between "radical evil" and "the banality of evil" is of course no coincidence. There is a historical trend connecting Kant's insights in *Religion* with the emergence of totalitarianism. Historian Richard Gawthrop claims that the extreme discipline and control developing in Pietist institutions represents an antecedent to the totalizing views of modern ideologies, in which people voluntarily subordinate themselves to the rules sanctioned by a group believed to hold the truth.[88] This discipline was seen as a way to bridge the gap between worldly deeds and the inscrutable moral worth of the inner self. This development, together with the emergence of censorship and of public regulations regarding profession of faith with Fredrick Wilhelm II, which seemed to subordinate inner conviction to public displays, turned Kant aware of a dangerous tendency in his time to fetichize external signs of virtue. His emphasis on the inscrutability of our disposition and his analysis of religious delusion can therefore be read as warnings against the development of a proto-totalitarian mentality, in which individuals hold fast to whatever moral rituals they are given at a certain point. The identification with these rituals conceals responsibility, because the person acts as if she had no choice but to comply with the rules that are given to her.

[88] According to Gawthrop, "while the 'door for betterment' was open for all, closing the threshold did not mean that the individual was free to pursue his or her spiritual development without restriction. In fact, although Francke and his co-workers did not believe in predestination, they none the less regarded themselves as an elite group with the power to direct and control the souls entrusted to them… They claimed to possess, as many modern ideologues have done, a type of knowledge that was simultaneously empirical and absolute" (*Pietism and the Making of Eighteenth-Century Prussia*, 152).

The historical trajectory connecting Kant's *Religion* with the social developments of the twentieth century did not go unnoticed within German philosophy. Arendt's analysis of totalitarianism was directly influenced by her reading of Kant, but also mediated by the interpretation of her two mentors, Heidegger and Jaspers. Both philosophers were concerned with uncovering a more original notion of responsibility that is not subordinated to pre-given external norms, but is rather the precondition for any norm. "Guilt," according to Heidegger and Jaspers, is not an effect of not complying with a norm, but rather a constitutive condition of our existence.[89] Of course, we do feel guilty for transgressing particular norms, but this would not be possible if we were not willing to assume responsibility for our own actions to begin with. To derive the experience of guilt from a specific norm is already a way to misunderstand it, for this way responsibility is displaced toward that which is external to us. That is why, in Jaspers's interpretation, Kant's analysis of radical evil shows that there is no objective distinction between good and evil. Such objectification, according to Jaspers's reading of Kant, is the first moral deception. This deception is operational in most philosophical systems, in which

> every time evil is almost looked over in that, by being turned into an object of observation (and even into a grandiose appearance), it becomes harmless. Why now this obstinacy to find possible objectifications of evil in all directions? Because each time one would miss, through the becoming palpable of evil, the chance to find the possibility of improvement in myself.[90]

According to Jaspers, the aim of Kant's practical philosophy consists in resisting this objectification by showing the limits of what we can know and control, thus opening up the

---

[89] See Heidegger's analysis of a primordial "being-Guilty" in *Being and Time*, trans. John Macquarrie and Edward Robinson (Oxford: Basil Blackwell, 1985), 326-335; and Jaspers's analysis of "metaphysical guilt" as the precondition for both moral and political guilt in *The Question of German Guilt*, trans. E. B. Ashton (New York: Fordham UP, 2000), 25-40.

[90] Karl Jaspers, "Das Radikal Böse bei Kant," in *Rechenschaft und Ausblick. Rede und Aufsätze* (Munich: R. Piper, 1958), 114-115; my translation.

possibility of an open-ended transformation. To objectify evil is already a way to externalize it, and thus to blur our own responsibility.

It should be clear that the concealment of one's responsibility behind visible signs has nothing to do with the concealment of a "true intention." This is the way Allison interprets the link between radical evil and the banality of evil. According to Allison, Eichmann is an example of the deception at play in Kant's understanding of evil, in which "immoral maxims appear to pass the universalizability test only because they ignore or obscure morally salient features of a situation."[91] In doing evil, Allison claims, one does not purposefully transgress the law, but rather deceives oneself regarding the features in a situation that determine whether one's actions comply or not with the moral law, that is, whether or not they pass the "universalizability test." This false justificatory procedure allows the person to ignore her "actual intention," which is "the morally salient factor, defining the very nature of the act."[92] The problem with this reading is that it ignores Kant's thesis on the inscrutability of our maxims and its effects (the "putting out of tune") on our ability to make moral judgments. According to Kant, there is no access to our "actual intention," which lies beyond anything given to us through our senses. There is, therefore, no empirical features that can serve as a test for grasping our moral maxims. As Alenka Zupančič has pointed out in response to Allison, the first deception consists precisely in taking empirical signs as a proof of moral virtue, as if our disposition could be turn into an object of experience.[93] Following Jaspers, we must understand the self-deception operative in radical evil as the attempt to turn empirical deeds into means of moral certainty, which relieve us from

[91] Henry E. Allison, *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy* (New York: Cambridge UP, 1996), 181.
[92] Ibid.
[93] Zupančič, *Ethics of the Real*, 93.

the need to judge each one of our deeds anew. In the words of Olivier Reboul's classical study, for Kant "evil is sanctimony, the fact of believing oneself justified by one's deeds, one's 'good works,' or of taking one's exterior non-culpability for innocence."[94] The thesis of radical evil is not a call to overcome deception and acknowledge our true intention, which can only lead to an even more radical delusion, akin to fanaticism, but rather to accept the inscrutability of our intentions as constitutive of our responsibility for action.

Conclusion

Eichmann claimed that, in complying with the law of the land as thoroughly as possible, he had followed a version of Kant's practical philosophy adapted to his context. According to Arendt, the one aspect of this "household Kant" that preserved his original spirit was the demand that one always sacrifice his own self-interest and make no exception to her compliance with duty. Because of the predominant view of evil as stemming from some sort of egoistic motivation, there seemed to be no viewpoint from which to hold Eichmann accountable for his deeds. Although Arendt took Kant's notion of radical evil as a contrast to "the banality of evil" that she attributed to Eichmann, the two are closely connected. Kant's notion emerged as a response to a phenomenon that resembled what Arendt observed in totalitarianism, namely, the concealment of one's own responsibility behind moral rituals and attachment to self-sacrifice. For Kant, the ground of responsibility does not stem from externally given rules or formulas, but rather from our capacity to choose the maxim of our actions. Given that this maxim is inscrutable, it can never be regulated by any sort of ritualized practices, which are always

[94] Olivier Reboul, *Kant et le problème du mal* (Montreal: Montreal UP, 1971), 103; my translation.

deceitful as signs of moral virtue. The attempt to overcome this inscrutability by holding fast to empirical signs of virtue, such as self-sacrifice, constitutes an active deception regarding our own responsibility. We can never know the disposition that underlies each one of our deeds, yet it is precisely because of this hiddenness that the demand for self-transformation is actualized with every new deed.

Part II

Judgment

Chapter 3: Kant on the Sublime and the Judgment of Action

The role of Kant's aesthetics in the context of his practical philosophy has been a persistent source of debate within Kant scholarship. With the publication of Hannah Arendt's *Lectures on Kant's Political Philosophy*, the issue transcended philosophical discussions and attracted the attention of political theorists. In Arendt's reading of Kant's *Critique of the Power of Judgment*, the judgment of taste constitutes a model for political judgment, because it stems from the capacity to communicate our particular perspective to others, as opposed to applying a transcendental law or principle. While the fruitfulness of this turn to Kant's *Third Critique* has been broadly acknowledged, a number of studies have criticized Arendt's attempt to read Kant's conception of the aesthetic judgment independently of his practical philosophy, arguing that this approach is unable to provide an account of the judgment's claim to assent by others.[1] Given this emphasis on the relationship between aesthetic and moral judgments in Kant, it is surprising that no study so far has addressed the role of this relationship in his analysis of the sublime in connection to Arendt's *Lectures*. As this chapter will show, the sublime plays a crucial role in addressing an issue that had been left open in Kant's practical philosophy until the publication of the third *Critique*, namely, the possibility of judging action on the basis of moral ideas. Therefore, an inquiry into the connection between judgment and action in Kant's analysis of the

[1] Robert J. Dostal, "Judging Human Action: Arendt's Appropriation of Kant," *The Review of Metaphysics* 37(4); Ronald Beiner, "Rereading Hannah Arendt's Kant Lectures," in *Judgment, Imagination, and Politics: Themes from Kant and Arendt*, eds. Ronald Beiner and Jennifer Nedelsky (Oxford: Rowman & Littlefield, 2001); Seyla Benhabib, "Judgment and the Moral Foundations of Politics in Hannah Arendt's Thought," in *Judgment, Imagination, and Politics: Themes from Kant and Arendt*; Alfredo Ferrarin, "Imagination and judgment in Kant's practical philosophy," *Philosophy & Social Criticism* 34(1-2). I discuss some of these criticisms in detail further down. Andrew Norris criticizes Arendt for neglecting the link between the judgment of taste and Kant's theoretical philosophy, which shows that the former is not an adequate model for political judgment, in "Arendt, Kant, and the Politics of Common Sense," *Polity* 29(2).

sublime will not only contribute to better understand its importance for his practical philosophy, but also clarify problematic aspects in Arendt's reading of Kant's aesthetics. As we will see, Kant's insight in his analysis of the sublime was closer to Arendt's views on the relationship between the actor and the spectator than many commentators acknowledge.

My reading of Kant's analysis of the sublime in connection to the problem of the judgment of action has two goals. First, I believe that this connection can help us clarify the source of the claim to approval by others in the judgment of action. Most of Arendt's remarks on the problem of the validity of judgment locate it in the capacity to take into account multiple perspectives, an idea that has drawn criticism on grounds that it makes validity dependent upon the empirical agreement of one's community. However, Arendt also considered the possibility that the validity of judgment stems from a universal idea, which she calls "the idea of mankind," and which connects the principle of judgment with the principle of action, allowing us to judge from the widest possible perspective. This idea, I will argue, is close to what Kant calls "the idea of humanity" in his analysis of the judgment of the sublime, and which is the ground of the claim to universal approval. The second goal of my reading is to contribute to the understanding of the role of the judgment of the sublime in Kant's practical philosophy. Against a number of studies that read the "Analytic of the Sublime" as situated exclusively within Kant's investigation of the nature of aesthetic judgment, I follow the view of Jean-François Lyotard, Paul Crowther, John H. Zammito, and Robert R. Clewis that the sublime represents an important development in Kant's practical philosophy.[2] One important element in this regard consists in clarifying how it is

[2] Jean-François Lyotard, *Lessons on the Analytic of the Sublime*, trans. Elizabeth Rottenberg (Stanford, CA: Stanford UP, 1994); Paul Crowther, *The Kantian Sublime: From Morality to Art* (Oxford: Clarendon Press, 1989); John H.

possible to cast a judgment based on moral ideas upon actions as they appear in experience, which seemed impossible in Kant's practical philosophy up to that point. This is the issue that will be the focus of my reading.

Before proceeding to the analysis, it is worth pointing out that both Arendt's interest in the judgment of taste and Kant's analysis of the sublime represent in part a response to their reflections upon evil. Arendt's *Lectures* were written shortly after *Eichmann in Jerusalem*, clearly in connection to Eichmann's lack of "representative thinking" and his "inability to judge"—two essential elements of what she called "the banality of evil." Kant's *Critique of the Power of Judgment* appeared two years before the publication of the first part of *Religion within the Boundaries of Mere Reason*, whose central topic is the problem of evil, and in which Kant acknowledges the deceptive effects of evil upon our capacity for moral judgment. As I will show further down, Kant suggests that the feeling of the sublime counters this deceptiveness, thus pointing to the judgment of the sublime as the proper site for the judgment of the morality of action. Therefore, there is a shared concern connecting Arendt's reading of the judgment of taste and Kant's analysis of the sublime, namely, the need to find a form of judgment that counters the propensity to moral deception.

Arendt's Turn to Kant's *Third Critique*

Arendt's turn to Kant's aesthetics in her *Lectures* is influenced by one of the central problems involved in *Eichmann in Jerusalem* and the notion of "the banality of evil," namely, the "inability to judge." In his classical study of the *Lectures*, Ronald Beiner identifies two

---

Zammito, *The Genesis of Kant's* Critique of Judgment (Chicago: The University of Chicago Press, 1992); Robert R. Clewis, *The Kantian Sublime and the Revelation of Freedom* (Cambridge: Cambridge UP, 2009).

reasons for Arendt's later concern with judgment: first, Eichmann's own lack of judgment showed the importance of the faculty of judgment for our public life; second, it was by no means evident how one was to judge Eichmann from a third-person standpoint—for example, the standpoint of the judges in Jerusalem.[3] Both reasons are part of a broader problem that concerned Arendt throughout her work, namely, the absence of stable, reliable standards for judgment in the modern world.[4] Eichmann showed to what extent the idea that, given the absence of such standards, we are dispensed with the need to judge altogether, can work as a mechanism of self-exculpation, while the difficulty of casting judgment upon him showed to what extent this mechanism is a public problem rather than a merely individual subterfuge. As Arendt claims in a later essay, "behind the unwillingness to judge lurks the suspicion that no one is a free agent, and hence the doubt that anyone is responsible or could be expected to answer for what he has done."[5] The refusal to judge our actions as well as the actions of others contributes to a generalized self-exculpation, in which no one feels personally responsible for what he does. The question is, then: how do we judge action in the absence of universally shared standards?

This question emerges in part from Arendt's observations regarding Eichmann's "thoughtlessness" and "inability to judge." Eichmann's thoughtlessness consisted in an "almost

---

[3] Ronald Beiner, "Interpretative Essay," in Hannah Arendt, *Lectures on Kant's Political Philosophy*, ed. Ronald Beiner (Chicago: The University of Chicago Press, 1992), 97-101. Leora Y. Bilsky analyzes both aspects of the problem of judgment in the context of Eichmann's trial in "When Actor and Spectator Meet in the Courtroom: Reflections on Hannah Arendt's Concept of Judgment," in *Judgment, Imagination, and Politics*.

[4] Especially relevant in this regard are Arendt's remarks on the difficulty of judging after totalitarianism in "Understanding and Politics," in *Essays in Understanding*, ed. Jerome Kohn (New York: Schocken Books, 1994), 316; and her claim in "The Crisis in Culture" that "inability to judge" is one of the characteristics of the modern "mass man" (in *Between Past and Future: Six Essays in Political Thought* [New York: The Viking Press, 1961], 199).

[5] Hannah Arendt, "Personal Responsibility Under Dictatorship," in *Responsibility and Judgment*, ed. Jerome Kohn (New York: Schocken Books, 2003), 19.

total inability ever to look at anything from the other fellow's point of view."[6] This was evident, according to Arendt, in the way he accounted for his role during the initial years as one of cooperation between he and the Jews in the process of emigration—that is, as if there had been a true communion of interest between the parties involved. This indicated that Eichmann was unable, both during the events and in his recollection of them during the trial, to assess his actions from any other viewpoint than the fulfillment of his bureaucratic duty. This "inability to think" was closely linked to his tendency to speak by means of a repetition of clichés and stock phrases, which worked as "the most reliable of all safeguards against the words and the presence of others," and as a consequence of which "no communication was possible with him."[7] Eichmann's second prominent trait, his "inability to judge," is not unconnected to this thoughtlessness. In Arendt's account, this becomes clear at the point in which Eichmann adapted to the new task of extermination as the person in charge of the Final Solution, of which he was informed during the Wannsee Conference in 1942. During those days, Eichmann witnessed how "the élite of the good old Civil Service" was more than willing to take part in the enterprise of mass murder. This was taken by him as a moral approval for his role, given that if people of higher social status had no guilty conscience, "*who was he to judge?*."[8] Thus both Eichmann's inability to think and his inability to judge played a crucial role in disregarding moral considerations upon his actions: the stock phrases which he and the people around him used to refer to their actions, on the one hand, and the views of the "respectable society" around him, on

[6] Hannah Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York: Penguin, 2006), 48.
[7] Ibid., 49.
[8] Ibid., 114.

the other, shielded him from the need to consider how his deeds would look like from the viewpoint of others—at least beyond his immediate social circle.

Eichmann's inability to think and to judge showed the importance of casting judgments independently of the standards of one's community. In the Postscript to *Eichmann in Jerusalem*, Arendt claims that Eichmann's trial demanded "that human beings be capable of telling right from wrong even when all they have to guide them is their own judgment which, moreover, happens to be completely at odds with what they must regard as the unanimous opinion of all those around them."[9] At the same time, however, Arendt does not believe that universal moral principles are a reliable ground for distinguishing right and wrong, which explains her insistence on separating the judgment of taste from what she calls "moral propositions."[10] In her view, holding fast to moral norms is no guarantee against adapting to a new situation in which those norms no longer hold. As she claims in "Personal Responsibility Under Dictatorship," we know after totalitarianism "that moral norms and standards can be changed overnight, and that all that then will be left is the mere habit of holding fast to something."[11] Arendt's distrust of our attachment to moral norms is influenced by her description of Eichmann as someone who had always been seeking an "Idea" for which he could sacrifice everything, and thus give meaning to his life. This attitude was also evident in Eichmann's claim that until his role in the Final Solution, he had lived his whole life according to Kant's definition of duty, only to switch to a new version (a "household version") of the categorical imperative when it seemed impossible to

[9] Ibid., 294-295.
[10] Arendt claims that "Kant does not believe that moral judgments are the product of reflection and imagination, hence they are not judgments strictly speaking" (*Lectures on Kant's Political Philosophy*, 72). The idea of "moral propositions" as stemming from reason as opposed to judgment appears on p. 10.
[11] Arendt, "Personal Responsibility Under Dictatorship," 45.

comply with the original one. In response to this attitude, Arendt claims that it is not those who cherish a system of moral values, but rather the "doubters and skeptics" who are "used to examine things and to make up their own minds," who preserve the capacity for autonomous judgment beyond the contingent standards of a given context. If judgment upon right and wrong is still possible in situation in which all customary standards are suddenly replaced, it is not because the faculty of judgment possesses immutable standards independent of the context, but rather because it "functions spontaneously, that is to say, is not bound by standards and rules under which particular cases are simply subsumed, but on the contrary, produces its own principles by virtue of the judging activity itself."[12]

It is precisely the idea that the power of judgment can produce its own principles, independently of either communitarian or transcendental rules, that led Arendt to Kant's *Critique of the Power of Judgment*. Arendt's reading of this text focuses on the judgment of taste, and especially on the notion of common sense, although her interpretation of it departs significantly from Kant's own framework. By contrast to cognitive or moral standards, which are objective because they are given to the power of judgment independently of the situation and of the presence of others, common sense is "intersubjective."[13] Arendt describes it as "that sense which fits us into a community with others, makes us members of it and enables us to communicate things given by our five private senses."[14] Unlike moral and cognitive judgments, which apply a general rule to a particular case, the judgment of taste makes communicable a private perspective of an object. This communicability stems from the combined operation of imagination and

[12] Ibid., 27.
[13] Arendt, "Some Questions of Moral Philosophy," in *Responsibility and Judgment*, 141.
[14] Ibid., 139.

common sense. Imagination, Arendt claims, "transforms an object into something I do not have to be directly confronted with," so that "one now has, by means of representation, established the proper distance, the remoteness or uninvolvedness or disinterestedness, that is requisite… for evaluating something at its proper worth."[15] By taking distance from an object through representation, we are able to transcend our immediate perception of it, thus acquiring what Arendt calls (following Kant) "impartiality." Then, common sense puts our own perspective in relation with the perspective of others, so that we "overcome our special subjective conditions for the sake of others."[16] This is done by considering what the object looks like from the viewpoint of others, so that the person who judges "claims the assent from others because in judging he has already taken them into account and hence hopes that his judgments will carry a certain general, though perhaps not universal, validity."[17] The communicability of our judgment does not stem from correctly applying a certain rule or principle independently of perspectives, but rather from considering other perspectives at the moment we judge. The more perspectives we take into account, the more valid our judgment will be.

Does common sense constitute a standard of validity that is independent of the empirical agreement of those around us? The difficulty of answering this question in the context of Arendt's analysis has drawn criticism on grounds that common sense is ultimately rooted in the standards of one's community. As we have seen, Arendt does not believe that the validity of our judgment is based on the empirical agreement of others—on the contrary, it is often the case that we must judge "against the unanimous opinion" of all around us. Indeed, taking into account

[15] *Lectures on Kant's Political Philosophy*, 67.
[16] Ibid.
[17] "Some Questions of Moral Philosophy," 140.

others when judging "does not mean that I conform in my judgment to their's," because "I still speak with my own voice and I do not count noses in order to arrive at what I think is right."[18] And although "one judges always as a member of a community, guided by one's community sense, one's *sensus community*," in the last analysis "one is a member of a world community by the sheer fact of being a human; this is one's 'cosmopolitan existence.'"[19] Common sense allows us to transcend our particular perspective so as to take into account to perspective of others, but it neither leaves behind our perspective nor does it stop at the perspective of our community. Even if we can consider how an object appears to other standpoints, we must also be able to consider standpoints beyond all those familiar to us, to the point of challenging the views of those around us. This leads us to the question of how common sense, the sense that "fits us into a community with others," can make us members of a "world community" to which we belong by the sheer fact of being human.

This problem has led a number of commentators to criticize Arendt's stark separation of the judgment of taste from the moral judgment. The criticism is shared by Ronald Beiner, Seyla Benhabib and Robert Dostal, who argue that by leaving aside Kant's concern with moral foundations, Arendt could not find a viable alternative between a strictly communitarian and a transcendental, universalist conception of judgment. According to Beiner, Arendt's attempt to completely separate Kant's analysis of the judgment of taste from his practical philosophy, something that a systematic reconstruction of the third *Critique* would certainly not allow, confuses rather than clarify the problem of the validity of judgment.[20] This is so for two

[18] Ibid., 141.
[19] *Lectures on Kant's Political Philosophy*, 75.
[20] Ronald Beiner, "Rereading Hannah Arendt's Kant Lectures," 98-99.

interconnected reasons: first, Kant conceived both the practical judgment and the judgment of taste under the ideal of autonomy, which means that the individual's judgment is independent from the community to which she belongs; second, the categories used by Kant to analyze the judgment of taste are transcendental rather than social, which means that its standards of validity are (again) independent from any empirical community. Beiner finds in John Rawls an example of how Kant's analysis of the aesthetic judgment could be read as a model for political judgment without disregarding its connection to his practical philosophy. Rawls's idea of an original position from which to evaluate the justice of actions and institutions, Beiner claims, "involves an experiment of moral reflection engaged in by one representative rational agent."[21] While this operation relies on representative thinking, in the sense of transcending one's particular viewpoint and imagining what the world looks like from the viewpoint of others, it is performed by an autonomous rational agent, whose moral standards are independent of the standards of the concrete community to which she belongs.

Benhabib's criticism of Arendt is connected to Beiner's, although she is more distrustful of the capacity of Kant's moral philosophy to solve the seemingly "communitarian" implications of Arendt's analysis. Like Beiner, Benhabib is critical of Arendt's attempt to disregard moral foundations in the judgment of action. Also like Beiner, she relies on Rawls's interpretation of Kant so as to argue that "there is a moral foundation to politics insofar as any political system embodies principles of justice."[22] Arendt's analysis of the judgment of taste in connection to representative thinking should therefore be read not as an alternative to moral foundations, but rather as a procedure by which to apply such foundations to concrete situations, thus preserving

[21] Ibid., 97.
[22] Seyla Benhabib, "Judgment and the Moral Foundations of Politics in Hannah Arendt," 200.

the "universalist-egalitarian kernel of Kantian morality."[23] On this point, Benhabib finds a deficit in Kant's practical philosophy, which holds that, because morality depends on the intentions driving the act, "we can never know if an action was morally virtuous in this sense at all, since the morally good defies embodiment in the phenomenal world."[24] Arendt's interpretation of the aesthetic judgment, Benhabib claims, compensates for this deficit by providing a procedure with which to overcome the gap between intentions and phenomenal action. The operation of common sense, which Benhabib translates into two formulas for judgment, works as a standard by which to describe the phenomenal action is such a way that its moral value is taken into consideration. This way, we can combine the universalist foundation of Kantian morality with the contextual thinking that makes it applicable to specific situations.

A similar attempt to turn the aesthetic judgment into a means by which to apply moral foundations to experience is developed by Robert Dostal. In his view, however, this procedure is implicitly developed by Kant himself. Like Benhabib, Dostal agrees with Arendt that the role of the imagination in the judgment of taste can be applied to the judgment of action, while rejecting the idea that this kind of judgment can be independent of practical reason. The challenge, then, is to connect the imagination with Kant's analysis of the practical judgment in the *Critique of Practical Reason*, which is concerned with pure moral principles rather than with their application to phenomenal actions. According to Dostal, this is achieved by supplementing the analysis of the practical judgment with the *Metaphysics of Morals*, in such a way that the latter becomes an exercise of the imagination by which to apply practical judgments to experience. Through this operation, "one comes to see what Kant himself did not clearly see—that Kant is

[23] Ibid., 197.
[24] Ibid., 193.

asking one to imagine reflectively a human world where such an action might occur."[25] For Dostal, the *Metaphysics of Morals* plays the same role that Rawls's original position plays for Beiner and Benhabib, namely, allowing us to imagine a situation that works as a guide for the application of moral principles to experience, in such a way that we can judge the moral value of empirical actions. This way, Dostal claims, the function of the imagination in moral matters is analogous to the function that Kant assigned to it in the *Critique of Pure Reason*—while here the imagination applies the concepts of the understanding to empirical objects, the *Metaphysics of Morals* is an example of how it can apply ideas of reason to empirical actions.[26]

The three criticisms of Arendt just described share the concern that, without a connection with moral foundations, the aesthetic judgment has no other principle of validity than the empirical agreement of others. If we lack a universal principle that we acknowledge independently of contingent circumstances, how are we to understand Arendt's claim that we must be capable of producing a judgment that is at odds with the views of those around us? How is common sense different from a communitarian ground for judgment? Arendt, I believe, was much more aware of these difficulties than her critical commentators allow, and she by no means considered this problem of secondary importance, as Linda Zerilli has proposed.[27] In Arendt's

[25] Robert J. Dostal, "Judging Human Action: Arendt's Appropriation of Kant," 736.

[26] Ibid., 736. Dostal's confusion regarding Kant's conception of practical judgment, as analyzed in "The Typic of Practical Judgment" in the second *Critique*, is evident in his claims that "the typic needs to be supplemented by this development of Kant's thought, for clearly in the typic of the *Critique of Practical Reason* Kant cannot quite make good the analogy with the schematism of the *Critique of Pure Reason* and the judgmental application of the pure concepts to sensibility." The reason why the practical judgment has a typic and not a schematism is precisely that the two are not identical, so it is unclear in what sense the latter does not "make good" the analogy.

[27] Linda M. G. Zerilli, "'We Feel our Freedom': Imagination and Judgment in the Thought of Hannah Arendt," *Political Theory* 33(2). According to Zerilli, Arendt "does not think that validity in itself is the all-important problem or task for political judgment—the affirmation of human freedom is" (164). Zerilli claims that Arendt does provide a response to the problem of the validity of judgment, but that it stems from (and therefore does not precede) "opinion formation" and "practices of freedom" in the public realm (p. 166). This says little about Arendt's concern with the need to judge even against the standards of one's own community.

view, the validity of the judgment of taste does not depend on the empirical agreement of others, but rather on the degree of generality attained by our *sensus communis*—a degree that need not remain within the boundaries of any given community.[28] Although Arendt says little about the "world community" to which we belong by the fact of "being human," she refers at one point in her *Lectures* to "the necessary condition for the greatest possible enlargement of the enlarged mentality."[29] This condition is "the idea of an original compact, dictated by mankind itself," which in the context of the third *Critique* refers to a state of perfect communication in which all our pleasures can be communicated to others.[30] Arendt's own use of this idea, however, departs significantly from Kant, and is closer to what he calls "the idea of humanity." Indeed, Arendt claims that it is by virtue of "the idea of mankind, present in every single man, that men are human, and they can be called civilized or humane to the extent that this idea becomes the principle not only of their judgments but of their actions."[31] Through this idea, "actor and spectator become united," because "the maxim of the actor and the maxim… according to which the spectator judges the spectacle of the world, become one."[32] Thus the idea of mankind both allows us to attain the greatest possible enlargement of our enlarged mentality, which is the

[28] Andrew Norris neglects Arendt's explicit references on this point when he claims that her understanding of "communicability" transforms "a transcendental principle into an empirical criterion" ("Arendt, Kant, and the Politics of Common Sense," 188).
[29] Arendt, *Lectures on Kant's Political Philosophy*, 74.
[30] Immanuel Kant, *Critique of the Power of Judgment*, ed. Paul Guyer, trans. Paul Guyer and Eric Matthews (Cambridge: Cambridge UP, 2000), 177; 5:297.
[31] Arendt, *Lectures on Kant's Political Philosophy*, 75.
[32] Ibid. This reference is surprisingly ignored by Richard Bernstein and Ronald Beiner, who claim that Arendt introduced and unbridgeable gap between the viewpoint of the actor and the viewpoint of the spectator, as well as by David Marshall and Jonathan P. Schwartz, who respond that bridging this gap is precisely the purpose of the judgment of taste (Richard J. Bernstein, *Philosophical Profiles: Essays in Pragmatic Mode* [Philadelphia: University of Philadelphia Press, 1986], Chapter 8; Ronald Beiner, "Rereiding Hannah Arendt's Kant Lectures," in *Judgment, Imagination, and Politics*; David Marshall, "The Origin and Character of Hannah Arendt's Theory of Judgment," *Political Theory* 38[3]; Jonathan P. Schwartz, *Arendt's Judgment: Freedom, Responsibility, Citizenship* [Philadelphia: University of Philadelphia Press, 2016], Chapter 5).

condition for the validity of our judgments, and bridge the gap between action and judgment. This idea, Arendt claims, is both a principle for action and a principle for judgment.

These brief remarks on the role of the idea of mankind for action and judgment bring us close to Kant's analysis of the judgment of the sublime. Indeed, as we will see further down, it is in this kind of judgment that principles of action and principles of judgment relate to one another, producing a judgment upon action that can legitimately claim universal approval. The sublime provides an alternative response to the relationship between aesthetic judgment and moral ideas to the ones proposed by Beiner, Benhabib, and Dostal.[33] While their interpretations subordinate the aesthetic judgment to the moral judgment, Kant shows that it is possible to relate the two kinds of judgment without subordinating one to the other. In this regard, Kant is as distrustful as Arendt of the application of moral foundations to the judgment of action. Unlike Arendt, however, Kant believes that this kind of judgment demands a relation to moral ideas, because they are the source of principles for action. The question is whether it is possible to relate moral ideas to experience without turning the imagination into an instrument with which to apply them, thus subordinating the aesthetic judgment to the moral judgment. In addressing this question by means of the judgment of the sublime, I depart from Arendt's strict separation between aesthetic and moral judgments, while avoiding the tendency to re-ground the former upon the latter present in Beiner's, Benhabib's and Dostal's approaches. Kant, as we will see, had both theoretical and practical reasons for avoiding this seemingly "Kantian" solution, which

[33] Given Beiner's reading of the sublime as an attempt to hold an "anthropological narcissism," in which nature is interpreted as having the human being as its end, it is not surprising that he neglects its importance for the problem of judging on the basis on moral ideas ("Kant, the Sublime, and Nature," in *Kant and Political Philosophy: The Contemporary Legacy*, eds. Ronald Beiner and William James Booth [Ann Arbor, MI: Yale UP, 1993]). Beiner simply disregards the fact that human beings are natural objects as well, which is the essential problem of conceiving the possibility of judging on the basis of moral principles, as we will see.

is why he developed an alternative in his analysis of the sublime. But before turning to this analysis, let us consider the problematic relationship between judgment and action in Kant's practical philosophy.

Judgment between the Empirical and the Intelligible

In the context of Kant's practical philosophy, the problem of judging action appears in the discussion of freedom as a cosmological idea in the *Critique of Pure Reason*. In the third antinomy of reason, Kant shows that the idea of freedom is not incompatible with the law of natural causality: as objects of experience, everything that happens is determined by a previous cause in time, while as "things in themselves," that is, as the things that are the source of appearances, freedom is at least possible. We cannot observe a free event in experience, yet for practical purposes, or for the purpose of adjudicating responsibility and judging action on the basis of it, we presuppose its existence. This means that the causality of objects endowed with freedom can be considered "from two sides" [*auf zwei Seiten*]: "as **intelligible** in its **action** [***intelligibel*** *nach ihrer* ***Handlung***] as a thing in itself, and as **sensible** in the **effects** of that action [***sensibel***, *nach den* ***Wirkungen*** *derselben*] as an appearance in the world of sense."[34] The actions of free beings must therefore be considered in a twofold manner, as caused at once by freedom (and so as "actions" properly speaking) and by a natural cause (and so as "effects"). Kant claims that every cause has "a law of its causality," which he calls a "character."[35] A free being has both an "empirical character, through which its actions, as appearances, would stand through and through in connection with other appearances in accordance with constant natural

[34] Immanuel Kant, *Critique of Pure Reason*, eds. and trans. Paul Guyer and Allen W. Wood (Cambridge: Cambridge UP, 1998), 535; A538/B566.
[35] Ibid., 536; A539/B567.

laws," and an "intelligible character, through which it is indeed the cause of those actions as appearances, but which does not stand under any conditions of sensibility and is not itself appearances."[36] The empirical character is constituted by the series of appearances that determine one another according to the law of natural causality, while the intelligible character determines appearances without being itself an appearance. The "law" of the intelligible character, Kant claims further down, is given by imperatives, which determine what a person "ought to" do in accordance with reason, independently of sensible determinations. Yet precisely because it is independent of sensible determination, the intelligible character does not appear—it is a "thing in itself."[37]

The gap between the empirical and the intelligible character is the source of an important difficulty in Kant's understanding of the judgment of action. Kant claims that the intelligible character, which he also calls the "transcendental subject," is "empirically unknown to us." The reason is that "although the effects of this thinking [*Denkens*] and acting [*Handelns*] of the pure understanding are encountered among appearances [*in den Erscheinungen angetroffen werden*], these must nonetheless be able to be explained perfectly from their causes in appearance [*Ursache in der Erscheinung*]… by following its merely empirical character [*bloß empirischen Charakter*]," while "the intelligible character [*intelligiblen Charakter*], which is the transcendental cause [*transzendentale Ursache*] of the former, is passed over as entirely unknown [*gänzlich als unbekannt vorbeigeht*], except insofar as it is indicated [*angegeben*] through the empirical character as only its sensible sign [*sinnliche Zeichen*]."[38] The idea that

[36] Ibid.
[37] Ibid.
[38] Ibid., 539; A546/B574.

empirical causes can work as a "sign" of our noumenal character is intriguing. A possible source for Kant's use of this term is Baumgarten's *Metaphysics*, where he defines a sign as "the means for knowing the existence of another thing."[39] For Baumgarten, however, this "other thing" is located in either the past, the present or the future, which cannot be the case with an intelligible character which, as a thing in itself, is a-temporal. This explains why we do not know the "mode of thought" [*Denkungsart*] of the intelligible character, even if Kant preserves the idea that somehow "it is indicated through appearances" [*wir… bezeichnen sie durch Erscheinungen*; ["we designate it through appearances"].[40] If this were not the case, the power of judgment would have nothing to hold on to in order to judge the morality of action. Yet Kant admits that the "judgment of imputation," which is the judgment concerned with adjudicating merit or blame to an act, cannot ultimately bridge the gap between empirical and intelligible character:

> The real morality of actions [*die eigentliche Moralität der Handlungen*] (their merit and guilt), even that of our own conduct [*unseres eigenen Verhalten*], therefore remains entirely hidden from us [*bleibt uns daher… gänzlich verborgen*]. Our imputations [*Zurechnungen*] can be referred [*bezogen werden*] only to the empirical character. How much of it is to be ascribed [*zuzuschreiben sei*] to mere nature and innocent defects of temperament [*unverschuldeten Fehler des Temperaments*] or to its happy constitution [*glücklicher Beschaffenheit*] (*merito fotunae*) this no one can discover [*ergründen*], and hence no one can judge [*richten*] it with complete justice.[41]

Although empirical acts are the only signs we have of our intelligible character, the latter remains nevertheless "entirely hidden" from us, and so no judgment can be completely just. It is always possible that a seemingly moral action is an effect of luck, rather than of the moral constitution of our intelligible character.

---

[39] Alexander Baumgarten, *Metaphysics*, in *Kant's Critique of Pure Reason: Background Source Materials*, ed. Eric Watkins (New York: Cambridge UP, 2009), 164.
[40] Kant, *Critique of Pure Reason*, 542; A551/B579.
[41] Ibid., 542; A551/B579.

Kant's skepticism regarding the judgment of imputation runs through his main writings on practical philosophy. In the *Groundwork of the Metaphysics of Morals*, Kant claims that the morality of action is determined by its maxim, that is, by the principle of volition on the basis of which it is performed. This means that a same empirical action may be moral or not depending on its underlying maxim. If the maxim motivating a seemingly moral act is compliance with the imperative of reason, or the moral law, it is performed "from duty" and is consequently moral, whereas if the maxim is some sort of self-interest, it is merely "in conformity with duty," and is not moral. Evidently, maxims cannot be observed, which is why it is seemingly impossible to differentiate an action that is performed "from duty" from one that is performed merely "in conformity with duty." Kant suggests that an acute power of judgment is aware of this distinction, and consequently distrustful of the real morality of action:

> One need not be an enemy of virtue but only a cool observer [*kaltblütiger Beobachter*], who does not take [*hält*] the liveliest wish [*lebhaftesten Wunsch*] for the good straight-away as its reality, to become doubtful [*zweifelhaft*] at certain moments (especially with increasing years, when experience has made one's judgment [*Urtheilskraft*; "power of judgment"] partly more shrewd [*gewitzigten*] and partly more acute in observation [*zum Beobachten geschärften*]) whether any true virtue is to be found in the world.[42]

As experience makes our power of judgment "more acute," and insofar as we are capable of separating our observation of actions from our wish for the good, we become doubtful regarding whether virtue even exists in the world. This doubtfulness, Kant believes, plays a positive role for morality insofar as it allows the power of judgment to distinguish its true ground, which is reason, from a spurious ground, which is the subordination of reason to some purpose prescribed by self-interest. Thus while an acute power of judgment shows that seemingly good actions are

[42] Immanuel Kant, *Groundwork of the Metaphysics of Morals*, in *Practical Philosophy*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge UP, 1999), 62; 4:407.

not truly moral if they are not grounded upon a maxim of reason, "reason by itself and independently of all appearances commands what ought to happen."[43] As a consequence, Kant claims, "actions of which the world has perhaps so far given no example [*kein Beispiel gegeben hat*], and whose very practicability [*Thunlichkeit*] might be very much doubted by one who bases everything on experience, are still inflexibly commanded by reason."[44] Insofar as morality is grounded upon reason, it is at odds with experience, and therefore any judgment based on experience cannot but play a negative function—namely, to make us doubtful regarding its correspondence with morality. Even if all that judgments upon experience show is that no moral action has ever existed, and most likely will never exist, reason still commands that it ought to exist.

The tension between experience and morality reappears in the *Critique of Practical Reason*, in the section titled "On the typic of pure practical judgment." This section represents Kant's first systematic attempt to explain the possibility of applying moral concepts (that is, good and evil) to phenomenal actions. The idea of the good is supersensible, which means that it has no corresponding object in experience. Actions, however, are events in the phenomenal world. As natural events, we can judge an action as the effect of a prior cause, because the imagination has a procedure, or what Kant calls a "schema," for applying the category of causality to sensible intuitions. It is important to note that, as Kant explains in the *Critique of Pure Reason*, the imagination plays an essential role in the constitution of objects of experience, because it mediates between concepts, which are given by the understanding, and intuitions, which are given by sensibility. The schema is the procedure by which the imagination applies a concept to

[43] Ibid.
[44] Ibid., 4:408.

intuitions—so that, for example, the concept of cause is applied to a concrete event that is then judged as the empirical cause of a certain effect. In the case of moral ideas, however, the imagination lacks a procedure, for "no intuition can be put under the law of freedom… and hence no schema on behalf of its application [*Anwendung*] *in concreto*."[45] The law of freedom (the moral law) cannot be applied to experience, because there is no intuition that corresponds to it. The question, then, is how a practical judgment is possible at all, namely, how it is possible to apply the ideas of good and evil to empirical actions.

In the absence of a schema guiding the imagination, Kant locates the possibility of applying ideas to concrete actions in what he calls a "type." While a schema mediates between concepts of the understanding and intuitions, the type mediates between ideas of reason and what ought to happen in intuition. The practical judgment is not concerned with what is, but rather with what ought to be, for "to appraise [*die Beurtheilung*; "the judging"] whether or not something is an object of *pure* practical reason is only to distinguish the possibility or impossibility of *willing* [wollen] the action by which… a certain object would be made real."[46] In order to know whether it is possible to will a concrete action on the basis of the moral law, we need a procedure by which we can apply this law to experience. Kant claims that this application is not performed by the imagination (through a schema) but rather by the understanding: "what the understanding can put under an idea of reason [*Idee der Vernunft*] is not a *schema* of sensibility [schema *der Sinnlichkeit*] but a law, such a law, however, as can be presented *in concreto* in objects of the senses [*Gegentänden der Sinnen*] and hence a law of nature, though only as to its form."[47] The

[45] Immanuel Kant, *Critique of Practical Reason*, in *Practical Philosophy*, 196; 5:69.
[46] Ibid., 186; 5:57.
[47] Ibid., 196; 5:69.

form of the moral law as can be presented in objects of experience is what Kant calls the "type of the moral law," and it is given by the law of nature. While the moral law determines what ought to be, independently of experience, the law of nature determines what is given in experience. In order to make the moral law applicable to experience, the understanding provides reason with the form of the law of nature so as to assess whether a certain action is in accordance with the moral law. This produces the formula of the practical judgment: "ask yourself whether, if the action you propose were to take place by a law of the nature of which you were yourself a part, you could indeed regard it as possible through your will."[48] In order to judge whether an action is morally good or not, that is, whether it is a case of an action under the moral law, the power of judgment borrows the form of the law from the understanding, and then asks whether we could will that this law exist—we thus see that, for example, we would not want to live in a world in which deception for the sake of self-interest were a law of nature, for we could not want to be deceived whenever it is in someone else's interest.

The analysis of the practical judgment shows that the power of judgment plays a role in the determination of action, but it does not address the possibility of judging existing actions, or what Kant calls "judgments of imputation" in the first *Critique*. Although the formula of the practical judgment is not entirely clear in this regard, Kant states that its object is not the action as such, but rather its underlying maxim, that is, the principle on the basis of which it is performed: "if the maxim of the action [*Maxime der Handlung*] is not so constituted [*beschaffen*] that is can stand the test as to the form of a law of nature in general, then it is morally impossible

[48] Ibid., 196; 5:69.

[*sittlich unmöglich*]."[49] The distinction, often neglected by commentators, is important because the maxim of an action is not observable.[50] According to Kant, every action has an underlying maxim, which is the general principle by which we decide what to do. A moral action is one whose underlying maxim is compliance with duty, while an evil action is one whose underlying maxim is the satisfaction of one's own appetites. But because maxims are not observable, we cannot unequivocally judge whether an action is good or evil. The practical judgment tells us what the form of a good maxim is, but it does not provide any means to assess the moral worth of an empirical action. Therefore, it is a mistake to read the type of the moral law, as Felicitas Munzel does, as a mediation between intelligible and empirical character, thus neglecting the difference between the type and the schema, and thus also between the imagination (which relates to intuitions) and the understanding (which relates to rules).[51] As Henry Allison points out, the practical judgment determines types of action, not particular actions as they appear in experience.[52]

The seeming impossibility of judging the morality of action becomes an important concern in *Religion within the Boundaries of Mere Reason*, in connection to the problem of evil. Kant explicitly states in this text that the maxim of an action is beyond observation, which is why "the judgment [*Urtheil*] that an agent [*Thäter*] is an evil human being cannot reliably be based on

---

[49] Ibid., 196; 5:69-70.

[50] Lewis White Beck's otherwise thorough analysis of the practical judgment says little about the distinction between action and maxim, and assumes that the type bridges the gap between what ought to be and what is (*A Commentary on Kant's* Critique of Practical Reason [Chicago: The University of Chicago Press, 1984], 157). This is at least problematic, if we consider that maxims for action do not clearly belong to either of the two realms.

[51] G. Felicitas Munzel, *Kant's Conception of Moral Character: The "Critical" Link of Morality, Anthropology, and Reflective Judgment* (Chicago: The University of Chicago Press, 1999), 126-132. Given Munzel's reading of the aesthetic judgment as one step in Kant's consistent view of the empirical character as cultivated by and expressive of the intelligible character, it is unsurprising that her analysis gives little room to the sublime.

[52] Henry E. Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge UP, 1990), 90.

experience [*nicht mit Sicherheit auf Erfahrung gründen*; "one cannot ground it upon experience with certainty"]."[53] This is not merely a theoretical problem, but a moral problem in its own right. According to Kant, the difficulties that the power of judgment encounters when assessing the morality of action can be manipulated so as to deceive oneself and others. An empirically good action ("in conformity with duty") can be used to conceal a maxim that is not moral—as when, for example, a person does not lie out of a concern with his own reputation, as opposed to compliance with moral duty. Kant links this deception to a form of "perfidy" [*Tücke*] which consists in "the human heart… deceiving itself [*selbst zu betrügen*] as regards its own good or evil disposition [*guten oder bösen Gesinnungen*] and, provided that its actions do not result in evil (which they could well do because of their maxims), in not troubling itself [*sich… nicht zu beunruhigen*] on account of its disposition but rather considering itself justified [*sich… gerechtfertigt zu halten*] before the law."[54] An empirically good action can be motivated by an evil maxim, in which case it is morally evil. Given that this maxim is inscrutable, we can distort our judgment so as to conceal an evil disposition under the guise of empirically good actions. This is one of the effects of what Kant calls "radical evil," which is the propensity to subordinate compliance with moral duty to self-interest as the supreme incentive in the maxims of our actions. Because this inversion of the incentives leads most of the times to empirically good actions (it is usually in our own self-interest to do what is good), radical evil "puts out of tune [*verstimmt*] the moral ability to judge [*moralische Urtheilskraft*; "moral power of judgment"] what to think of a human being, and renders any imputability [*die Zurechnung*; "the imputation"]

[53] Immanuel Kant, *Religion within the Boundaries of Mere Reason*, trans. George di Giovanni, in *Religion and Rational Theology*, eds. Allen W. Wood and George di Giovanni (Cambridge: Cambridge UP, 2001), 70; 6:20.
[54] Ibid., 84; 6:38.

entirely uncertain [*ganz Ungewiß macht*], whether internal or external."[55] The theoretical problem of judging the morality of action thus becomes a moral problem in its own right, for it is in our own self-interest to deceive the power of judgment so as to confuse actions "from duty" with action that are merely "in conformity with duty." And given that our moral disposition is beyond observation, there would seem to be no way to overcome the evil propensity to deceit, and thus the conclusion that judgments of imputation are impossible.

Kant addresses this problem by mentioning two ways to counter the evil propensity to deceit and to put the power of judgment "out of tune," as part of his reflections on moral education. The first one, which further develops his earlier insight in the *Groundwork*, is the use of examples in a negative way, namely, as a means to train the power of judgment to detect the impurity in the incentives of seemingly good actions, thus "allowing our apprentices in morality to judge [*beurtheilen*] the impurity of certain maxims [*Unlauterkeit mancher Maximen*] on the basis of the incentives actually behind their actions [*aus den wirklichen Triebfedern ihrer Handlungen;* "from the real incentives of their actions"]."[56] The judgment that apparently moral actions stem from impure incentives gradually becomes an "attitude of mind," according to which "*duty* merely for itself begins to acquire in the apprentice's heart a noticeable importance." In this negative use of judgment, the person becomes aware that duty is important "merely for itself" by separating it from impure incentives. According to Kant, "even the most limited human being is capable of all the greater a respect for a dutiful action [*größeren Achtung für eine pflichtmäßige Handlung*] the more he removes [*entzieht*] from it, in thought, other incentives which might have

[55] Ibid., 85; 6:38.
[56] Ibid., 93; 6:48.

influence upon its maxim through self-love [*durch die Selbstliebe*]."[57] By judging the impurity behind virtuous actions, we acquire respect for the idea of acting from duty. And given that an action from duty is merely an idea, we can only use examples negatively, namely, as a means to distinguish this idea from non-moral incentives.

The second way to counter the evil propensity to deceive the power of judgment is positive and more enigmatic. Kant claims that "there is one thing in our soul [*eines ist in unserer Seele*] which, if we duly fix our eye on it [*wenn wir es gehörig ins Auge fassen*], we cannot cease viewing with the highest wonder [*mit der höchsten Verwunderung zu betrachten*], and for which admiration [*Bewunderung*] is legitimate and uplifting [*seelenerhebend*; "elevating of the soul"] as well. And that is the original moral predisposition in us, as such [*die ursprüngliche moralische Anlage in uns überhaupt*]."[58] The idea of "fixing our eye" upon "the original moral predisposition in us" is intriguing, since such predisposition is beyond observation. Kant adds that

> The very incomprehensibility [*Unbegreiflichkeit*; impossibility to conceptualize] of this predisposition, proclaiming as it does a divine origin, must have an effect on the mind, even to the point of exaltation [*muß auf das Gemüth bis zur Begeisterung wirken*], and must strengthen it for the sacrifices [*Aufopferungen*] which respect for duty [*Achtung für seine Pflicht*] may perhaps impose upon it. Often to arouse [*rege zu machen*] this feeling of the sublimity of our moral vocation [*dieses Gefühl der Erhabenheit seiner moralischen Bestimmung*] is especially praiseworthy [*vorzüglich anzupreisen*] as a means of awakening moral dispositions [*Mittel der Erweckung sittlicher Gesinnungen*], since it directly counters [*entgegen wirkt*] the innate propensity to pervert [*Hange zur Verkehrung*] the incentives in the maxims of our power of choice.[59]

[57] Ibid., 92-93; 6:48.
[58] Ibid., 93; 6:49.
[59] Ibid., 93-94; 6:50.

Arousing a feeling of the sublimity of our moral vocation is a means to counter the propensity to pervert the incentives for action, or what Kant calls "radical evil." The question, then, is what arouses this feeling of sublimity, considering Kant's repeated assertion that nothing in experience corresponds to moral ideas. This is a central issue in Kant's "Analytic of the Sublime" in the third *Critique*.

A Sublime Judgment

The relationship between the sublime and Kant's practical philosophy has been the subject of debate in recent years. One problematic aspect is that while Kant presents the judgment of the sublime as a kind of aesthetic judgment, its interrelation with moral categories seems to put its purely aesthetic nature into question. In a classical study, Paul Crowther argues that Kant provides insufficient grounds for distinguishing the aesthetic feeling of the sublime from the moral feeling of respect for the moral law, and develops an alternative formulation of the sublime that accounts for its aesthetic (by contrast to moral) nature.[60] While more recent studies by Patricia M. Matthews, Henry E. Allison, and Rodolphe Gasché have explicitly or implicitly challenged Crowther's criticism of Kant, they share the concern with emphasizing the aesthetic, as opposed to moral character of the judgment of the sublime.[61] However, in their attempts to make Kant's text amenable to a stark distinction between aesthetic and moral judgments, these studies give little or no room to Kant's discussion of human actions and affects as cases of sublimity, with the consequence that the relationship between sublimity and morality is not

[60] Paul Crowther, *The Kantian Sublime: From Morality to Art* (Oxford: Clarendon Press, 1989), 131-133.
[61] Patricia M. Matthews, "Kant's Sublime: A Form of Pure Aesthetic Reflective Judgment," *The Journal of Aesthetics and Art Criticism* 54(2); Henry E. Allison, *Kant's Theory of Taste: A Reading of the* Critique of Aesthetic Judgment (Cambridge: Cambridge UP, 2001), Chapter 13; Rodolphe Gasché, *The Idea of Form: Rethinking Kant's Aesthetics* (Stanford, CA: Stanford UP, 2002), Chapter 5.

subject to a systematic scrutiny. A recent study by Robert R. Clewis addresses this issue by focusing on the relationship between the sublime, morality, and enthusiasm, stressing the importance of the sublime for the experience of moral ideas.[62] My reading of Kant's analysis of the sublime shares Clewis's approach, although with disagreements that I will address further down. I follow more closely Milton Nahm's suggestion (in an essay from 1956) that the sublime represents a response to the problematic status of the judgment upon action in Kant, in part because it concerns the relationship between the two faculties that must be involved in a judgment of this kind: imagination and reason.[63] By exploring the relationship between these two faculties in the judgment of action, which will be the focus on my reading, we can better understand the practical implications of the sublime.

One of the main innovations of Kant's conception of the aesthetic judgment in the third *Critique* is the role of the imagination. We have seen that in practical judgments the imagination plays no role whatsoever. Moral ideas belong to reason, and nothing corresponds to them in experience, which means that there is no possible presentation of a good action, or an action that corresponds with the idea of the moral law. In cognitive judgments, the imagination plays a subordinate role: the understanding provides the concepts with which to constitute objects of experience, and the imagination gives form to intuitions in accordance to these concepts. In the aesthetic judgment, by contrast, the imagination becomes relatively autonomous in relation to the other faculties. The reason is that, like in a cognitive judgment, the aesthetic judgment is

---

[62] Robert R. Clewis, *The Kantian Sublime and the Revelation of Freedom* (Cambridge: Cambridge UP, 2009).

[63] Milton C. Nahm, "'Sublimity' and the 'Moral Law' in Kant's Philosophy," *Kant-Studien* 48. Zammito explains that the reason behind Kant's "ethical turn" in the composition of the third *Critique*, which led to the inclusion of the "Analytic of the Sublime," "had to do with his desire to reconcile his phenomena-noumena theory of freedom with the problem of actualizing the moral good. He needed, or felt he needed, to reformulate and strengthen the analogue to schematism which he had developed in the *Second Critique*" (*The Genesis of Kant's* Critique of Judgment, 264).

concerned with objects given in experience, but unlike the former, there is no concept under which to subsume them. As a consequence,

> the power of judgment, which has no concept ready for the given intuition, holds the imagination (merely in the apprehension of the object) together [*zusammenhält*] with the understanding (in the presentation [*Darstellung*] of a concept in general [*eines Bregriffs überhaupt*]) and perceives [*wahrnimmt*] a relation of the two faculties of cognition which constitutes the subjective, merely sensitive [*empfindbare*] condition of the objective use of the power of judgment in general (namely the agreement [*Zusammenstimmung*] of those two faculties with each other).[64]

The description corresponds to the judgment of taste, rather than the sublime. The important point is that the role of the imagination is not to apprehend an object on the basis of a determined concept, but rather on the basis on of a "concept in general." As a consequence, imagination and understanding reach an agreement without either one of the two subordinating the other: the imagination attempts to give form to intuitions as if there were a concept, and the understanding provides "a concept in general," or a concept without a determinate content. This agreement is the "subjective condition" of the objective use of the power of judgment, that is, of the determinative or cognitive judgment. Cognitive judgments produce the agreement between the two faculties by subordinating the imagination to the concepts of the understanding. In the aesthetic judgment, by contrast, there is no concept ready to guide the imagination, but only "a concept in general." Later on, in the second Introduction to the *Critique*, Kant describes the agreement between the imagination and the understanding as a "free play," in which "no determinate concept restricts them to a particular rule of cognition."[65] Given that there is no determinate concept ruling the imagination, yet the imagination strives to give intuitions a certain

[64] Kant, *Critique of the Power of Judgment*, 26; 20:223-224.
[65] Ibid., 102; 5:217.

form, it must establish a play with the understanding without the latter providing a rule on how to proceed.

The agreement between the faculties explains why the judgment of taste produces a pleasure that demands the agreement of others. Indeed, a second innovation in Kant's analysis of aesthetic judgment concerns the role of the faculty of pleasure and displeasure, which until that point was considered to be entirely private, in the sense of depending on purely subjective conditions. In the third *Critique*, Kant distinguishes between the "agreeable" and the "beautiful": while the former pleases our inclinations and appetites, the latter pleases by virtue of the "state of mind" [*Gemüthszustand*] that it produces. Given that this state of mind is an effect of the accord between our cognitive faculties, which are shared by all human beings, the pleasure occasioned by the beautiful demands universal agreement: "it is the universal capacity for the communication of the state of mind [*die allgemeine Mittheilungsfähigkeit des Gemüthszustandes*] in the given representation which, as the subjective condition of the judgment of taste, must serve as its ground and have the pleasure in the object as a consequence."[66] When we judge an object as beautiful, we are not merely saying that it pleases us individually, but rather that it pleases us by virtue of producing a state of mind that we can expect to be shared by others who contemplate the same object. This state of mind is the same as the one reached in cognitive judgments, with the difference that in the aesthetic judgment there is no determinate concept regulating the agreement. Therefore, there is no a priori rule determining whether an aesthetic judgment is communicable or not, and thus no "proof" regarding its

[66] Ibid.

validity.[67] However, it is possible to compel the agreements of others on the basis of the state of mind produced in us by the object.

The judgment of the sublime resembles the judgment of taste in that the imagination relates to indeterminate concepts, but in this case they are concepts of reason. This introduces an important difference in the role of the imagination, as well as in the structure of its relationship with the other faculty involved. Concepts of the understanding are concepts of experience, that is, concepts whose sole purpose is to give form to intuitions. By contrast, concepts of reason, or what Kant more commonly calls "ideas," have no experiential content. Rational ideas like the world as a whole (the totality of objects in the infinity of space and time) or freedom (a cause that is not in its turn the effect of previous cause in time), do not have a corresponding object in intuition. Therefore, while "the beautiful in nature concerns the form of the object, which consists in limitation [*Begränzung*]… the sublime, by contrast, is to be found in a formless [*formlosen*] object insofar as limitlessness [*Unbegränzheit*] is represented in it, or at its instance [*durch dessen Veranlassung*], and yet it is also thought [*hinzugedacht wird*] as a totality."[68] Because concepts of the understanding give form to intuitions, the free play between the imagination and the understanding involves a limited, yet undetermined form. Ideas of reason, by contrast, have no form, and so the imagination can only relate to reason by encountering objects that are formless, and consequently allow for limitlessness to be represented in them. Evidently, there is no such thing as a formless object in intuition. However, as we will see, there are objects that challenge the imagination's capacity to give form to intuitions. It is these objects that awaken the sublime feeling.

[67] Ibid., 216; 5:340.
[68] Ibid., 128; 5:244.

Kant analyzes two kinds of sublime: mathematical and dynamical. While it is in the analysis of the latter that the relationship between judgment and action comes to the fore, the explanation of the mathematically sublime lays out the foundations for understanding the judgment of the sublime in general. The mathematically sublime is that which is "absolutely great," that is, great beyond comparison. Kant suggests that this is the meaning of the word "great" in a judgment like "the man is great," 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."[69] Greatness, however, is not a concept for determining an object, but rather a "comparative concept" [*Vergleichungsbegriff*]. The judgment of magnitude requires a unit of measurement for determining how great something is, and this unit requires in turn something else to be compared with—a kilometer is great in comparison with a meter, but small in comparison with a mile. The question, then, is what in the contemplation of the object produces the idea of greatness independently of comparison. On this point, Kant claims:

> Nothing that can be an object of the senses [*Gegenstand der Sinnen*] is… to be called sublime. But just because there is in our imagination a striving to advance to the infinite [*ein Bestreben zum Fortschritte ins Unendliche*], while in our reason there lies a claim [*Anspruch*] to absolute totality, as to a real idea, the very inadequacy [*Unangemessenheit*] of our faculty for estimating the magnitude of the things of the sensible world [*Dinge der Sinnenwelt*] awakens [*ist… die Erweckung*] the feeling of a supersensible faculty [*übersinnliches Vermögens*] in us; and the use that the power of judgment naturally makes in behalf of the latter [*zum Behuf des letzteren*] (feeling), though not the object of the senses, is absolutely great [*schlechthin groß*], while in contrast to it any other use is small.[70]

The feeling of the sublime is awakened by the inadequacy between imagination and reason. The imagination strives to present the infinite on behalf of reason's claim to absolute totality. But this

[69] Ibid., 132; 5:248.
[70] Ibid., 134; 5:250.

is impossible, for there is no possible presentation of infinity. Yet in its very striving for the infinite, the imagination awakens the feeling of a supersensible faculty in us, which is 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 is occasioned by objects that challenge the imagination's capacity to estimate their magnitude. Kant explains that this estimation depends on two operations: apprehension and comprehension. In apprehension, the imagination presents an individual sensuous perception to the mind. In comprehension, it retains the series of presentations so as to form an object. The sublime is awakened by a disruption in comprehension,

> for when apprehension [*Auffassung*] has gone so far that the partial representations [*Theilvorstelungen*] of the intuition of the senses [*Sinnenanschauung*] that were apprehended first already begin to fade [*erlöschen*] in the imagination as the latter proceeds [*fortrückt*] on to the apprehension of further ones, then it loses on one side as much as it gains on the other, and there is in the comprehension [*Zusammenfassung*] a greatest point [*ein Größtes*] beyond which it cannot go.[71]

In the presence of certain objects, the imagination finds itself unable to comprehend all partial apprehensions. Kant's examples 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 these objects, it encounters difficulties retaining apprehensions as it must proceed to new ones. Ultimately, given that the imagination is finite, it encounters a point beyond which it cannot go, and comprehension fails. In other words, the imagination fails to give the object a form, which thus becomes "formless."

[71] Ibid., 135; 5:252.

The failure of the imagination in representing the object produces the thought of infinity, thus arousing the feeling of respect. Reason is a supersensible faculty, which means that it lies beyond the limitation of the senses. Consequently, in its demand to present all the elements in a series, it "does not exempt [*nicht ausnimmt*] from this requirement [*Forderung*] even the infinite (space and past [*verflossene*] time), but rather makes it unavoidable [*unvermeindlich macht*] for us to think of it [*sich dasselbe… zu denken*; "that it be thought"] (in the judgment of common reason [*Urtheile der gemeinen Vernunft*]) as **given entirely** [*ganz… gegeben*] (in its totality [*seiner Totalität nach*])."[72] In its relation with the imagination, reason demands what is impossible for the latter, namely, to comprehend the infinity of space and time. It is the failure of comprehension produced by this demand that makes it unavoidable for us to think of this totality, even if it cannot be represented. Thus what is a failure from the viewpoint of presentation becomes a victory from the viewpoint of thinking: we cannot present the totality of space and time, but the very attempt to do so shows that we are capable of thinking it. It is this capacity that awakens the feeling of "respect" [*Achtung*], which Kant defines as "the feeling of the inadequacy [*Unangemessenheit*] of our capacity for the attainment [*Erreichung*] of an idea **that is a law for us**."[73] In its effort to present an idea of reason, the imagination meets its inadequacy, but "at the same time its vocation [*Bestimmung*] for adequately realizing that idea as a law," which shows that "the feeling of the sublime in nature is respect for our own vocation [*Achtung für unsere eigene Bestimmung*], which we show to an object in nature through a certain subreption [*Supreption*] (substitution of a respect for the object instead of for the idea of humanity [*Idee der*

[72] Ibid., 138; 5:254.
[73] Ibid., 140; 5:257.

Menschheit] in our subject)."[74] Although the failure of the imagination is felt as displeasure, it also "arouses [*rege macht*] 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 [*Ideen der Vernunft*; "ideas of reason"]."[75] Therefore, even in their contrast, imagination and reason find a sort of harmony in disharmony, for the latter is felt as having a purpose, namely, to arouse the feeling of the supersensible vocation of our mind.

Before proceeding to the implications of this analysis for the judgment of action, let us note one important contrast between the feeling of the sublime and the feeling of the beautiful. The feeling of the beautiful, as we have seen, stems from the harmony achieved in the interaction between the imagination and the understanding. The feeling of the sublime is more complex. On the one hand, the judgment of the sublime involves harmony between the imagination and reason, even though this harmony is mediated by disharmony. But the feeling of the sublime is not simply the feeling of this harmony—according to Kant, 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 the voice 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."[76] Commentators who emphasize the continuities between the beautiful and the sublime tend to leave aside these and other descriptions of the sublime feeling as dependent on the (moral) feeling of respect, and overlook that the former,

[74] Ibid., 141; 5:257.
[75] Ibid.
[76] Ibid., 172; 5:292.

unlike the feeling of pleasure in the beautiful, does not result from the harmony between the faculties.[77] Therefore, I believe that Paul Guyer is right in describing the harmony between the faculties in the sublime as a "psychological state" (although it is rather a "state of mind," a *Gemütszustand*) that characterizes the aesthetic judgment in general, which does not imply that the phenomenological structure of the sublime feeling or the source of the claim to universal approval of the judgment (which I will address in the next section) is the same in both kinds of judgment.[78] Regarding these two elements, Kant is clear that the feeling of the sublime, unlike the beautiful, depends on our capacity to feel the supersensible vocation of our mind, which is why it has a moral foundation.[79] It is this interrelation between the sublime feeling and the feeling for rational ideas that comes to the fore in the analysis of the dynamically sublime.

A Supersensible Power

Kant comes closer to the analysis of the judgment of action in his analysis of the dynamically sublime and in the general remarks that follow it. While the mathematically sublime is related to the faculty of cognition, the dynamically sublime is related to the faculty of the

[77] Patricia Matthews claims that "the feeling of sublimity… is a feeling based on the relationship of the faculties of reason and imagination, not a direct response to reason" ("Kant's Sublime: A Form of Pure Aesthetic Reflective Judgment," 172), a definition that makes the feeling of the sublime similar to the feeling of the beautiful, but leaves aside Kant's own description of the former as a feeling for the supersensible vocation of our mind. A similar neglect for this difference between the feeling of the beautiful and the feeling of the sublime is present in Rudolf Makkreel's claims that "reflection on beauty leads us to hope for a greater harmony and systematic order in nature; the sublime points to the possibility of an overall integration of our faculties of mind" (*Imagination and Interpretation in Kant: The Hermeneutical Import of the* Critique of Judgment [Chicago: The University of Chicago Press, 1994], 83).

[78] Paul Guyer, "Kant's Distinction between the Beautiful and the Sublime," *The Review of Metaphysics* 35(4), especially p. 766. A similar point, with a greater emphasis on the contrast between the beautiful and the sublime, is made by Eva Schaper in "Taste, sublimity, and genius: The aesthetics of nature and art," in *The Cambridge Companion to Kant*, ed. Paul Guyer (Cambridge: Cambridge UP, 2006).

[79] The feeling of the sublime, according to Peter Fenves, needs a "supplement," which explains why Kant presents the "Analytic of the Sublime" as a "mere appendix" to the *Critique of the Aesthetic Power of Judgment* ("Taking Stock of the Kantian Sublime," *Eighteenth-Century Studies* 28[1], 78).

desire, which is the faculty that determines action.[80] It is therefore no coincidence that in the analysis of the dynamically sublime human actions and affects are introduced as examples of sublimity. This does not mean, however, that actions belong to a specific kind of sublime, as Crowther and Clewis have proposed in different ways. Rather, the analysis of the sublime in general shows that a judgment of action is possible. This is evident if we consider that already in the analysis of the mathematically sublime Kant mentions "the man is great" as an example of a judgment of the sublime, and refers to "the magnitude of a certain virtue" as one possible object of this kind of judgment.[81] The view that the judgment of actions and affects constitutes a special kind of sublime not only departs from Kant's text, but also, and more importantly, neglects the tension between the aesthetic and the moral components of the judgment of the sublime, which will be the focus of my reading in what follows.

While the mathematically sublime emerges out of the attempt to represent magnitude, the dynamically sublime emerges out of the attempt to represent power. Kant introduces this idea as follows:

> **Power** [*Macht*] is a capacity that is superior to great obstacles. The same thing is called **dominion** [*Gewalt*] 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.**[82]

In the dynamically sublime, nature is considered as power, yet as not powerful enough to have dominion over us. As a consequence, the dynamically sublime arouses admiration for our own power over nature. This is an effect 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

[80] Kant, *Critique of the Power of Judgment*, 131; 5:248.
[81] Ibid., 133; 5:249.
[82] Ibid., 143; 5:260.

were threatening cliffs, thunder clouds towering up into the heavens, bringing with them flashes of lightening and crashes of thunder, volcanoes with their all-destroying violence…, make our capacity to resist [*Vermögen zu widerstehen*] into an insignificant trifle [*unbedeutenden Kleinigkeit*] in comparison with their power."[83] Even if these objects are fearful, Kant claims, "the sight of them [*ihr Anblick*] only becomes all the more attractive [*anziehender*] the more fearful [*furchtbarer*] it is," given that "they elevate [*erhöhen*] the strength of our soul [*Seelenstärke*] above its usual level, and allow us to discover within ourselves a capacity for resistance of quite another kind, which gives us the courage [*Muth*] to measure ourselves against the apparent all-powerfulness [*Allgewalt*] of nature."[84] In contemplating a violent volcano, we experience a power that makes our capacity to resist insignificant. Yet even if we are afraid of it, we can "**think of** the case in which we might wish to resist it [*ihm etwa Widerstand thun wollten*]."[85] And even if resisting the power of a volcano would be futile, the very thought of this resistance elevates our soul to experience a different kind of resistance—one that does not depend on physical power. The thought that we may want to resist nature, with its apparent all-powerfulness, shows that there is a power in us of a different kind, unconstrained by the boundaries of physical power.

The contrast between nature and its resistance resembles Kant's contrast between our empirical and our intelligible character. We have seen that while our empirical character is subject to the law of nature, our intelligible character is independent from it. Therefore, just like

[83] Ibid., 144; 5:261.
[84] Ibid., 144-145; 5:261.
[85] Ibid., 143; 5:260.

the mathematically sublime revealed a superiority of reason over nature by virtue of the latter's incapacity to match its ideas,

> likewise the irresistibility of its power certainly makes us, considered as natural beings, recognize our physical powerlessness, but at the same time it reveals [*entdeckt*] a capacity for judging ourselves as independent of it and a superiority over nature [*Überlegenheit über die Natur*] on which is grounded a self-preservation [*Selbsterhaltung*] of quite another kind than that which can be threatened [*angefochten*] and endangered [*in Gefahr gebracht*] by nature outside us, whereby the humanity in our person remains undemeaned [*die Menschheit in unserer Person unerniedrigt bleibt*] even though the human being must submit to that dominion.[86]

While our self-preservation can always be threatened by nature "outside of us," there is a self-preservation of another kind that remains undemeaned 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 are universally communicable.[87] By raising the imagination to a case in which we might resist nature in all its power, we 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," [*höchste Grundsätze*] which stem from the moral law.[88] Somewhat paradoxically, the irresistibility of nature "outside us" rises the imagination to a power over nature inside us, because no natural incentive is powerful enough to make us disregard the moral law as an object of admiration. As with the mathematically sublime, every standard for assessing the value of a natural incentive for action is trivial in comparison with the moral law, which calls us to resist such incentives.

---

[86] Ibid., 145; 5:261-262.
[87] I follow here Kant's definition of humanity in the third *Critique*, p. 229; 5:355.
[88] Ibid., 145; 5:262.

The connection between the representation of nature and the idea of a resistance to it discloses a new possibility for the power of judgment, which is crucial for the possibility of relating it to actions. We have seen that all objects of experience, including human deeds, are ruled by the law of natural causality. This means that as objects of experience all actions are determined by the inclinations, and so actions performed out of moral duty simply do not exist in experience. However, if there are objects which, by virtue of their resistance to our capacity to represent them as a purely natural power, arouse the feeling of a vocation of the mind that is beyond natural necessitation, then a new possibility opens up for the representation of action. As objects of experience, all actions are ruled by the law of natural causality. But there are actions that challenge our ability to represent them as a power of nature, and thus arouse the idea of a supersensible power. Kant mentions the warrior as "an object of the greatest admiration [*grösten Bewunderung*]," whom aesthetic judgment determines as deserving respect.[89] By appearing to sacrifice all inclinations and putting all self-interest at risk, the warrior surpasses our capacity to represent his actions as determined by nature, that is, by his inclinations, thus arousing the idea of a supersensible power. Therefore, Kant claims, the mentality of the people conducting war is "all the more sublime, the more dangers it has been exposed to and before which it has been able to assert its courage."[90]

Kant's reference to human actions seem to depart from his own definition of the dynamically sublime as a power that stretches the imagination to its limit, which has led commentators to cast it as a new type of sublime. According to Crowther, the example of the warrior constitutes "a subcategory of dynamic sublimity whose origin is to be found in our

[89] Ibid., 146; 5:262
[90] Ibid., 146; 5:263.

affective response to rationally significant *human deeds*."[91] This affective response does not consist in stretching the imagination to the limit of what it can present, as Kant defines the dynamically sublime, but rather in leading the imagination "to envisage possibilities of moral action."[92] In a similar vein, Clewis claims that the references to actions constitute a "moral sublime" that is different from both the mathematically and the dynamically sublime, and which consists in "a response to the moral law (or a representation or embodiment thereof) and more directly reveals the human capacity for morality."[93] Both interpretations neglect the fact that there is no such thing as an action that embodies or represents compliance with duty, and there is no reason to assume that Kant is being inconsistent on this point. An action that overcomes great obstacles of sensibility is not an example of moral action, but rather an object that, by virtue of surpassing the capacity of the imagination to represent it as natural power, arouses 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 great 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)."[94]

This consideration is important for understanding the precise sense in which a judgment of action is possible. A morally good action, as we have seen, is an action performed on the basis of compliance with the moral law, which is an idea of reason that lacks any possible representation

---

[91] Crowther, *The Kantian Sublime*, 116. Crowther's misreading of the role of deeds in Kant's analysis of the dynamically sublime in an example of his tendency to quickly assume breaks and inconsistencies both within this analysis and in its relationship with Kant's broader conception of the sublime—which is curious given his claim that "after the labyrinthine intricacies of the mathematical mode, Kant's discussion of the dynamical sublime comes as something of a relief, in so far as it contains one of the clearest progressions of argument in any one section of the third *Critique*" (108).
[92] Ibid., 118.
[93] Clewis, *The Kantian Sublime and the Revelation of Freedom*, 87.
[94] Kant, *Critique of the Power of Judgment*, 147; 5:264.

in experience. In the second *Critique,* Kant explain that we are aware of the moral law through the feeling of respect, which "is produced by a purely intellectual ground," and is therefore independent of sensuous perceptions or sensations.[95] The feeling of the sublime is aesthetic, which means that it is connected to sensuous perceptions, but is also dependent of ideas of reason, for otherwise the imagination would not feel the demand to overstep its own limitations. This is why Kant claims that the feeling of the sublime has its foundation "in the predisposition [*Anlage*] to the feeling for (practical) ideas, i.e., to that which is moral," which "is the ground for the necessity of the assent [*Beistimmung*; "approval"] of the judgment of other people concerning the sublime to our own."[96] In order to judge an object as sublime, we need to feel respect for the moral law, for otherwise we would not experience the failure of the imagination in the representation of an object as serving a purpose for the mind that is universally communicable. Consequently, while on the one hand the feeling of the sublime is awakened by an object that challenges our capacity to represent it, on the other hand it requires the feeling of respect. This means that the representation of the morally good has both an aesthetic and an intellectual "side," as Kant explains in the following passage:

> The object of a pure and unconditioned intellectual satisfaction [*reinen und unbedingten intellectuellen Wohlgefallens*] is the moral law in all its power [*in seiner Macht*], 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 [*ästhetisch kenntlich*] only through sacrifices [*Aufopferungen*]…, the satisfaction on the aesthetic side [*von der ästhetischen Seite*] (in relation to sensibility) is negative, i.e., contrary to this interest, but considered from the intellectual side [*von der intellectuellen*] it is positive, and combined with an interest. From this it follows that the intellectual, intrinsically purposive (moral) good, judged aesthetically [*ästhetisch beurtheilt*], must not be represented

[95] Kant, *Critique of Practical Reason,* 199-200; 5:73.
[96] *Critique of the Power of Judgment,* 149; 5:265.

> [*vorgestellt*] so much as beautiful but rather as sublime, so that it arouses [*erwecke*; "awakens"]… the feeling of respect (which scorns charm).[97]

Although the moral law is the object of a purely intellectual satisfaction through the feeling of respect, its power is only aesthetically knowable through sacrifices, that is, through the overcoming of sensible incentives for action—which, as Kant claims in the second *Critique*, are the "subjective antagonist" of the moral law. The aesthetic representation, however, is only negative, and works by means of awakening the feeling of respect for the moral law, which is purely intellectual. Therefore, the representation of the morally good is never purely aesthetic, for without the feeling of respect, sacrifices would not arouse the idea of the moral law. The two sides of the sublime feeling, aesthetic and intellectual, are necessary for the representation of the morally good. The judgment of action, like action itself (according to Kant's description of it in the first *Critique*, as we have seen), has two "sides": one purely intellectual and one accessible to the senses.

It is this twofold nature of the judgment of the sublime that allows us to judge on the basis of ideas, without disregarding their incommensurability with anything given in experience. The imagination cannot "apply" ideas to intuitions and bridge the gap between our noumenal and our empirical character, because there is no correspondence between the realm of freedom and the realm of experience. But objects in experience can arouse ideas of reason by driving the imagination to the limit of what it can represent. It is the "arousal" or "awakening" of respect for ideas by virtue of this stretching of the imagination that constitutes the sublime feeling. The judgment of the sublime thus relates ideas and actions while preserving their heterogeneity: no action can represent compliance with duty, but an action can arouse the feeling of a power that is

[97] Ibid., 153; 5:271.

beyond natural necessitation, and thus respect for the idea of duty. Superiority over nature is the negative side of the sublime feeling, but by itself it is not enough to ground a judgment claiming universal approval. The source of this claim is the positive side, which is the feeling of respect for the moral law. There is no middle term between these two sides because, as we have seen, the sublime feeling emerges out of the inadequacy between two faculties, and it is this inadequacy that puts them in relation to one another. As Lyotard puts it, "the differend cannot be resolved. But it can be felt as such, as differend. This is the sublime feeling."[98] The experience of a discordance between the object and our capacity to represent it awakens in us the feeling for ideas, which is the basis to making a judgment upon action claiming universal approval.

Between Enthusiasm and Sublimity

While Kant had theoretical reasons for stressing the tension between ideas and the imagination in the judgment of the sublime, practical concerns play a significant role in his analysis. These concerns become apparent in his remarks on enthusiasm. As we have seen, the judgment of the sublime has an "aesthetic side," which is the feeling of the superiority over obstacles of sensibility, and an "intellectual side," which is respect for the moral law. Enthusiasm [*Enthusiasmus*] is a state of mind that is produced by the feeling of the sublime without a connection to the feeling of respect, so that superiority over obstacles of sensibility becomes an end in itself. Kant describes enthusiasm as "the idea of the good with affect [*mit Affekt*]," adding that "this state of mind seems [*scheint*] to be sublime." The explanation of why this state of mind only "seems" to be sublime is the following:

[98] Lyotard, *Lessons on the Analytic of the Sublime*, 234.

> Now, however, every affect is blind [*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 [*Bewegung des Gemüths*] that makes it incapable of engaging in free consideration of principles [*freie Überlegung der Grundsätze*], in order to determine itself in accordance with them. Thus it cannot in any way merit a satisfaction [*Wohlgefallen*] of reason. Nevertheless, enthusiasm is aesthetically sublime [*äesthetisch gleichwohl ist der Enthusiasmus erhaben*; "aesthetically nevertheless enthusiasm is sublime"], because it is a stretching of the powers through ideas [*eine Anspannung der Kräfte durch Ideen*], which give the mind a momentum [*Schwung*] that acts [*wirkt*] far more powerfully [*mächtiger*] and persistently [*dauerhafter*] than the impetus given by sensory representations [*Sinnenvorstellungen*].[99]

The movement of the mind that makes one capable of engaging in "free consideration of principles" requires the feeling of respect for the moral law. Without this free consideration of principles, the aesthetic feeling of sublimity (the feeling of superiority over nature) remains "blind," in the sense that superiority over sensible obstacles is not linked to the pure idea of the moral law. Kant's example is "indignation, as anger," which suggests an affect that shares with the idea of the good the disregard of sensuous self-interest, but which does not necessarily involve a free consideration of principles for action and their implementation in accordance with reason. A sense of indignation can predispose us to overcome our inclinations, but this predisposition does not necessarily involve respect for the idea of the law, which is why enthusiasm "cannot merit any satisfaction of reason." Considered aesthetically, enthusiasm is sublime because it produces a momentum in the mind that exceeds the power produced by any sensory representations. But because it is not related to the purely intellectual regard for the law, and thus to a free consideration of principles, it lacks the non-aesthetic component that is essential for the judgment of the sublime. This enthusiasm, Kant claims, can be compared to a

[99] Ibid., 154; 5:272.

"delusion of sense" [*Wahnsinn*; literally "madness"], that is, to the idea that sensuous representations are equivalent to that which they represent—so that superiority over the inclinations is not taken to be a representation of morality, but rather to be morality as such.

Kant's concern with enthusiasm shows the importance of the twofold nature (both aesthetic and intellectual) of the judgment of the sublime and of the sublime feeling. In *Religion within the Boundaries of Mere Reason,* Kant identifies "delusion of sense" as part of the deceptive religious practices that he calls "delusion of religion" [*Religionswahnn*], which consists in "the habit of taking a mere representation [*eine bloße Vorstellung*] (of the imagination) for the presence of the thing itself [*die Gegenwart der Sache selbst*], and to value [*würdigen*] it as such."[100] In religion, the "mere representation" are the practices that express our moral disposition, while the "thing itself" is the disposition as such. Significantly, Kant mentions rituals of self-sacrifice as one of the delusional practices which we often confuse with morality, because we take its aesthetic component (the overcoming of our inclinations) as the thing itself. The judgment of the morality of action only in terms of its aesthetic component, without a free consideration of principles, becomes a source of delusion, that is, an attachment to the appearance of virtue without a regard for the moral ideas that underlie it. As we have seen, this delusion is not merely a theoretical mistake, but rather a manipulation of the power of judgment stemming from what Kant calls "radical evil." The attempt to make the morality of action directly accessible to sensibility cannot but deceive the power of judgment, erasing the gap between sensible representations and moral ideas. As Kant puts in in his remarks on affects:

> Even tumultuous movements of the mind [*stürmische Gemüthsbewegungen*], whether they be associated with ideas of religion, under the name of edification,

[100] Kant, *Religion within the Boundaries of Mere Reason*, 188; 6:168.

> 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 [***erhabenen*** *Darstellung*], if they do not leave behind a disposition of mind [*Gemüthsstimmung*] 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 [*intellectuelle Zweckmäßigkeit*] (the supersensible).[101]

No matter how noble the idea we are presenting (some of Kant's examples are courage, indignation, and humility), it is only sublime if it makes us conscious of our supersensible capacities. Otherwise, it remains at the level of enthusiasm, and potentially leads to delusion.

The contrast between enthusiasm and the sublime brings us back to the distinction between the attachment to a given system of values and the capacity to judge independently of them. Arendt, as we have seen, believes that the attempt to ground our judgments upon a given system of values leads to the mere habit of holding fast to something, that is, to whatever standards are given to us at any given moment. This view is influenced by Eichmann's "idealism," which consisted in a desire to find an "Idea" for which he could sacrifice everything. This strange idealism, which finds satisfaction in the mere representation of virtue and disregards its true moral ground, is precisely what Kant calls enthusiasm. Sacrifice itself, according to Kant, merits no satisfaction of reason, for it lacks a relation to the purely intellectual source of moral ideas. The sublime feeling, by contrast, relates sacrifices to respect for the moral law, thus preserving its true ground. There is therefore a subtle but crucial difference between the state of mind in enthusiasm and in the sublime feeling: the former finds satisfaction in sacrifice as an end in itself, while the latter finds satisfaction in sacrifice only insofar as it arouses respect for the idea of duty. Because this idea is incommensurable with anything in experience, the state of mind in

[101] Kant, *Critique of the Power of Judgment*, 155; 5:273.

the sublime feeling involves an inadequacy between our faculties, between the aesthetic representation of virtue by means of the imagination and the intellectual respect for ideas by means of reason. The judgment of the sublime does not solve the inadequacy, but rather stems from it: it is the tension between the two sides of the sublime feeling that produces a state of mind that is communicable.

We can now respond to the problem of the universal validity of the judgment of action. The judgment of the sublime has certainly no "validity" in the sense of a procedure for adjudicating whether a phenomenal action corresponds or not with an idea of reason. However, if we follow Kant's insight in the third *Critique* that a state of mind can be communicable without the mediation of concepts, and thus of rules that we apply to intuitions, then the judgment of the sublime is valid insofar as it stems from the feeling of respect for the idea of humanity. In the contemplation of an act that awakens this feeling, we expect that everyone else will feel the same way. It is the state of mind of the spectator, rather than anything in the object as such, that is communicable to others. And it is because we presuppose that we all feel respect for moral ideas that we expect that others will agree with our judgment. If, as Arendt claims, we can judge from the viewpoint of a "world community" to which we belong by "the sheer fact of being human," it is because humanity means nothing but the capacity to act on the basis of ideas, and thus to be receptive to them. Even if we cannot represent ideas, we can feel aroused to think about them in the presence of a certain object. It is this arousal that makes possible a relationship between actions and ideas, thus opening up the possibility of a judgment claiming universal approval.

Conclusion

Arendt's turn to Kant's aesthetics represents an attempt to find an alternative between a communitarian and an objective standard for the judgment of action. The focus of this turn is the notion of common sense, which constitutes an "intersubjective" standard for judgment, in the sense that it stems from the capacity to take into account multiple standpoints. This conception leaves open an important problem, namely, the capacity of our judgment to be valid beyond all viewpoints familiar to us. In order for our judgment upon action to claim universal approval, it must have a ground in universal ideas, as Arendt suggests in her references to the "idea of mankind" and our "cosmopolitan existence." Kant's analysis of the sublime shows how it is possible to judge on the basis of ideas, thus avoiding a purely communitarian standard for judgment, but also an objective standard that would work as a fixed procedure with which to judge. Like Arendt, Kant is distrustful of our attachment to representations of moral virtue, because they tend to conceal the true ground of morality, which is the purely intellectual respect for the moral law. The judgment of the sublime represents an alternative to this attachment, because it does not consider the object as a representation of an idea, but rather as an awakening of respect for the idea of humanity in those who contemplate it. It is on the basis of this feeling that we expect the approval of others, without the need for objective rules grounding knowledge of the object.

Chapter 4: Lyotard on Good and Evil in Postmodernity

Does it still make sense to talk about good and evil in a context of postmodernity, where absolute standards for judgment seem to have been lost? But, if it does not, does it mean that evil simply does not exist anymore? Would this not amount to a kind of "anything goes" in politics? On this point, one of the main political theorists of postmodernity, Jean-François Lyotard, claims that "politicians cannot have the good at stake, but they ought to have the lesser evil." As Lyotard defines it, evil is the "incessant interdiction of possible phrases, a defiance of the occurrence, the contempt for Being."[1] The terms "phrase," "occurrence," and "Being" refer to that which arrives unpredictably, without following a pre-existing rule. Yet all phrases and occurrences fall into a web of discourses, each of which seeks to regulate them under its rules, thus doing wrong to other discourses. The political good, according to Lyotard, consists in minimizing this wrong, in such a way that the conflict between discourses for the regulation of phrases does not become a complete obliteration, as if all phrases had to belong to a single discourse. Attempting to regulate this lack of regulation under an absolute set of rules, as the philosophical discourse (as Lyotard understands it) has often sought to do, would be self-defeating, for such rules would necessarily wrong other rules. However, if politics is still to have an orientation, and not fall into complete skepticism, it is necessary to have an idea of what the "lesser evil" would be. It would be necessary, in other words, to know how to minimize the wrong that takes place with each occurrence, and with the ensuing conflict regarding how to regulate it—that is, the conflict around how to speak about the occurrence.

[1] Jean-François Lyotard, *The Differend: Phrases in Dispute* (Minneapolis, MN: Minnesota UP, 1989), 140.

The goal of this chapter is to show, following Lyotard, that good and evil are still operative political concepts in a context of postmodernity. Lyotard's conception of postmodernity has been repeatedly and mistakenly interpreted as advocating for political skepticism or nihilism, according to which political judgment should abandon all aspirations to communicability and universality, and limit itself to local struggles and narratives.[2] While it is true that Lyotard rejects totalizing narratives and universal standards for judgment, this does not entail the complete abandonment of the universalist aspiration of political judgment and action. Political action can still be judged as "just" or "good," on condition that we do not ground justice and goodness upon a pre-given set of norms that we can simply apply to a multiplicity of situations. Rather, a just or a good political action will be one that invents new means of communication in a situation that seems deprived of them, thus establishing a link between heterogeneous parties. The universal appeal of such action does not stem from a rule that grounds this linkage, but rather from the very need to communicate, even when there are no means by which to do so. The model for this kind of universality is Kant's judgment of the sublime. In Lyotard's view, the judgment of the sublime exposes the general procedure underlying Kant's critical philosophy, which is in itself a model for politics. By focusing closely on Lyotard's interpretation of Kant, this chapter will show that universality and postmodernity are not mutually exclusive—rather, postmodernity, properly understood, exposes what universality truly is.

[2] Some examples of this one-sided reading are: Linda M. G. Zerilli, "'We Feel Our Freedom': Imagination and Judgment in the Thought of Hannah Arendt," *Political Theory* 33(2), 181-182; Claude Piché, "The Philosopher-Artist: A Note on Lyotard's Reading of Kant," *Research in Phenomenology* 22(1), 158; James Williams, *Lyotard and the Political* (New York: Routledge, 2000), 120. I discuss these interpretations further down.

The Problem of Legitimacy and the Turn to Kant

In *The Postmodern Condition*, Lyotard succinctly defines "postmodern" as "incredulity toward metanarratives."[3] Given that narratives play a crucial role in the legitimation of knowledge, this incredulity leads to a crisis of legitimacy. Modernity, according to Lyotard, is characterized by the attempt to legitimize scientific knowledge by means of a metadiscourse that explains its rules. This is the function of philosophy, which ever since Plato has provided a narrative, such as the Allegory of the Cave, that accounts for the validity of scientific statements: "scientific knowledge cannot know and make known that it is the true knowledge without resorting to the other, narrative, kind of knowledge, which from its point of view is no knowledge at all."[4] The implications of this interrelation between narrative and knowledge, however, go far beyond the realm of science. Once knowledge has created a metadiscourse that legitimates it, this metadiscourse can be used to legitimate ethical and political norms. Philosophical narratives emerge originally from the need to legitimate scientific knowledge but, after the Enlightenment, "narrative knowledge makes a resurgence in the West as a way of solving the problem of legitimating the new authorities."[5] When these narratives lose credibility, legitimacy is in crisis, for there is no single source that regulates the distinction between the true and the false, the just and the unjust, the beautiful and the ugly, the productive and the unproductive, and so on. As a consequence, there is a dispersion of "language games," each producing its own standards of legitimacy without a meta-language that mediates between them.

[3] Jean-François Lyotard, *The Postmodern Condition: A Report on Knowledge*, trans. Geoffrey Bennington and Brian Massumi (Minneapolis, MN: Minnesota UP, 1984), xxiv.
[4] Ibid., 29.
[5] Ibid., 30.

This linguistic dispersion creates a new problem for the legitimacy of judgments. A language game can determine whether a specific statement is legitimate or not according to its own criteria, but there is no higher standard determining which criteria should apply in each case. In the absence of a metadiscourse that organizes the relationship between judgments, social pragmatics "is a monster formed by the interweaving of various networks of heteromorphous classes of utterances (denotative, prescriptive, performative, technical, evaluative, etc.)."[6] As a consequence, we cannot expect that there will be consensus regarding the rules that determine the legitimacy of each utterance, or what Lyotard calls "metaprescriptives": "there is no reason to think that it would be possible to determine metaprescriptives common to all of these language games or that a revisable consensus… could embrace the totality of metaprescriptions regulating the totality of statements circulating in the social collectivity."[7] Consensus, Lyotard claims, is a dated idea, for it presupposes that the heterogeneous language games that constitute social pragmatics could agree on common rules for regulating all utterances. In the absence of a credible metadiscourse, such agreement is simply not possible. This does not mean, however, that universal principles for judgment and practice must be abandoned: "justice as a value is neither outmoded nor suspect. We must thus arrive at an idea and practice of justice that is not linked to that of consensus."[8] But how can we determine what justice is if we lack a consensus in this regard? What language game could provide a definition of justice that is not invalid from the viewpoint of a different language game?

[6] Ibid., 65.
[7] Ibid.
[8] Ibid., 66.

Somewhat paradoxically, justice stems precisely from the recognition of the fact that there is no absolute answer to the question of what justice is. Lyotard develops this idea in the dialogue *Au juste*, translated as *Just Gaming*, where he claims that the "Idea of Justice" consists "in preserving the purity of each game, that is, for example, in insuring that the discourse of truth be considered as a 'specific' language game, that narration be played by its 'specific' rules."[9] This is a contrast to what Lyotard calls "terror," which consists in one language game seeking to impose its rules upon all others. While a just judgment acknowledges the heterogeneity between the rules of different language games, the terrorist judgment seeks to obliterate this heterogeneity. This is why, according to Lyotard, justice can only be an Idea in the Kantian sense: it is not the referent of a judgment, but a principle that orients how we judge. To judge on the basis of the Idea of Justice is to keep the "game of justice" open, in such a way that another judgment may disagree on what justice is. Terrorism, by contrast, consists in judging in such a way that one assumes that one knows what justice is, thus de-legitimizing any judgment that disagrees: "any attempt to state the law, for example, to place oneself in the position of enunciator of the universal prescription is obviously infatuation itself and absolute injustice, in point of fact."[10] This is why the just judgment partakes in the "game of justice": it does not presuppose a definitive concept of justice, but listens to what other judgments have to say about this, in such a way that "one speaks only insasmuch as one listens, that is, one speaks as a listener, and not as an author."[11] Justice is a "game" in the sense that its meaning is defined by

[9] Jean-François Lyotard and Jean-Loup Thébaud, *Just Gaming*, trans. Wlad Godzich (Minneapolis, MN: Minnesota UP, 1985), 96.
[10] Ibid., 99.
[11] Ibid., 72.

the mutual responses between the "players." Injustice, by contrast, precludes these responses, thus withdrawing from the game.

This conception of justice is essential for Lyotard's understanding of political judgment. Politics, according to Lyotard, presupposes an idea of justice, for it is this idea that supports prescriptions regarding what society should be like: "there is no politics if there is not at the very center of society… a questioning of existing institutions, a project to improve them, to make them more just."[12] Now justice, as we have seen, is a game rather than a set of principles. This means that political judgments do not rely on an absolute conception of justice, as if they possessed the standards for determining what justice is, but rather respond to other judgments in an open-ended "game of justice." The game is played between different language games with specific rules, which is why the only "rule" of the game of justice, if it is not to become terror, is to respect the heterogeneity of rules. At the moment of uttering a judgment, one cannot presuppose any rule: "I judge. But if I am asked by what criteria do I judge, I will have no answer to give." Otherwise, Lyotard claims, "it would mean that there is actually a possible consensus on these criteria between the readers and me; we would not be then in a situation of modernity, but in classicism" (the contrast between modernity and classicism here corresponds to the contrast between postmodernity and modernity in later writings).[13]

Lyotard's interpretation of the Idea of justice as a ground for judgment has been criticized as secretly falling back into a philosophical universalism. David Ingram claims that Lyotard's denial of absolute standards for judgment brings about a new "communitarian ideal," according to which "the *plurality of voices* (or language games) would be preserved without the violence of

[12] Ibid., 23.
[13] Ibid., 15.

hegemony."[14] In other words, Lyotard's idea of a judgment lacking absolute foundations would lead back to a new foundation under the form of a regulative idea—namely, that of a world in which all voices are respected. A similar argument is presented by Samuel Weber in his Afterword to *Just Gaming*, where he claims that "by prescribing that no game, especially not that of prescription, should dominate the others, one is doing exactly what it is simultaneously claimed is being avoided: one is dominating the other games in order to protect them from domination."[15] Jean-Luc Nancy further pursues this line of critique in his response to *Just Gaming*, claiming that "Lyotard says that justice must intervene in order to purify the games that are impure, such as narrations infiltrated by prescriptions: but where are the 'purities' in question determined from?"[16] According to Nancy, this question leaves Lyotard with two equally problematic options: either the purity of each discourse must be determined empirically, without a connection to the idea of justice, or the idea itself has to be determined by a particular judgment, in which case it would cease to be a universal. Thus, Lyotard's idea of justice would do either too little or too much: either it prescribes that particular judgments should follow empirical laws, which can hardly lead to justice, or it constitutes a new universal standard for judgment, which may be in its turn contaminated by the particular. Counterintuitively, according to these three critiques, Lyotard's postmodern conception of judgment is not as postmodern as it claims to be, but rather carries the essential elements of the philosophical universalism that it seeks to overcome.

---

[14] David Ingram, "The Postmodern Kantianism of Arendt and Lyotard," in *Judging Lyotard*, ed. Andrew Benjamin (New York: Routledge, 1992), 136.
[15] Samuel Weber, "Afterword," in *Just Gaming*, 105.
[16] Jean-Luc Nancy, "Dies irae," in *La faculté de juger*, by Jacques Derrida et al. (Paris: Éditions de Minuit, 1985), 14; my translation.

Lyotard's later work, where ideas (in the Kantian sense) play a more prominent and systematic role, contains a response to this criticism, as we shall see in what follows. Lyotard's attempt to translate his diagnosis of a postmodern condition into a theory of political judgment proceeds through a reading of Kant, and especially of his *Third Critique*. The aesthetic judgment serves as a model for political judgment, because it does not presuppose any pre-given rule mediating between heterogeneous discourses. This does not mean, however, that it lacks any sort of orientation. Rather, the political judgment stems from the feeling that one ought to judge even in the absence of standards by which to do so, thus inventing new standards by the very act of judging. It is this feeling, of which the feeling of the sublime in Kant serves as a model, that connects the power of judgment with universal ideas, such as justice or (as we will see) "humanity." In order to understand this relationship between heterogeneity, political judgment, and ideas, let us turn to Lyotard's reading of Kant.

Critical, Political, and Reflective Judgment

"There is an affinity between the critical (the 'tribunal' of critique, the 'judge' who examines the validity of the claims of various phrase families…) and the politico-historical: each has to make judgments without having a rule for making judgments, as opposed to the politico-juridical."[17] These opening words of *Enthusiasm*, Lyotard's collection of studies on Kant, summarize the link between political judgment and critical judgment. Critique, according to Lyotard, presupposes no doctrine, that is, no set of rules regulating a discourse, because its aim is to inquire into the legitimacy of all rules. Evidently, a rule for inquiring into the legitimacy of

[17] Jean-François Lyotard, *Enthusiasm: The Kantian Critique of History*, trans. Georges Van Den Abbeele (Stanford, CA: Stanford UP, 2009), xvii.

rules would lead to an infinite regress. While there are specific doctrines for specific realms of experience, such as knowledge and ethics, the critique of such doctrines, that is, the inquiry into their validity, cannot itself be a doctrine. Given Lyotard's conception of society as a web of language games without a meta-discourse, it is not difficult to understand in what sense there is an affinity between the political and the critical. The institution of society is determined by a plurality of language games interacting with one another, without any discourse regulating this interaction. Despite this lack of regulation, it is necessary to judge, which is why there are judgments that lack a prior regulation, that is, critical judgments. As we have seen, the political judgment must proceed without rules, acknowledging the heterogeneity of language games as irreducible to a common ground. Therefore, Lyotard suggests, "perhaps the critical… is the political in the universe of philosophical phrases, and perhaps the political is the critical… in the universe of sociohistorical phrases."[18] Just like critique emerges out of the need to adjudicate the validity of heterogeneous faculties upon their respective realms, politics emerges out of the relationship between heterogeneous language games (science, art, economics, popular narrative, etc.).

According to Lyotard, Kant's fundamental analysis of how critical judgments proceed in the absence of rules takes place in the *Third Critique*, whose object is what Kant calls the "reflective judgment." The reason is that "the critical is determined in general as reflexive," in the sense that "it does not arise from a faculty, but from a quasi- or 'as if' faculty (the faculty of judgment, sentiment) inasmuch as its rule for determining which universes are pertinent to it entails some

[18] Ibid.

indeterminacy (the free play of the faculties among themselves)."[19] The reflective judgment does not proceed from a specific realm of objects (natural, moral, or aesthetic), but from interaction between universes of phrases. As a consequence, the "faculty" of judgment is not really a faculty, because it does not regulate any specific realm; it emerges rather from the play between faculties. This is why critique is essentially reflective: it does not proceed from a faculty, but rather from the need to find new rules that relate the faculties to one another. In the absence of a rule for judgment, one must judge on the basis of the indeterminate relationship, or the "free play," between different sets of rules.

It is in order to understand the structure of the critical judgment, which proceeds by inspecting the legitimacy of judgments, that Lyotard turns to Kant's *Third Critique*. The procedure for legitimating cognitive judgments was the central concern of the *Critique of Pure Reason*, while the procedure for legitimating moral judgments was the central concern of the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason*. The *Critique of the Power of Judgment*, Lyotard claims, has "the mission of unifying the field of philosophy" by "making manifest… the reflexive manner of thinking that is at work in the critical text as a whole."[20] The reflective judgment is not just another kind of judgment, but the judgment that makes manifest the manner in which critical thinking proceeds in general. Lyotard is careful to stress Kant's distinction between "manner" and "method": while the latter "follows definite principles," the former "possesses no standard other that than the *feeling* of unity in the

[19] Ibid.
[20] Jean-François Lyotard, *Lessons on the Analytic of the Sublime*, trans. Elizabeth Rottenberg (Stanford, CA: Stanford UP, 1994), 8.

presentation."[21] A method provides thinking with definite principles by means of which to unify its presentations and produce a judgment upon objects, while a manner provides nothing but a feeling of this unity, without a rule for producing a judgment. This is precisely what makes thinking critical: it lacks any fixed resting point that would put a halt to its quest for a judgment's legitimation.[22]

In the absence of rules that legitimate a judgment, critical thinking is oriented by the "aesthetic feeling" which, in the *Third Critique*, is the feeling of pleasure and displeasure. The reason is that only this feeling accompanies all the states of consciousness, and thus it is "able to pass through the different spheres of thought that the critique distinguishes."[23] Whenever we think of something, whether an object of experience or an idea, we have a feeling of our own thinking. Lyotard calls this feeling "tautegorical," by which he means that it is a sign of thought to itself. Cognitive judgments are based on the one hand on the object to which they refer and, on the other, on principles that legitimate them (the categories and their deduction, as shown in the *First Critique*). Aesthetic judgments also refer to an object, but their legitimation is a subjective feeling. When I say "the flower is beautiful," I have nothing to rely on in order to legitimate the judgment, other than the feeling of pleasure. Properly speaking, the judgment says nothing about the object, but rather refers to the state of the mind in the presence of the object. This state of mind does not need any rule or principle to be known: if I feel pleasure, I am aware of it immediately. This is why the "two aspects of judgment, referentiality and legitimacy, are but one

[21] Ibid., 6.
[22] Ibid.
[23] Ibid., 13.

in the aesthetic."[24] In the aesthetic judgment, the referent and that which legitimates my judgment upon the referent are the same. The aesthetic judgment is therefore reflective: its object is the judging subject itself. The aesthetic feeling makes this judgment possible by informing the judging subject of its own state, which does not need the mediation of rules.

The central question of the aesthetic judgment concerns the communicability of this state. How can an inherently private sensation of pleasure be "universal" and "necessary," so that it demands the agreement of others? According to Kant, communicability stems from the accord between our faculties. Both theoretical and moral judgments are valid, and can therefore demand universal agreement, insofar as our faculties relate to one another following universally valid principles. In the theoretical judgment, the principles are given by the understanding, which legislates over the imagination, while in the practical judgment they are given by reason, which legislates over the understanding. In the reflexive judgment, such principles are missing. However, if the judgment still demands to be communicated and be universally valid, it means that it stems from the same disposition of thinking present in the other two kinds of judgment.[25] Consequently, the demand to make aesthetic pleasure communicable reveals that there is a "unity of the faculties," which precedes "the unison required of other individuals by the individual who judges."[26] In making an aesthetic judgment, thinking feels the same disposition as when it makes a cognitive or practical judgment. Unlike these judgments, however, it lacks a definite principle that determines the accord between the faculties. Therefore, no "disputation" is possible regarding the legitimacy of an aesthetic judgment, because there are no objective grounds upon

[24] Ibid., 12.
[25] Ibid., 82.
[26] Ibid., 88-89.

which to decide whether it is valid or not. However, it is possible to "argue" about taste, which implies the hope to reach universal agreement without objective proofs.[27]

In the absence of a principle that regulates the interaction among the faculties, their agreement has to emerge spontaneously. But how can the faculties establish an accord without a rule that regulates it? There are two ways: the beautiful and the sublime. The agreement between our faculties in the beautiful, and the corresponding demand of communication, relies on what Lyotard calls "the supersensible." This idea appears as a solution to the antinomy of taste, according to which there is either a concept determining what is beautiful and what is not, and therefore taste is objective, or there is no concept of the beautiful at all, and therefore it is entirely subjective. The judgment of taste, however, demands to be communicated, which means that it is not merely subjective, but it does not proceed according to concepts, and is therefore not objective either. Kant's solution relies on what he calls an "indeterminate concept," that is, a concept that does not correspond to any intuition, which amounts to what he calls an idea.[28] This idea is precisely "the supersensible." According to Lyotard's reconstruction, the idea appears in the antinomy of each of the faculties, as an effect of their striving for the unconditioned condition of their use: "the unconditioned of knowledge cannot be known. The absolute law of the faculty of desire cannot be desired. The supersensible principle that founds the demand for the universal communication of taste is not the object of an aesthetic pleasure."[29] The unconditioned in each of the faculties is a "limit-Idea," that is, an idea after which each faculty can and must strive without ever reaching it. In this striving, however, the faculties encounter an

[27] Ibid., 208-209; Immanuel Kant, *Critique of the Power of Judgment*, ed. Paul Guyer, trans. Paul Guyer and Eric Matthews (Cambridge: Cambridge UP, 2000), 214; 5:338.
[28] Lyotard, *Lessons*, 210. Kant, *Critique of the Power of Judgment*, 216; 5:341.
[29] Lyotard, *Lessons*, 214.

affinity among each other, which "is revealed by way of the similarity of their respective inconsistencies."[30] There is an affinity between heterogeneous faculties, not because there is a higher principle or faculty regulating their interaction, but because they are equally unable to totalize their respective fields.

The beautiful provides a model for the commonality within heterogeneity that characterizes what Kant calls *sensus communis*. In order to establish an accord between the faculties, it is first necessary to deny their striving for totalization. This cannot be done by means of a higher, totalizing principle that would determine the limit of each faculty, but only by means of each faculty immanently encountering its own limit. The discordance experienced by each faculty is a sign of its affinity with the other faculties, and so "one must think excessively, until one reaches a discordance, in order to hear the voice of concordance."[31] This discordance signals the need that each faculty moderate itself, so that "each lends itself to the play of the other." It is this free play that produces the aesthetic delight by which "thought succeeds in putting itself in the place of the other."[32] The communication of taste thus stems from thought experiencing its own heterogeneity. Without this experience, the free, indeterminate accord between the faculties, on which the aesthetic feeling depends, would be impossible. By reaching its own limit, thought becomes critical, in the sense of questioning the legitimacy of its claims (cognitive, moral, or aesthetic) and lending itself to an indeterminate argumentation. But does each faculty acknowledge its own limitation independently of the encounter with other faculties? Or does this

[30] Ibid., 215.
[31] Ibid., 216.
[32] Ibid., 222.

acknowledgement stem from the disaccord produced by the interaction itself? It is this latter possibility that Lyotard's reading of the sublime explores.

The Sublime and the Differend

The sublime represents a commonality within heterogeneity that is different from that of the beautiful. The faculties do not communicate here by means of their free play with one another, but rather by means of the subjection of one to the other. This subjection is not mediated by rules, as in the theoretical and practical judgments. In a way, the sublime proceeds according to an absolute violence, to the point that Lyotard describes it as a sort of rape.[33] Reason, which is the faculty of ideas, subdues the imagination, which is the faculty of presentation. The heterogeneity is here extreme, for ideas are inherently non-presentable, and thus reason's command is impossible to fulfill. The imagination attempts to present absolute ideas, such as the absolutely great and the absolutely forceful, but it fails, for every presentable magnitude or force can be surpassed by a greater one—the ideas of reason are infinite, while the imagination is finite. The failure is felt as displeasure, but it gives raise to the feeling of pleasure as we become aware of our mind's vocation for the supersensible:

> The feeling of the sublime is thus a feeling of displeasure from the inadequacy of the imagination in the aesthetic estimation of magnitude for the estimation by means of reason, and a pleasure that is thereby aroused at the same time from the correspondence of this very judgment of the inadequacy of the greatest sensible faculty in comparison with ideas of reason, insofar as striving for them is nevertheless a law for us.[34]

[33] Ibid., 180.
[34] Kant, *Critique of the Power of Judgment*, 141; 5:257.

The heterogeneity between reason and the imagination lends itself to no agreement, not even one mediated by a disagreement. Rather, it is in the disagreement itself that the sublime feeling emerges, which is why aesthetic pleasure here is necessarily mixed with displeasure.

The discordance experienced in the sublime feeling is at the root of critical thought. This kind of thought, as we have seen, does not follow any fixed set of rules—if it did, it would be unable to inquire into the legitimacy of these rules. This is what happens in cognitive judgments, where we determine an object under a concept and, provided that we have followed the right procedure, we demand that everyone else agrees. In the judgment of the sublime, by contrast, no cognition is involved, but merely a feeling of pure thinking, that is, of thinking prior to any of its determinations, such as sensible forms, schemas, or concepts: "'before' all of this, thinking is the power to think…, irrelative, 'raw,' that comes from nothing other than itself."[35] Critical thinking, which is thinking prior to any determined form, is awakened by the dissonance that takes place when it encounters incommensurable forms. As a consequence, resistance is an essential component of the sublime feeling: if the imagination did not resist, by virtue of its finitude, the demands of reason, the heterogeneity between the two faculties would not be felt. This dissonance "is essential for thought to feel reflexively its heterogeneity when it brings itself to its own limits."[36] The dissonance produced by the judgment of the sublime is felt as a disposition of thought for which no standard for judgment is legitimate a priori, that is, as critical thought. This is why critical thought has no specific location: it is not one more discourse among others, but rather the feeling that takes place at the limit between discourses. The critical judgment is

[35] Lyotard, *Lessons*, 122.
[36] Ibid., 150.

oriented by this feeling, which cannot be translated into a rule because it stems from the conflict between rules.

The heterogeneity of the faculties in the sublime serves as a model for any situation in which two heterogeneous discourses interact with one another, thus producing what Lyotard calls a "differend." The differend is the elementary political situation, which takes place when "something which must be able to be put into words cannot yet be."[37] The reason why something cannot be put into words is that it lacks an "idiom," that is, a discourse that is able to communicate it. One of Lyotard's examples is a Martinican who is considered a French citizen under French law. By being considered a French citizen, the Martinican suffers a wrong, for he does not consider himself a French citizen.[38] Yet any complaint that he can bring to the French authorities will be taken as a litigation under French law. The wrong thus lacks an idiom to express itself, for it is always translated into a discourse that is foreign to it. The question that follows is: if the wrong lacks an idiom in which to express itself, can it be communicated at all? How do we even know that it is there? It is on this point that the political import of the sublime feeling becomes evident.

The wrong suffered by one of the parties in a differend produces both a silence and a "call upon phrases which are in principle possible."[39] The differend is a state in which something cannot be put into phrases, yet this "something" is felt as a call to invent new phrases by means of which it can be expressed. Lyotard describes this feeling through an analogy that resembles the sublime:

---

[37] Jean-François Lyotard, *The Differend: Phrases in Dispute*, trans. George Van Den Abbeele (Manchester: Manchester UP, 1988), 13.
[38] Ibid., 27.
[39] Ibid., 13.

> Suppose that an earthquake destroys not only lives, buildings, and objects but also the instruments used to measure earthquakes directly and indirectly. The impossibility of quantitatively measuring it does not prohibit, but rather inspires in the minds of the survivors the idea of a very great seismic force. The scholar claims to know nothing about it, but the common person has a complex feeling, the one aroused by the negative presentation of the indeterminate.[40]

The feeling that something is to be phrased is a negative presentation, which emerges from the impossibility of presenting an experience under any available idiom. The example of the earthquake refers to the impossibility of properly phrasing the wrong produced by the Nazi extermination camps: given that the means to prove the magnitude of the crime were destroyed, the feeling of the crime is incommensurable with any demonstration of it. Yet this failure of demonstration is felt as the presence of something that demands to be phrased. The impossibility of phrasing is felt as displeasure, while pleasure accompanies the invention of new phrases.[41]

The sublime feeling could be called the "feeling of heterogeneity." Heterogeneity can only be felt subjectively because there is no objective judgment that could designate it. A certain discourse can be "objectified" by another discourse so as to explain the rules that govern it. But there is no rule governing the interaction between heterogeneous discourses, just like there is no rule that determines how reason should relate to the imagination, which is why the relationship takes the form of an absolute violence. Similarly, there is no rule determining how French law should relate to someone who does not acknowledge it as his or her law, and so violence ensues. The incommensurability between the two discourses cannot be shown, for this would presuppose a meta-discourse that knows the rule for translating one language into the other. In the absence of such a meta-language, the gap separating one discourse from another can only be experienced as

[40] Ibid., 56.
[41] Ibid., 13.

a feeling. Unlike a judgment, the feeling has no determined object. It is rather a sign that there is heterogeneity and, as an effect of it, a wrong that demands to be phrased. The political challenge consists in turning the feeling into a new idiom that can bring the two heterogeneous discourses into communication with one another, as opposed to subordinating one to the other.

The invention of new idioms does not overcome the heterogeneity between different discourses, but establishes connections between them. Lyotard describes this procedure through the symbol of an archipelago: "each genre of discourse would be like an island; the faculty of judgment would be, at least in part, like an admiral or like a provisioner of ships who would launch expeditions from one island to the next, intended to present to one island what was found… in the other."[42] The faculty of judgment moves from one genre of discourse to another, with the aim of presenting to one genre what it has found in another. In order to do so, it invents "passages." Lyotard identifies a model of passage in what Kant calls the "type" in the second *Critique*, which establishes a link between ethical and cognitive judgments so as to constitute the categorical imperative. The type, according to Lyotard, makes an object of reason (the moral law) presentable to the understanding, whose objects are ruled by the law of natural causality. The goal of this passage is to make the moral law effective in the world, which is achieved through the formula of the categorical imperative: "act as if the maxim of your action were to become by your will a *universal law of nature*." Here, the object of one faculty is presented to a different faculty "as if" it belonged to it. Through the "as if," the faculty of judgment acknowledges the heterogeneity between ethical and cognitive judgments, and "neither hollows out nor fills in the abyss" that separates them, but rather "passes or comes to pass over it, and

[42] Ibid., 130-131.

takes it therefore into consideration."[43] The passage thus allows two different discourses or "islands" to communicate with one another without disavowing their heterogeneity—that is, without relying on an underlying commonality that subsumes their difference.

Lyotard's focus on the sublime shows that the ground of communicability is not a universally shared standard or rule, but rather the experience of dissensus. It certainly seems paradoxical that the invention of new means of communication stems from an incommunicable feeling, or of the feeling that there is no communication. Yet this is precisely why communication depends on *invention* rather than on the *discovery* of any procedure. For Lyotard, the encounter with an other with whom we cannot communicate and the feeling that there is a wrong summons us to "institute idioms which do not yet exist."[44] This institution has nothing to do with the discovery of a common ground mediating two standpoints. If this were the case, the experience of dissensus would be a means to discover an underlying consensus, which would serve as an "Archimedean viewpoint" from which to assess the validity of the conflicting judgments. If, by contrast, dissensus is irreducible to a higher consensus, then it is up to each party to produce new means of communication *from nowhere*—or, following Kant's term in his analysis of the sublime, from the "abyss" that a party encounters when it reaches the limit of what it can present.[45] It is on the basis of feeling this abyss that heterogenous discourses may invent passages that allow them to communicate with one another, as opposed to attempting to subordinate the other to a given procedure.

---

[43] Ibid., 123.
[44] Lyotard, *The Differend*, 13.
[45] Kant, *Critique of the Power of Judgment*, 141; 5:258.

The sublime feeling is the ground of a political judgment that presupposes dissensus rather than agreement. Politics, according to Lyotard, is not a specific genre of discourse with its own specific rules, but rather it "bears witness to the nothingness which opens up with each occurring phrase and on the occasion of which the differend between genres of discourse is born."[46] When a new phrase occurs, a differend opens up between genres of discourse seeking to regulate it—as in the above-mentioned example, an identity claim by a Martinican will be described differently by the traditions of the nationals and by French law, and there is no rule to make the description of one discourse presentable to the other. Politics bears witness to this differend between genres of discourse and attempts to invent a passage by which the claim that is made in terms of local traditions can be presentable to French law. There is no rule guiding the invention of this passage, other than the principle that "heterogeneity ought to be respected in an affirmative manner."[47] Thus the political actor, like the critical philosopher, remains always between genres of discourse, refusing to accept any principle of legitimation without considering the wrongs that they produce. Politics cannot get out of the archipelago, but only navigate it indefinitely, wondering at the heterogeneity between islands and enjoying each invention of passages.

The notion that one ought to respect the differend certainly makes it sound like Lyotard has replaced one transcendental idea for another, as his critiques point out: instead of a law for subsuming differends under a rule, we have a law stating that we ought to respect differends. But let us note that there is no rule stating *how* we ought to respect them. The feeling of the differend (the sublime feeling) summons us to respect the differend and to create a new idiom by which the parties may communicate with one another, but it does provide us with a rule for doing so.

[46] Lyotard, *The Differend*, 141.
[47] Ibid., 133.

As Geoffrey Bennington puts it, "the obligation to judge justly does not project an achieved state of justice as the end of history, but encourages the critical watcher… to discover ever more *torts* [wrongs] and *différends*, in the very effort made to find idioms to phrase those *torts* and *différends* already discovered."[48] The critical or political judgment is not ruled by a regulative idea (as Ingram suggests), but stems rather from the felt need to communicate what remains incommunicable, in an open-ended process with no preestablished end. If one judges on the basis of the idea of justice, it is only in the sense that one respects the heterogeneity of rules for judgment, and invents new rules by which to communicate the heterogeneous parties. These new rules are not neutral: they create new wrongs that invite new political judgments. Consequently, as Bennington points out, the idea of justice has no end: it does not guide judgment to an ultimate state in which everything is communicable, but keeps us alert to the emergence and displacement of differends, and thus to the open-ended displacement of the boundary between the communicable and uncommunicable.

As it is clear, Lyotard is not simply rejecting the possibility of agreement so as to affirm disagreement, but rather rethinking the relationship between agreement and disagreement altogether. Linda Zerilli misses this when she criticizes "Lyotard's celebration of the Kantian sublime…: the affirmation of a *differend* or *Widerstreit*, that is, a conflict that permits no resolution whatsoever," claiming that this reading "tends to foreclose any possibility of a politically mediated agreement about community whatsoever."[49] Lyotard's point, however, is not that agreement is not possible, but rather that the principles by which we reach agreement are not external to disagreement. In other words, the principles that ground agreement are not waiting to

[48] Geoffrey Bennington, *Lyotard: Writing the Event* (New York: Columbia UP, 1988), 177.
[49] Zerilli, "'We Feel Our Freedom,'" 181.

be discovered, but are invented on the tension experienced in each particular disagreement. This is the central lesson of the sublime, in Lyotard's interpretation: the parties have no common language by which to communicate, yet the very fact that they relate to one another without a common language is felt as a sign that communication is nevertheless possible. It is this feeling aroused by the disagreement that orients the invention of new idioms by which the parties may agree. Yet because every new idiom creates new differends, the dividing line between agreement and disagreement is never clear-cut: disagreements generate the possibility of new agreements, while agreements generate new disagreements. Therefore, politics consists neither in seeking to overcome disagreement by finding unquestionable principles of agreement, nor in celebrating disagreement for its own sake. Rather. it consists in navigating disagreements with the aim of inventing partial agreements.

We are now in a position to understand the distinction between good and evil in a context of postmodernity. Clearly, any rule for determining what good and evil are would lead to terrorism, for such rule would wrong other rules. Indeed, judging those who judge differently as "evil" is perhaps the most extreme form of terrorism, for it equates different perspectives with moral faults. This does not mean, however, that the distinction between good and evil has ceased to exist. Political judgments institute new tribunals by which the parties of a differend can communicate with one another—and yet by inventing new rules that regulate differends, these judgments will in its turn produce new differends: "it is impossible that the judgments of the new tribunal would not create new wrongs, since they would regulate (or think they were regulating) differends as though they were litigations."[50] Thus, given that there is no way out of differends,

[50] Lyotard, *The Differend*, 140.

"politicians cannot have the good at stake, but they ought to have the lesser evil. Or, if you prefer, the lesser evil ought to be the political good."[51] In postmodernity, the distinction between good and evil is not absolute, but it is still operative. The political good consists in minimizing evil, namely, the wrong suffered by discourses that lack the means to communicate. Yet this minimization cannot follow a pre-stablished procedure, for such procedure would necessarily wrong other procedures. The good, as a consequence, consists in remaining attentive to the open-ended displacement of the distinction between good and evil, in such a way that political actors minimize wrongs while acknowledging the creation of new wrongs.

Universality as a Sublime Sign

If politics takes place amid an archipelago, are universalist political judgments possible? In other words: if politics consists in the invention of passages that communicate heterogeneous discourses, can we make claims that appeal to everyone, beyond those involved in a particular differend? Lyotard's answer is yes, on condition that we properly understand what this "everyone" means. On this point, Lyotard is critical of a universalist politics that presupposes the existence of a universal subject, such as "humanity," which could be the referent of a judgment. Commenting on the Declaration of the Rights of Man, Lyotard claims that the Declaration fell into an unsolvable confusion between its addressor, that is, the French people, and its addressee, namely, humanity. Through this confusion, "the members of the Constituent Assembly would have been prey to a 'transcendental appearance' and even perhaps to a *dementia*…. They hallucinated humanity within the nation."[52] Here, Kant's distinction between concepts and ideas

[51] Ibid.
[52] Ibid., 147.

plays a crucial political role. For both Lyotard and Kant, "humanity" is an idea, that is, a concept without a corresponding object in experience. The "transcendental illusion" consists precisely in confusing the idea with a concept, in such a way that something in experience can correspond to it. By claiming to speak in the name of humanity, the Constituent Assembly confused the idea of humanity with the French Nation, thus opening the door to imperialism and terror. This illusion is common to all revolutionary politics, which "confuses what can be presented as an object for a cognitive phrase with what can be presented as an object for a speculative and/or ethical phrase."[53]

The critique of revolutionary politics does not imply, however, that universal claims should be abandoned. It is a usual mistake to interpret Lyotard as advocating for a particularist politics that disregards universal ideas. Claude Piché, for example, claims that for Lyotard the philosopher must develop an "aesthetic sensitivity" as he observes political events, departing from universalist aspirations: "the feelings of pleasure and pain which are intertwined in the sublime do not have to be shared by everyone…. Fragmentation and particularism become the lot of philosophy."[54] But this particularism is impossible after the emergence of modern, cosmopolitan discourses structured upon the idea of a common humanity. To abandon the idea of this community would entail returning to a pre-modern state in which different communities do not share anything in common, and are therefore authorized to destroy one another. If there is an obligation, whatever it may be, to people from different communities, it is because all people belong to a cosmopolitan community called "humanity." The problem with this idea is that it has no referent—no one can show what "humanity" is, and no one can speak in its name. However,

[53] *Enthusiasm*, 22.
[54] Piché, "The Philosopher-Artist," 158.

as Kant shows, the fact that ideas have no referent in experience does not mean that they have no possible connection to it. The political challenge consists in acting and judging on the basis of this idea, while avoiding the transcendental illusion that one can know it.

Lyotard finds an example of the relationship between political events and the idea of humanity in Kant's famous interpretation of the French Revolution in *The Conflict of the Faculties*, and more specifically, in the notion of the "sign of history." Kant introduces this notion as part of his attempt to respond to the question of whether the human race is morally progressing. This question, Kant claims, cannot be answered by means of observation, because we cannot foresee free actions, which are the actions that stem from a moral disposition. In order to know whether human beings will act morally in the future (and thus that they are morally progressing), we need to know their moral disposition, which is impossible. However, Kant claims,

> There must be some experience in the human race which, as an event, points to the disposition and capacity of the human race to be the cause of its own advance toward the better, and (since this should be the act of a being endowed with freedom), toward the human race as being the author of this advance.[55]

This occurrence, Kant adds, could not be considered in itself as the cause of history, "but only as an intimation, a historical sign… demonstrating the tendency of the human race viewed in its entirety."[56] The historical sign is then an occurrence in history that demonstrates that the human race has a good moral disposition, and is therefore on the path of moral progress. But how is this demonstration possible? Kant's solution to this problem consists in shifting the demonstration from the act itself to the "mode of thinking" of the spectators. Indeed, the occurrence that shows

[55] Immanuel Kant, *The Conflict of the Faculties*, trans. Mary J. Gregor and Robert Anchor, in *Religion and Rational Theology*, eds. Allen W. Wood and George Di Giovanni (Cambridge: Cambridge UP, 2001), 301; 7:84.
[56] Ibid.

that the human race is progressing is not a deed, but rather "the mode of thinking of the spectators which reveals itself *publicly* in this game of great revolutions, and manifests such a universal yet disinterested sympathy for the players on one side against those of the other, even at the risk that this partiality could become very disadvantageous for them if discovered."[57] In the presence of certain deeds, the spectators experience a disinterested sympathy for one of the parties, and this sympathy is universal, in the sense that they expect that all spectators will participate in it. Given that the revolution does not serve the interests of the spectators, the only cause for this way of thinking is their moral disposition. Thus, the universal and disinterested sympathy for one of the parties in a revolution is the "historical sign" that demonstrates that the human race is progressing.

Lyotard interprets Kant's brief remarks on the French Revolution through an analysis of the sublime. In the *Critique of the Power of Judgment*, Kant claims that "aesthetically, enthusiasm is sublime, because it is a stretching of the powers through Ideas, which gives the mind a momentum that acts far more powerfully and persistently than the impetus given by sensory representations."[58] Observing a certain event, the imagination of the spectators is stretched to the point of arousing a feeling of that which lies beyond representation, namely, Ideas. But unlike the transcendental illusion, which mistakenly makes Ideas representable, enthusiasm only feels that which lies beyond representation, and remains at the threshold between the representable and the non-representable. The sign, therefore, does not produce a unity between representation and ideas, but rather remains in a zone of indeterminacy between the two: "the 'passage' does not take place; it is a 'passage' in the course of coming to pass, and its course, its motion, is a

[57] Ibid., 302; 7:85.
[58] Kant, *Critique of the Power of Judgment*, 154; 5:272; translation modified J.B.

kind of agitation in place, within the impasse of incommensurability, over the abyss."[59] One cannot legitimately claim that a certain historical event is caused by a good moral disposition, because there is an abyss between ideas and historical events. However, the agitation produced by certain events makes the spectators feel the abyss. Given that this feeling is only possible if the imagination is aroused to go beyond its limit by Ideas of reason, the feeling is in itself a sign of the presence of Ideas, even if we cannot "present" them.

The sublime, in this interpretation, constitutes a model for a specifically political kind of universality. According to Lyotard, there is an affinity between politics and aesthetics. Just like the importance of Kant's philosophy of the beautiful and the sublime "resides in the de-realizing of the object of aesthetic feelings, and at the same time, in the absence of a faculty of aesthetic cognition," the same goes "for the historico-political object, which has no reality in and of itself, and for a faculty of political cognition, which must remain nonexistent."[60] Both in politics and in aesthetics, we make universal claims regarding a certain object, without a rule that determines whether the claims are valid. Strictly speaking, we do not make judgments about the object, but rather about how we feel when contemplating an object—which is why the object of the aesthetic (or political) feeling "has no reality." This changes the sense in which the claim is "universal." The sublime feeling, Lyotard claims, "is not universal the way a well-formed and validated cognitive phrase can be; a judgment of cognition has its determining rules set 'before it,' while the sublime phrase judges in the absence of rules."[61] Therefore, while the cognitive judgment applies a rule that is given to it beforehand, the judgment of the sublime has a "rule in

[59] Lyotard, *Enthusiasm*, 32.
[60] Ibid., 21.
[61] Ibid., 34.

waiting, a rule with the 'promise' of universality."[62] The judgment of the sublime is universal only in the sense that those who judge expect that their judgment will be shared by others, without possessing a rule that grounds this expectation.

This promised universality constitutes a community that cannot be presented, or what Kant calls a *sensus communis*. In order to present an object in a way that can legitimately claim the agreement of others, one must possess beforehand the universal rule that grounds the presentation. *Sensus communis*, by contrast, is a communicability without a rule, whose ground is the aesthetic feeling. In the case of the beautiful, the feeling stems from the relationship between the understanding and the imagination. In the sublime, by contrast, the feeling stems from the relationship between the imagination and reason, which is the source of moral ideas. This is why only the sublime can be a sign of historical progress which, as we have seen, is moral progress. The capacity of the spectators to feel enthusiasm for certain events is a sign that they are "susceptible to Ideas," and this susceptibility is already moral. Thus, even if the sublime "is indicative only of a free causality," because it only indicates that those who feel it are susceptible to the Idea of morality, "it nonetheless has 'proof' value for the phrase that affirms progress, since the spectating humanity must already have made cultural progress in order to be able to make this sign, by its 'way of thinking' the revolution."[63] The very capacity to have the sublime feeling proves that we are susceptible to moral ideas. And given that this susceptibility is the aim of cultural progress, the sublime feeling also proves that humanity is progressing. Humanity is thus the name of an ethical community on the path to moral progress, whose only "proof" is the sublime feeling aroused by certain historical occurrences.

---

[62] Ibid.
[63] Ibid., 39.

The analysis of *sensus communis* in connection to the French Revolution provides the link between the differend and universality. Lyotard explain this link through an analysis of Marxism and, more specifically, in response to the question: "Marxism has not come to an end, but how does it continue?" According to Lyotard, Marx, like most modern revolutionary thinkers, fell prey to the transcendental illusion of confusing the Idea of humanity with a concrete subject. Class struggle is a differend, for the proletarians suffer as an effect of capital, and their suffering does not find an idiom in which it can be expressed. As a consequence, "Marx tries to find the idiom which the suffering due to capital clamors for…. He thinks he hears the demand of the proletariat, which is the object of an Idea, an ideal of reason, namely an emancipated working humanity."[64] Now, as we know at this point, this idea has no referent, which means that it cannot be represented. Yet Marx fell into the transcendental illusion of confusing the enthusiasm felt by those who contemplate the proletarian struggles with a request by the proletarians to take part in a common being: "The referent of the Idea of communism is transcribed as a subject (addressor) who prescribes communism. The common being wants itself."[65] In other words, by virtue of the enthusiasm that proletarian struggles arouse in the spectators, the proletariat is mistakenly understood as a representative of the Idea of humanity. From this illusion, a second illusion follows, namely, that of giving this subject a political organization, which is the task of the party.

Is it possible to preserve the universal appeal of the proletarian's suffering, without turning the proletarian into the representative of a universal Idea? The sublime provides the answer. The feeling of enthusiasm that the spectators feel when contemplating the struggle of a wronged party does not stem from their common participation in a common being—as if, by virtue of the wrong

[64] Lyotard, *The Differend*, 171.
[65] Ibid., 172.

suffered by the proletarian, everyone were being wronged. Rather, enthusiasm is a sign of the differend, which summons thought to be phrased. Lyotard expresses this, in a very condensed way, in the following complex passage on the wrong suffered by the proletarians:

> The wrong is expressed through the silence of feeling, through suffering. The wrong results from the fact that all phrase universes and all their linkages are or can be subordinated to the sole finality of capital… and judged accordingly. Because this finality seizes upon or can seize upon all phrases, it makes a claim to universality. The wrong done to phrases by capital would then be a universal one. Even if the wrong is not universal…, the silent feeling that signals a differend remains to be listened to. Responsibility to thought requires it. This is the way in which Marxism has not come to an end, as the feeling of the differend.[66]

There are two ways in which the suffering of the proletarians can be universalized. The first one is that, because the wrong stems from a discourse (capital) that is itself universal in its aspirations, such wrong concerns all groups and discourses. The second and more fundamental one is that, even if capitalism had no claim to universality, "responsibility to through" requires that we listen to the differend. This responsibility is signaled by a "silent feeling," a "feeling of the differend," which as we know by now, is the sublime feeling. But what is this "responsibility to thought" that the sublime feeling signals to?

We find a hint of an answer in Lyotard's essay "Tomb of the Intellectual," where Lyotard addresses the role of the intellectual in a context of post-modernity, in which there are no universal ideas that guide thinking. Addressing the question of whether there is such thing as a responsibility to thought that transcends the particular responsibilities assigned by particular discourses, Lyotard claims:

> The decline, perhaps the ruin, of the universal idea can free thought and life from totalizing obsessions. The multiplicity of responsibilities, and their independence

[66] Ibid., 171.

> (their incompatibility), oblige and will oblige those who take on those responsibilities, small of great, to be flexible tolerant, and svelte. These qualities will cease to be the contrary to rigor, honesty, and force; they will be their signs. Intelligences do not fall silent, they do not withdraw into their beloved work, they try to live up to this new responsibility, which renders the "intellectuals" troublesome, impossible.[67]

Postmodernity, which is characterized by the ruin of totalizing systems of thought, does not render thought as such useless, as if intellectuals had no other responsibility than withdrawing into one specific discourse. The "multiplicity of responsibilities" contained in the multiple discourse, which are often incompatible to one another, produces a "new responsibility," which compels intelligence not to remain silent. In postmodern conditions, thought consists in listening to the differend, that is, to the silence and suffering that is felt by virtue of the impossibility of phrasing a wrong. As we have seen, this is what "critical thought" exemplary does. Critical thought does not look for the rules by which to regulate the differend, as if one or both parties ignored them and needed to be enlightened by a third discourse. Rather, it acknowledges the heterogeneity of discourses and the wrong that emerges from it. Responsibility to thought is therefore responsibility to that which summons critical thinking, namely, the differend.

Now we can understand how the sublime links the differend to a universal judgment. The enthusiasm that the spectators feel when they contemplate the French Revolution or the suffering of the proletarian signals that there is a wrong, namely, a damage that does not find a discourse by which it can be phrased. The wrong summons thinking precisely because this is what it means to think: to witness the heterogeneity of phrases and to invent passages by which to communicate them with one another. The transcendental illusion consists in mistaking the feeling of the

---

[67] Jean-François Lyotard, *Political Writings*, trans. Bill Readings and Kevin Paul Geiman (London: University College London Press, 1993), 7.

differend with a universal subject that could be the referent of a phrase. If the suffering of one of the parties summons all thinking, it is not because it is the representative of a common being, but rather because we think by virtue of being confronted with an irreducible heterogeneity—just like the feeling of the sublime arouses a "way of thinking" stemming from the tension between two incommensurable faculties. If the capacity to think is universal, it is because we are all capable of witnessing differends, which announce themselves thought the sublime feeling. This feeling is prior to any regulation, which necessarily creates new wrongs and is therefore no longer universal. The feeling as such, however, demands to be communicated to everyone, insofar as everyone can witness differends.

It should be clear by now that Lyotard's conception of politics is postmodernity has nothing to do with a sort of political nihilism or skepticism. James Williams misunderstands the role of ideas in Lyotard's conception of politics when he claims that "the principal desire in the politics of the differend is a negative one: to negate Ideas of reason as ways of bridging differends."[68] As we have seen, however, political judgments do not simply "resist ideas," as Williams asserts, in order to protect heterogeneity from the threat of universal claims. Rather, they resist the transcendental illusion that ideas can be represented by a subject that can be located in history. Following Kant, ideas are not the referents of judgments—they are rather what announces itself in the sublime feeling, which is the feeling of the differend. If this feeling demands to be communicable, it is because we expect everyone else to partake in a universal community. Although the community cannot be represented, it is presupposed by every attempt to make communicable what remains incommunicable, namely, the wrong. The feeling that one must

[68] Williams, *Lyotard and the Political*, 120.

judge when there are no existing shared standards for judgment signals the idea of humanity, namely, the idea of a world community in which all judgments are communicable. It is by virtue of this idea that we are capable of uttering universal judgments, even when there is no rule that explains this universality. As Kant shows, it is not in any describable procedure, but rather in the feeling that the judgment arouses in the spectators, that the idea of humanity announces itself.

Conclusion

Let us go back to the opening question of this chapter: is it possible to talk about good and evil in a context of postmodernity? The answer, as we have seen, is yes, on condition that we understand that postmodernity entails a reformulation of the very concepts of good and evil. According to Lyotard, there is no absolute rule that allows us to distinguish between good and evil, or justice and injustice. But this unruliness does not mean that politics should abandon these concepts altogether. Rather, the acknowledgement that the heterogeneity of discourses is irreducible, that is, that any rules regulating the distinction between good and evil does wrong to other rules, leads to a politics of the "lesser evil." This politics does not seek the absolute good, but rather acknowledges that any regulation of the distinction between good and evil will be evil to some extent, insofar as it wrongs alternative regulations. The "lesser evil" does not consist in eliminating evil, but rather in preserving the indeterminacy of the distinction between good and evil, so that those groups and discourses that are wronged by it may be potentially listened to. This is what "the good" means in a context of postmodernity: to be responsible to the heterogeneity of discourses, that is, to listen to the wrong that is produced when heterogeneous discourses interact with one another, and to invent passages that make it possible to communicate this wrong.

Conclusion

Totalitarianism, according to Arendt, is a new word coined to designate a new form of evil in politics. The goal of this dissertation has been to understand the novelty behind the word, as a way to better understand the nature of action and judgment in modern politics. Kant, Arendt, and Lyotard have illuminated different aspects of evil in modernity. Let us simply list them: 1) Evil as the refusal to establish relationships with others through action and thinking, thus foreclosing uncertainty and unpredictability; 2) Evil as the mechanization of action, as a means to withdraw from the responsibility for self-transformation; 3) Evil as the regulation of the moral power of judgment under fixed standards, as a means to escape the uncertainty inherent to the judgment of action; 4) Evil as the absolutization of standards for judgment, as a way to master heterogeneity and disavow the differends that stem from it. The elements in this list have a characteristic in common, namely, the link between evil and the elimination of uncertainty and unpredictability. At the most general level, the common thread running through Kant's, Arendt's, and Lyotard's thoughts on evil is the idea that evil stems from a desire to eliminate that which exposes action and judgment to uncertainty and unpredictability—not only regarding their outcome, but also their meaning. Yet action and judgment presuppose uncertainty, which is why evil consists in a sort of anti-action and anti-judgment. An evil action destroys the act's uncertainty by turning it into a mere mechanical event, while an evil judgment destroys the judgment's uncertainty by turning it into the application of a rule. In both cases, the subjective aspect of action and judgment, namely, that which makes the person responsible for them, is replaced by an objective determination. The paradox is that the adoption of this objective determination stems from a

subjective choice, in such a way that the person acts and judges as if she were simply the instrument of a higher process.

Each chapter of this dissertation has explored a specific aspect of this kind of evil. Chapter 1 inquired into the uniqueness of totalitarian evil, which for Arendt consists in "making human beings superfluous." To be superfluous consists in eliminating the capacity to act and to think, in such a way that one becomes the instrument of an overarching process and lacks personal responsibility. Chapter 2 traced the problem of the elimination of responsibility back to Kant's thesis of radical evil. Responsibility for action, according to Kant, stems from the capacity to transform ourselves, which is why the subordination of action to a fixed set of rules eliminates responsibility. Chapter 3 showed that, for Kant, evil corrupts the power of judgment by attempting to regulate the gap separating the moral law from phenomenal actions. Relying on Kant's insight, Lyotard, as seen in chapter 4, defines evil as the attempt to regulate the heterogeneity of discourses under a meta-discourse that disavows the heterogeneity of rules. In each case, evil stems from the attempt to eliminate the core of uncertainty that is essential to action and judgment, replacing it with a set of rules. While Kant acknowledged the subjective disposition involved in modern evil-doing, Arendt and Lyotard noted the link between this disposition and the public world, where action and judgment are intertwined with the actions and judgments of others. Consequently, while Kant locates the root of evil in a disavowal of our internal discordance (between our phenomenal actions and our moral disposition, between the empirical and the moral power of judgment), Arendt and Lyotard locate it in the disavowal of our relationship to others. In all cases, however, evil stems from the attempt to eliminate a fundamental uncertainty that is constitutive of action and judgment.

Evidently, this approach to the problem of evil does not tell us how to draw the distinction between good and evil—if anything, it tells us that we cannot draw the distinction in any unequivocal way. This does not mean, however, that the ideas of good and evil play no role in the way we act and judge. Rather, we must rethink these ideas in a way that takes into account the fundamental uncertainty regarding their manifestation. Whenever we subordinate our actions to an unquestionable procedure or process, disavowing all other viewpoints from which they may be judged as inherently illegitimate, we do evil, independently of the outcome of what we do. Conversely, good actions are not the ones that follow a certain procedure that guarantees their goodness, but rather those that expose themselves to the re-actions and judgments of others. This means that, as Kant, Arendt, and Lyotard show in different ways, pure goodness is simply not possible, for an absolutely good action would need to be unaffected by other actions and judgments in order to perfectly execute the rule that determines what "the good" is. In order to act freely, that is, independently of a pre-determined law, we must insert ourselves into a web of actions and judgments that will determine the goodness and evilness of what we do. If the good, as Lyotard points out, consists in doing "the lesser evil," it is because doing good demands first of all an acknowledgement that there is no "Archimedean viewpoint" that guarantees that our actions are good. By exposing itself to a web of actions and judgments, a good action runs the risk of eventually being judged to be "evil," thus demanding forgiveness.

It is important here to distinguish between two levels of evil. There are certainly actions that we normally judge to be evil in the usual sense of the term, that is, in the sense of transgressing a norm driven by selfish or impulsive motivations. At this level, as seen especially in chapters 1 and 2, the line between good and evil is always unstable: our apparently good actions may be

ultimately driven by egoistic motives, and we can never be certain what our motivations are. What we thought was good at a certain point may turn out to be evil (or a "trespassing," to use the word Arendt took from Jesus), which is why we can ask for forgiveness or accept punishment. Radical evil, by contrast, is beyond ambiguity, because the action or judgment is seen as stemming from an absolute process that admits no alternatives. At this level, forgiveness and punishment are not possible, because the actor admits no viewpoint for judgment other than that of the law of which she is an executioner. If good actions, as Lyotard claims, are those that seek the "lesser evil," it is because any attempt to absolutely escape from evil cannot but lead to more evil—a war to end all wars, for example, cannot but be a total war that accepts no limitations, just like a revolution to end all injustice cannot but accept all injustices to achieve its goal. In order to avoid radical evil, political action must seek the "lesser evil" which, evidently, always entails a degree of evil. As long as we live with others, there is no way out of evil, but only a constant renegotiation of the distinction between what counts as good and what counts as evil. The attempt to escape this open-ended play leads either to a withdrawal from the web of human relations or, even worse, to its subordination to a rule or process that "secures" the meaning of all actions—that is, to radical evil.

The open-ended play between good and evil does not entail the elimination of all standards for deciding between one and the other. Without an attempt to do the "lesser evil," there would be no reason to engage in political action. Even though there is no rule for determining what the "lesser evil" is, the very need to judge actions in a way that is communicable to others provides an orientation. In politics, good actions are not those whose motivations are pure, nor those that disregard the distinction between good and evil altogether, but rather those that invite the re-

actions and judgments of others. We do not know how our actions will be judged, but their exposure to the potentially infinite web of re-actions and judgments is a sign of goodness. Thus, the good often goes against all existing standards for distinguishing between good and evil—not because we possess an absolute standard that invalidates them, but rather because these standards preclude or constrain other re-actions and judgments, such as those of other communities in different places and times. As seen in chapters 3 and 4, we judge an action to be good not because we possess a rule for determining that it is good, but rather because we want to communicate the feeling that it arouses in us to everyone, even beyond those around us. It is this feeling that we need to judge actions in a way that is communicable to everyone, even if we find no rule by which to do it, that orients our capacity to draw a distinction between good and evil.

These conclusions intersect with contemporary debates around the moral foundations of political action and judgment. Specifically, this dissertation contributes to recent efforts to move away from theories that conceive political action and judgment under the model of autonomy, that is, as based on universal procedures that can be applied to concrete actions. One prominent example in this regard in John Rawls's *A Theory of Justice*. Relying on an interpretation of Kant, Rawls claims that to act autonomously means to act on principles of justice that one would choose independently of contingent desires and social position. Rawls deduces these principles by means of what he calls the "veil of ignorance," which consists in subtracting from individuals any knowledge of their natural endowments, specific ends in life, and position in society. Under these conditions, individuals would choose to ground institutions and actions upon principles that do not favor certain individuals over others based on contingent circumstances. Because

individuals under the veil of ignorance are completely equal, they would all agree on these principles. Although evidently we are never behind the veil of ignorance, Rawls believes that to act on the basis of principles of justice expresses our capacity to act as free and equal rational beings, independently of natural and social contingencies: "thus men exhibit their freedom, their independence from the contingencies of nature and society, by acting in ways they would acknowledge in the original position."[1] To act justly is therefore to act on the basis of principles that all individuals in a position of complete equality would agree with. To act unjustly, by contrast, is to let ourselves belong to a "lower order," "as though we were a creature whose first principles are decided by natural contingencies."[2]

The problem with Rawls's perspective is that it disregards the problematic relationship between the noumenal and the phenomenal realms in Kant, which I have analyzed in chapters 2 and 3. Rawls explicitly abandons what he calls the "dualism" between the two realms so as to fill in the "missing part" in Kant's understanding of the relationship between the noumenal and the phenomenal self, namely, the "concept of expression." When we act on the basis of principles of justice, Rawls claims, we express our noumenal self, that is, our self as a free and equal person, better than when we act on the basis of unjust principles. Thus, the principles of justice deduced by means of the "veil of ignorance" provide a "procedural interpretation of Kant's conception of autonomy and the categorical imperative within the framework of an empirical theory," and as a consequence, "no longer are these notions purely transcendent and lacking explicable connections with human conduct."[3] Chapters 2 and 3 have shown Kant's own inquiries into the

[1] John Rawls, *A Theory of Justice* (Cambridge: Harvard UP, 1999), 225.
[2] Ibid
[3] Ibid., 226.

problem of the link between the noumenal and the phenomenal self. We have seen that Kant insistently denied that there could be such thing as a "procedure" for connecting one to the other, because the moral choice that constitutes our noumenal character cannot be subordinated to a rule. This has important practical implications. Rawls's attempt to identify freedom with an empirical procedure overlooks the true ground of freedom, which is the choice of whether or not to comply with a given rule for action, rather than the rule itself. The view that freedom stems from following a certain rule leads us to disavow those instances where our moral character is put into question, and where self-transformation is more decisive than following a procedure.

A more nuanced, yet also problematic conception of political action in terms of autonomy is presented by Rainer Forst in *The Right to Justification*. Unlike Rawls, Forst locates justice in the practice of justifying our actions by explaining the reasons that underlie them, as opposed to complying with a fixed set of reasons. This produces a more dynamic interpretation of Kant's notion of autonomy, according to which "autonomous human beings formulate their moral and political judgments independently and critically evaluate them with the practice; at the same time, they are also required to justify those judgments."[4] By acknowledging their duty to justify their actions, human beings emancipate themselves from arbitrary rule, and contribute to build a society that is more just. Yet while Forst seeks to ground justice upon the practice of justification, rather than upon fixed principles of justice, he relies at the end on fixed rules for validating judgments of justification. In order to distinguish between the justifiable and the unjustifiable, Forst appeals to the principles of reciprocity ("no one may refuse the particular demands of others that one raises for oneself") and generality ("reasons for generally valid basic

[4] Rainer Forst, *The Right to Justification*, trans. Jeffrey Flynn (New York: Columbia UP, 2007), 7.

norms must be sharable by all those affected").[5] Clearly, these principles for validating practices of justification once again fill in the ground of action and judgment with a procedure. If the morality of action depends on its justification, we need a standard for distinguishing the justifiable from the unjustifiable, and thus once again we subordinate action to a set of rules. The question is, again, how to know that we are complying with the right rules, and for the right reasons—that is, that we are not simply adapting to whatever rules are available to express our "autonomy."

The problem with theories of justice in terms of justification is that they assume a correspondence between action and the moral character that underlies them. Chapters 2 and 3 have shown that a consistent reading of Kant does not allow for such correspondence. There is no procedure for determining whether an action is moral or not, because the moral character that determines the act is inaccessible to experience. The attempt to overcome this gap by means of a procedure is not only theoretically misleading, but also practically dangerous. The idea that a certain procedure can guarantee that we are acting justly undermines the need to examine the moral worth of our actions, and to be prepared to transform ourselves. As we have seen in chapter 2, it is this capacity for self-transformation that grounds our sense of responsibility. We never know whether we have acted justly for the right reasons and, consequently, we never know our moral character, nor that of others. This lack of knowledge, or what we could call moral uncertainty, is not a practical problem to be overcome by means of philosophical analysis, but rather a constitutive characteristic of action and judgment. Therefore, the theoretical challenge is

[5] Ibid., 6.

not to find a procedure that overcomes this uncertainty, but rather to understand how political actors respond to it.

Theorists of modern democracy have been more attentive to the constitutive uncertainty of action and judgment than theorists of justice and justification. Claude Lefort famously linked modern democracy to the "dissolution of the markers of certainty," as a consequence of which societies experience a fundamental indeterminacy regarding the basis of power, knowledge, and law. Democracy, according to Lefort, is not merely an institutional arrangement, but rather a new symbolic order in which "the locus of power becomes *an empty space*."[6] This means that the space of power is not occupied by any individual or group, but rather left indeterminate, so that "the exercise of power is subject to the procedures of periodical redistribution."[7] Importantly, this implies a "disentangling of the sphere of power, the sphere of law and the sphere of knowledge," by which the exercise of power, the generation of truth, and the legitimation of law become independent from one another. Thus, democracy implies a fundamental indeterminacy that affects all spheres of life, and which is actively protected and preserved by institutions that assure that the place of power is not occupied. Totalitarianism, Lefort claims, emerges as a reaction to this uncertainty, which leads to the "quest for a substantial identity, for a social body which is welded to its head, for an embodying power, for a state free from division."[8] The distinction between democracy and totalitarianism thus overlaps with the distinction between uncertainty and (fantasized) certainty.

---

[6] Claude Lefort, *Democracy and Political Theory*, trans. David Macey (Cambridge: Polity, 1988)., 17.
[7] Ibid.
[8] Ibid., 20.

Lefort's interpretation of totalitarianism connects in part to the one proposed in this dissertation, but with important differences. Lefort is right that totalitarianism represents a reaction against uncertainty, with the ensuing attempt to restore a lost unity. But his analysis remains at the level of the totalitarian ideology, focusing on party leaders who actively seek to occupy the place of power. At a general level, however, the strive for certainty does not necessarily take the form of a systematic ideology with a totalizing social view characteristic of totalitarianism. As Arendt shows in her analysis of Eichmann's trial, as well as in her later writings on moral philosophy, totalitarianism relies on functionaries who are relatively indifferent to the ideas motivating their actions, provided that these ideas grant them a sense of moral worth. Those who refuse uncertainty are the most willing to believe in any ideology that provides them with a rule that distinguishes good and evil. The reason why totalitarianism can mobilize entire populations of otherwise ordinary people into massive projects of systematic violence and extermination is that many of these people are simply looking for something to believe in, that is, something that regulates the uncertainty that they experience with each action and judgment. On this point, Simona Forti is right in linking totalitarianism to what she calls "the normality of evil," namely, an evil that is complicit with the everyday norms of mass societies, and which does not require any belief or ideological identification. In Forti's words, "we internalize imperatives that dictate 'what ought to be,' and we adapt seamlessly to the norms, so much so that… our lives now seem programmed to become prey to the petty seduction of consumerism and the society of the spectacle."[9] However, as Forti aptly clarifies, this internalization is an effect of our own desire for normality, for "if the new normativity for the

[9] Simona Forti, *New Demons: Rethinking Evil and Power Today*, trans. Zakiya Hanafi (Stanford, CA: Stanford UP, 2015), 318.

optimization of life has managed to prevail along with new forms of acquiescence and indifference, this is because we not only accept it, we seek to implement it at all cost."[10] Indeed, as we have seen in chapters 1 and 2, the line separating political indifference and ideological extremism is much thinner than it looks at first sight. At the level of everyday citizens and functionaries, extremism does not stem from the attachment to one set of beliefs, but rather from the desire for any set of beliefs whatsoever that regulates our actions and judgments.

The problem of "the banality of evil," which has concerned us throughout this dissertation, points to the generalized acquiescence and indifference mentioned by Forti. Although the analysis of parties and social identities is obviously important for understanding totalitarianism, an analysis of action and judgment at its everyday level is essential. In order to understand not only the emergence of totalizing ideologies, but also the acquiescence of ordinary people to these ideologies, it is necessary to understand why people become acquiescent to begin with. One reason, as we have seen, is that the uncertainty inherent to action and judgment produces a desire to withdraw from action and judgment altogether. With the "dissolution of the markers of certainty," this uncertainty becomes an everyday experience, for there is no unquestionable authority that determines whether we act justly or unjustly. One possible response to this uncertainty consists in covering it up with rules and procedures that make the meaning of what we are doing immediately accessible to us. At this level, belief in an ideology is much less important than thoughtless compliance with it. As Arendt shows in her analysis of totalitarianism, the perfect totalitarian individual is not the one who consents with the totalitarian ideology, but rather the one who complies without even thinking about it—consensus implies

[10] Ibid.

freedom; thoughtless compliance does not. This means that those who thoughtlessly submit to totalitarianism are the ones who refuse the uncertainty that is inherent to action and judgment, in such a way that they recognize no spontaneity in what they do. Uncertainty, that is, the fact that we have no direct access to the meaning of our actions, is essential to freedom and responsibility: if there were a complete explanation for our actions, it would replace their freedom with the procedure that we use to explain them. Consequently, the attempt to overcome uncertainty amounts to the undermining of our own freedom, and thus of our sense of responsibility.

The question that follows from this analysis is: what are we to do with this uncertainty? Does the acknowledgement that we cannot judge actions with certainty amount to an impossibility to judge altogether? Chapters 3 and 4 have shown that this is not the case. Even if there is no rule by which the judge whether an action is good or evil, the very feeling aroused by the need to judge when we contemplate certain actions works as a ground for judgment. Evidently, this is not a ground in the sense of a procedure that guarantees that a judgment is valid, provided that it is applied correctly. Rather, the ground of validity is the expectation that our judgment will be shared by everyone who has a sense of morality. As we have seen in chapter 4, the very desire to communicate our judgment to everyone is a sign that it is valid, even if there is no possible confirmation of this validity. This is only a problem if we believe that only a procedure can work as a standard of validity. If, by contrast, the validity of judgments is decided by their communicability to a potentially infinite public, then the very expectation of agreement can work as a sign of validity. Unlike the procedure, the sign does not stem from any rule that can be objectified and applied, and therefore there is no definitive way to settle the dispute on whether the judgment is valid or not. As long as we feel the need to judge, however,

we can expect that others will share our feeling. The only measure of the judgment's validity is the infinite communicability of this feeling, the fact that it does not stop with those around us, but addresses everyone who is receptive to moral ideas. The very feeling that we must judge the French Revolution as good, or those who protected people from Nazi persecution, or the fight against the Apartheid, and that we must communicate our judgment to everyone, is already a sign that it is valid, even if we possess no unequivocal rule to determine that those who acted are morally good.

This conception of the universality of judgment departs from the ones proposed by some prominent democratic theorists. Without specifically addressing the problem of judgment, Jacques Rancière claims that universality is an effect of the political contestation of the order of the police, namely, of the established relationships between social groups. This order produces a wrong by excluding certain groups from the field of visibility, that is, by not counting them as part of the community to begin with. Politics, according to Rancière, consists in the contestation of the police order through the demand of equality by "the part of those who have no part": "politics is the practice whereby the logic of the characteristic of equality takes the form of the processing of a wrong, in which politics becomes the argument of a basic wrong that ties in with some established dispute in the distribution of jobs, roles, and places."[11] This processing of a wrong by contesting the police order institutes what Rancière calls a "singular universal," by means of "tying the presentation of equality, as the part of those who have no part, to the conflict between parts of society."[12] Given that this "singular universal" stems from the contestation of

[11] Jacques Rancière, *Disagreement: Politics and Philosophy*, trans. Julie Rose (Minneapolis, MN: Minnesota UP, 1999), 35.
[12] Ibid., 39.

existing regulations, it cannot be derived from a regulation. Therefore, Rancière claims, "the universal is always at stake here in a peculiar way, in the form of cases whereby its very existence and pertinence are in dispute."[13] Paradoxically, universality is never given as an object of consensus, but always as an object of dispute. In other words, the universal is the dispute on what is universal, on the principle that regulates the parts that are counted in a community.

Rancière is right in linking political universality to singular disputes as opposed to fixed procedures, but his analysis neglects the problem of judgment. The contrast between police and politics suggests that any act that challenges the established order of identities is by itself universal, because it is tied to the presentation of equality. But is the dispute of the existing distribution of identities by itself a struggle for equality? Here, we run the risk of simply inverting the relationship between consensus and heterogeneity, shifting universality from the former to the latter, as if any act of dis-identification were by itself universal. In order to make a universally valid judgment upon an action, it is not enough that the action be disruptive of the order of established identities. It is also necessary, as we have seen in chapters 3 and 4, that the action arouses moral ideas in the spectators, in such a way that they feel confidence in the moral constitution of humanity. This is why it is important, following both Arendt and Lyotard, to consider the relationship between action and judgment. As Rancière points out, a universalist action, that is, an action that demands universal approval, is necessarily disruptive, because ideas cannot be represented by any objective procedure. But whether the disruption is a sign of ideas can only be decided in the judgment of the spectators. Universalist actions summon the spectators to judge on the basis of their moral feeling, and there is no standard (not even the

[13] Ibid.,56.

absence of a standard) other than this feeling by which to judge. Consequently, the universal validity of an action can never be decided by the action itself, but only by the open-ended relationship between action and judgment

On this point, it is helpful to consider Shalini Satkunanandan's distinction between "calculable" and "incalculable" responsibility. Calculable responsibility stems from specific moral rules that define what we ought to do, while incalculable responsibility stems from our attentiveness to moral rules. According to Satkunanandan, moral considerations always involve a certain rule-based calculus by means of which we judge our practical endeavors. But in order to engage in moral considerations to begin with, it is necessary that we give attention to moral rules. Through readings of Plato, Kant, Nietzsche, Heidegger, and Weber, Satkunanandan shows that even if moral rules are necessary, the experience by which we give attention to these rules and become responsible is different than the procedure of applying them to actions. This does not mean, however, that we should disregard morality altogether, as if somehow we were capable of dwelling in our moral attentiveness without ever relying on rules: "rules themselves are not the problem. The problem is the relationship between rule-following and attentiveness."[14] In order to remain attentive, that is, to not lose our attention in the strict compliance with rules, we must give attention both to rules and to that which is concealed by rules: "confronted by calculable responsibility, we do not refuse or resist it but rather remain on the lookout for a detail that is overlooked, a nuance that is flattened, a need or a claim that is denied or cut short."[15] Interpreting Arendt's analysis of Eichmann under this light, Satkunanandan claims that

[14] Shalini Satkunanandan, *Extraordinary Responsibility: Politics Beyond the Moral Calculus* (Cambridge: Cambridge UP, 2015), 189.
[15] Ibid.

Eichmann's "thoughtlessness" is precisely a lack of attentiveness, namely, an incapacity to give attention to anything that cannot be framed within the application of rules. Therefore, in order to avoid doing evil, it is important that we cultivate attentiveness and avoid falling into blind compliance with rules.

Satkunanandan's analysis is enlightening in many regards, but it is important to avoid considering attentiveness as a kind of faculty that we may either cultivate or lose. If this were the case, we would fall back into the image of evil as deficiency, as if lacking a mental faculty would lead us to "inadvertently" do evil. If we are unavoidably responsible for evil, as we saw in the first part of this dissertation, it is because whatever we seem to lack is an effect of our own choice. Attentiveness is not an asset that we may either cultivate or otherwise lose, but rather an effect of choice. Each action implies a choice regarding the rule that will determine it, and it is this choice cannot in its turn be derived from a rule. However, one can choose to conceal one's own choice, that is, to act as if one's actions were nothing but the effect of rules that are given from without. It is at this level that modern ideologies seek the active complicity of individuals determined to escape the uncertainty that comes with responsibility. And it is this active complicity with the destruction of one's own responsibility, rather than an incapacity to take responsibility, that makes extreme evil possible. Therefore, the fight against political evil cannot consist primarily in preserving or cultivating a faculty or a capacity. Before this cultivation is even possible, one must choose whether to face the uncertainty that underlies any action and any judgment, or otherwise to hold fast to rules. To face uncertainty is to throw ourselves into the web of actions and judgments, accepting that who we are will be decided by the relationship we establish with others. Whenever we choose a course of action, we do not know whether it is the

right one, yet insofar as we accept this uncertainty, we also accept the responsibility that comes with action. To accept responsibility for this choice does not guarantee that we will do the right thing, but it makes unlikely that we will do the worst thing.

Bibliography

Adorno, Theodor W. *Negative Dialectics*. Translated by E. B. Ashton. New York: Routledge, 2004.

Allison, Henry E. *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy*. New York: Cambridge UP, 1996.

———. *Kant's Theory of Freedom*. New York: Cambridge UP, 1990.

———. *Kant's Theory of Taste: A Reading of the Critique of Aesthetic Judgment*. Cambridge: Cambridge UP, 2001.

Arendt, Hannah. "Basic Moral Propositions," n.d. Courses - The University of Chicago, Series: Subject File, 1949-1975. The Hannah Arendt Papers, Manuscript Division, Library of Congress, Washington D.C.

———. *Between Past and Future*. New York: Penguin, 2006.

———. *Denktagebuch*. Edited by Ursula Ludz and Ingeborg Nordmann. Munich: Piper, 2002.

———. *Eichmann in Jerusalem: A Report on the Banality of Evil*. New York: Penguin, 2006.

———. *Essays in Understanding*. Edited by Jerome Kohn. New York: Schocken, 1994.

———. *Lectures on Kant's Political Philosophy*. Edited by Ronald Beiner. Chicago: The University of Chicago Press, 1992.

———. *On Revolution*. New York: Penguin, 1965.

———. "Political Theory of Kant," n.d. Subject File, 1949-1975, Hannah Arendt Papers, Manuscript Division, Library of Congress, Washington D.C.

———. *Responsibility and Judgment*. Edited by Jerome Kohn. New York: Schocken, 2003.

———. *The Human Condition*. Chicago: The University of Chicago Press, 1998.

———. *The Jewish Writings*. Edited by Jerome Kohn and Ron H. Feldman. New York: Schocken, 2007.

———. *The Life of the Mind*. Orlando, FL: Harcourt, 1978.

———. *The Origins of Totalitarianism*. New York: Harcourt Brace & Company, 1976.

———. *Vita activa, oder Vom tätigen Leben*. Stuttgart: R. Piper, 1960.

Arendt, Hannah, and Karl Jaspers. *Correspondence: 1926-1969*. Edited by Lotte Kohler and Hans Saner. New York: Harcourt Brace Javanovich, 1992.

Aristotle. *Ethica Nicomachea*. In *The Basic Works of Aristotle*, edited by Richard McKeon. New York: Random House, 2001.

Beck, Lewis White. *A Commentary on Kant's Critique of Practical Reason*. Chicago: The University of Chicago Press, 1984.

———. *Early German Philosophy: Kant and His Predecessors*. Cambridge: Harvard UP, 1969.

———. “Five Concepts of Freedom in Kant.” In *Stephan Körner – Philosophical Analysis and Reconstruction*, edited by Jan T. J. Srzednicki. Dordrecht, the Netherlands: Martinus Nijhoff, 1987.

Beiner, Ronald. “Interpretative Essay.” In *Lectures on Kant’s Political Philosophy*, by Hannah Arendt, edited by Ronald Beiner. Chicago: The University of Chicago Press, 1992.

———. “Kant, the Sublime, and Nature.” In *Kant and Political Philosophy*, edited by Ronald Beiner and William James Booth. Ann Arbor, MI: Yale UP, 1993.

———. “Rereading Hannah Arendt’s Kant Lectures.” In *Judgment, Imagination, and Politics: Themes from Kant and Arendt*, edited by Ronald Beiner and Jennifer Nedelsky. Oxford: Rowman & Littlefield, 2001.

Benhabib, Seyla. “Judgment and the Moral Foundations of Politics in Hannah Arendt’s Thought.” In *Judgment, Imagination, and Politics: Themes from Kant and Arendt*, edited by Ronald Beiner and Jennifer Nedelsky. Oxford: Rowman & Littlefield, 2001.

———. *The Reluctant Modernism of Hannah Arendt*. Lanham, MD: Rowman & Littlefield, 2003.

Bennington, Geoffrey. *Lyotard: Writing the Event*. New York: Columbia UP, 1988.

Bernstein, Richard J. “Did Hannah Arendt Change Her Mind? From Radical Evil to the Banality of Evil.” In *Hannah Arendt: Twenty Years Later*, edited by Larry May and Jerome Kohn. Cambridge: Cambridge UP, 1999.

———. *Philosophical Profiles: Essays in Pragmatic Mode*. Philadelphia: University of Philadelphia Press, 1986.

———. *Radical Evil: A Philosophical Interrogation*. Cambridge: Polity, 2008.

Bilsky, Leora Y. “When Actor and Spectator Meet in the Courtroom: Reflections on Hannah Arendt’s Concept of Judgment.” In *Judgment, Imagination, and Politics: Themes from Kant and Arendt*, edited by Ronald Beiner and Jennifer Nedelsky. Oxford: Rowman & Littlefield, 2001.

Brecht, Martin. “August Hermann Francke und der Hallische Pietismus.” In *Geschichte des Pietismus*. Göttingen, Germany: Vandenhoeck & Ruprecht, 1993.

Canovan, Margaret. *Hannah Arendt: A Reinterpretation of Her Political Thought*. Cambridge: Cambridge UP, 1992.

Cassirer, Ernst. *Kant’s Life and Thought*. Translated by James Haden. New Haven, CT: Yale UP, 1981.

Clark, Christopher. *Iron Kingdom: The Rise and Downfall of Prussia, 1600-1947*. Cambridge: Harvard UP, 2006.

Clewis, Robert R. *The Kantian Sublime and the Revelation of Freedom*. Cambridge: Cambridge UP, 2009.

Cloutier, Sophie. “La question du mal chez Hannah Arendt : rupture ou continuité?” *PhaenEx* 3, no. 1 (2008): 82–111.

Connolly, William E. *Why I Am Not a Secularist*. Minneapolis, MN: Minnesota UP, 1999.

Crowther, Paul. *The Kantian Sublime: From Morality to Art*. Oxford: Clarendon Press, 1989.

Despland, Michel. *Kant on History and Religion*. London: McGill - Quenn's University Press, 1973.

Dietz, Mary. "Arendt and the Holocaust." In *The Cambridge Companion to Hannah Arendt*, edited by Dana Villa. New York: Cambridge UP, 2002.

Diprose, Rosalyn. "Arendt and Nietzsche on Responsibility and Futurity." *Philosophy and Social Criticism* 34, no. 6 (n.d.).

Dostal, Robert J. "Judging Human Action: Arendt's Appropriation of Kant." *The Review of Metaphysics* 37, no. 4 (1984): 725–55.

Fenves, Peter. *Late Kant: Towards Another Law of the Earth*. New York: Routledge, 2003.

———. "Taking Stock of the Kantian Sublime," *Eighteenth-Century Studies* 28, no. 1 (1994): 65–82.

Ferrarin, Alfredo. "Imagination and Judgment in Kant's Practical Philosophy." *Philosophy and Social Criticism* 34, no. 1–2 (2008): 101–21. https://doi.org/10.1177/0191453707084276.

Forst, Rainer. *The Right to Justification*. Translated by Jeffrey Flynn. New York: Columbia UP, 2007.

Forti, Simona. *New Demons: Rethinking Evil and Power Today*. Translated by Zakiya Hanafi. Stanford, CA: Stanford UP, 2015.

Fulbrook, Mary. *Piety and Politics. Religion and the Rise of Absolutism in England, Wüttemberg and Prussia*. Cambridge: Cambridge UP, 1983.

Gasché, Rodolphe. *The Idea of Form: Rethinking Kant's Aesthetics*. Stanford, CA: Stanford UP, 2002.

Gawthrop, Richard. *Pietism and the Making of Eighteenth-Century Prussia*. Cambridge: Cambridge UP, 1993.

Guyer, Paul. "Kant's Distinction between the Beautiful and the Sublime." *The Review of Metaphysics* 35, no. 4 (1982): 753–83.

Heidegger, Martin. *Being and Time*. Translated by John Macquarrie and Edward Robinson. Oxford: Basil Blackwell, 1985.

Henrich, Dieter. "Der Begriff der sittlichen Einsicht und Kants Lehre vom Faktum der Vernunft." In *Kant: Zur Deutung seiner Theorie von Erkennen und Handeln*, edited by Geold Prauss. Koln, Germany: Krepenheuer & Witsch, 1973.

Honig, Bonnie. *Political Theory and the Displacement of Politics*. Ithaca, NY: Cornell UP, 1993.

Hunter, Ian. *Rival Enlightenments: Civil and Metaphysical Philosophy in Early Modern Germany*. Cambridge: Cambridge UP, 2003.

Ingram, David. “The Postmodern Kantianism of Arendt and Lyotard.” In *Judging Lyotard*, edited by Andrew Benjamin. New York: Routledge, 1992.

Jaspers, Karl. *Rechenschaft und Ausblick. Rede und Aufsätze*. Munich: R. Piper, 1958.

———. *The Question of German Guilt*. Translated by E. B. Ashton. New York: Fordham UP, 2000.

Kant, Immanuel. *Critique of Practical Reason*. In *Practical Philosophy*, edited by Mary J. Gregor. New York: Cambridge UP, 1999.

———. *Critique of Pure Reason*. Edited by Paul Guyer and Allen W. Wood. New York: Cambridge UP, 2000.

———. *Critique of the Power of Judgment*. Edited by Paul Guyer. Translated by Paul Guyer and Eric Matthews. Cambridge: Cambridge Uu, 2000.

———. *Groundwork of the Metaphysics of Morals*. In *Practical Philosophy*, edited by Mary J. Gregor. New York: Cambridge UP, 1999.

———. *Lectures on Ethics*. Edited by Peter Heath and J. B. Schneewind. New York: Cambridge UP, 1997.

———. *Religion Within the Boundaries of Mere Reason*. In *Religion and Rational Theology*, edited by Allen W. Wood and George Di Giovanni. New York: Cambridge UP, 2001.

———. *The Conflict of the Faculties*. In *Religion and Rational Theology*, edited by Allen W. Wood and George Di Giovanni, translated by Mary J. Gregor and Robert Anchor. Cambridge: Cambridge UP, 2001.

Kateb, George. *Hannah Arendt: Politics, Conscience, Evil*. Totowa, NJ: Rowman and Allanheld, 1984.

Kawamura, Katsutoshi. *Spontaneität und Willkür: Der Freiheitsbegriff in Kants Antinomienlehre und seine historischen Wurzeln*. Stuttgart: Frommann Holzboog, 1996.

Kuehn, Manfred. *Kant: A Biography*. New York: Cambridge UP, 2001.

Lefort, Claude. *Democracy and Political Theory*. Translated by David Macey. Cambridge: Polity, 1988.

Leibniz, G. W. *Theodicy: Essays on the Goodness of God, the Freedom of Man and the Origin of Evil*. Edited by Austin Farrer. Translated by E. M. Huggard. New York: Biblio Bazaar, 1985.

Leibniz, G. W., and Samuel Clarke. *Correspondence*. Edited by Roger Ariew. Indianapolis, IN: Hackett, 2000.

Lestition, Steven. “Kant and the End of the Enlightenment in Prussia.” *The Journal of Modern History* 65, no. 1 (1993): 57–112.

Lyotard, Jean-François. *Enthusiasm: The Kantian Critique of History*. Translated by Georges Van Den Abbeele. Stanford, CA: Stanford UP, 2009.

———. *Lessons on the Analytic of the Sublime*. Translated by Elizabeth Rottenberg. Stanford, CA: Stanford UP, 1994.

———. *Political Writings*. Translated by Bill Readings and Kevin Paul Geiman. London: University College London Press, 1993.

———. *The Differend: Phrases in Dispute*. Translated by Georges Van Den Abbeele. Minneapolis, MN: Minnesota UP, 1989.

———. *The Postmodern Condition: A Report on Knowledge*. Translated by Geoffrey Bennington and Brian Massumi. Minneapolis, MN: Minnesota UP, 1984.

Lyotard, Jean-François, and Jean-Loup Thébaud. *Just Gaming*. Translated by Wlad Godzich. Minneapolis, MN: Minnesota UP, 1985.

Makkreel, Rudolf. *Imagination and Interpretation in Kant: The Hermeneutical Import of the* Critique of Judgment. Chicago: The University of Chicago Press, 1994.

Marshall, David. "The Origin and Character of Hannah Arendt's Theory of Judgment." *Political Theory* 38, no. 3 (2010): 367–93. https://doi.org/10.1177/0090591709359595.

Mathewes, Charles T. *Evil and the Augustinian Tradition*. Cambridge: Cambridge UP, 2004.

Matthews, Patricia M. "Kant's Sublime: A Form of Pure Aesthetic Reflective Judgment." *The Journal of Art and Criticism* 54, no. 2 (1996): 165–80.

Menke, Christoph. "Auf der Grenze des Rechts: Hannah Arendts Revision des Eichmann-Prozesses." *Merkur* 67 (n.d.): 227–56.

Michalson, Gordon E. *Fallen Freedom: Kant on Radical Evil and Moral Regeneration*. Cambridge: Cambridge UP, 1990.

Munzel, G. Felicitas. *Kant's Conception of Moral Character: The "Critical" Link of Morality, Anthropology, and Reflective Judgment*. Chicago: The University of Chicago Press, 1999.

Nahm, Milton C. "'Sublimity' and the 'Moral Law' in Kant's Philosophy." *Kant-Studien* 48 (1956): 502–24.

Nancy, Jean-Luc. "Dies irae." In *La faculté de juger*, by Jacques Derrida, Vinvent Descombes, Garbis Kortian, Philippe Lacoue-Labarthe, Jean-François Lyotard, and Jean-Luc Nancy. Paris: Éditions de Minuit, 1985.

———. *L'impératif catégorique*. Paris: Flammarion, 1983.

———. *The Experience of Freedom*. Translated by Bridget McDonald. Stanford, CA: Stanford UP, 1988.

Neiman, Susan. *Evil in Modern Thought: An Alternative History of Philosophy*. Princeton, NJ: Princeton UP, 2004.

Norris, Andrew. "Arendt, Kant, and the Politics of Common Sense." *Polity* 29, no. 2 (1996): 165–91.

Ophir, Adi. "Between Eichmann and Kant: Thinking on Evil after Arendt." *History and Memory* 8, no. 2 (1996): 89–136.

Piché, Claude. “The Philosopher-Artist: A Note on Lyotard’s Reading of Kant.” *Research in Phenomenology* 22, no. 1 (1992): 152–60. https://doi.org/10.1163/156916492X00133.

Pitkin, Hanna Fenichel. *The Attack of the Blob: Hannah Arendt’s Concept of the Social*. Chicago: The University of Chicago Press, 1998.

Plato. *The Republic of Plato*. Translated by Allan Bloom. New York: Basic Books, 1991.

Plotinus. *The Enneads*. Edited by John Dillon. Translated by Stephan MacKenna. New York: Penguin, 1991.

Rancière, Jacques. *Disagreement: Politics and Philosophy*. Translated by Julie Rose. Minneapolis, MN: Minnesota UP, 1999.

Rawls, John. *A Theory of Justice*. Cambridge: Harvard UP, 1999.

Reboul, Olivier. *Kant et le problème du mal*. Montreal: Montreal UP, 1971.

Reinhold, Carl Leonhard. “Erörterung des Begriffs von der Freiheit des Willens.” In *Materialen Zu Kants “Kritik Der Praktischen Vernunft,”* edited by Rudiger Bittner and Konrad Cramer. Frankfurt am Main: Suhrkapf, 1975.

Ruffing, Margit, ed. *Kant-Bibliographie 1945-1990*. Frankfurt am Main: Vittorio Klostermann, 1999.

Saint Augustine. *City of God*. Translated by Henry Bettenson. London: Penguin, 2003.

Satkunanandan, Shalini. *Extraordinary Responsibility: Politics Beyond the Moral Calculus*. Cambridge: Cambridge UP, 2015.

———. “The Extraordinary Categorical Imperative.” *Political Theory* 39, no. 2 (2011): 234–60. https://doi.org/10.1177/0090591710394124.

Schaller, Klaus. “Pietismus und moderne Pädagogik.” In *Pietismus und moderne Welt*, edited by Kurt Aland. Witten, Germany: Luther-Verlag, 1974.

Schaper, Eva. “Taste, Sublimity, and Genius: The Aesthetics of Nature and Art.” In *The Cambridge Companion to Kant*, edited by Paul Guyer. Cambridge: Cambridge UP, 2006.

Schneewind, J. B. *The Invention of Autonomy: A History of Modern Moral Philosophy*. New York: Cambridge UP, 1998.

Schupmann, Benjamin A. “Thoughtlessness and Resentment: Determinism and Moral Responsibility in the Case of Adolf Eichmann.” *Philosophy and Social Criticism* 40, no. 2 (2014): 127–44. https://doi.org/10.1177/0191453713518323.

Schwartz, Jonathan P. *Arendt’s Judgment: Freedom, Responsibility, Citizenship*. Philadelphia: University of Philadelphia Press, 2016.

Shell, Susan Meld. *Kant and the Limits of Autonomy*. Cambridge: Harvard UP, 2009.

Spinoza, Baruch. *The Letters*. Translated by Samuel Shirley. New York: Hackett, 1995.

Spinoza, Benedict de. *Theological-Political Treatise*. Edited by Jonathan Israel. Translated by Michael Silverthrone and Jonathan Israel. Cambridge: Cambridge UP, 2007.

Spener, Philipp Jackob. *Der neue Mensch*. Edited by Hans-Georg Feller. Stuttgart: J. F. Steinkopf, 1966.

Villa, Dana R. “Beyond Good and Evil: Arendt, Nietzsche, and the Aestheticization of Political Action.” *Political Theory* 20, no. 2 (1992): 274–308.

———. *Politics, Philosophy, Terror: Essays on the Thought of Hannah Arendt*. Princeton, NJ: Princeton UP, 1999.

Wallmann, Johannes. “Was ist Pietismut?” *Pietusmus und Neuzeut* 20 (1994): 11–27.

Watkins, Eric, ed. *Kant’s* Critique of Pure Reason*: Background Source Materials*. New York: Cambridge UP, 2009.

Williams, Garrath. “Disclosure and Responsibility in Arendt’s *The Human Condition*.” *European Journal of Political Theory* 14, no. 1 (2014): 37–45. https://doi.org/10.1177/1474885114522187.

Williams, James. *Lyotard and the Political*. New York: Routledge, 2000.

Wood, Allen W. *Kant’s Moral Religion*. Ithaca, NY: Cornell UP, 1970.

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

Zerilli, Linda M. G. “‘We Feel Our Freedom’: Imagination and Judgment in the Thought of Hannah Arendt.” *Political Theory* 33, no. 2 (2005): 158–88. https://doi.org/10.1177/0090591704272958.

Zupančič, Alenka. *Ethics of the Real: Kant, Lacan*. London: Verso, 2000.