From Trilemma Towards Tolerance:
A Comprehensive Approach to
Containing Unreasonableness
in Political Liberalism

Author: Marinus Leonard Eigenhuis
Date: June 16th, 2024
Supervisor: Dr. Jos Philips
Department: Philosophy & Religious Studies
Faculty of Humanities
Utrecht University

From Trilemma Towards Tolerance:
A Comprehensive Approach to Containing Unreasonableness in Political Liberalism

Marinus Leonard Eigenhuis

Utrecht University

Abstract

John Rawls' political theory has profoundly shaped the discourse in contemporary political philosophy. However, recognizing the recent shifts in political trends and increasing illiberal challenges to liberal democracy, I present in this thesis project a critical re-evaluation of Rawlsian political liberalism. I explore the rise of unreasonable political attitudes among citizens in liberal democracies, which poses a challenge to political liberalism. I survey existing responses within liberal political philosophy—conciliation, transformation, and marginalization—and highlight their limitations in addressing what I call 'the trilemma of pluralism': balancing the ideals of legitimacy, stability, and inclusivity given an enduring and significant presence of unreasonableness in liberal society. Building on this analysis, I propose a multifaceted and comprehensive strategy to contain the political influence of unreasonable doctrines and citizens while addressing the complexities of the trilemma.

In particular, I will seek to advance the ongoing academic debate by emphasizing seven points specifically in this thesis project. Firstly, that the epistemic elements of reasonableness warrant the inclusion the additional basic aspect modeled of the 'duty of prudence'. Secondly, that recent findings from moral psychological research imply that political liberalism should postulate simple pluralism instead of reasonable pluralism. Thirdly, that the challenges posed by rising unreasonableness are interconnected as modeled by the trilemma and should not be solved separately but in an integrated manner. Fourthly, that the current binary model of unreasonable citizens is too reductive and ought to be refined with four significant subcategories: 'uncooperative,' 'unequitable,' 'anti-reasonable,' and 'unaware' citizens. Fifthly, that an approach of involvement towards partially reasonable citizens can enhance the inclusivity of political liberalism. Sixthly, that wholly unreasonable citizens can be provided with an adequate reason to dismiss their challenges to the legitimacy of the liberal order. Finally, that a 'duty of vigilance' is essential for bolstering the stability of liberal democracy, encouraging heightened civic engagement and awareness. Despite the serious challenges facing liberal democracy, I ultimately conclude that Rawlsian principles can provide valuable and practicable insights, thus remaining relevant and adaptable to contemporary political issues.

Acknowledgments

This thesis project could not have been completed without the invaluable help of many others. In particular, I would like to express my deep gratitude to the following people:

… Jos Philips, for his insightful feedback, his flexibility, and his steady guidance as the supervisor of this thesis project, and for his stimulating philosophical rigor while teaching one of the courses that inspired its topic.

… Pooyan Tamimi Arab, for having provided me with numerous thought-provoking interdisciplinary perspectives while teaching the other course that inspired this topic.

… Leon and Paulineke Eigenhuis, for their ceaseless support, faith, and encouragement throughout my entire educational journey, which has culminated in this treatise.

… Sam Eigenhuis, without whose inspiration I would not have started this thesis project.

… Minwhi Heo, without whose inspiration I could not have finished it.

Overview

Introduction

PART I: Establishing the Trilemma

    Chapter 1: Theoretical Framework & Rawls' Political Liberalism

    Chapter 2: Empirical Framework & Contemporary Moral Psychology

    Chapter 3: Critical Framework & Trade-offs Within Pluralism

PART II: Criticizing Current Responses

    Chapter 4: Conciliating the Unreasonable

    Chapter 5: Transforming the Unreasonable

    Chapter 6: Marginalizing the Unreasonable

PART III: Resolving the Trilemma

    Chapter 7: Inclusivity & Involvement

    Chapter 8: Legitimacy & Intelligibility

    Chapter 9: Stability & Vigilance

Conclusion

Bibliography

Table of Contents

Introduction

John Rawls is arguably the most influential political philosopher of the twentieth century, as can be illustrated in several ways.[1] His publications on justice and liberalism have garnered widespread and enduring academic attention, making him the most cited author in his field. His theories revitalized the analytic tradition of political philosophy and their influence spread to legal discussions. His technical vocabulary has been widely adopted by both advocates and critics. Furthermore, two of his pupils went on to become leaders of the oldest liberal democracy in the world. As Robert Nozick stated: *"Political philosophers must either work within Rawls's theory or explain why not."*[2] In light of this, this thesis project will operate within Rawlsian theory.

However, writing two decades after Rawls passed away, it should also be noted that the general political tide has shifted over time. Rawls wrote about liberal democracy during an era marked by several global waves of democratization and a sense of liberal optimism. This stands in stark contrast to the political developments of the new millennium, as historical momentum seems to have shifted. Contemporary liberal democracies face significant and growing challenges – not just externally by resurgent authoritarian states but also internally due to institutional decline and democratic backsliding.[3] Since the current era is characterized by a pervasive sense that liberalism is faltering, it is worth asking, as Katrina Forrester phrases it: *"For five decades Anglophone political philosophy has been dominated by the liberal egalitarianism of John Rawls. With liberalism in crisis, have these ideas outlived their time?"*[4] Given the present state of liberal democracy, Rawls' political philosophy is in need of a re-evaluation.

Considering the contemporary challenges to liberalism, a good place to start this re-evaluation is the concept and phenomenon of unreasonable citizens. Foundational to traditional liberal thought is the idea that, generally speaking, citizens have a capacity to become *reasonable*: to engage with each other on the basis of shared social norms, civility, and equal respect while keeping an open mind. However, as liberal democracies struggle with illiberal influences, an increasing share of citizens dismisses the value of being reasonable. These are the citizens that Rawls calls *unreasonable*. Their growing political presence poses a significant challenge to Rawls as it undermines the foundational assumptions upon which his theory operates – but moreover, it is interconnected with the institutional process of democratic backsliding.

An increasingly unreasonable citizenry is intertwined with the decline of constitutional liberal democracy in several ways.[5] When individual citizens reject the foundational principles of liberal democracy, they contribute to the erosion of those social norms that sustain civil discourse and democratic institutions. As democratic backsliding is linked with the weakening of the rule of law, this trend emboldens unreasonable citizens to become more assertive in promoting illiberal ideologies, thereby challenging social practices of tolerance and compromise. This fuels partisan responses that further entrench liberal-democratic decline as the space shrinks for constructive public

---

[1] Rasmussen & Den Uyl, "Did Rawls Restore Political Philosophy?"
Clark, "John Rawls revisited: Politics Behind the Veil".
[2] Nozick, *Anarchy, State, and Utopia*, p. 183.
[3] Freedman, "The Crisis of Liberalism", p. 37-43.
[4] Forrester, 'The Future of Political Philosophy'.
[5] Levitsky and Daniel Ziblatt, *How Democracies Die*, p. 1-11.

deliberation and reasonable consensus-building. Thus, democratic guardrails are tested on whether they weaken over time, which would allow unreasonable views to gain even more traction. Once the share of unreasonable citizens has reached a tipping point, illiberal-populist leaders gain serious opportunities to mobilize unreasonable movements directly against democratic institutions.

The interconnection between democratic erosion and a growing share of unreasonable citizens can be illustrated by referring to Rawls' home country: the United States was unable to accomplish a peaceful transfer of power following the 2020 presidential election for the first time in its history when a crowd of unreasonable citizens stormed the seat of its legislature.[6] This event exemplifies how a rising presence of unreasonableness can disrupt stability even in the most established liberal-democratic states. Thus, in light of the phenomenon of a growing share of unreasonable citizens, traditional theorists of liberal democracy have increasingly focused on a question that can be phrased from their perspective as: 'What to do with *them*?'[7]

In contemporary political philosophy, the debate over how to address the resurgence of unreasonable influences is ongoing. At the turn of the millennium, Erin Kelly and Lionel McPherson claimed that these movements pose little threat to stability and that efforts to manage them contradict the spirit of political liberalism.[8] Later on, the discussion developed and philosophers like Jonathan Quong argued for a more assertive stance, suggesting that the liberal order should restrict the basic rights of unreasonable citizens in particular circumstances, such as when they engage in hate speech.[9] Conversely, other scholars such as Fuat Gurzoslu have advocated for a transformational approach and a focus on education.[10] Despite these evolving perspectives, a consensus has not yet emerged. Recently, Aaron Ancell even remarked that those within the Rawlsian tradition *"have hardly begun to address"* the fundamental issue: achieving a just and stable liberal order in a pluralistic society with significant unreasonable influences.[11] Thus, my aim with this thesis project is to join this ongoing discussion and contribute by formulating a comprehensive response to the rising phenomenon of unreasonableness within the framework of political liberalism.

In general, my thesis project makes both theoretical and methodological contributions to the present academic debate. On the theoretical front, I offer an extensive review of current philosophical literature on unreasonableness, presenting a scope that has not been previously available. I also introduce a conceptual model that integrates the primary challenges currently dominating the critical discourse in the form of the 'trilemma of pluralism'. Methodologically, I conduct a thorough conceptual analysis focusing on reasonableness, and I also provide an interdisciplinary perspective through a literature overview of relevant moral psychological findings. Ultimately, this project concludes with a comprehensive strategy for engaging with unreasonable citizens and doctrines that has the potential to inform decision-making at the practical level of politics and initiatives to restore trust in liberal-democratic institutions.

---

[6] Crothers & Burgener, "Assessing the January 6 Attack on the U.S. Capitol", p. 129.
[7] Although, this question is by no means a new one. See Spinoza's discussions on the "*vulgus*" in *Theologico-Political Treatise*, for example on p. 176-177.
[8] Kelly and McPherson, "On Tolerating the Unreasonable", p. 38-55.
[9] Quong, *Liberalism Without Perfection*, p. 290-314.
[10] Gursozlu, "Political Liberalism and the Fate of Unreasonable People", p. 35-56.
[11] Ancell, "The Fact of Unreasonable Pluralism", p. 410-428.

My approach is grounded in the analytic tradition of political philosophy, following the example of the approach of Rawls himself. Throughout the thesis project, I utilize several philosophical methods. The first is conceptual analysis, as I will break down reasonableness into its basic aspects and analyze its key features. The second involves an interdisciplinary perspective, where I will summarize existing literature reviews in moral psychology. In addition, I use hypothetical scenarios to draw upon intuition for answers when direct observational data are lacking. I also engage in close reading of Rawls' canonical texts and address potential objections whenever I propose an argument that goes beyond the current framework. Finally, the overarching method of this thesis project is reflective equilibrium, which aims to achieve conceptual coherence by going back and forth between our practical judgments, moral intuitions, and reasoned principles that apply to the moral-political issue at hand. This combination of methods ensures a balanced and well-grounded analysis.

With the context, relevance, and methods of this thesis project now clarified, I turn to its core focus: providing a philosophical analysis of the extent to which political liberalism can successfully contain unreasonableness in a pluralistic society governed by a liberal order. Accordingly, the central inquiry of this thesis is the following:

> *How should a liberal order navigating the fact of simple pluralism achieve the containment of unreasonableness while remaining normatively consistent with political liberalism?*

The goal of this research is to determine the general principles of the strategies by which political liberalism can properly manage enduring unreasonableness in a pluralistic society while remaining true to its core ideals of legitimacy, stability, and inclusivity. To explore this question systematically, I have organized this thesis project into a total of nine chapters. These are divided among three parts that address a different dimension of the central research question, ensuring that each relevant aspect can be analyzed in detail. I will now provide a general overview of how the rest of this thesis project will unfold.

Part I develops the core concepts of the central inquiry and sets out the conceptual framework that will guide the subsequent discussions. I start in *Chapter 1* by explaining Rawls' main concepts in his theory of political liberalism, most notably his understanding of reasonableness. I also introduce the new concept of the 'duty of prudence' and argue that reasonableness should include it explicitly as a basic aspect. Then, in *Chapter 2*, I examine recent moral psychological research and its implications for Rawls' theory. I discuss how psychological constraints may hinder citizens from fully embodying a reasonable attitude, and why there is reason to believe that a significant share of citizens will likely persist in their rejection of reasonableness. Thereafter, I analyze in *Chapter 3* the core ideals of political liberalism in the form of legitimacy, stability, and inclusivity. I assert that recognizing the fact of simple pluralism entails certain restrictions on political liberalism that cause a liberal order to navigate the competing demands of its three core ideals, producing a trilemma. Nevertheless, I contend that Rawls' concept of reflective equilibrium can provide an initial suggestion on how to resolve the trilemma. In this way, *Part I* establishes the trilemma of pluralism and how it challenges political liberalism.

Part II features a critical review of various suggestions within political philosophy on how a liberal order can address the problem of unreasonable influences. I explore in

particular the options of conciliation, transformation, and marginalization – each with its own strengths and flaws. I begin by delving into the conciliatory approach in *Chapter 4*, which seeks to settle tensions between the liberal order and the unreasonable through dialogue and concessions. I also evaluate Rawls' suggestion of reasoning by conjecture in the context of engaging with unreasonable interlocutors. Next comes *Chapter 5*, where I examine the strategy of transformation: an approach that aims to inculcate in unreasonable citizens a reasonable attitude so that they will ultimately come to endorse the public justifications of power. I consider in particular applications of this idea to government institutions, educational initiatives, and public technologies. Subsequently, I investigate in *Chapter 6* the approach of marginalization, which aims to contain unreasonable movements by restricting some rights of unreasonable citizens in certain circumstances. I will analyze the practical and ethical implications of this approach in both its reactive and preventive forms. Overall, *Part II* demonstrates that current responses to the trilemma run into a decisive challenge to either their stability, legitimacy, or inclusivity that ultimately renders their proposed strategy to contain the unreasonable inadequate.

Part III works towards a multifaceted strategy that can successfully contain unreasonable doctrines in a liberal society while also overcoming any of the three challenges posed by the trilemma. To achieve this, I first introduce in *Chapter 7* a new typology of unreasonable citizens that rejects the current binary model. Instead, I identify four significant unreasonable subcategories: uncooperative, unequitable, anti-reasonable, and unaware citizens. By recognizing these distinctions, I aim to tailor effective strategies for engaging and persuading each group toward greater reasonableness. I also argue in favor of an approach of involvement concerning citizens, which improves the inclusivity of the liberal order. Then, in *Chapter 8*, I advocate in favor of an argument that can provide anti-reasonable citizens an adequate reason why they are excluded from further deliberation as members of the constituency of public justification. This bolsters the legitimacy of the liberal order by demonstrating the inherent incompatibility between unreasonable proposals and meaningful participation in public discourse. Finally, I propose in *Chapter 9* that reasonableness should include a new moral responsibility: the 'duty of vigilance'. This duty emphasizes the active preservation of fair terms of cooperation among citizens, fostering a culture of civil vigilance that reinforces stability for the right reasons in a liberal society. Proceeding like this, *Part III* develops three separate responses that address different challenges but remain mutually compatible.

In the conclusion of this thesis project, I suggest a multifaceted strategy to contain unreasonable doctrines in a liberal society that can successfully navigate the complexities caused by the trilemma of pluralism. I will integrate promising insights from the current philosophical literature on unreasonableness with my own proposals in order to articulate a strategy for containment that is comprehensive, cohesive, and effective. Even though there are many severe challenges confronting liberal democracy, a re-examination of Rawls' political liberalism suggests that there are also myriad possibilities in which they may potentially be overcome.

PART I: Establishing the Trilemma

In Part I of this thesis project, I aim to develop the terms of the main research question and the conceptual framework for the rest of the discussion. Since the subsequent treatise is informed and inspired by the Rawlsian tradition in political philosophy, it is appropriate to provide at the start some general context about Rawls and his work, beginning with some brief biographical information.

John Rawls was born in Baltimore, United States, in 1921. He studied philosophy at Princeton, and upon his graduation in 1943, he served in the army during the Second World War in the Pacific Theater. Afterward, Rawls successfully pursued a career in philosophy, fascinated by the question of whether it is politically possible to envision a social world in which the collective life of human beings would be worthwhile – a question he experienced as urgent in light of his life experiences up until that point. Rawls taught for thirty years at Harvard University, published three majorly influential philosophical books, and remained active in academic philosophy after his retirement. He passed away in 2001.[12]

Rawls' work is characterized in part by the introduction of many specialized neologisms that function as terms of art in his political theory. Although these novel concepts have had a deep and enduring influence on the vocabulary of academic discourse in liberal political philosophy, they might occasionally be challenging for the general reader – not just due to their quantity, but also because some of these specialized meanings are markedly different from the colloquial use of the phrases. Throughout this thesis project, I will introduce the relevant terms and recap them if necessary; still, it is important to keep the technical definitions in mind while reading.

In general, Rawls took a systematic and holistic approach to philosophy. Since most elements of his theory of political liberalism are interconnected and interdependent, this affects how we can engage with his work critically. While this thesis project focuses specifically on unreasonable citizens, mere isolated criticisms are insufficient as Rawls' concepts are deeply intertwined and supported by a wider theoretical structure. Thus, a constructive critique of Rawls must consider the implications of any proposed adjustments on the larger framework to ensure that any changes align with the overall coherence and integrity of his philosophical system. This comprehensive approach allows for a more accurate and meaningful engagement with Rawls' ideas.

With these general remarks laid down, let us now continue with Part I and develop the conceptual framework of this thesis project. In Chapter 1, I explore the foundational ideas of John Rawls' political theory as presented in Political Liberalism. In Chapter 2, I examine recent moral psychological research and its implications for Rawls' theory. In Chapter 3, I analyze the core ideals of political liberalism and the trilemma posed by simple pluralism. Together, these chapters clarify the main terms of the central inquiry of this thesis project: *How should a liberal order navigating the fact of simple pluralism achieve the containment of unreasonableness while remaining normatively consistent with political liberalism?* Before answering this question, it is necessary to develop a deeper understanding of the challenges that unreasonableness presents for political liberalism.

---

[12] Pogge, *John Rawls: His Life & Theory Of Justice*, p. 3-27.

Chapter 1: Theoretical Framework & Rawls' Political Liberalism

Thesis I: That reasonableness as conceptualized by Rawls implies and ought to entail a third basic aspect in the form of the duty of prudence: a moral responsibility to exercise one's intellect conscientiously and to attempt to address personal prejudice, bias, fallacies, and informational blindness.

In this first chapter, I will explain the pluralist assumptions that are the foundation of the political theory of Rawls as presented in his book *Political Liberalism*. Additionally, I will clarify how Rawls conceptualizes what it means to be 'reasonable' and what 'ideal theory' entails. Moreover, I will introduce a new interpretation of Rawls and propose that an implied third fundamental aspect of reasonableness ought to be made explicit.

1.1 Rawls & Pluralism

Before a political theorist can engage in normative analysis, they must first posit certain facts about human nature, psychology, and sociology. Likewise, Rawls has constructed his theory around a core set of facts, namely the following:[13]

First, the *fact of simple pluralism*.[14] Rawls observes that the establishment of free institutions typically leads to the emergence of a diverse variety of doctrines about religious, political, and moral issues. This variety is permanent, and it persists not only due to the divergent self-interests of individual citizens but also due to the inherent human inclination towards narrow rather than broad points of view. Some of these doctrines within the cacophony of pluralism evolve into *comprehensive doctrines:* belief systems that contain an integrated and extensive account of religion, politics, and morality.[15] A comprehensive doctrine is a grand family of ideas that, as a package, attempt to provide a definitive answer to what the good life entails.

Second, the *fact of oppression*.[16] Rawls accepts that separate comprehensive doctrines will inevitably end up in lasting and irreconcilable conflict with each other. Moreover, even if there is free discussion and all participants are equitable and conscientious, then citizens will still be unlikely to arrive at a consensus through reason alone. Only through oppressive state power can a political community curb pluralism and achieve a sustained and unified adherence to a single comprehensive doctrine. Rawls – and liberal political philosophy in general – rejects the notion that political relations must be governed by power and coercion alone.[17]

Third, the *fact of coalition*.[18] Rawls asserts that the durability and stability of a liberal-democratic order depends on a critical mass of politically involved citizens giving their free and willing support – a considerable majority is necessary at least. Public support

---

[13] Rawls first mentions these facts in his paper "The Domain of the Political and Overlapping Consensus" (p. 233-235), and later reiterates them in *Political Liberalism* (p. 35-38).
[14] Rawls sometimes refers to this as "pluralism as such" (Ibid., p.35), but elsewhere he calls it "simple pluralism" (p. xvii).
[15] Rawls, *Political Liberalism*, p. 58-60.
[16] Rawls notes that he borrows this term from Sanford Shieh.
[17] Ibid., p. lx.
[18] Since Rawls leaves this fact unnamed, I have assigned it this label for purposes of consistency.

is necessary, even though the citizenry remains pluriform and highly diverse. Thus, a liberal democracy has no alternative but to draw its support from a variety of comprehensive doctrines that will have significant internal ideological conflicts.

Fourth, the *fact of reasonable pluralism*.[19] Rawls assumes that the set of comprehensive doctrines denoted by simple pluralism is sufficiently large that it will contain a significant subset of *reasonable* doctrines. In this context, 'reasonable' specifically indicates compatibility with the fundamental principles of a liberal-democratic order: a willingness to cooperate under fair terms even in case of ideological difference, plus a disinclination to impose one's own doctrine on others.[20] Rawls emphasizes that reasonable pluralism is not primarily motivated by group- or self-interest, but is instead the natural outcome of human reason operating freely within just institutions.[21]

In light of these four facts, Rawls introduces a distinction between *liberalism* as a comprehensive doctrine and *political liberalism* as a political conception of justice.[22] An example of comprehensive liberalism is the political philosophy of Kant, who valorizes autonomy as a regulatory ideal in all dimensions of life. In contrast, political liberalism does not propagate a specific view on what constitutes the good life, nor does it advocate a single religious, political, or moral perspective as the ultimate truth.

Instead, Rawls limits his focus to providing merely a *political conception of justice*: a unified system of principles that can serve as the foundational structure of a constitutional democratic regime.[23] Though it functions independently from any particular comprehensive doctrine, it incorporates three fundamental intuitive ideas that are usually implicit in any democratic public political culture: that each citizen is *free*, that all citizens are *equal*, and that social cooperation should be *fair*.[24] In this way, the framework of political liberalism offers a thin superstructure that can encompass a diverse range of reasonable comprehensive doctrines, with each of them offering their support for reasons particular to their own doctrine. The coalition based upon a political conception of justice becomes an *overlapping consensus* once it reaches a critical mass of support among the citizens of its society.[25]

In short, political liberalism as imagined by Rawls is not primarily concerned with truth, but ought to be understood as a limited political framework, a module that is compatible with and can fit into a plurality of reasonable comprehensive doctrines. When reasonable citizens, as a clear majority, endorse the same fundamental principles for fair cooperation among free and equal individuals, this generates an overlapping consensus that can serve as the foundation of a just, liberal-democratic society.

---

[19] Rawls himself provides the following definition (*Political Liberalism*, p. xxxvi, fn. 3): "*This is the fact that a plurality of reasonable comprehensive doctrines, religious, philosophical, and moral, is the normal condition of democratic culture given its free institutions.*"
[20] Rawls, *Political Liberalism*, p. xlii.
[21] Ibid., p 144: "*[Political liberalism sees the diversity of reasonable pluralism] as the long-run result of the powers of human reason within an enduring background of free institutions.*" Also, p. xxiv.
[22] Ibid., p. xxvii: "*Political liberalism is not comprehensive liberalism.*" See also p. 99.
[23] Rawls, "Domain of the Political", p. 240. "*For a conception to be even partially comprehensive it must extend beyond the political and include nonpolitical values and virtues.*"
[24] Because there are many reasonable interpretations of 'free,' 'equal', and 'fair,' there are also many possible liberal political conceptions of justice.
[25] Rawls, *Political Liberalism*, p. 10-11:

1.2 Reasonableness & Its Requirements

Still, Rawls admits that even if a political conception of justice achieves such an overlapping consensus, it can only appeal to citizens who are themselves reasonable.[26] The concept of reasonableness serves a crucial function within the theoretical framework of political liberalism, and though its colloquial meaning is somewhat opaque, it has a technical meaning in the theory of Rawls. To start, 'reasonable' is in the Rawlsian sense not synonymous with but complementary to 'rational'.[27] *Rationality* refers to the ability to effectively pursue one's self-interest and typically involves logical reasoning to maximize utility.[28] In contrast, *reasonable* describes an attitude where one displays two traits simultaneously: abiding by the *criterion of reciprocity* and accepting the *burdens of judgment*. Let us now discuss these two basic aspects in turn.

1.2.1 Abiding by the Criterion of Reciprocity

The first trait of reasonableness is a willingness to abide by *the criterion of reciprocity*. For Rawls, this means *"to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so."*[29] In other words, it is a disposition to initiate and honor *fair* terms of cooperation, motivated by the belief that other citizens are *equal* and *free* persons. What "fair", "equal", and "free" mean specifically is somewhat dependent on the norms and values per political culture,[30] but their general meaning ought to remain consistent with several characteristics.

First, Rawls considers cooperation as *fair* when it occurs under conditions that individuals would agree to from a position behind a 'veil of ignorance.' [31] This veil is a thought experiment in which individuals lack knowledge of their own characteristics (such as their social status, wealth, talents, or beliefs) ensuring impartiality; it is also assumed the individuals are rational and want to secure the optimal outcome. For Rawls, whatever principles for social cooperation are chosen under these conditions can justifiably be called fair for all citizens.

Second, Rawls considers citizens as *equal* when they have three essential powers to such a degree that they can participate fully in societal cooperation.[32] These powers are their capacity to reason, their capacity for a sense of justice, and their capacity for a conception of the good. The sense of justice entails understanding and acting in accordance with the principles of fairness that underpin social cooperation, while also being inclined to act in ways that others can publicly endorse. The conception of the good allows individuals to formulate, adjust, and pursue their evolving understanding

---

[26] Ibid., p. 35: *"These are the [reasonable comprehensive] doctrines that reasonable citizens affirm and that political liberalism must address."*
[27] Ibid., p. 52. *"Neither the reasonable nor the rational can stand without the other. Merely reasonable agents would have no ends of their own they wanted to advance by fair cooperation; merely rational agents lack a sense of justice and fail to recognize the independent validity of the claims of others."*
[28] Ibid., p. 48-50. Later (on p.176-178), Rawls further develops this notion into his concept of 'goodness as rationality' in order to specify the contents of the principles of justice.
[29] Ibid., p. 49.
[30] Rawls, "Domain of the Political", p. 235. *"The political culture of a reasonably stable democratic society normally contains, at least implicitly, certain fundamental intuitive ideas from which it is possible to work up a political conception of justice suitable for a constitutional regime."* Since the ideas are intuitive and implicit, reasonable diversity in interpretation is possible.
[31] Rawls, *Political Liberalism*, p. 22-28.
[32] Ibid., p. 18-22.

of personal fulfillment and what is valuable in human life as they mature. In short, citizens are equal because they are in a similar way fundamentally free.

Third, Rawls regards a citizen as *free* if they meet three requirements.[33] Firstly, they recognize their own moral agency, allowing them to shape and revise their conception of the good over time without compromising their identity as free individuals (even despite changes in personal beliefs, such as religious conversions). Secondly, they view themselves as autonomous sources of valid claims, entitled to advocate for their conceptions of the good within the boundaries set by the public conception of justice. Thus, the possibility of institutions such as slavery and serfdom is removed from the agenda of political discussion. Lastly, they are affirmed in their personal responsibility for their life goals, assessing their claims within the framework of societal fairness and adjusting their aspirations accordingly. This aligns with the concept of society as a fair system of cooperation, where citizens understand the need to balance their desires with what is reasonably achievable within a broader society.

To recap: the first requirement of reasonableness is a willingness to abide by *the criterion of reciprocity*. This entails proposing and honoring fair terms of cooperation while regarding other citizens as equal and free. Let us now discuss the other aspect.

1.2.2 Accepting the Burdens of Judgment

The second trait of reasonableness is a willingness to accept *the burdens of judgment*.[34] A prerequisite for this willingness is the acknowledgment that even in deliberation with other citizens who similarly abide by the criterion of reciprocity, disagreement on fundamental issues will persist – especially in the realm of religion, philosophy, and morality. Since the complex questions regarding the good life are so difficult that even conscientious and capable people will not converge on some consensus, Rawls asserts that this disagreement can still be reasonable, as it arises naturally due to each citizen exercising their reason freely in combination with human reason being limited. Those limitations in reasoning are what constitute the 'burdens of judgment'. Rawls describes them as *"the hazards involved in the correct and conscientious exercise of our powers of reason and judgment in the ordinary course of political life".*[35]

What are these burdens specifically? Rawls lists some of the sources of reasonable disagreement that he thinks are most obvious.[36] These include the complex nature of empirical evidence (which can on occasion even be contradictory); the necessity of trade-offs to decision-making (justly prioritizing moral-political values is often non-obvious and thus open to reasonable disagreement); the inherently abstract nature of moral-political concepts(which can sometimes be ambiguous and underdetermined); and the varying ways in which different individuals assess evidence differently (made unavoidable by the impact of our total life experience on our judgment). Still, Rawls also explicitly excludes some sources – most notably prejudice, bias, and self-interest.

What are the implications of 'accepting' the burdens of judgment? First, one posits these limitations as the primary explanation of lasting moral-political disagreement between reasonable and rational citizens. This attitude exists in contrast with the

---

[33] Ibid., p. 29-35.
[34] Rawls, *Political Liberalism*, p. 54-58. Not to be confused with 'the burden of proof'. In publications before 1993, Rawls calls it the 'burdens of *reason*', not 'judgment'.
[35] Ibid., p. 56.
[36] Ibid., p. 56-57. Also mentioned in "Domain of the Political", p. 237.

temptation for citizens to dismiss those with whom they disagree on moral-political matters as absolutely mistaken; rather, they would regard them as reasonable agents who make diligent use of their moral powers, and subsequently respect their rights and perspectives. Second, recognizing the limits of human judgment (combined with a commitment to fairness and reciprocity) entails a willingness to justify their political positions within the wider context of their pluralistic society using shared public values rather than relying solely on their own comprehensive doctrines.[37]

Rawls refers to the set of principles, values, and arguments that individuals use to justify their positions and decisions in the public sphere within a pluralistic society as *public reason*.[38] Its domain is limited to the public political forum (most notably the discourse of judges, legislators, government officials, and candidates for public office). The wider social background culture (such as churches, universities, and the nonpublic political culture as expressed in the media) remains outside of its scope. Still, when deliberating in a public political forum, reasonable individuals ought to refrain from invoking religious beliefs or controversial contentions, and instead focus on public or political concepts already present in their political culture and are accessible and acceptable to all other reasonable citizens.

In short, reasonableness is best understood as the political virtue where a citizen abides by the *criterion of reciprocity* (engaging in fair cooperation with others that they acknowledge as equal and free) and accepts *the burdens of judgment* (refraining from imposing their comprehensive doctrine on others).

1.3 The Duty of Prudence

Rawls makes it clear that he regards his conception of reasonableness as having just two basic aspects – the criterion of reciprocity and the burdens of judgment – and this twofold characterization has been accepted in the secondary literature.[39] Simultaneously, Rawls admits that he does not directly provide a strict definition of 'reasonable' – thus, those operating within his framework have some leeway to build upon his conceptualization. Hence, I propose that Rawlsian reasonableness entails an additional third basic aspect: the *duty of prudence.* Explicitly including a basic aspect that covers the epistemic elements of reasonableness is necessary to complement the criterion of reciprocity and the burdens of judgment, since intellectual humility and conscientious reasoning are essential for fair and informed public discourse.

1.3.1 Definition

I define the concept of fulfilling the 'duty of prudence' as such: *the moral responsibility of exercising one's intellect conscientiously during and practicing critical self-reflection ahead of engaging in political deliberation, specifically to remedy a) personal prejudice, b) cognitive bias, c) logical fallacies, d) informational blindness.* It does *not* apply to self-interest, group-interest, and willfulness – for these sources of disagreement are removed when one abides by the criterion of reciprocity. Rawls writes that: *"Being reasonable is not an epistemological idea (though it has epistemological elements)."*[40] The duty of prudence expresses the ongoing effort to realize those epistemic elements: e.g. reflecting upon

---

[37] Ibid., p. 66-68.
[38] Ibid., p. 220.
[39] See: Wenar, *Stanford Encyclopedia of Philosophy*; and Talisse, *On Rawls*, p.64-65.
[40] Rawls, *Political Liberalism*, p. 62.

one's presumptions, reasoning carefully and coherently, assessing all available evidence, and considering all available reasons and perspectives. Respecting this process is more important than any particular outcome of reasoning – in this context, what matters is the core attitude of valuing the conscientious exercise of intellect.

Alternatively, the duty of prudence can be explained like this: it is similar to the virtue practiced by legal experts who work within the domain of the *colorable*. 'Colorable' arguments are plausible to some experts – despite others being in disagreement – when *"the claim is strong enough to have a [realistic] chance of being valid if the legal basis is generally correct and the facts can be proven in court. [T]he claim need not actually result in a win."*[41] The burdens of judgment apply to the discussion between colorable positions as in that context, the sources of reasonable disagreement mentioned before by Rawls come into play: such as the fact that admissible evidence can be contradictory, the great difficulty involved in determining an equitable prioritization of relevant moral-political values, and the abstract and therefore occasionally ambiguous and underdetermined nature of moral-political concepts. The duty of prudence functions complementary to accepting the burdens of judgment: it does not primarily pertain to reasonable discussion itself, but it helps to ensure that all participants can rightly be assumed to meet the conditions that make reasonable discussion possible.

1.3.2 Textual Evidence

In my reading, Rawls already presupposes the duty of prudence in all three dimensions of reasonableness,[42] and thus there are several reasons why it is necessary and warranted to make it explicit. Firstly, because it is clearly implied in Rawls' description of the burdens of judgment for individual *citizens*. He writes:

> We do not, of course, deny that prejudice, bias, [and] blindness, play their all too familiar part in political life. But these sources of unreasonable disagreement stand in marked contrast to those compatible with everyone's being fully reasonable."[43] […]

In other words, it is not possible to be both fully reasonable and severely biased or prejudiced or underinformed; and since reasonable discussion is eroded by pointing out the cognitive defects in the reasoning of our interlocutors and arouses hostility,[44] all participants must exercise their intellect conscientiously and have practiced critical self-reflection before entering in deliberation. The duty of prudence still allows people to make mistakes or defend positions that turn out incorrect – the point is that they have done their best to avoid such scenarios and remain alert to prevent them.

Secondly, Rawls makes a similar implication in his description of reasonable comprehensive *doctrines*. He writes that *"a reasonable comprehensive view [normally] draws upon a tradition of thought and doctrine"* which, although stable over time, *"tends to evolve slowly in the light of what, from its point of view, it sees as good and sufficient reasons."*[45] Put differently, comprehensive doctrines must be sufficiently thorough and theoretically rigorous to be classified as reasonable. For instance, somebody who ardently advocates the doctrine of astrology, which is incoherent, cannot be reasonable, regardless of whether that individual abides by the criterion of reciprocity or

---

[41] Legal Information Institute, "Colorable Claim", June 2022.
[42] Ibid., p. 48: *"[T]he idea of the reasonable, whether applied to persons, institutions, or doctrines"*.
[43] Ibid., p. 58. I omitted the categories that were already excluded in the definition of the duty.
[44] Rawls, "The Domain of the Political", p. 238.
[45] Ibid., p. 59.

acknowledges that consensus on moral-political matters is impossible. This shows why the duty of prudence is needed in political liberalism, for it prevents an over-zealous interpretation of the burdens of judgment: a radical skepticism that regards all ways of viewing the world as being equally unsound.[46] Besides, historically speaking, unreasonable comprehensive doctrines have displayed a tendency to justify anti-egalitarian views and oppressive hierarchies with pseudo-scientific rhetoric that serves as a legitimizing myth.[47] The duty of prudence could provide an apt safeguard against disparaging tropes contaminating reasonable discussion.

Thirdly, Rawls directly mentions the contents of the duty of prudence when he discusses what *institutions* are necessary for political liberalism to develop:

> *"A. Public financing of elections and ways of assuring the availability of public information on matters of policy. The statement of these arrangements (and of those below) merely hints at what is needed [...] to provide the knowledge and information upon which policies can be formed and intelligently assessed by citizens using public reason.*
> *B. A certain fair equality of opportunity, especially in education and training. Without these opportunities, all parts of society cannot take part in the debates of public reason or contribute to social and economic policies."*[48]

Here, Rawls underscores that reasonable discussion depends on adequate knowledge and reliable information, and also emphasizes the importance of cultivating reason through education and training. The duty of prudence does not necessitate that every citizen becomes an expert in all areas of political discourse and has mastered every controversial dossier, nor does it mandate that all participants possess a university diploma. Still, reasonable discussion is contingent upon a minimum level of knowledge and intellect, which ensures that citizens can connect and engage in political deliberation with a shared understanding of the relevant facts and concepts.[49]

Moreover, if one does not explicitly include the duty of prudence as a basic aspect of reasonableness, this has unfortunate and likely unintentional implications. As mentioned before, Rawls rejects that reasonable disagreement can result from prejudice, bias, or logical errors: *"These explanations are too easy and not the kind we want."*[50] Yet, another Rawlsian assumption is that *"the political culture of a reasonably stable democratic society normally contains, at least implicitly, certain fundamental intuitive ideas from which it is possible to work up a political conception of justice suitable for a constitutional regime"*.[51] Therefore, it is conceivable that public reason remains distorted by pseudo-science, cultural stereotypes, or ethnically structured group interests. The duty of prudence makes it explicit that these influences are unwelcome.

1.3.3 Addressing Possible Objections

Since I could not find any precedent for my proposed exegesis in the secondary literature, it is fitting to discuss several possible objections against my proposal.

---

[46] Rawls disavows that the burdens of judgment suggest a radical skepticism on ibid., p. 62-63: *"Skepticism may seem to be suggested by the account of the burdens of judgment. [...] The account of the burdens of judgment does none of these things."*
[47] See Quist & Resendez, "Social dominance threat: Examining social dominance theory's explanation of prejudice as legitimizing myths", p. 287-293.
[48] Rawls, *Political Liberalism*, p. lvi-lvii.
[49] Ibid., p. 139.
[50] Ibid., p. 55.
[51] Rawls, "The Domain of the Political", p. 235.

The first objection could be that adding the duty of prudence is superfluous because it is already covered by accepting the burdens of judgment. After all, reasonable citizens accept the burdens of judgment as they share a common human reason.[52] However, the fact that interlocutors both have the capacity for prudent reasoning does not warrant the assumption that they also actualize it in their discussion. Then again, this objection could insist that Rawls mentions that his list of sources of reasonable disagreement is incomplete. Still, this approach does not work since Rawls explicitly stipulates that prejudice, bias, and blindness are distinct from the burdens of reason.[53] Alternatively, this objection could invoke the passage of Rawls where he writes: *"If there is time to expose through discussion the falsehoods and fallacies to avert the evil by the processes of education, the remedy to be applied is more speech, not enforced silence."*[54] Why introduce the duty of prudence and practice critical self-reflection in advance of discussion, if any shortcomings in reasoning can simply be pointed out during reasonable deliberation? The answer is that this solution ought to be deployed carefully and infrequently. As Rawls points out: *"Such accusations [of prejudice and bias] arouse resentment and hostility, and block the way to reasonable agreement."*[55] The duty of prudence prevents that public correction gets out of hand and erodes reasonableness.

Finally, this first objection could insist that the duty of prudence is included whenever Rawls talks about the *conscientious* exercise of reason.[56] I think that such an approach bites the bullet since the dual meaning of 'conscientious' in its colloquial use is likely to generate confusion and needs clarification. Its first meaning is 'wishing to perform one's duty well and thoroughly'; the second is 'being governed by the dictates of conscience.' By introducing the duty of prudence, it would be evident that 'conscientious' in the Rawlsian sense involves both meanings of the word.

The second objection could be that postulating the duty of prudence is superfluous because it is already included in the *duty of civility*, defined by Rawls as such:

> *"The ideal of citizenship imposes a moral, not a legal, duty—the duty of civility—to be able to explain to one another on those fundamental questions how the principles and policies they advocate and vote for can be supported by the political values of public reason. This duty also involves a willingness to listen to others and a fairmindedness in deciding when accommodations to their views should reasonably be made."*[57]

Though the duty of civility is certainly related to the duty of prudence, it remains distinct: civility is focused on other citizens, whereas prudence applies to oneself. Without the duty of prudence, one could be civil and make accommodations in response to a view that asks to rectify social marginalization, but nevertheless retain unexamined prejudices against that socially marginalized group. This would not be reasonable, and thus the duty of civility does not encompass prudence.

The third objection is that postulating the duty of prudence is unwarranted because it links reasonableness with rationality, which Rawls rejects insistently. I think that this response has some merit since exercising one's intellect conscientiously and practicing

---

[52] Ibid., p. 55.
[53] Ibid., p. 58.
[54] Ibid., p. 352.
[55] Rawls, "Domain of the Political", p. 238.
[56] For instance: Rawls, *Political Liberalism*, p. 54, 58, 119, 154.
[57] Ibid., p. 217.

critical self-reflection has a rational element indeed – in particular in the case of logical fallacies and informational blindness. Still, two things should be noted. Firstly, Rawls insists on reasonableness and rationality as distinct concepts because he is adamant that being reasonable cannot be conclusively derived from being rational.[58] Yet, the duty of prudence does not engage in such an attempt. Secondly, Rawls notes that rationality is characterized by self-interest, concerned with its own ends, and that the reasonable is public in a way the rational is not.[59] The duty of prudence does not focus on enhanced reasoning about one's own ends and interests but is instead primarily aimed at improving the flow and quality of public discussion. All in all, it seems unfair to classify the duty of prudence as a rational rather than a reasonable virtue.

Thus, since none of the possible objections seem persuasive, I maintain that reasonableness as understood by Rawls requires that a third basic aspect is made explicit in the form of the duty of prudence. This addition is warranted as it addresses several needs: epistemic completeness by underscoring the importance of conscientious reasoning in public discourse, avoidance of cognitive defects by promoting critical self-reflection, and a knowledgeable public by ensuring that citizens participate based on adequate information. Now that we have deepened our understanding of reasonableness, it is time to turn our attention to unreasonableness.

1.4 Unreasonableness & Ideal Theory

The ideal of reasonableness as imagined by Rawls is far from common practice as a civic virtue and political attitude. It is important to emphasize that even when a person lacks just one basic aspect of reasonableness, Rawls will classify them as *unreasonable*. In political practice, an unreasonable attitude is ubiquitous: refusing cooperation under fair terms, professing unwarranted moral certainty, engaging in distorted reasoning, selective blindness to cognitive and emotional biases, partisan resistance to new information, and ethnocentric reactions to immigrants – just to name a few examples.[60] Rawls acknowledges this regrettable dimension of politics:

> "[P]eople are often irrational and not very bright, and this mixed with logical errors leads to conflicting opinions. But while such explanations explain much, they are too easy and not the kind we want. We want to know how reasonable disagreement is possible, for we always work at first within ideal theory.[61] [...] Of course, every society also contains numerous unreasonable doctrines. Yet in this essay I am concerned with an ideal normative conception of democratic government, that is, with the conduct of its reasonable citizens and the principles they follow, assuming them to be dominant and controlling. How far unreasonable doctrines are active and tolerated is to be determined by the principles of justice and the kinds of actions they permit."[62]

In short, Rawls acknowledges that every society must find a way to cope with unreasonable citizens,[63] but thinks that the framework of *ideal theory* justifies his

---

[58] Ibid., p. 51-52.
[59] Ibid., p. 53-54.
[60] See: *Oxford Handbook of Political Psychology*, p. 4.
[61] Rawls, *Political Liberalism*, p. 55.
[62] Ibid., p. 441 fn 3.
[63] Ibid., p. 39: "[U]nreasonable comprehensive doctrines (these, we assume, always exist)".
Also, ibid., p. 126: "An overlapping consensus of reasonable doctrines may not be possible under many historical conditions, as the efforts to achieve it may be overwhelmed by unreasonable and even irrational (and sometimes mad) comprehensive doctrines."

assumption that a substantial majority of citizens are reasonable and that unreasonable perspectives do not exert significant influence in public discourse. That gives rise to the question: what does Rawls mean by 'ideal theory?'

As an ideal theory, political liberalism develops the principles and structures of an ideal liberal society, operating under the best foreseeable conditions that are conducive to justice and stability. Specifically, ideal theories make two assumptions.[64] First, that whatever the principles chosen in the original position, everybody will strictly obey them: the conception of justice can count on full compliance. Second, that the relevant society is both economically and socially sufficiently developed that it can achieve justice, maintaining favorable conditions necessary for sustaining a just social order.

These ideal circumstances do not reflect current realities, but rather, they model what could be achievable under optimal circumstances. Thus, ideal theory engages with counterfactual assumptions, due to the vast disparity between present institutions and ideal conditions. Despite the disparity between the ideal and the current reality, Rawls wants his approach to be *realistically utopian*: *utopian* because it uses political ideals to specify a reasonable and just society, yet *realistic* because what is imagined must be feasible and might actually exist, if not now, then at some future time under more fortunate circumstances.[65] This means that even ideal theory must operate, as Rawls phrases it, within the *"fixed constraints of human life"* that include *"the general facts of moral psychology."*[66] For instance, Rawls regards the aforementioned fact of reasonable pluralism as one such fixed constraint. But he also admits that if his moral-psychological assumptions are not justified, this poses a serious challenge to his theory.[67]

1.5 Conclusion

In this first chapter, I have specified that *the fact of reasonable pluralism* lies at the heart of political liberalism; I have explained the traditional two basic aspects of reasonableness (*the criterion of reciprocity* and the *burdens of judgment*); I have argued that Rawls' notion of being reasonable implies an additional third basic aspect in the form of the *duty of prudence*; and I have clarified that the conditions of *ideal theory* must remain realistically utopian, constrained by the general facts of moral psychology.

If my account so far is accepted as valid, then this has important implications for the theory of political liberalism; for Rawls grants that moral psychology imposes boundaries on which conceptions of persons and ideals of citizenship are viable. Although moral psychology cannot be prescriptive and dictate what particular vision of politics must be adopted, it can rule out those accounts that remain unable to meet the practical needs of political life. Therefore, we should consider how contemporary moral psychological research relates to the basic aspects of reasonableness, and examine whether Rawls can feasibly postulate the fact of reasonable pluralism.

---

[64] Rawls, *Theory of Justice*, p. 8 & 215; and Rawls, *Law of Peoples*, p. 11-12. Also, see Valentini, "Ideal vs. Non-Ideal Theory: A Conceptual Map", p. 655.
[65] Rawls, *Law of Peoples*, p. 12-14.
[66] Rawls, *A Theory of Justice*, respectively p. 216 & 126.
[67] Rawls, *Political Liberalism*, p. 252. *"Should these [moral-psychological] assumptions be mistaken, there is a serious problem with justice as fairness as I have presented it."*

Chapter 2: Empirical Framework & Contemporary Moral Psychology

Thesis II: That recent moral psychological research indicates that ordinary citizens face significant cognitive difficulties in adopting the three basic aspects of reasonableness, which suggests that, even within the framework of ideal theory, it is more appropriate to postulate the fact of simple pluralism rather than reasonable pluralism.

In the previous chapter, I established that being reasonable in the Rawlsian sense entails three basic aspects. Now, I will consider them individually and link them to relevant moral-psychological concepts that may hinder citizens from reliably embodying a reasonable attitude. My objective is to examine which limitations are imposed on Rawls' political liberalism by the fixed constraints of the human condition. At the end of the chapter, I discuss several potential objections to my argument that political liberalism should postulate simple instead of reasonable pluralism. But first, I discuss psychological complications to the three distinct aspects of being reasonable.

2.1 Criterion of Reciprocity & Minimal Groups

The first basic aspect of being reasonable is abiding by the criterion of reciprocity. This describes the virtue of being willing to initiate and honor fair terms of cooperation, motivated by the belief that other citizens are equal and free.[68] Its opposite attitude is regarding some fellow citizens as belonging to an out-group and engaging in cooperation under unfair terms. Such an attitude is highlighted within moral psychology by the *minimal group paradigm* (MGP), which proposes *that "the minimal condition for group biases – like favoritism towards your own group and prejudice towards other groups – is simply being a member of a group."*[69] MGP has three main characteristics: (a) the creation of new groups on arbitrary grounds (b) full anonymity among group members (3) no direct advantages to the participant in the outcome.[70]

The body of research using MGP reveals that biases of ingroup preferences and outgroup disfavor can arise without any actual competition, conflict, or significant differences between the groups.[71] Even without any interaction or shared history, participants displayed a preference for their in-group.[72] The effects of minimal group dynamics have been observed across various measures relevant to political deliberation such as resource distribution,[73] and have been confirmed by meta-analytical studies.[74] The debate among researchers about its theoretical grounds is

---

[68] Rawls, *Political Liberalism*, p. 49: *"Persons are reasonable in one basic aspect when, among equals, they are ready to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so."*
[69] Psychology Glossary, "Minimal Group Paradigm."
[70] See: Schmidt & Drake, "Minimal Group Procedures and Outcomes."
[71] Brown, "Origins of the Minimal Group Paradigm", p. 371.
[72] Hewstone et al., "Intergroup bias", p. 575.
[73] Dunham, "Mere membership", p. 780.
[74] Balliet et al., "Ingroup favoritism in cooperation: A meta-analysis"; Fischer & Derham, "Is in-group bias culture-dependent? A meta-analysis across 18 societies"; Mullen et al., "Ingroup bias as a function of salience, relevance, and status: An integration."

ongoing; still, MGP is regarded by experts as a valuable and reliable model.[75] Within evolutionary psychology, it has been established that groups – even trivial ones – are central to how people define themselves and that this phenomenon can be observed in every culture.[76] Hence, the Group-Instincts-Hypothesis[77] has been defended as the most plausible explanation of human commitment to groups,[78] which suggests that humans actually are naturally predisposed toward their psychology being marked by thinking in terms of in-group and out-group. [79]

Thus, the moral-psychological tendency described by the minimal group paradigm is a notable challenge to citizens abiding by the criterion of reciprocity. Moreover, since the basis of group differences can be arbitrary, ideal theory is unlikely to help with remedying this issue. For instance, reducing social and economic inequality cannot address biases stemming from political partisanship.

2.2 Burdens of Judgment & Moral Conviction

The second basic aspect of being reasonable is accepting the burdens of judgment. This means that one recognizes the inherent difficulties in making correct judgments on moral-political matters, and therefore one adopts a willingness to grant others the freedom to hold different reasonable views and to express themselves in publicly accessible terms when appropriate. The opposite of accepting the burdens of judgment is an attitude of absolute certainty on matters related to the question of the good life. Within moral psychology, this attitude is called *moral conviction*: *"the belief that a given attitude is a reflection of one's core feelings or beliefs about fundamental issues of right and wrong."* [80] Its research paradigm asserts that a moral conviction is psychologically very different from related attitudes like preference (a subjective, personal inclination) and convention (a social norm with defined boundaries).[81]

The body of research on moral convictions reveals that people tend to perceive their attitudes of moral conviction as objective (i.e., as facts that are grounded in fundamental truths about reality) and as universal (i.e., as mandates that apply across time, place, and cultural context).[82] It suggests that strong polarization in many social debates stems not only from differing attitudes on those moral issues as such but also from differing perceptions of whether these issues are moral or nonmoral in nature.[83] Additionally, it finds that people are willing to condone various illicit means, including deceit and violence, so long as they achieve morally preferred ends.[84] As one paper on the social and political implications of moral conviction states:

---

[75] Otten, "The Minimal Group Paradigm", p. 85
[76] For instance, as argued in the first chapter of *The Power of Us* by Van Bavel & Packer. Also, see *Our Moral Fate* by Allen Buchanan.
[77] The original moniker is "Tribal Instincts Hypothesis." However, the word 'tribe' carries colonial connotations; which is quite ironic since colonialism is marked by in-group/out-group thinking.
[78] Richerson & Boyd, 'The Evolution of Subjective Commitment to Groups: A Tribal Instincts Hypothesis", p. 184.
[79] Van Vught & Park, 'The Tribal Instincts Hypothesis: Evolution and the Social Psychology of Intergroup Relations', p. 27.
[80] Skitka et al., "The Psychology of Moral Conviction", p.348.
[81] Nucci, *Education in the moral domain*, p. xviii. One key assumption is that morality is conceptualized as a meta-perception that people can access and report.
[82] Skitka et al. "The Psychology of Moral Conviction", p. 352.
[83] Wright et. al., "The Cognitive and Affective Dimensions of Moral Conviction," p. 1475.
[84] Mueller, "Liars, damned liars, and zealots: the effect of moral mandates on transgressive advocacy acceptance", p. 184.

"The normative implications of these and other findings are both reassuring (moral conviction can protect against obedience to potentially malevolent authorities) and terrifying (moral conviction is associated with rejection of the rule of law and can provide a motivational foundation for violent protest and acts of terrorism)."[85]

These worries are exacerbated by the fact that there is no conclusive evidence on what approaches are effective in de-moralizing attitudes of conviction.[86] Furthermore, some research seems to indicate that the association between moral conviction and intolerance may be primarily automatic rather than controlled.[87]

In short, attitudes of moral conviction constitute a serious hurdle that citizens must overcome before they can accept the burdens of judgment. Besides, since people tend to regard their own moral convictions as objective and universal, one cannot assume that their intensity will decrease under conditions of ideal theory.

2.3 Duty of Prudence & Motivated Reasoning

The third basic aspect of being reasonable is fulfilling the duty of prudence. This moral obligation entails addressing personal prejudices, cognitive biases, logical fallacies, and informational blind spots in order to conscientiously apply one's intellect before participating in political deliberation. In opposition to this mindset is the tendency to selectively gather, interpret, and assess information in a way that aligns with one's preexisting beliefs and preferences. In the context of moral psychology, this approach is referred to as *motivated reasoning*: the phenomenon where *"motivation affects reasoning through reliance on a biased set of cognitive processes: strategies for accessing, constructing, and evaluating beliefs."*[88] To illustrate this concept, imagine a courtroom with a judge and two attorneys. A judge reasons 'forward': they attempt to objectively evaluate facts and normative principles to arrive at a conclusion without bias or preconception. In contrast, the two attorneys reason 'backward': they advocate for a predetermined conclusion for reasons independent of truth and will marshal any arguments in support to convince others. Motivated reasoning represents this second approach.

The body of research on motivated reasoning reveals that *"in many judgment situations, particularly those that involve people and issues we care about deeply, people act more like lay attorneys than lay judges"*.[89] It is very challenging to avoid motivated reasoning since we often lack introspective awareness and we can rarely find bias in our own reasoning whenever when we make the attempt.[90] Some moral psychologists speak of 'the illusion of objectivity' since individuals nonetheless commonly perceive themselves as impartial and objective even while they (subconsciously) engage in the manipulation of arguments and evidence to align with their predetermined conclusions.[91]

Additionally, neuroscientific research has suggested that motivated reasoning is qualitatively distinct from cold, unemotional reasoning; it engages neural circuitry

---

[85] Skitka & Morgan, "The Social And Political Implications Of Moral Conviction", p. 95.
[86] Skitka et al. "The Psychology of Moral Conviction", p. 361.
[87] Baumgartner & Morgan, "Mindfulness and cognitive depletion", p. 31.
[88] Kunda, "The Case for Motivated Reasoning". Moreover, motivation is conceptualized as *"any wish, desire, or preference that concerns the outcome of a given reasoning task"* (p. 480).
[89] Ditto et al., "Motivated Moral Reasoning", p. 310.
[90] Pronin et al. "The Bias Blind Spot: Perceptions of Bias in Self Versus Others", p. 369.
[91] Irwin & Real, " The Illusion of Objectivity", p. 1.

linked to emotion and threat response rather than pure cognition.[92] The tendency for motivated reasoning is theorized to create reinforcing neural pathways that further ingrain an individual's reasoned beliefs over time through 'affective contagion.'[93] Though it can vary in degree across individuals, motivated reasoning seems to arise from core psychological needs and neural processes inherent to human reasoning and perception.[94] As psychologists Lodge & Taber conclude in their literature review: *"At this juncture then we are skeptical of the ability of citizens to reliably and veridically access the sources of their beliefs, the reasons for their attitudes, and their past or future intentions and actions. […] People are typically more rationalizing than rational."*[95]

Thus, the practice of motivated reasoning makes it significantly more difficult for citizens to fulfill the duty of prudence. Also, since motivated reasoning apparently arises from fundamental human needs (such as the compulsion to decrease cognitive dissonance), one cannot assume that it will vanish under conditions of ideal theory.

2.4 Anti-reasonableness & RWA-Personality

Until now, I have discussed the basic aspects of reasonableness and associated moral-psychological obstacles separately from each other. However, it also seems that a sizeable group of citizens does not merely struggle with adopting a specific part, but instead appears recalcitrantly opposed to the attitude of reasonableness as a whole. I introduce the concept of *anti-reasonable* to describe such a position, given its *a priori* dismissal of reasonableness. One particularly striking example of a deeply ingrained and pre-political attitude of unreasonableness is provided by the psychological profile of the Right-Wing Authoritarian personality (RWA).[96] People who score high on RWA tests are characterized by a desire for order, obedience to authority, and a willingness to enforce social norms by punishing deviance.[97] They also tend to valorize moral absolutism, regard those who hold different views as a threat to social order, and reject cooperative or collaborative efforts with them.[98] All in all, the RWA personality indicates a hostile attitude towards reasonableness.

Available research suggests that a sizeable portion of citizens has a right-wing authoritarian personality even in the population of a longstanding liberal democracy. For instance, Bob Altemeyer estimated that around a quarter of the US population has a right-wing authoritarian personality,[99] and these findings were replicated by the research bureau Morning Consult, indicating that 25% of the US population score high on the RWA scale.[100] When the occurrence of RWA was measured in the populations of other prominent liberal democracies, scores were significantly lower: for instance, 10% in the UK and 13% in Australia. On the whole, it seems that a base of roughly 1 in 9 citizens displaying High-RWA traits remains consistent across countries. Of course, it is important to note that psychological generalizations do not apply strictly to all

[92] Westen et al., "Neural bases of motivated reasoning", p. 1947.
[93] Lodge & Taber, *The Rationalizing Voter*, p. 20, 134.
[94] Kahan, "The Politically Motivated Reasoning Paradigm."
[95] Lodge & Taber, "The Rationalizing Voter: Unconscious Thought In Political Information Processing", p. 42.
[96] Altemeyer, *The Authoritarians*, p. 25-32.
[97] Osborne et al., "Psychological Causes And Societal Consequences Of Authoritarianism", p. 220.
[98] See also: Altemeyer, *The Authoritarians*, p. 62-65.
[99] Ibid., 82-83.
[100] Morning Consult, "International Study on Right-Wing Authoritarianism".

individuals with an RWA personality; individual differences exist within the group. Additionally, it is important to recognize that political attitudes and behaviors are shaped by a variety of social, cultural, and historical factors. Still, research using twin studies has suggested that half of the variance in RWA can be explained by genes.

In short, there is good reason to believe that (a) a significant portion of the population – a percentage of at least double digits – is highly likely to be hostile to reasonableness; and that (b) this group will persist even within liberal democracies that have been firmly established for generations with strong and inclusive institutions. Thus, the research on RWA warrants the assumption that a sizable share of citizens will remain actively opposed to reasonableness even under circumstances of ideal theory.

2.5 Possible Objections

So far, I have examined the fundamental components of reasonableness individually and proposed that citizens encounter significant moral-psychological obstacles on the path to achieving reasonableness. Human unreasonableness in disposition, judgment, and reasoning that is so frequently on display in moral-political matters cannot fully be attributed to the fact that our current society remains unjust and nonideal in many ways. In addition, I have argued that a considerable percentage of citizens are anti-reasonable, as evidenced by the prevalence of the Right-Wing Authoritarian personality. In rejecting all basic aspects of being reasonable, anti-reasonable citizens pose a persistent challenge even within longstanding liberal democracies. In light of this, I assert that the fact of reasonable pluralism, though not impossible to achieve, faces too many challenges to be postulated in ideal theory without further arguments and evidence in support. I will now discuss two potential objections to this claim.

2.5.1 Cherry-picking

The first objection is that my characterization of contemporary moral psychology is not properly representative of the field, and therefore engages in cherry-picking. I recognize that any interdisciplinary research must address this potential issue.

Still, such a charge would be unwarranted in my view for several reasons. First, I have cited multiple sources by multiple authors on each moral-psychological concept. Second, these sources were primarily literature reviews, not isolated experiments. Third, I engaged with the most frequently cited relevant literature, not obscure studies. Fourth, the conclusions of these sources acknowledge quite some nuance. Research on minimal groups found that it can help to overcome ethnic divides with the right boundary conditions;[101] a meta-analysis on moral conviction found an association with political engagement and activism;[102] and one literature review on motivated reasoning writes that *"affect may play an important role both in promoting bias and restraining it."*[103] Lastly, there are other moral psychological paradigms that would also challenge

---

[101] Van Bavel & Cunningham, "Self-categorization with a novel mixed-race group moderates automatic social and racial biases", p. 321.
[102] Skitka et al., "Political Orientation And Moral Conviction: A Conservative Advantage Or An Equal Opportunity Motivator Of Political Engagement?", p. 711.
[103] Ditto et al., "Motivated Moral Reasoning", p. 333.

reasonable attitudes, even beyond what I have selected.[104] Therefore, I think that the objection of cherry-picking is not applicable.

2.5.2 Interpretation of Rawls

The second objection is that the position which I advocate does not really contradict Rawls. Indeed, I am in agreement with his acknowledgments of several issues mentioned so far. Firstly, that one can be both reasonable and incorrect: *"As with any form of reasoning in public reason, the reasoning may be fallacious or mistaken."*[105] Secondly, that there will always be unreasonable elements in society: *"There are doctrines that reject one or more democratic freedoms is itself a permanent fact of life, or seems so."*[106] Thirdly, that unreasonableness can destabilize the liberal order: *"[Unreasonable views] may not be strong enough to undermine the substantive justice of the regime. That is the hope; there can be no guarantee."*[107] Fourthly, that reasonable pluralism is not always achievable: *"An overlapping consensus of reasonable doctrines may not be possible under many historical conditions, as the efforts to achieve it may be overwhelmed by unreasonable and even irrational (and sometimes mad) comprehensive doctrines."*[108]

Nevertheless, I maintain that moral psychology imposes constraints on political liberalism that have problematic implications. Rawls acknowledges this: *"It's clear, however, that should these [moral-psychological] assumptions be mistaken, there is a serious problem with justice as fairness as I have presented it."*[109] Likewise, the fragility of reasonableness: *"What if it turns out that the principles of justice as fairness cannot gain the support of reasonable doctrines, so that the case for stability fails? Justice as fairness as we have stated it is then in difficulty."* Then, he continues: *"I do not pursue this inquiry but assume, on the basis of a number of plausible considerations, that the case for the stability of justice as fairness, or some similar conception, goes through."*[110] I differ from Rawls in that assumption, and the rest of this thesis project can be regarded as one such inquiry.

Contra Rawls, I assume that human psychology displays certain cognitive features that pose challenges to becoming reasonable and may be overcome occasionally, but not reliably. Since these difficulties arise naturally, being either inherent to human moral psychology or intrinsic to political deliberation, there is no reason to believe that they will disappear in an ideal liberal society. The sources of unreasonableness function, I think, similarly to the burdens of judgment: both inevitably contribute to simple pluralism – they are not insurmountable in every instance, yet often enough to prevent the resolution of many disagreements on how to live. Of course, social progress is not impossible, and neither is occasional consensus nor convergence on truth. Rather, the point is that one cannot assume without further reasons that a political liberal order under ideal conditions will reliably transform the fact of simple pluralism into

---

[104] For instance, the criterion of reciprocity could also be undermined by the fundamental attribution error, the burdens of judgment by the false consensus effect, and the duty of prudence by moral dumbfounding. Their relation to reasonableness could be the topic of further research.
[105] Rawls, *Political Liberalism*, p. liv, fn 32.
[106] Ibid., p 64, fn 19. Elsewhere, he writes similarly: *"That a democracy is marked by the fact of pluralism as such is not surprising, for there are always many unreasonable views"* (p. 63-64). Also: *"I noted in the beginning the fact that every actual society, however dominant and controlling its reasonable citizens may be, will normally contain numerous unreasonable doctrines that are not compatible with a democratic society"* (p. 488-489).
[107] Ibid., p. 65.
[108] Ibid., p. 126.
[109] Ibid., p. 252.
[110] Ibid., p. 65-66.

reasonable pluralism. If political liberalism wants to remain realistically utopian, then postulating *simple* pluralism would be appropriate – not reasonable pluralism.

2.6 Conclusion

In this second chapter, I have discussed how moral psychology imposes limitations on Rawls' political liberalism and argued that citizens face significant cognitive difficulties before they can adopt the three basic aspects of reasonableness. In addition, I have postulated that at least one in nine citizens in a liberal democracy can be expected to adopt an anti-reasonable attitude. Therefore, I conclude that political liberalism must operate on the assumption of the fact of simple pluralism, not reasonable pluralism. This change in the empirical framework requires that political liberalism makes adaptations as a political theory. Therefore, let us now consider which potential responses are available to those theorizing in the tradition of Rawls, and examine the trade-offs involved.

This page has been intentionally left blank

Chapter 3: Critical Framework & Trade-offs Within Pluralism

Thesis III: That political liberalism aspires to the three core ideals of legitimacy, stability, and inclusivity, which cannot be fully realized simultaneously under the fact of simple pluralism, thus causing a trilemma.

Thus far, it has been established that reasonableness plays an essential role in the political theory of John Rawls, that there are substantial constraints imposed on this attitude by moral psychology, that a significant share of citizens in a liberal democracy will likely continue to adopt an anti-reasonable attitude, and that political liberalism must make adjustments if it can only postulate the fact of simple (and not reasonable) pluralism. In this third chapter, I will start by clarifying three core ideals of Rawls' theory. Then, given that political liberalism must now operate within the constraints of simple pluralism and thus deal with considerable challenges by the unreasonable, I posit that a liberal order can only endure if it develops a successful strategy to contain unreasonable doctrines and citizens. Next, I will assert that, given the necessity of containment, there is no obvious way for political liberalism to escape making a difficult and complicated trade-off between these three ideals within the limitations of simple pluralism. Hence, it is faced with a trilemma. Then, I explain the concept of reflective equilibrium and how this epistemic device opens up a way forward.

Rawls's core ideals are identified in the central question of political liberalism: *"How is it possible that there may exist over time a stable and just society of free and equal citizens profoundly divided by reasonable though incompatible religious, philosophical, and moral doctrines?"*[111] In other words, Rawls expresses a deep desire for the state to be stable (which he thinks is neglected but still essential in philosophy),[112] legitimate (since the use of coercive state power must be justified to the citizens),[113] and inclusive (because all citizens are regarded as equal and free).[114] I will first develop these three ideals one by one and then consider the tensions of how they interrelate under simple pluralism.

3.1 Legitimacy & Justification

The colloquial meaning of 'political legitimacy' is often descriptive or sociological in kind, indicating the *de facto* acceptance of political authority by the public population. In contrast, Rawls desires a normative account – for any sufficient definition ought to include what standard of acceptance is actually appropriate. Thus, he defines legitimacy as the proper moral justification by the state for the exercise of political and coercive power over its citizens.[115] For a liberal order in particular to be legitimate, Rawls asserts that it must abide by the following principle:

---

[111] Rawls, *Political Liberalism*, p. xviii.
[112] Ibid., p. xvii.
[113] Ibid., p. xliv.
[114] Ibid., p. 5-6.
[115] Ibid., p. 137. *"Seeing political power as the power of citizens as a collective body, [...] our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason."*

"The liberal political ideal is that [political] power should be exercised, when constitutional essentials and basic questions of justice are at stake, only in ways that all citizens can reasonably be expected to endorse in the light of their common human reason."[116]

In other words, for Rawls, political legitimacy is derived from a reasonable expectation of agreement among free and equal citizens on the public justification of political power. A legitimate political order is one that can be endorsed by all reasonable citizens from within their own comprehensive doctrines. [117] Rawls sees legitimacy as crucial for ensuring that cooperation in a pluralistic society remains stable and sincere without resorting to oppression or inquisition by a dominant comprehensive moral doctrine. Thus, legitimate political power would respect citizens as free and equal in its execution, and in its substantiation would appeal to public reasons that could be broadly and reasonably acceptable to all citizens that form the constituency of public justification.

Rawls provides several comments on what legitimacy would look like in practice. First, since it applies to procedures, laws, and policies, legitimacy is institutional.[118] It is also multilayered: for instance, a constitution can be legitimate on basic matters of justice while it supports some specific statute that is illegitimate.[119] Additionally, it involves having the right pedigree (i.e. having come about in accordance with established rules and traditions).[120] Furthermore, when there is a clash between political values and other non-political values in society, a legitimate regime has primacy in resolving conflicting values and their ideals ought to prevail.[121] Though Rawls' conceptualization of legitimacy involves intersubjective deliberation among the citizens, institutions carry the main responsibility of providing the conditions for endorsement or reconciliation by the citizenry.[122]

3.2 Stability & Modus Vivendi

Political stability is usually understood descriptively, indicating the situation of a political system that achieves continuity of government, consistency of policy, public order, social cohesion, and state security while operating without major disruptions for a prolonged period of time. However, a normative account is not just interested in the effects of stability, but also in the nature of the forces that secure it.[123] Thus, Rawls makes a twofold distinction concerning stability. First, there is stability as a mere *modus vivendi*. This occurs when citizens accept the constitutional essentials only as a pragmatic compromise to avoid instability and conflict – not because they affirm the principles as based upon their own comprehensive doctrines. Rawls regards this as a

---

[116] Ibid, p. 139-140.
[117] See Langvatn, "Legitimate, But Unjust; Just, But Illegitimate". Legitimacy does not require justice as fairness. Instead, Rawls assumes that reasonable pluralism applies – besides comprehensive doctrines – also to political liberal views of justice. Thus, the standard for political legitimacy becomes adequate justice in relation to the family of reasonable political conceptions of justice, guided by an ideal of public deliberation concerning constitutional essentials.
[118] Rawls, *Political Liberalism*, p. 427: *"Though there is of course an essential connection with justice."*
[119] Ibid., p. 393.
[120] Ibid., p. 427.
[121] Ibid., 137-138.
[122] Rawls's principle of legitimacy is less voluntaristic than typical social contract theories, for it requires that reasonable citizens *could* endorse the public justifications, but it does not require actual consent from citizens.
[123] Ibid.,. p. 147.

truce, and not a sufficient basis for a stable society. He points out: *"When social consensus is founded on self- or group interests, or on the outcome of political bargaining, social unity is only apparent, as its stability is contingent on circumstances remaining such as not to upset the fortunate convergence of interests."*[124] In short, the endurance of a modus vivendi depends on the relative balance of forces and on happenstance – there is nothing about such a political system intrinsically that ensures robust stability.

In contrast, Rawls advocates stability *for the right reasons*.[125] This occurs when citizens willingly comply with the constitutional order for personal reasons unique to their own comprehensive doctrines, and when they affirm the political conception of justice not merely as a compromise of power but for its own sake.[126] The resulting overlapping consensus is what secures true stability in a just society. The difference between oppressive and coercive political power consists in whether it is justified in terms that reasonable citizens are actively and eagerly endorsing, not (only) for pragmatic but also moral reasons.[127] Whereas legitimacy is made primarily possible by institutions, stability for the right reasons relies squarely on the citizens: it is contingent upon their actualized willingness to uphold the principles of justice over time in a well-ordered society governed by those principles. This means that any feasible political conception of justice should be able to generate its own support; for *"if a conception [of justice] fails to be stable, it is futile to try to realize it."*[128]

Challenges to stability for the right reasons might come in various forms.[129] For instance, technological development plays an important role in changing human society's system of cooperation.[130] Another example is cultural drift, which is relevant since religion evolves over time in unpredictable ways and informs one's imagination of what justice entails.[131] Moreover, there is the important challenge that is posed by each succeeding generation: how can young citizens born into a liberal society reliably be persuaded to willingly endorse its values and contribute to the flourishing of its institutions?[132] Still, these matters are regarded as external influences on the societal system of cooperation and thus separate from the challenge of unreasonableness, which is an internal one.

To what extent does the stability of political liberalism require the exclusion of unreasonable citizens and doctrines? Of course, every society must exclude some ways of life,[133] and, as Rawls accepts, it can never be assumed that all individuals will always

---

[124] Ibid., p. 147.
[125] Ibid., p. xl: *"As always, stability means stability for the right reasons."*
[126] Rawls, "The Domain of the Political", p. 250: *"It must not be political in the sense of merely specifying a workable compromise between existing interests, nor in looking to the particular comprehensive doctrines known to exist in society and then being tailored to gain their allegiance."*
[127] To clarify Rawls' position: stability for the right reasons is voluntarist – legitimacy is not.
[128] Rawls, *Political Liberalism*, p. 142.
[129] Ibid., p. 273.
[130] Still, as Murphy and Potts point out in *Culture and Technology* (p. 21): *"The relationship between technology and society cannot be reduced to a simplistic cause-and-effect formula. It is, rather, an intertwining, whereby technology does not determine but operates, and are operated upon in a complex social field."*
[131] For instance, as theorized by Robert Wright in *The Evolution of God*.
[132] Rawls, *Political Liberalism*, p. 198.
[133] Ibid., p. 197: *"Social influences favoring some doctrines over others cannot be avoided by any view of political justice. No society can include within itself all forms of life. But these social necessities are not to be taken for arbitrary bias or injustice."*

accept the justification for their own coercion.[134] Still, one could object that Rawls cannot coherently achieve both (a) the primacy of the moral power to have a conception of the good[135] and (b) an overlapping consensus of comprehensive doctrines affirming that primacy *"for its own sake"* rather than just as a pragmatic compromise.[136] However, the tension highlighted here does not necessarily mean incoherency. Concerning (a): citizens are not necessarily bound to pursue the specific conception of the good they currently endorse; they are capable of revising their conception based on reasonable and rational grounds.[137] Political liberalism assumes that it can reliably and persuasively provide such justifications. Regarding (b): a reasonable citizen supports the overlapping consensus for its own *sake* (i.e. endorsing the political outcome as an end, not means) but still for their own *reasons* (i.e. justified based on the moral foundation of their personal comprehensive doctrine). Political liberalism assumes that its account is sufficiently freestanding that each reasonable citizen can answer individually how the module of the conception of justice fits into and is supported by their comprehensive doctrine, thereby accommodating their liberty of conscience.[138]

3.3 Inclusivity & Constituency

In addition to the ideals of legitimacy and stability, Rawls deeply values regarding all citizens as free and equal. This third ideal, I will argue, is best described as *inclusivity*. It combines the relevant parts of three interrelated concepts in Rawls' work that, though normative and aspirational, are also each limited in their own particular ways.

First, egalitarianism. Rawls stated explicitly and repeatedly that he saw his political theory as an egalitarian one,[139] which clearly shows in multiple aspects of his body of thought. To start, there is his insistence on regarding all citizens as free and equal persons. Additionally, there is the veil of ignorance,[140] the difference principle,[141] and the idea of society as a fair system of cooperation.[142] Still, his egalitarianism leaves some gaps: for instance, he treats the family as a "special institution" in his writings, and does not provide a thorough account of justice within the family.[143]

Second, universality. On the one hand, Rawls seeks to provide a political conception of justice that is universal in the sense of being impartial and applicable to all citizens

---

[134] Ibid., p. 136: *"This [coercive] power is regularly imposed on citizens as individuals and as members of associations, some of whom may not accept the reasons widely said to justify [the constitution], or when they do accept [it], they may not regard as justified many of the statutes enacted by the legislature to which they are subject."*
[135] Ibid., p. 30.
[136] Ibid., p. 148. This objection comes from Talisse, "Rawls on Pluralism and Stability", p. 188-191.
[137] Rawls, *Political Liberalism*, p. 30.
[138] Ibid., p. 140.
[139] Ibid., 5-7. Also, Ibid., p. 7, fn 6. *"Some have thought that [political liberalism has] meant giving up the egalitarian conception of Theory. [...] I think the surmise has no basis."*
[140] Rawls argues that in the original position behind the veil of ignorance, rational self-interested individuals would choose principles that maximize the position of the worst-off, since they could end up in that position themselves.
[141] It states that social and economic inequalities are only justified if they benefit the least advantaged members of society.
[142] Ibid., p. 15.
[143] Ibid., p. 469: *"Political principles do not apply directly to its internal life."* As Walsh points out in "Private and Public Dilemmas: Rawls on the Family": *"Women's historical situation reveals the impossibility of separating political autonomy from human autonomy (and political autonomy from familial autonomy)."*

regarded as free and equal persons, regardless of their particular circumstances.[144] On the other hand, he argues that his political conception of justice is not derived from or justified by appeal to a universal moral order antecedent to and independent of particular societies and traditions.[145] Overall, his political constructivism aims to avoid the pitfalls of both moral intuitionism and relativism by providing a reasonable basis for objectivity within a shared political culture.[146] Thus, Rawls does value universality – though not as a metaphysical but only as a political ideal.[147]

Third, consensus. Rawls prefers that, if appropriate, the basic matters of justice and constitutional essentials are settled not by a simple majority vote (which could be slim), but rather by reaching a general agreement.[148] The consensus principle strives to maximize the share of parties that have coalesced around a political arrangement. It is contrasted with unanimity, which requires full agreement and can rarely be achieved on basic matters of justice. For instance, the idea of an overlapping consensus does not require unanimity on all political issues, but rather a convergence on a shared political conception of justice that respects certain basic rights and liberties. Reasonable disagreement on many non-political questions is still expected.

Since these three concepts each imply a valorization of citizens as being free and equal but also have their own particular limitations, I introduce the term *inclusivity*: this describes the ideal of making all citizens, being free and equal, part of the *constituency of public justification*.[149] The constituency of public justification is an idealized construct that represents the hypothetical pool of citizens to whom political liberalism ought to justify its principles and constitution, thus functioning as an external constraint on the theoretical possibilities.[150] Inclusivity is motivated by the fundamental liberal ideal to treat human persons as ends in themselves and thus part of this constituency; its major limitation is that some citizens will be unreasonable and beyond the reach of liberal public reason and are thus unavailable for justification. Nevertheless, liberal democracy ought not merely to mean the rule by a subset of society whose democratic commitments are restricted to its own ranks. Inclusivity retains the dimension of its colloquial use that indicates broad participation, representation, and access – specifically involving marginalized groups in society. Hence, the unreasonable must be incorporated and accounted for in some way. As Rawls writes:

> "There is not one account of toleration for reasonable doctrines and another for unreasonable ones. Both cases are settled by the appropriate political principles of justice and the conduct those principles permit."[151]

---

[144] Rawls, *Political Liberalism*, p. xix-xx.
[145] Ibid., p. 22-23.
[146] Ibid., p. 89-98.
[147] Ibid., p. 10.
[148] Ibid., p. 423-424.
[149] I borrow this latter term from Quong, *Liberalism Without Perfection*. This is necessary because Rawls himself uses this concept but does not provide it with a moniker.
[150] Ibid., p. 5-6: *"The constituency is not defined by its acceptance of any particular liberal principles or values, rather the validity of liberal principles and values is contingent on being justifiable to a constituency of persons that has been independently identified."* Ideally, the constituency of public justification is the same as the constituency of public reason: in that case, it consists of citizens who are reasonable and, in a public forum, express their moral-political reasons in terms that are accessible and acceptable to all other reasonable citizens.
[151] Rawls, *Political Liberalism*, p. 489.

3.4 The Trilemma

To recap: Rawls assumes the fact of reasonable pluralism, which includes the existence of a compelling majority of reasonable citizens. Then, he claims that: (a) *inclusivity* is achieved when all citizens, being regarded as free and equal, are part of the constituency of public justification; (b) *legitimacy* is achieved when the liberal order provides justifications for its coercive power that citizens can reasonably be expected to endorse; and (c) *stability* is achieved when the compelling majority of reasonable citizens forms an overlapping consensus by willingly supporting it for its own sake but for their own reasons. Rawls mentions (only twice and briefly) that political liberalism has a practical task to 'contain' unreasonable doctrines; however, he does not get much more specific about what that would entail.[152]

In contrast to Rawls, we assume merely the fact of simple pluralism, which contains a significant share of unreasonable elements. If left unchecked, the unreasonable will: (a) wield their influence as part of the constituency of public justification to (b) challenge the justifications of the liberal order and (c) undermine the overlapping consensus. Therefore, political liberalism ought to become much more explicit on what constitutes a proper policy of containment of the unreasonable, and how to restrict the spread of doctrines or reduce the influence of citizens who reject the essential principles of liberal democracy. In other words, the fact of simple pluralism gives rise to the problem of the unreasonable, which causes the issue of containment.

Applied to the trilemma, the obvious avenues for containment are either by (a) excluding unreasonable citizens from the constituency of public justification, (b) inculcating in them the public justifications of power, or (c) coercively restricting certain rights of the unreasonable if they threaten the overlapping consensus. Even if successful, these approaches would still mean deprioritizing either inclusivity, legitimacy, or stability. Hence, I assert that these three ideals are forced into an uneasy triangle given the significant presence of unreasonable doctrines and citizens in society. The result is the following trilemma:

> *Given the fact of simple pluralism, there is no obvious way for a liberal order based upon political liberalism to contain unreasonableness and simultaneously achieve ideal inclusivity, legitimacy, and stability: one of these ideals must be deprioritized for the proper realization of the other two to become possible.*[153]

Now, let us consider more in detail what each option of the trilemma would look like.

3.4.1 Prioritizing Inclusivity & Legitimacy over Stability

In the first option, the liberal order valorizes inclusivity and legitimacy. This means that it includes all citizens in its constituency of justification regardless of reasonableness and that it justifies its power in accordance with public reason. However, it leaves a sizable share of unreasonable citizens who do not support the overlapping consensus for the right reasons – or not at all. Rawls writes: *"Under many historical conditions, the efforts to achieve it may be overwhelmed by unreasonable and even irrational (and sometimes mad) comprehensive doctrines."*[154] This outcome could occur once

---

[152] Ibid., p. xvi-xvii & p.64 fn.19.
[153] Similar to Rawls, the trilemma assumes a closed society, and strict compliance by reasonable citizens with the political conception of justice, and sufficient economic and social development.
[154] Ibid., p. 126.

a liberal democracy has become a low-trust society. For an extreme example, consider the scenario where a group of unreasonable citizens overruns the seat of government of a long-standing liberal democracy and prevents a peaceful transition of power. This would entail that the overlapping consensus has deteriorated into a modus vivendi.

3.4.2 Prioritizing Inclusivity & Stability over Legitimacy

In the second option, the liberal order valorizes inclusivity and stability. This means that it includes all citizens in its constituency of justification and that the vast majority of citizens willingly support the overlapping consensus. However, it does not merely offer its public justifications of power to unreasonable citizens, but it indoctrinates them. For instance, by mandating that all basic education instills liberal values in children, by imposing strict censorship on illiberal media, and by strictly prohibiting seditious speech.[155] These methods are rejected by Rawls, who writes: *"Those coerced by law must be able to endorse the society's fundamental political arrangements freely, not because they are dominated or manipulated or kept uninformed."*[156] Making unreasonable citizens support the overlapping consensus for the right reasons requires a radical, perhaps excessive intervention in their lives that has insufficient potential to be endorsed by enough reasonable citizens. Thus, indoctrination of the unreasonable has no clear claim to liberal legitimacy.

3.4.3 Prioritizing Legitimacy & Stability over Inclusivity

In the third option, the liberal order valorizes legitimacy and stability. This means that it justifies its power in accordance with public reason and that the vast majority of citizens willingly support the overlapping consensus. However, unreasonable citizens are excluded from the constituency of public justification, thereby reinforcing liberal legitimacy and stability at the cost of inclusivity. This implies that citizens can only be regarded as equal and free citizens and as full members of the constituency of public reason if they have their moral powers *"to the requisite minimum degree to be fully cooperating members of society."*[157] For example, consider an exclusive democracy as envisioned by epistocratic theorists – they propose to make a citizen's right to vote conditional upon them successfully demonstrating their competency.[158] This suggestion would entail compromising inclusivity and reducing the constituency of public justification.

3.5 Reflective Equilibrium

It has been established that the trilemma provides three possible options, each of which seem unfavorable in some way. Thus, the question arises whether political liberalism can remain a plausible theory at all. It is here that the concept of *reflective equilibrium* becomes relevant: an epistemic device that involves a back-and-forth process of adjusting our intuitive judgments and theoretical principles until the moral-political theory has been solidified into a coherent set of moral principles.[159] Rawls maintained that political theories ought to be judged according to this standard. The goal of the reflective equilibrium is to resolve the discrepancies between considered judgments

---

[155] This restricts, respectively, the freedom of education, association, and speech. As a non-ideal example, Singapore implemented such measures while remaining a high-trust society.
[156] Ibid., p. 446.
[157] Ibid., p. 19.
[158] Brennan, *Against Democracy*. For an opposing view, see Gunn, "Against Epistocracy".
[159] Rawls, *Political Liberalism*, 95-97.

about particular cases and general principles that arise from the innate human sense of justice into a stable harmony that can then provide consistent practical guidance.

Practicing the method of reflective equilibrium opens up a way forward since it consists of reasoning towards a dynamic coherence and instead of reasoning from unrevisable foundations. Thus, the problem posed by the trilemma is not how political liberalism can simultaneously achieve all three of its core ideals *fully* – instead, the question is how these ideals can be integrated *coherently* into one vision with conceptual harmony, acceptable trade-offs, practical plausibility, and clear action-guidance. After all, none of Rawls' core ideals can be charitably interpreted as absolute. Though we value inclusivity and maximize the constituency of public justification, nobody proposes that we let young children vote. Though we value legitimacy and the public justification of power to *anyone*, in an imperfect world it can never be acceptable to *everyone*.[160] Though we value stability for the right reasons, it can only be fully realized by an inquisition that Rawls' project disavows at the very start.[161] Moral principles are inherently underdetermined,[162] so we ought to explore the possibilities made available by leeway, adaptability, and even compromise.[163]

3.6 Conclusion

In this third chapter, I have clarified that the political theory of Rawls values three core ideals in the form of legitimacy, stability, and inclusivity. Next, I asserted that the fact of simple pluralism and the necessity of containing the unreasonable forces these ideals into an uneasy triangle, allowing for three options that are each unfavorable for deprioritizing a different ideal. Still, Rawls' principle and epistemic device of reflective equilibrium show a way out of the trilemma: to defend the plausibility of political liberalism by bringing its three core ideals into a coherent harmony with each other, going back and forth between our practical judgments, moral intuitions, and reasoned principles until a comprehensive strategy of containment can be devised.

This concludes Part I, where I have sought to establish multiple things. First, I have argued that reasonableness as conceived by Rawls includes a third basic aspect in the form of the duty of prudence, which allows us to make the epistemic responsibilities of being reasonable more explicit. Second, I have examined the limitations imposed upon political liberalism by moral psychology and asserted that postulating the fact of simple pluralism is significantly more warranted than the fact of reasonable pluralism. Finally, I have asserted that the persistent presence of unreasonable citizens poses a significant trilemma for Rawls' political liberalism, which in recognition of the fact of simple pluralism needs to reconsider the cohesion of its ideals of legitimacy, stability, and inclusivity. Now, as I proceed towards Part II of this thesis project, I will consider currently available responses to the challenges posed by trilemma.

---

[160] For instance, Quong points out that murderers are highly disincentivized to accept the arguments offered to them when their liberty is restricted by the state, but this does not render that restriction illegitimate. See *Liberalism without Perfection*, p. 312.
[161] Rawls, *Political Liberalism*, p. 37.
[162] Bader, "Taking Pluralism Seriously: Arguing for an Institutional Turn", p. 379-381.
[163] Rawls, *Political Liberalism*, p. 163.

PART II: Criticizing Current Responses

Thus far, Part I has clarified the main concepts of the central inquiry of this thesis topic: *"How should a liberal order navigating the fact of simple pluralism achieve the containment of unreasonableness while remaining normatively consistent with political liberalism?"*
In particular, Part I established that a significant and enduring presence of unreasonable citizens and doctrines in liberal society causes a *trilemma* for Rawls' political liberalism: it must somehow preserve its ideals of inclusivity, legitimacy, and stability while also achieving containment of unreasonable movements so that they do not overwhelm liberal institutions or undermine the unity and justice of society. Throughout Part II, I will provide a critical review of several suggestions in political philosophy on how a liberal order can address this problem. My focus will be specifically on the issue of determining the principles underlying a containment strategy of the unreasonable, which must be effective in practice but also normatively consistent with the philosophical commitments of political liberalism.

I define *containment* as any policy with the primary aim of restricting the spread of doctrines that reject the essential principles of liberal democracy, reducing the institutional threat of politically influential citizens who ardently adhere to such doctrines,[164] or restoring the desired equilibrium where a compelling majority of the population consists of reasonable citizens. A doctrine or a citizen who adheres to that doctrine is considered unreasonable if they fail to meet any of the following criteria:[165] they (1) believe that political society should be a fair system of social cooperation for mutual benefit, (2) regard all citizens as free and equal persons, (3) accept the burdens of judgment, (4) are responsive to reasons,[166] and (5) assign these ideals deliberative priority when they engage in practical reasoning. Also, it should be stipulated that containment does not simply mean regular application of the political conception of justice – in whatever policy form, it should remain distinct from protecting of basic rights (though it may produce that effect).

The subsequent discussion will prioritize those philosophical publications that are directly relevant to the problem that unchecked unreasonableness poses for a liberal order. As a result of this narrow focus, not all tangentially related discussions will be addressed. In particular, two omissions warrant a short explanation. Although these theories acknowledge the fact of simple pluralism, their potential suggestions to circumvent the trilemma do not fit well in the context of this thesis project.

First, liberal nationalism. Most notably advocated by David Miller,[167] it aims to reconcile the liberal values of individual rights, democracy, and equality with nationalist sentiment, to provide its institutions with social solidarity. Its hope is that incorporating moderate nationalist ideals can bring enough unreasonable citizens into the broader liberal tent. However, I will not include liberal nationalism in this thesis

---

[164] In theory, containment is primarily aimed at unreasonable doctrines. Unfortunately, practical application of containment policies will often directly involve unreasonable citizens. Yet, strictly speaking, it is not their person which is addressed but their unreasonable claims and actions. A citizen is unreasonable due to their behavior and doctrine, not because they belong to a distinct class of persons in society.
[165] This definition is inspired by the dominant definition in the debate, provided by Quong in *Liberalism Without Perfection*, p. 299. However, it is altered and expanded for greater accuracy.
[166] Freeman, *Rawls*, p. 346.
[167] Miller, "On Nationality".

project for several reasons. Firstly, it is not compatible with political liberalism: it replaces the moral-political primacy of a shared political conception of justice with that of the nation. Secondly, because its own liberal character is questionable: in practice, its criteria apply equally to conservative accounts of nationalism.[168] Thirdly and most importantly, available empirical evidence suggests that promoting nationalism – even when done in the name of liberalism – comes at the unacceptable cost of increased xenophobic sentiments among the population.[169] Hence, I think that its proposed strategy of dealing with the unreasonable is not relevant for our purposes.

Second, agonism. This democratic theory regards conflict as a political good, reconceptualizing political struggle as a core feature of vibrant democracies. As Chantal Mouffe, its most famous advocate, writes: *"In a democratic polity, conflicts and confrontations, far from being a sign of imperfection, indicate that democracy is alive and inhabited by pluralism."*[170] Agonism suggests that the unreasonable should not be contained, but endlessly contested in the public arena. Nevertheless, I have decided to omit the agonistic perspective for several reasons. Firstly, because it is incompatible with political liberalism: it rejects the idea of an overlapping consensus and instead embraces engaging with the unreasonable as adversaries. But more importantly: agonism has no real advantage over political liberalism, as it faces its own parallel problem caused by the fact of simple pluralism. Whereas political liberalism relies on citizens being reasonable, agonism postulates widespread mutual respect among the citizenry (which is what separates enemies from adversaries and verbal conflict from violent conflict). Just as political liberalism seems to have no reliable way to achieve a compelling majority of reasonable citizens, it is unclear how agonism can develop widespread civic respect out of mere antagonism.[171]

Rather, my discussion in Part II will focus on the main options for a liberal order to achieve containment of unreasonableness suggested in the current academic literature of political theory. In short, these come down to either conciliation, transformation, or marginalization of the unreasonable – each with its own strengths and flaws. *Conciliation* attempts to settle tensions between the liberal order and the unreasonable in good faith by friendly engaging in mutual dialogue, and prioritizes concessions over escalation into conflict; it will be discussed in Chapter 4. Alternatively, *transformation* attempts to convert unreasonable citizens into supporting public justifications of power by establishing societal structures that reorient the fundamental moral-psychological condition towards greater reasonableness – thus trying to develop the fact of simple pluralism into the fact of reasonable pluralism. Chapter 5 will elaborate on this. Finally, *marginalization* attempts to reduce the destabilizing capacity of unreasonable citizens by revoking some of their rights when appropriate and counters the spread of unreasonable doctrines by taking them off the agenda of public discourse. This notion will be further developed in Chapter 6.

Exploring these options will enable us to evaluate what constitutes a successful strategy for containing the unreasonable both in principle and in practice. I will now continue and examine the three selected options individually.

---

[168] Daniel, "The liberal/conservative nationalism divide: A distinction without a difference?"
[169] Hjerm, "Defending liberal nationalism—at what cost?"
[170] Mouffe, *Deliberative Democracy Or Agonistic Pluralism*, p. 34.
[171] Knops, "Debate: Agonism as Deliberation?", p. 115.

Chapter 4: Conciliating the Unreasonable

Thesis IV: That a conciliatory approach to containment, as exemplified by a method of appeasement or conjecture, is an ineffective strategy that does not properly address the fundamental problem of unreasonableness.

How can a liberal order, faced with the trilemma caused by simple pluralism, appropriately contain unreasonableness? In this fourth chapter, I examine two suggestions from political philosophy that advocate a conciliatory stance towards unreasonable citizens. Since both proposals prioritize inclusivity and legitimacy over stability, they seem to prefer the first option provided by the trilemma. The first approach, argued by Kelly and McPherson, downplays the problem of unreasonableness and maintains that full inclusivity is warranted without further caveats concerning stability. The second approach, implied by Rawls himself and developed by Clayton and Stevens among others, contends that the method of conjecture can persuade unreasonable citizens to support the overlapping consensus for the right reasons, and thereby ultimately enhance stability. I will now explain and examine these two approaches in turn.

4.1 Conciliation by Appeasement

Some political theorists, such as Erin Kelly and Lionel McPherson,[172] dismiss the idea that unchecked unreasonableness poses a noteworthy problem for political liberalism at all. In support of this position, they argue three main points.

First, they introduce a distinction between philosophical and political unreasonableness. *Philosophical unreasonableness* refers to privately adhering to a doctrine while being unable to provide 'good arguments' in support that others can evaluate rationally.[173] *Political unreasonableness* refers to a public unwillingness to abide by fair terms of cooperation or to treat other citizens as free and equal. In other words: the philosophically unreasonable lack the duty of prudence in private,[174] whereas the politically unreasonable lack the criterion of reciprocity in public. Though Kelly and McPherson grant that philosophically unreasonable doctrines are likely irrational,[175] they claim that this alone does not indicate any threat, provided that its adherents remain politically reasonable and respect the liberal order in their political behavior.

Second, Kelly and McPherson contend that most unreasonable people are merely so in the philosophical sense and do not actually threaten the liberal order. In their view, most unreasonable people usually withhold their unreasonable views from the public sphere out of fear of social sanction and separate these views from their public political behavior when they see the need to cooperate with the liberal order. Additionally, Kelly and McPherson imply that unreasonable doctrines are inherently limited in their

---

[172] Kelly and McPherson, "On Tolerating the Unreasonable," p. 38-55.
[173] Ibid., p. 44: *"Good arguments are arguments that plausibly rely on empirical evidence -when relevant), give compelling and clear accounts of the basis of their claims, and offer nondogmatic interpretations of: how to understand key concepts, the priority ranking of values, the appropriate weight of various considerations, and the ways in which experience supports value judgments."*
[174] Kelly & McPherson see philosophical unreasonableness as 'implicit' in the burdens of judgment. Thus, it is appropriate for our purposes to take them to mean the duty of prudence.
[175] Here, 'irrational' indicates being fallacious, prejudiced, biased, or informationally blind.

ability to gain widespread traction because, once subjected to public scrutiny, they will rarely persuade the reasonable majority. In short, their view is that for the vast majority of unreasonable citizens, their unreasonable beliefs do not translate to unreasonable behavior that could destabilize liberal institutions.

Third, Kelly and McPherson assert that any approach of containment by political liberalism would itself be unreasonable. Supposedly, efforts to contain those who are philosophically unreasonable in their views but still politically reasonable in their behavior would undermine core principles of political liberalism, specifically toleration, equal justification to all citizens, and avoiding unreasonable exclusions. The only coherent course of action, they maintain, is to involve unreasonable people in the constituency of justification.

In this way, Kelly and McPherson oppose containment as unnecessary and unreasonable, instead contending that unreasonable people ought to be included outright within the constituency of public justification. Let us now evaluate their case.

4.2 Evaluating Appeasement

Kelly and McPherson provide some worthwhile insights – for instance, the idea that unreasonableness may consist of certain subcategories and the view that political liberalism ought to remain theoretically coherent and consistent. Yet, their position may be aptly described as appeasement: they propose to attempt to pacify the unreasonable by making concessions to their demands while downplaying them as a threat. In response, I argue that each of their three major contentions is deeply flawed.

Firstly, the concept of philosophical unreasonableness relies on multiple serious misconceptions. In the Rawlsian sense, reasonableness is a public, political attitude. Yet, Kelly and McPherson's concept of philosophical unreasonableness seems to also describe a private, epistemological status.[176] Thus, their distinction provides more conceptual confusion than clarity. Another misconception occurs when Kelly and McPherson propose that the constituency of justification includes all citizens who abide by the criterion of reciprocity – thereby leaving out the burdens of judgment in their account of political reasonableness. As pointed out by Jonathan Quong,[177] this leaves a key question unanswered: what moral reason could politically reasonable citizens have to willingly support the political conception of justice?[178] Since these persons reject the burdens of judgment, it should be expected that they support the liberal order only as a mere modus vivendi, opportunistically waiting for the right moment to seize power and impose their comprehensive doctrine. In short, I find the distinction between philosophical and political unreasonableness to be untenable.

About the second main claim of Kelly and McPherson – that most unreasonable citizens frequently and willingly separate their private beliefs from their public behavior – I think several criticisms are in order. First of all, in the age of social media,

---

[176] As illustrated by their examples that discuss reasonableness in the context of a *"private all-male club"* (Ibid., p 39.) or *"metaphysical tenets"* (p. 47). It is unclear how these examples relates to reasonableness in a Rawlsian sense – since Rawls explicitly disavows that public reason applies to the background culture or has a metaphysical dimension.

[177] Quong, *Liberalism Without Perfection*, p. 296.

[178] Kelly and McPherson only address this issue as an aside: *"There presumably are those, few as they may be, with philosophically unreasonable views who believe that the values of toleration and equal citizenship are politically fundamental."* ("On Tolerating the Unreasonable," p. 54)

this claim has the burden of proof and must be updated with extensive additional argumentation to remain plausible. Moreover, social sanction against the unreasonable requires that a clear majority of reasonable citizens is reasonable – and this very assumption is thwarted by the fact of simple pluralism. Besides, when unreasonable citizens feel that their beliefs and actions are contradictory, the resultant cognitive dissonance could cause psychological discomfort to such an extent that it will outweigh other considerations, like social scorn from the reasonable.

However, I think that the most important objection on this point concerns Kelly and McPherson's misunderstanding of unreasonable discourse. Their argument relies upon the assumption that unreasonable doctrines have an inherently limited public appeal because these views do not hold up to public scrutiny and will therefore rarely persuade the reasonable majority. I contend that this implication is problematic in multiple ways. Firstly, unreasonable discourse undermines the liberal order not by relying on persuasion, but primarily by creating confusion and instilling apathy.[179] Secondly, unreasonableness does not need to achieve a majority itself before it can overwhelm liberal institutions – to induce instability, it merely needs to decrease the majority of reasonable citizens from compelling to slim. Thirdly, there are ways in which unreasonable doctrines can avoid public scrutiny while still infiltrating the public discourse.[180] Fourth, public scrutiny is itself not an infallible antidote against unreasonableness and can even prove counterproductive.[181] Fifth, even just 'rarely' persuasive unreasonable doctrines only need to successfully overwhelm liberal institutions once to succeed in their goal; thus, a lower level of instability is not equivalent to true stability. Together, these objections reveal that the vulnerabilities of Kelly and McPherson's second claim stem from their tenuous conceptualization of the unreasonable in terms that are marked by a fundamentally reasonable perspective.

Thirdly, Kelly and McPherson claim that any attempt at containment by political liberalism would itself be unreasonable. However, this assertion is unsound.[182] To paraphrase Quong: the moral force of public reason extends beyond merely those who endorse its premises.[183] First, it should be noted that the discussion about including the unreasonable in the constituency of public justification is separate from the question of their political rights. Rawls is very clear that, even if unreasonable citizens are excluded from the constituency of public justification, they still retain the benefits of citizenship: political liberalism affirms that unreasonable citizens receive the full protection of the system's basic rights and liberties.[184] The exclusion regards a hypothetical construct, not their moral status as free and equal persons. Thus, Kelly and McPherson's insistence that unreasonable citizens are *"owed a justification, just like*

---

[179] See: Hinck et al., "Authoritarians Don't Deliberate".

[180] For instance, by using *dogwhistles*: coded language that conveys a secondary, more insidious meaning to a specific target audience, while appearing innocuous or having a different literal meaning to the general public.

[181] For example, fact-checking is only partially effective. See: Nathan et al., *"Fact-checking: A meta-analysis of what works and for whom"*, p. 350.

[182] This third objection is derived from Quong, *Liberalism Without Perfection*, p. 293-299 & 312-314.

[183] Ibid., p. 293: *"The original position is not really a social contract, nor is it an actual agreement between parties. [...] It is crucial to always keep in mind that the original position is meant as a device of representation from which we draw conclusions about justice."*

[184] Rawls, *Political Liberalism*, p. 489: *"There is not one account of toleration for reasonable doctrines and another for unreasonable ones. Both cases are settled by the appropriate political principles of justice and the conduct those principles permit."*

their reasonable counterparts are" is misplaced.[185] In any case, *all* citizens will receive *a* justification for the use of coercive state power – the problem is that unreasonable citizens reject the *public* justification on offer as illegitimate.[186]

Any political conception of justice developed by reasonable citizens will be unacceptable to unreasonable citizens, but political liberalism does not have to abide by the improbable principle that political power must gain unanimous acceptance to become legitimate. For example, criminals may not agree with the reasoning behind the state's decision to limit their freedom, yet this disagreement does not render such restriction wrongful or illegitimate. It is not that political liberalism offers reasonable citizens some justification and unreasonable citizens none. Instead, as Quong writes:

> "***All*** *persons are offered a justification for the exercise of political power that they could endorse* ***in their capacity as*** *free and equal citizens committed to the idea of society as a fair system of social cooperation.*"[187]

After all, Rawlsian legitimacy is not voluntarist: it requires that all reasonable citizens could endorse the public justifications, but it does not require actual consent from citizens. Thus, rejection of fundamental liberal values by unreasonable citizens does not automatically mean that a liberal order treats them disrespectfully or unfairly when its justifications of political power are based on the principles of freedom, equality, and fairness. Indeed, unreasonable citizens lie beyond the realm of liberal public reason as they do not recognize the moral authority of its constitutional essentials. That fact does not invalidate political liberalism or the political conception of justice any more than the existence of criminals should undermine faith in the value of laws. Instead, it is necessary for the viability of political justice that the unreasonable are contained somehow; being reasonable towards the unreasonable would not increase reasonableness but unreasonableness. Exclusion does not mean that unreasonable citizens are denied any of their basic civic rights, but only their status as members of the constituency which deliberates on what those basic civic rights ought to be. Hence, contrary to what Kelly and McPherson claim, containment of the unreasonable is consistent with being reasonable and political liberalism as a whole.

Because of the aforementioned objections, appeasement of the unreasonable falls short of a satisfying approach under the fact of simple pluralism. However, conciliation with the unreasonable might be achieved via the alternative method of conjecture.

4.3 Conciliation by Conjecture

Another conciliatory approach to contain unreasonableness without coercion relies upon public reason and the burdens of judgment, hoping to persuade the unreasonable over time through continual public debate and reasoned dialogue. Rawls suggests that, if citizens cannot be brought to endorse a political account of tolerance, they could conceivably be motivated by alternative accounts that are expressed from within one's own comprehensive doctrine. Here is – in full – what he writes on *conjecture*:

---

[185] Kelly and McPherson, "On Tolerating the Unreasonable", p. 43.
[186] For an opposing view, see: Friedman, "John Rawls and the Political Coercion of Unreasonable People". I agree with Quong, Gurzoslu, and others in their assessment that Friedman's interpretation of Rawls is uncharitable and incorrect. Thus, I will not discuss it further in the context of this thesis project.
[187] Quong, *Liberalism Without Perfection*, p. 313. Emphasis by Quong.

"[This is] what I call reasoning from conjecture. In this case we reason from what we believe, or conjecture, may be other people's basic doctrines, religious or philosophical, and seek to show them that, despite what they might think, they can still endorse a reasonable political conception of justice. We are not ourselves asserting that ground of toleration but offering it as one they could assert consistent with their comprehensive doctrines. [...] Conjecture is defined thus: we argue from what we believe, or conjecture, are other people's basic doctrines, religious or secular, and try to show them that, despite what they might think, they can still endorse a reasonable political conception that can provide a basis for public reasons. The ideal of public reason is thereby strengthened. However, it is important that conjecture be sincere and not manipulative. We must openly explain our intentions and state that we do not assert the premises from which we argue, but that we proceed as we do to clear up what we take to be a misunderstanding on others' part, and perhaps equally on ours."[188]

In other words, conjecture is a method where one engages in open conversation with unreasonable citizens and offers them reasons from within their comprehensive doctrine to endorse a reasonable political conception of justice. It is an attempt to make unreasonable citizens more reasonable in their attitude by providing them reasons that they have not considered yet, but could still accept without radically changing their conceptual framework. Provided that the effort is genuine, Rawls hopes that conjecture can appeal to unreasonable citizens even when the shared premises of public reason seem beyond reach.[189]

Rawls' account could benefit from further development, and various political theorists have elaborated on the concept.[190] Yet, their attempts differ significantly in how they image the purpose of conjecture. For instance, Alessandro Ferrara views conjecture primarily as a means to legitimize reasonable liberal democratic institutions,[191] whereas Micah Schwarzman focuses on ethical considerations and addresses communitarian objections to political liberalism.[192] Conversely, Matthew Clayton and David Stevens specifically mention containment of the unreasonable as the main reason to engage in conjectural reasoning.[193] Their account is centered on the identity of conjecturers and whether they need to embody proximity to the belief systems of their interlocutors. In contrast to Rawls, who posits that conjecturers argue from premises that they do not share, Clayton and Stevens argue that reasonable conjecturers should largely share the same doctrinal beliefs as their interlocutors. Their reason is that the criterion of sincerity is much more likely to be met and the effort at persuasion much more likely to succeed if the unreasonable interlocutor regards the conjecturer as part of their in-group, which improves trust and decreases suspicion.

4.4 Limits of Conjecture

Though the method of conjecture may still prove useful for other purposes, I argue that there are at least three reasons to believe that its potential to make a substantial contribution to systematically containing unreasonable doctrines is limited.

---

[188] Rawls, *Political Liberalism*, p. 461-461 & 465-466.
[189] Since conjecture does not presuppose shared premises, it is distinct from public reason. Still, it remains a form of reasoning since it involves developing and evaluating arguments – thus, it is not merely a rhetorical device.
[190] For instance, see: March, *Islam and Liberal Citizenship*, p. 13-33.
[191] Ferrara, *The Democratic Horizon*, p. 71–72.
[192] Schwarzman, "The Ethics of Reasoning from Conjecture", p. 523.
[193] Clayton and Stevens, " Political Liberalism and Unreasonable Religions," p. 73.

4.4.1 Practices of Deradicalization

Containment by persuasion is in some ways quite similar to attempts at deradicalization. Not only do both approaches share the same goal of dissuading citizens from adopting anti-liberal views incompatible with liberal democracy, but they also deploy the same method of reframing the interlocutor's own perspectives to make it compatible with liberal principles, rather than simply imposing an external liberal view. However, beyond these similarities, the decidedly doctrinal approach of conjecture stands in stark contrast with the actual practices of counter-radicalization that have been found effective in contemporary liberal democracies. This gap between the theory of conjecture and the practice of counter-radicalization is symptomatic of the fundamental shortcomings of conjectural reasoning.

There exists a wide variety of effective approaches to deradicalization.[194] Some are holistic and centered around the principle of inclusion, offering participants wide-ranging services: not only counseling, but also healthcare, housing, and employment. This approach is complemented by a trained personal mentor in whom participants can confide about their struggles and who can help them find paths of social inclusion in their daily lives. Rather than examining participants' religious beliefs directly, the approach aims to transform the totality of personal, social, cultural, and political elements into an attitude of reasonable citizenship.[195] Other approaches have advocated a community-driven approach to deradicalization that focuses on social cohesion[196] or on recontextualizing what citizenship means for a marginalized group by highlighting its social contributions in the history of that country.[197]

Together, these examples illustrate that actual practices of deradicalization can differ significantly from conjecture. While each of these initiatives emphasizes engagement and dialogue, they lack the central focus on *doctrinal* engagement that characterizes the conjectural framework. Instead, if discussion of doctrine is included at all, it serves as one smaller facet of the total deradicalization program, not as its core component. Thus, the first reason to be skeptical of conjecture as an effective countermeasure to unreasonableness is the currently available empirical evidence on deradicalization.

However, I argue that there are at least three reasons why conjecture should not be rejected completely as a worthwhile notion to contain unreasonable doctrines. Firstly, many deradicalization practices do deploy it in some manner, even if they do not emphasize it. Secondly, the non-conjectural approaches to deradicalization could also be understood as attempts to restore the conditions that must be present before citizens become open to reasonable dialogue at all. In that sense, deradicalization approaches are maybe not opposed to conjecture but rather serve as its precursor. Thirdly, not all unreasonable citizens are radicalized: it is conceivable that those who abide by a reasonable conception of justice merely as a modus vivendi would be more susceptible to conjecture than those who oppose the overlapping consensus outright. In short, the empirical evidence clarifies the conditions and limits of conjecture, rather than demonstrating its inefficacy.

---

[194] Popp et al., "Common Characteristics of Successful Deradicalization Programs of the Past".
See also Doosje et al. "Terrorism, radicalization and de-radicalization", p. 79-84.
[195] Bertelsen, "Danish Preventive Measures and Deradicalization Strategies", p. 241–53.
[196] Ahmad, "Youth Deradicalization", p. 119-168.
[197] Shabi, "Deradicalising Britain".

4.4.2 Available Persuasive Conjecturers

Besides empirical considerations, conjecture faces practical caveats. For instance, Gabriele Badano has highlighted that there are probably few available persuasive conjecturers.[198] After all, reasoning from conjecture is quite difficult – it is insufficient to merely possess a vague understanding of the relevant comprehensive doctrine or to have only passing familiarity with the unreasonable interlocutor. Substantial expertise is required: the conjecturer must be knowledgeable about the important texts, influential interpretations, historical developments, cultural traditions, and relevant authorities of the relevant doctrine. Therefore, most politicians will not be suitable conjecturers, as they cannot reasonably expected to possess the requisite skills. Similarly, ordinary citizens are ruled out as likely candidates – for they are wont to simply refer to their doctrinal authorities to provide satisfying answers on moral-political issues. Furthermore, even if somebody possesses all the relevant knowledge, they will need to adjust their conjectural arguments to the specific concerns of their interlocutor in order to be persuasive. This means that most scholars are also excluded since they often lack the individualized insights necessary for effective conjecture. Ultimately, it seems that the only people who are suitably positioned to engage in conjecture are reasonable and educated leaders of doctrinal communities. Even assuming that conjecture is an effective strategy for containment, only a few citizens can bring it into practice.

However, I think this point by Badano is a caveat rather than an objection for several reasons. Firstly, even if the potential for persuasive conjecture is limited across doctrines, it could still remain effective within one's own doctrine and cultural community. After all, reaching out to other doctrines only becomes truly necessary if that doctrine internally has no reasonable community leaders at all. Under the fact of simple pluralism, there is no clear majority of reasonable citizens among the whole population; however, there could perhaps still be a clear reasonable majority among just the group of educated community leaders. Secondly, the limited pool of available persuasive conjecturers could be mitigated by amplifying their voices using modern technology. In the digital age, technological advancements such as online education platforms, virtual dialogue forums, and social media outreach campaigns could be used to facilitate engagement with conjectural reasons on a larger scale. Thirdly, perhaps it is possible to train prospective conjecturers in whatever skills or knowledge they currently lack. It is an open question whether the pool of available persuasive conjecturers can be expanded by the implementation of such training programs – educational initiatives aimed at equipping individuals with the necessary skills and expertise for effective conjectural engagement. All in all, though there are substantial practical challenges to conjecture, they need not be insurmountable.

4.4.3 The Dilemma of Doctrine

Aside from these challenges, I argue that the biggest problem with implementing conjecture as an effective containment strategy arises due to theoretical problems. This objection can be expressed in what I call the *dilemma of doctrine*. It asks the following question: "Has the unreasonable interlocutor who will potentially be swayed by conjectural reasoning based their attitude in a significant way upon their

---

[198] Badano, "The Limits of Conjecture", p. 302-303.

comprehensive doctrine or not?" Whether the answer is yes or no, conjecture seems to encounter significant trouble.

The first horn of the dilemma is that the interlocutor is indeed unreasonable because of doctrinal reasons. If this is the case, then the reasonable conjecturer will likely be unable to overcome the burdens of judgment in their appeal. To recap, the burdens of judgment entail the notion that even if all persons engaged in discussion are intelligent, well-informed, and well-intentioned, disagreements about moral-political issues are still inevitable. In light of this, Badano has pointed out that conjecture, as a method seeking consensus on reasonable political conceptions, must inherently rely on certain contentious moral-political assumptions or religious premises for its attempt at persuasion to make sense – and by definition, the conjecturer cannot share all assumptions with the interlocutor, or they would already have seen the consistency between their comprehensive doctrine and a reasonable political attitude.[199]

Moreover, conjecture involves a multi-step process; multiplying the argumentative steps that need to be taken also multiplies the opportunities for the burdens of judgment to interfere with the reasoning process. After all, the argument advanced by the conjecturer must necessarily deal with vague concepts, complex evidence, and conflicting considerations at each stage. Due to the cumulative effects of the burdens of judgment, even adept conjecturers cannot realistically be expected to reliably persuade unreasonable citizens toward reasonable political ideals.

The second horn of the dilemma is that the interlocutor is unreasonable mostly because of reasons that are independent of their comprehensive doctrine. If this is the case, reasoning from conjecture is misplaced and will likely prove a fruitless endeavor. For instance, consider the phenomenon that moral psychology refers to as *moral dumbfounding*, which occurs in the following scenario.[200] First, a person makes some strong moral judgment, usually condemning certain behavior as morally wrong. Then, when that person is pressed to explain the reasons behind their judgment, they subsequently struggle to provide coherent justifications that actually support their stance. Nevertheless, in the case of moral dumbfounding, this person remains firmly committed to their initial moral condemnation and refuses to revise their judgment – even after admitting that they cannot provide valid reasons in justification.

There are all kinds of reasons why individuals would overrule reasons related to their comprehensive doctrine and persist in their unreasonable attitude. For example, they might have developed deeply ingrained mental habits during their upbringing, or regard any changes to their moral-political dogmas as tantamount to betraying their group of origin. Alternatively, they might desire power for its own sake and have a cynical view of moral-political reasons, only engaging with doctrine as rationalization. Moreover, the unreasonable interlocutor could be unmoved by reasoned appeals to their moral-political doctrine because that doctrine lacks sufficient theoretical cohesion.[201] For instance, it is nearly impossible to persuade somebody towards greater

[199] See Badano, "The Limits of Conjecture", p. 303-305.
[200] Jacobsen, "Moral Dumbfounding & Moral Stupefaction", p. 289.
[201] Such an accusation should not be made lightly – it applies mostly to conspiracy theories that grow into comprehensive doctrines, not esoteric theories in general. Still, this scenario is relevant, given recent historical examples of unreasonable forces overwhelming a reasonable institute.

reasonableness if they have already adopted a conspiratorial mindset and view of the world: any dislodging attempt would be further proof of the supposed conspiracy.[202]

In short, the dilemma of doctrine says that either doctrine dictates unreasonableness, in which case attempts at persuasion run into the burdens of judgment, or that unreasonableness dictates doctrine, in which case conjecture is misguided. All things considered, the method of conjecture has some merit but nevertheless is too limited in its potential to significantly impact the systematic and successful containment of unreasonable doctrines in a liberal society, especially once they are included in the constituency of public justification.

4.5 Conclusion

In this fourth chapter, I have examined the merit of a conciliatory approach to contain the unreasonable. In the form of appeasement, as argued by Kelly and McPherson, I contend that such an attempt is counterproductive. I have argued that the distinction between philosophical and political unreasonableness is untenable, that it is a misconception to assume that the destabilizing capacity of unreasonable doctrines or citizens is inherently limited, and that containment of the unreasonable is consistent with being reasonable. Alternatively, conciliation could take the form of conjecture, as argued by Clayton and Stevens. Besides noting that the empirical evidence on deradicalization is cause for a skeptical attitude concerning its effectiveness and that there are substantial practical challenges to overcome in finding suitable persuasive conjecturers, I also argued that conjecture cannot solve the dilemma of doctrine. For these reasons, I maintain that conjecture is ultimately an insufficient answer to the question of how the unreasonable ought to be contained. Still, though the unreasonable are unlikely to be placated by direct persuasion, perhaps they can be transformed in a more indirect approach.

---

[202] Napolitano, *"Conspiracy Theories and Evidential Self-insulation"*, p. 82-106.

52

Chapter 5: Transforming the Unreasonable

Thesis V: That a transformational approach that incorporates institutional, educational, and technological options – although not sufficiently effective on its own – is an essential strategy to address the problem of unreasonable doctrines rejecting the political conception of justice.

How can a liberal order, given the fact of simple pluralism, achieve the containment of unreasonable doctrines? Perhaps, unreasonable citizens can over time be brought towards reasonableness by means other than appeasement or conjecture. In this fifth chapter, I examine several suggestions from political philosophy that advocate the transformation of the basic moral-psychological status of unreasonable citizens, and, instead of attempting direct persuasion, aim to indirectly nudge their attitude towards greater reasonableness. These responses, by prioritizing inclusivity and stability over legitimacy, seem to prefer the second option provided by the trilemma. In particular, three proposals are discussed: the institutional account suggested by Rawls, comprehensive education programs, and technological trellises. I will now examine these one by one.

5.1 Rawls & Just Institutions

5.1.1 Transformation by the Habituation of Fair Cooperation

In his political theory, Rawls acknowledges the enduring presence of unreasonable elements in liberal society and affirms the complexity of the issue of containment. In light of this, he writes: *"How far unreasonable doctrines may be active and are to be tolerated in a constitutional democratic regime does not present a new and different question, [but] are settled by the appropriate political principles of justice and the conduct those principles permit."*[203] Then, he provides two suggestions for how doctrines that are in conflict with the constitutional essentials of the liberal order and citizens that reject reasonableness as a virtue could be reformed.[204] These suggestions are not primarily framed in terms of persuasion or limiting rights – instead, they are focused on the transformative experience and educative effects of enduring and just institutions on all its participating citizens. Legitimacy is thus improved not by offering new justifications to the unreasonable, but by making them agree with the current, public reasons.

Rawls' first point is that: *"The liberties of the intolerant may persuade them to a belief in freedom. This persuasion works on the psychological principle that those whose liberties are protected by and who benefit from a just constitution will, other things equal, acquire an allegiance to it over a period of time."*[205] In other words, he suggests that when unreasonable citizens experience the basic rights and freedoms afforded by a just constitution, they are likely to be transformed over time and gradually come to appreciate the essential principles of the political conception of justice.[206]

---

[203] Rawls, *Political Liberalism*, p. 489.
[204] In *A Theory of Justice* §35, and in *Political Liberalism* V §6.2.
[205] Rawls, *A Theory of Justice*, p. 192.
[206] An important caveat follows on ibid., p. 193: *"Whether the liberty of the intolerant should be limited to preserve freedom under a just constitution depends on the circumstances. [...] The natural strength of free institutions must not be forgotten, nor should it be supposed that tendencies to depart from them go unchecked and always win out."*

Rawls' second point is that: *"There is no social world without loss: that is, no social world that does not exclude some ways of life that realize in special ways certain fundamental values. The nature of its culture and institutions proves too uncongenial. But these social necessities are not to be taken for arbitrary bias or injustice."*[207] Political liberalism, while neutral in its goals, refrains from favoring any specific comprehensive doctrine or aiding its adherents, but cannot and ought not guarantee neutrality in terms of its societal impact. According to Rawls, its basic structure has the unintentional but permissible side-effect of transforming unreasonable doctrines into more reasonable variants.

Elsewhere, Rawls is more explicit about this specific mechanism.[208] His notion is that a stable political conception of justice has the power to influence and shape comprehensive doctrines, gradually nudging and guiding them away from unreasonableness and towards reasonableness. When a political conception appears effective and cooperation proves fruitful and enduring institutions prove just, then citizens will naturally develop trust in political systems and each other via habituation, thereby fostering over time an internalized allegiance to the liberal order. This transformative process would take quite some time to unfold; yet, given the proper conditions and sufficient patience, Rawls thinks that *"simple pluralism moves toward reasonable pluralism."*[209]

Several scholars are in agreement with him, such as Fuat Gurzoslu, who writes: *"This larger political dynamic in political liberalism supports the account of containment as transformation. It shows that containment as transformation is not an exception, but part of a larger dynamic that lies at the heart of political liberalism."*[210] Similarly, Steven Macedo thinks that liberalism forms a regime that is and ought to be profoundly influential in the way that it shapes its citizens' lives, indirectly molding their commitments and behaviors.[211] Still, the question is not whether political liberalism has an impact on the political attitudes of its citizens, but rather: can it successfully and reliably transform unreasonable citizens so that the majority of reasonable citizens reaches a critical mass?

5.1.2 Criticism of Habituation

In response to Rawls, I argue that the habituation of fair cooperation among citizens and eventual internalization of a reasonable attitude – being made possible by just institutions over time – is by itself not enough to let political liberalism reach a critical mass of reasonable citizens. This argument starts from the observation that, although liberal democracy certainly provides greatly valuable and desirable outcomes to its population, these benefits are of a certain kind. Consider some of the most valued products of liberal democracy: peace, human rights, economic growth, rule of law, or public healthcare. These things are observable not as a distinct presence, but instead, they are noteworthy because they signify a prolonged absence of certain causes of suffering.[212] Most benefits of political liberalism come in the form of things that make

---

[207] Ibid., p. 197.
[208] Ibid., p. 163-164, 246.
[209] Ibid., p. 164.
[210] Gursozlu, "Political Liberalism and the Fate of Unreasonable People," p. 54.
[211] Macedo, *Liberal Virtue*, p. 59.
[212] For each of these examples, its opposite is far more demanding of human attention. Contrast, respectively, the harmony of peace versus the horrors of war; exercising freedom of speech versus imprisonment after protest; the autonomy afforded by luxury versus the suffering of poverty; occasionally appealing for judicial help versus frequently having to bribe officials; and receiving incidental vaccinations versus becoming ill on a regular basis.

daily life more effortless and more carefree; they are usually rarely remarked upon – until they disappear.[213] The psychological term for this cognitive process is *hedonic adaptation*: the human tendency to quickly return to a relatively stable level of happiness despite major positive or negative events or life changes.[214] This can be worsened by pervasive negativity bias: the tendency for negative events, emotions, or information to have a greater impact on our psychological state and behavior compared to positive ones.[215] Hedonic adaptation could render the unreasonable unaware of the benefits of liberal democracy and negativity bias could subsequently fuel a one-sided preoccupation with its flaws.

In short, the unreasonable are likely immune to liberal internalization through institutional habituation. This is because the benefits of an established liberal order are of a kind that has an inherently limited capacity to create a narrative that transforms the unreasonable, as these benefits are mainly characterized by what they are not. Due to features inherent to human cognition, the same mechanism that allows citizens to internalize liberal values could also cause them to take the fruits of fair cooperation for granted. Instead of being persuaded, unreasonable citizens are equally likely to become desensitized to the precious achievements that liberal democracy has brought about. Thus, neither mechanism pointed out by Rawls can be counted upon to adequately transform the unreasonable.

5.2 Political Liberal Education

An alternative suggestion to transform the unreasonable citizens into being reasonable – and thereby safeguard that political liberalism can prosper over time – is imposing a mass education initiative that instills in each new generation a willing support of the liberal order. Shawn Rosenberg, a psychologist who thinks that the current trajectory of liberal democracy is one of decline and headed toward collapse, writes:

> *"The alternative is to create the citizenry that has the cognitive and emotional capacities democracy requires. This would entail a massive educational initiative, one that would have to be premised on recognizing the dramatic failure of prior efforts along these lines. Perhaps in this way, democratic forms of governance may yet prevail."*[216]

Academic analysis of the issue of education in political liberalism has produced a vast literature.[217] In the context of this thesis project, I will focus specifically on the issue of whether education has the potential to transform the basic fact of simple pluralism into the fact of reasonable pluralism: to what extent can education initiatives ensure that a compelling majority within each new generation is reasonable? The following discussion examines three approaches of liberal education: as focused on values and attitude, on skills and knowledge, or on exposure and social familiarization.

---

[213] In general, this is somewhat inevitable due to the nature of negative liberty as freedom from external restraint and interference – i.e. freedom of *not* being intruded upon.
[214] Brickman, "Hedonic Relativism and Planning the Good Society", p. 287.
Also, consider the phenomenon of 'abundance denial' – see: Easterbrook, *The Progress Paradox: How Life Gets Better While People Feel Worse*.
[215] Rozin and Royzman, "Negativity Bias, Negativity Dominance, and Contagion", p. 296.
[216] Rosenberg, "Democracy Devouring Itself", p. 31.
[217] See: Fernández & Sundström, "Citizenship Education and Liberalism: A State of the Debate", p. 363-384; Sant, "Democratic Education: A Theoretical Review", p. 655-696; Surridge, "Education and Political Liberalism: Pursuing the Link", p. 146-164; Podschwadek, *Educating the Reasonable: Political Liberalism & Public Education*, p. 1-20; Bader, *Secularism or Democracy?*, p. 263-288.

5.2.1 Education as Values & Attitude

Since the goal of transformational liberal education is to spread reasonableness, and since being reasonable is an attitude contingent upon certain values, the most straightforward approach is to inculcate all children with liberal values and to cultivate a reasonable attitude directly. For instance, Rawls writes: *"Their education should prepare them to be fully cooperating members of society and enable them to be self-supporting; it should also encourage the political virtues so that they want to honor the fair terms of social cooperation in their relations with the rest of society."*[218] Inspired by political liberalism, Stephen Macedo has developed Rawls' notions into a model of citizenship education called 'civic liberalism' which emphasizes a transformative agenda.[219] According to Macedo, liberal institutions have a responsibility to mold young citizens in such a way that they come to endorse liberal principles, thereby ensuring that they willingly support and contribute to the liberal order. By prioritizing the promotion of liberal values –fairness, freedom, equality, and, to at least some extent, autonomy – citizenship education fosters a critical yet reasonable attitude that consequently will extend beyond the educational sphere and spread throughout society.

Concerning what kind of citizenship education is required by liberalism, Rawls claims that political liberalism demands *"far less"* than comprehensive liberalism – suggesting rather modest requirements for citizenship education.[220] However, this claim has been heavily disputed by other liberal political theorists: with regard to education, political liberalism becomes nearly indistinguishable from comprehensive liberalism. For instance, Amy Gutmann points out that teaching students to achieve political autonomy is in practice essentially the same as fostering ethical autonomy.[221] Similarly, Eamonn Callan contends that political liberalism cannot successfully distinguish itself from comprehensive liberalism if it requires teaching autonomy – political and ethical autonomy are not sufficiently distinct.[222] Likewise, Richard Dagger asserts that political liberal education favors a particular conception of the good life, and therefore violates its own norm of being freestanding from any particular comprehensive doctrine.[223]

Indeed, Rawlsian educational theorists who attempt to clarify the minimum of education required by political liberalism – not only Macedo but also scholars like Victoria Costa[224] – advocate an interpretation that still appears quite robust; modest variants are unlikely to effectively instill the necessary civic virtues and capacities in its new generations. The key issue seems whether political liberalism ought to propose citizenship education that develops students' autonomy – the ability to critically scrutinize and potentially revise their ethical beliefs and values.[225] Including autonomy would create a welcome exit option for citizens raised in an unreasonable private environment. Yet, it could also form an imposition that itself threatens political legitimacy and alienates citizens from otherwise reasonable comprehensive doctrines. Thus, political liberalism contains a major internal tension between demands for a robust citizenship education that develops key civic capacities on the one hand, and on

[218] Rawls, *Political Liberalism*, p. 199-200.
[219] Macedo, *Civic Education in a Multicultural Democracy*, p. 15.
[220] Rawls, *Political Liberalism*, p. 199.
[221] Gutmann, "Civic Education and Social Diversity", p. 557.
[222] Callan, "Political Liberalism and Political Education,", p. 40.
[223] Dagger, *Civic Virtue: Rights, Citizenship, and Republican Liberalism*, p. 190.
[224] Costa, *Rawls, Citizenship, and Education*, p. 64-65.
[225] Neufeld, "Political Liberalism, Autonomy, and Education", p. 35.

the other hand constraints around state neutrality concerning conceptions of the good life which limits how far liberal education can go in promoting student autonomy to question ethical frameworks. This internal tension prevents the liberal order from relying solely on value education to transform the unreasonable.

5.2.2 Education as Skills & Knowledge

Alternatively, transformational education can be conceptualized not as focusing on values but emphasizing primarily the acquisition of skills and knowledge, aiming for cognitive liberation of the unreasonable. In this view, the attitude of reasonableness emerges naturally once citizens are imbued with sufficient critical thinking skills and social-political knowledge. Such an approach would depart from Rawls, who writes: *"The reasonable cannot be derived from the rational. [...] The serious attempts to [do so] do not succeed."*[226] After all, being reasonable is independent of having true assumptions or correct reasoning.[227] Still, this approach to education is worth exploring since its implementation can likely count upon support from the unreasonable.

In the academic literature on liberal education, multiple voices have advocated a reorientation away from values and towards knowledge and its prerequisite skills. For instance, Robert Scott laments what he perceives to be a lack of knowledge among influential officials and states that *"citizens cannot be fully responsible unless they are more knowledgeable about and sensitive to the differences in culture."*[228] Likewise, Zongyi Deng contends that the primary role of schools, a matter of social justice, ought to be that all students have access to disciplinary knowledge not available to them at home.[229] Another example is Piet van der Ploeg, who proposes that civic education eschews any values and restricts itself to merely providing insights about different models of citizenship.[230] Of course, this perspective has a kernel of truth: Rawls himself stipulates that effective deliberation is possible without the citizenry having sufficient knowledge of the matter at hand,[231] and knowledge about one's constitutional and civic rights is crucial to education for political liberalism.[232]

The implications of this principle – that adequate knowledge is indispensable to good citizenship – have been taken to the extreme by other political theorists, such as David Estlund and Jason Brennan, who assert that uninformed voters cause the main pathologies of liberal democracy.[233] They argue in favor of *epistocracy*, proposing a political system where one's right to vote is dependent upon that citizen having demonstrated a sufficient level of knowledge about political issues to participate in decision-making. In contrast, the aforementioned educational theorists do not equate being uninformed with being unreasonable. Rather, they argue that cultivating knowledge and skills ought to take primacy over instilling values in education.

However, the available empirical evidence seems to indicate that education focused on skills and knowledge does not lead to increased reasonableness, but actually makes

---

[226] Rawls, *Political Liberalism*, p. 53.
[227] Respectively: ibid., p. 58; ibid., p. liv, fn 31.
[228] Scott, "The Meaning of Liberal Education", p. 25.
[229] Deng, "Bringing Knowledge Back In: Perspectives From Liberal Education", p. 335.
[230] Van der Ploeg, "Different Citizenship Education", p. 296.
[231] Rawls, *Political Liberalism*, p. 449.
[232] Rawls, *Political Liberalism*, p. 199.
[233] Estlund, "Why Not Epistocracy?", p. 52-69; Brennan, "Giving Epistocracy A Fair Hearing", p. 35-49. For an opposing view, see Gunn, "Against Epistocracy", p. 26-82.

citizens more prone to partisanship and an unreasonable attitude. There appears to be a positive correlation between the competencies acquired by schooling, and partisan reasoning and distorted judgment.[234] This correlation applies both to skills (e.g. critical reflection,[235] logical reasoning,[236] and mathematical ability[237]) and to knowledge (e.g. concerning science,[238] economics,[239] and politics[240]). The underlying issue is that generally speaking, increased knowledge and reasoning skills often serve to bolster citizens in defensively rationalizing their political claims. Alan Abramowitz even goes so far as to predict that when a population becomes more educated, partisan polarization and unreasonableness will also escalate.[241] Regardless, an increase in skill or knowledge does not help the unreasonable to fulfill the duty of prudence – on the contrary, it increases their opportunity for motivated reasoning. Hence, education as focused on skill and knowledge can likely count on unreasonable support but can probably not achieve their transformation.

5.2.3 Education as Exposure & Familiarization

There is an additional, third approach to education that warrants a brief discussion in the context of transforming the unreasonable. According to this approach, as argued by Rob Reich, liberal citizenship education can be achieved by giving priority to teaching skills and knowledge or instilling values and attitudes, but by letting students experience the practice of the *common school*.[242] The common school – opposed to a common curriculum or common doctrine – prioritizes the notion that children are exposed on a daily basis to peers from different backgrounds and comprehensive doctrines, who will thereby indirectly develop a familiarity with and toleration of other ways of life. The underlying idea is that it is much harder to otherize divergent doctrines when you have grown up alongside the children of its adherents. Thus, the common school aims to cultivate in its students greater reasonableness not just by fostering mutual respect among students, but also by having them immediately practice that respect in an educational setting where students interact with peers from diverse backgrounds, in a student body that is representative of the diversity of society and encompasses several communities with different comprehensive doctrines.

However, I argue that this approach runs into several problems of its own. Firstly, in order to function properly it needs to ensure that each student body is actually diverse. In practice, this diversity would need to be imposed by the state. Despite its difference from autonomy, the value of diversity is still unlikely to count on parental consent by

---

[234] Ezra Klein, "How Politics Is Making Us Stupid": *"Research shows that the more information partisans get, the deeper their disagreements become. [...] Being better at math made partisans less likely to solve the problem correctly when solving the problem correctly meant betraying their political instincts. [...] Among people who were already skeptical of climate change, scientific literacy made them more skeptical of climate change."*
[235] Kahan, "Ideology, Motivated Reasoning, and Cognitive Reflection", p. 407.
[236] Cavojová et al., "Myside Bias in Reasoning about Abortion", p. 656.
[237] Kahan et al., "Motivated Numeracy and Enlightened Self-Government", p. 54.
[238] Drummond & Fischhoff, "Individuals with greater science literacy and education have more polarized beliefs on controversial science topics," p. 9587.
[239] Prior, Sood, & Kanna, "The impact of accuracy incentives on partisan bias in reports of economic perceptions", p. 489-518.
[240] Baekgaard, "The Role of Evidence in Politics", p. 1117.
[241] Abramowitz, *The Disappearing Center*, p. 127.
[242] Reich, "Common Schooling and Educational Choice as a Response to Pluralism", p. 21-40. See also: Callan, *Creating Citizens: Political Education and Liberal Democracy*, p. 162-195. For an opposing view, see: Neufeld, "Political Liberalism and Citizenship Education", p. 781-797.

the unreasonable, whose opposition stems from their refusal to regard all citizens as free and equal who cooperate under fair terms. Furthermore, the effect of exposure on unreasonableness could be less than anticipated. For instance, some unreasonable students may compartmentalize the mental effect, and keep their positive judgment of a friend from a different background at their school separate from their personal antipathy against the associated doctrine. Besides, exposure is in the long run not a full substitute for direct instruction in liberal values. Consider that students can only be exposed to social groups that existed at the time of their education and that their reasonableness arises out of local habituation – not out of reflection and induction. Thus, if a new social group and doctrine emerges later in their lives, it remains an open question whether the now former students will extend to them the reasonable attitude practiced in their youth, or show unreasonableness, perhaps even explicitly on the grounds that they lack personal familiarity with them. All in all, transforming the unreasonable by education seems to require more than mere exposure to divergent doctrines in society.

5.2.4 Education and General Limits

Apart from these three discussions on various views on education in political liberalism, I argue in favor of four additional caveats on the transformational efficacy of education in general. Firstly, unreasonable parents might thwart the educative process at home. After all, parents often have emotional primacy to their children compared to society as a whole, and becoming reasonable is partly an affective process. Secondly, even if education has the desired effect of making the unreasonable more reasonable, this effect may naturally fade with time. Also, it could even be overridden once students move on to other social contexts that encourage attitudes and instill values that are not naturally compatible with liberal democracy. After all, becoming unreasonable can and often does occur later in life: the effect of education should not be considered permanent. Thirdly, education can only gradually change the political attitudes of a population. Thus, if the unreasonable pose an acute threat to liberal institutions, new educational initiatives may take effect too late to make a difference.

Finally, I argue that any demanding educational initiative enacted by the liberal order might be more polarizing among the *reasonable* than expected. George Lakoff, using cognitive linguistics to analyze conceptual metaphors in politics, has pointed out that progressives and conservatives are split among two models of child development that mirror their split in politics.[243] Progressives abide by the *nurturant parent model*, which emphasizes communication, mutual respect, reasoning, and self-fulfillment in parenting. In contrast, conservatives adhere to the *strict parent model,* which emphasizes discipline, adherence to authority, obedience, and self-reliance during upbringing. This is relevant to political liberalism since it regards education as fostering civic engagement and fair cooperation, not obedience and self-reliance. Hence, it is closer to the nurturant parent model, which could lead to significant opposition among conservative-leaning yet reasonable citizens.

---

[243] Lakoff, *Moral Politics*, p. 65-142. This characterization occurs on group-level – there will of course be some individual families that do not conform to these descriptions.
For empirical evidence, see Ohl et al. "Lakoff's Theory of Moral Reasoning in Presidential campaign advertisements, 1952–2012", p. 488-507.

All things considered, while civic education is a key issue in any liberal democratic theory and likewise in political liberalism, there remain normative tensions and practical challenges that cast doubt on how fully fundamentally unreasonable citizens can be transformed through educational means alone.

5.3 Technological Trellises

5.3.1 Transformation by Technology

Instead of habituation or education, perhaps technology can contribute to transforming the unreasonable elements in liberal society into willing members of a compelling and enduring majority of reasonable citizens. For instance, Yuval Noah Harari and Felipe Arocena are proponents of such a view.[244] I will start this discussion on transformational technology by mentioning recent philosophical research on autonomy by Joel Anderson, which reveals some relevant similarities with how this thesis project understands reasonableness.

Like autonomy, reasonableness is understood as a liberal value and virtue that is both action-guiding and dependent on social recognition.[245] Anderson proposes that the capacity of autonomy can be supported by forms of scaffolding – e.g. social technologies. Inspired by that idea, I examine in this section whether some technologies can function as *trellises* – permanently established scaffolds of reasonableness – that transform the unreasonable and support their capacity to become reasonable. Political liberalism cannot avoid engaging with matters of technology since certain technological artifacts and systems can embody specific forms of power, and thus potentially enhance or undermine the values essential to reasonableness.[246] For our discussion, examples of technology become relevant specifically when they seem inherently political: they are strongly compatible with specific political attitudes and they make it difficult to achieve alternative arrangements.

The following exploration – necessarily brief – does not advocate technological determinism and should not be interpreted as such. Neither does it claim that some specific technology is irredeemably unreasonable. Rather, it serves to highlight the dynamic interconnection between some technologies that are foundational to modern life and the social-political system in which they are embedded. With that in mind, let us examine three technologies in particular that seem inherently political and influence the conditions that make widespread reasonableness possible. These examples concern architecture, transportation, and information.

I will start with architecture, since, in the words of Winston Churchill: *"First, we shape our buildings; thereafter they shape us."*[247] Certain characteristics of urban planning and building design can unintentionally increase social isolation and prevent a vibrant community life. For instance, the absence of parks, plazas, community centers, and other public gathering spaces limits opportunities for spontaneous social interaction and community building. When public spaces discourage people from lingering and

---

[244] Harari, "Why Technology Favours Tyranny".
Arocena et al., "Technological disruption and democracy in the twenty-first century", p. 3.
[245] Anderson, "Scaffolding and Autonomy", p. 158-166.
[246] Winner, "Do Artifacts Have Politics?", p. 121-136.
[247] Love, "A Sense of Architecture in the Past", p. 213.

interacting on purpose, this can even be called 'hostile architecture'.[248] Urbanist literature points out another example, single-use zoning, which diminishes a vibrant sense of community and perpetuates socioeconomic segregation.[249] This is relevant to our discussion of the unreasonable since available empirical research on political radicalization suggests that social isolation and estrangement from one's surrounding community constitute an important causal factor of extremism.[250] Hence, I assert that if architecture indeed influences social cohesion, cooperative attitudes, and political preferences, then political liberalism ought to take an interest in designing not just its constitution but also its residential environment upon the liberal principles of equality, freedom, and fairness.

A second relevant area for transformational technology is transportation. It is contingent upon infrastructure that requires massive public investment and funding decisions, and its forms of availability can reinforce or challenge existing power dynamics between groups.[251] For instance, car-centric transportation policy designs cities around the private automobile at the expense of other mobility options, which often significantly undermines social cohesion in multiple ways.[252] First, by imposing mobility challenges for citizens who are unable to drive – such as the elderly, youth, and lower-income groups. This unfairly restricts their access to important services, facilities, and activities. Second, it leads to urban sprawl, which fragments residential enclaves and leaves them disconnected from other parts of the community. public transport. Thirdly, cars provide an insulated space that becomes regarded as a private realm where only one's close social circle is allowed to enter. Thus, the car facilitates a sense of individualism – even if its use is wholly contingent upon public systems. As Loren Lomasky writes: *The automobile is for twentieth-century American society the quintessential bastion of privacy.*[253] Also, it should be mentioned that car travel is the most dangerous mode of transportation.[254] In contrast, public transportation options such as the train or metro are open to all citizens, encourage living density instead of urban sprawl, provide opportunities to encounter and engage with other citizens during travel, and guarantee better safety. Thus, I contend that, given that transportation facilitates civic cooperation, political liberalism cannot profess neutrality on issues of transportation since overreliance on car-centric policies is incompatible with the liberal ideal of all citizens as free and equal.

Third, a discussion concerning the intersection of politics and technology must include the issue of digital social media platforms. Since its rise to ubiquity, there have been growing concerns regarding its compatibility with liberal democracy. Here, I focus on three in particular. First, the ease with which misinformation can spread virally on social media distorts people's understanding of facts and reality. This impairs reasoned deliberation based on accurate information, which is crucial for Rawlsian principles of justice. Second, the incentives and dynamics of social media tend to amplify

---

[248] New York Times, "Hostile Architecture: How Public Spaces Keep The Public Out".
[249] Resseger, "The Impact of Land Use Regulation on Racial Segregation".
[250] Pfundmair, "How Social Exclusion Makes Radicalism Flourish: A Review", p. 341-359. Stolberg, "Extended Loneliness Can Make You More Vulnerable to Extremist Views."
[251] Klein et al., "Political Partisanship and Transportation Reform", p. 163.
[252] Keith & Del Rio: "New Urbanism, Automobile Dependency and Sense of Community", p. 59. See also: Marohn, *Strong Towns*.
[253] Lomasky, "Freedom and the Car", p. 21.
[254] National Safety Council, "Deaths by Transportation Mode".

outrageous, controversial content that evokes strong emotions. In practice, this often pushes people towards partisan rather than reasonable attitudes. Third, social media lacks sufficient regulation norms that can effectively constrain hate speech, which violates the norm of mutual respect that underlies Rawls' theory. These points stand in stark contrast with the institution and associated technology of the editorialized newspaper, which now faces considerable decline. Therefore, I posit that political liberalism is unlikely to succeed in containing unreasonable doctrines unless it addresses the challenges posed by informational technologies that deeply inform political attitudes but have a contentious relationship with core liberal values.

Together, these three examples illustrate the point that certain policies on technology are friendly to the proliferation of reasonableness, while others seemingly undermine its preconditions. In light of this, a technological approach to transformation insists that unreasonable citizens can and ought to be nudged into being reasonable by having the liberal order favor some policies and forms of technology over others. Ideally, the unreasonable would continually engage with these nudging technologies and over time internalize the liberal values that they contain. In this way, the unreasonable, provided with trellises that permanently scaffold their ongoing process of becoming more reasonable, could maybe come to endorse the liberal constitutional essentials.

5.3.2 Limitations of Nudging Trellises

In general, I largely agree with the analysis of the technological argument of transformation – in my view, this approach could make a significant and necessary contribution to containing unreasonableness in a liberal society. Nevertheless, there are several reasons why I think that it will not suffice as the sole or even the main strategy of containment. To start, the transformative power of technology in shaping reasonableness can easily be overestimated. Although technology undoubtedly influences social dynamics, its impact on a deeply ingrained attitude like unreasonableness is likely modest – especially compared to other factors such as education, cultural norms, and economic conditions. In addition, the suggested approach is vulnerable to neglecting the agency of individuals in determining their attitudes and behaviors. While technology can contribute to a conducive environment for fostering reasonableness, ultimately, individuals retain the autonomy to embrace or reject reasonableness regardless of trellises.

Besides, its implementation as policy is likely unfeasible. The case for trellises is complex, and since the relevant ubiquitous technologies are usually regarded as politically neutral, the burdens of judgment come into play. Thus, the liberal order faces a dilemma. If it does not politicize these technologies, it cannot convince and mobilize a reasonable majority coalition. But if it does politicize these technologies, it alerts the unreasonable to their purpose as nudging trellises. This probably defeats the purpose, for the unreasonable can resist adopting these technologies in their daily life. All in all, the technological approach to transformation is limited due to the many challenges involved in attempting to engineer greater societal reasonableness.

5.4 Conclusion

In this fifth chapter, I have delved into various strategies aimed at transforming unreasonable citizens into more reasonable ones. I have examined the case by Rawls that just institutions can transform the unreasonable by habituation and argued that this effect is undermined by the specific kind of benefits that liberal democracy produces which causes them to be easily taken for granted. Additionally, I have discussed three different approaches to transformational education. When focused on values, it leads to political liberalism becoming comprehensive liberalism; when focused on skills and knowledge, it is ineffective; and when focused on exposure, it faces considerable normative and practical problems. Furthermore, I have briefly explored the transformative potential of several technologies, which seems promising, though its effect is likely modest. Ultimately, while each strategy of transformation could make some significant contribution to a successful containment of the unreasonable (albeit to varying degrees) – none is sufficient by itself: the inherent complexity of the task of transformation requires a multifaceted approach.

However, even an attempt at transformation that combines institutional habituation, comprehensive education, and technological trellises would still have some notable flaws. First, the successful implementation of just institutions, thorough education, and supporting technologies is likely contingent upon the condition it aims to bring about, i.e. a compelling majority of reasonable citizens. Second, the transformational approach takes effect gradually and is therefore not appropriate when the liberal order faces an acute threat by the unreasonable. Third, not all unreasonable citizens may be equally susceptible to transformation. Even if some of the unreasonable end up transformed, it is an open question whether this allows the critical mass of a compelling majority of reasonable support to be reached. Possibly, beyond the reach of transformation, there remains a significant faction of unreasonable citizens with a deep-rooted antipathy against liberal democracy.

Yet, even if unreasonableness cannot be sufficiently contained by conciliation or transformation, then perhaps its spread can be halted by marginalization: the restriction of some of the rights of the unreasonable when necessary to preserve the continuation of the liberal order.

64

Chapter 6: Marginalizing the Unreasonable

Thesis VI: That a marginalizing approach, combining reactive and preventive intervention that restricts the rights of the unreasonable in exceptional cases, is necessary for successful containment; although, due to normative challenges, it should not function as the primary strategy.

How can a liberal order, under conditions of simple pluralism, succeed in containing unreasonable movements? If conjecture or transformation proves insufficient in the face of an acute unreasonable threat, maybe stronger and more decisive measures should be considered. In this sixth chapter, I examine two suggestions from political philosophy that defend the view that in exceptional circumstances, the liberal order sometimes has no choice but to restrict certain rights of unreasonable citizens – like the freedoms of expression, education, or association –before liberal institutions are overwhelmed. An appropriate state intervention avoids two extreme scenarios: neither being so indecisive that the threat of the unreasonable is not actually addressed nor being so invasive that it violates the liberal principles that it is supposed to protect.[255] Since this marginalizing approach would prioritize legitimacy and stability over inclusivity, it seems to prefer the third option provided by the trilemma. In particular, I will discuss two ways in which a liberal order can contain unreasonableness. This twofold distinction arises from Rawls' brief comment on containment:

> *"That there are doctrines that reject one or more democratic freedoms is itself a permanent fact of life, or seems so. This gives us the practical task of containing them—like war and disease—so that they do not overturn political justice."*[256]

Although Rawls, after this introduction, does not develop containment any further, the analogy with war and disease still provides diverging ways to interpret this passage: containment could either react to an imminent escalation of challenges by the unreasonable or, additionally, prevent the emergence of such issues. Nussbaum adheres to the first interpretation, and takes Rawls to mean that restricting the rights of unreasonable citizens may occur *"only in the sort of emergency that amounts to a constitutional crisis."*[257] In contrast, Quong thinks that there are at least two *"additional situations that liberal democracies are likely to face where the behavior of unreasonable citizens might be justifiably restricted by appealing to the argument for containment."*[258] I will now examine these two accounts,[259] starting with containment as de-escalation in exceptional circumstances.

---

[255] I call these approaches 'marginalizing' to emphasize that very tension: while unreasonable doctrines ought to be at the margins of liberal society, nevertheless, unreasonable citizens ought not to be treated in an unjust way that violates their innate moral worth as human beings.

[256] Rawls, *Political Liberalism*, p. 64 fn.19. Rawls only mentions 'contain' with this meaning in one other place. Ibid., p. xvi-xvii: *"Of course, a society may also contain unreasonable and irrational, and even mad, comprehensive doctrines. In their case the problem is to contain them so that they do not undermine the unity and justice of society."*

[257] Nussbaum, *Perfectionist Liberalism and Political Liberalism*, p. 23.

[258] Quong, *Liberalism Without Perfection*, p. 300-301.

[259] For alternative reading of Rawls, see: Fumagalli, "Can Rawlsian Containment of Hateful Viewpoints Be Effective?". Fumagalli insists on another two-fold distinction: 'harsh' or 'soft' containment. This is based upon who does the containing of the unreasonable: the liberal order or its reasonable citizens. However, I think this interpretation is improbable in light of Rawls'

6.1 Reactive Containment

According to the first approach of containment as marginalization, the liberal order is justified in restricting certain rights of unreasonable citizens only in exceptional cases when its institutions are in clear and present danger, facing an imminent threat. In other words, a just government ought to postpone its intervention until the unreasonable movement is on the verge of overpowering liberal democratic institutions. This is likely the position of Rawls, judged by his remarkably protective position on free speech.[260] This position is also endorsed by Cass Sunstein, [261] who emphasizes that practically all preventive intervention is ruled out by the principle of clear and present danger – a very high bar to clear.[262] For instance, precautionary measures that block the dissemination of unreasonable doctrines by deplatforming, even when such policies have been found to be effective, are rejected under this principle.[263] Instead, as Nussbaum points out, a liberal order protects basic civic rights by enshrining them and its associated major liberties the constitution. [264] In this way, aggressive unreasonable doctrines cannot easily abolish liberal rights by the regular legislative process.

An important argument in favor of this approach, besides appeals to tolerance and neutrality of the liberal state, what Kristian Skagen Ekeli calls 'the argument of distrust.' [265] In short, it expresses the concern that those empowered to suppress unreasonable doctrines may misuse that power, and unjustifiably restrict whatever dissenting views may challenge their authority. Governments often have strong incentives to silence critics, making it essential to provide rights like freedom of expression with extra protection. Abiding by the principle of clear and present danger would ensure a heavy burden of proof on whomever advocates for containment as suppression. Relaxing these requirements could increase abuses of power – after all, it is quite complex to determine whether precautionary measures are warranted. So intervention carries the risk of being motivated by political concerns rather than genuine threats to stability. To paraphrase Rawls: political speech may by its nature often appear to be dangerous.[266]

6.2 Objections to Reactive Intervention

Despite these benefits of the reactive approach to containment, it has also some considerable downsides. I argue in favor of four objections in particular.

---

analogy with war and disease: concerning both armed conflict and pandemics, it is the state who has the primary responsibility and ability of containment, not its citizens.

[260] Rawls, *Political Liberalism*, p. 355. *"For free political speech to be restricted, a constitutional crisis must exist requiring the more or less temporary suspension of democratic political institutions, solely for the sake of preserving these institutions and other basic liberties. Such a crisis did not exist in 1862—64 [during the American Civil War]."*

[261] Sunstein, *Laws of Fear*, p. 219–220.

[262] Rawls and Sunstein use the phrase 'clear and present danger' referring to a judicial decision by the U.S. Supreme Court in the case *Brandenburg v. Ohio*.

[263] Thomas & Wahedi, "Disrupting hate: The effect of deplatforming hate organizations." Jhaver et al. "Evaluating the effectiveness of deplatforming", p. 1.

[264] Nussbaum, "Perfectionist Liberalism and Political Liberalism", p.24.

[265] Ekeli, "Liberalism and Permissible Suppression of Illiberal Ideas", p. 176-177.

[266] Rawls, *Political Liberalism*, p. 354.

6.2.1 Timing of Intervention

My first objection is that the liberal order cannot afford to react so late to threats by unreasonable doctrines. Although there are normative reasons to delay interventions, this extra time usually comes at a serious cost: liberal institutions are generally stable but, being gradually eroded by the unreasonable, can only absorb so much. Thus, there is a trade-off between time and social harm, and eventually, additional time will yield only marginal benefits at increasing institutional costs. Since it is very difficult to determine with certainty when exactly an unreasonable doctrine becomes a clear and present danger, and since reactive containment encourages an attitude of reluctance, state intervention may likely come too late. In that worst-case scenario, the institutional damage has been fatal. Therefore, given that the cost of failure is incredibly high, a liberal order can hardly be blamed if it intervenes earlier as a precaution.[267] Since the ultimate goal is to preserve liberal democratic institutions, the trade-off may swing in favor of intervention earlier than Rawls and Sunstein anticipated.

Moreover, delayed intervention may be counterproductive. Once an increasingly powerful unreasonable movement qualifies as a clear and present danger, it will probably be already so widespread and influential that protecting the liberal order becomes nigh impossible. At that point, state intervention would require such extreme measures in order to have a plausible chance to prevent the unreasonable from overwhelming the liberal state, that these measures themselves could very well be the decisive blow that irreparably harms the liberal democratic institutions. For instance, Ekeli endorses this view.[268] Conversely, an intervention that faces a less powerful unreasonable faction would require less extreme measures to achieve containment.

6.2.2 Unity of Society

As a second objection, I argue that reactive containment would insufficiently preserve the unity of society. In Rawls' other comment on containment, he writes:

> "Of course, a society may also contain unreasonable and irrational, and even mad, comprehensive doctrines. In their case the problem is to contain them so that they do not undermine the unity and justice of society."[269]

In other words, containment is not just about protecting the stability of liberal institutions, but it should also preserve social harmony. Rawls says: "Social unity is based on a consensus on the political conception; and stability is possible when the doctrines making up the consensus are affirmed by society's politically active citizens".[270] Since social unity depends on the overlapping consensus, that consensus is disrupted when a significant group of unreasonable citizens rejects it, which already warrants containment even if the liberal order remains respected as a modus vivendi. In other words, the principle of clear and present danger interprets stability merely descriptively, but Rawls is committed to the normative view that "as always, stability

---

[267] Or, as Cato the Younger commented while the Roman Senate debated their response to discovering an imminent coup d'etat: "Other crimes you may punish after they have been committed; but as to this, unless you prevent its commission, you will, when it has once taken effect, in vain appeal to justice. When the republic is taken, no power is left to the vanquished." Sallust, Conspiracy of Catiline.

[268] Ekeli, "Liberalism and Permissible Suppression of Illiberal Ideas", p. 178.

[269] Rawls, Political Liberalism, p. xvi-xvii.

[270] Ibid., p. 134.

means stability for the right reasons."[271] When an unreasonable movement becomes powerful enough to threaten the liberal order, that clear potential danger already has reduced the civil peace into a mere civil truce, and thereby caused sufficient harm to warrant state intervention. Reactive containment is insufficient since unreasonable groups can already be quite effective at undermining the stability of institutions long before they attempt to deliver the final blow to the liberal order.

Besides, measures of containment by the liberal order depend for their efficacy largely on the unity of society. If the population has low trust in its government, then it is less likely to abide by its policies. Containment is contingent upon the very things it wants to protect. Hence, I think that it is peculiar that under the principle of clear and present danger, intervention is only allowed once the unity and justice of society have already been seriously eroded by unreasonable doctrines and citizens. Rather, once a politically unreasonable movement has gained sufficient strength to overwhelm the institutions of the liberal order, this fact itself indicates already that the unreasonable have 'broken through' and thus that containment has failed.

6.2.3 Institutional Incentivization

My third objection is that abiding by the principle of clear and present danger provides perverse incentives.[272] Stability for the right reasons is in part contingent upon widespread mutual respect among citizens and reciprocal toleration. These qualities in turn rely upon social trust between all different reasonable citizens, doctrines, and groups. Mutual connections evolve into civic bonds of trust not only because of good intentions but also because of institutional guardrails that regulate and discipline those who violate fair terms of cooperation. Citizens trust, but the institutions verify. Hence, if the liberal order keeps delaying its intervention against rising unreasonableness, this emboldens the unreasonable and demoralizes reasonable citizens who still believe in the ideal of society as a system of fair cooperation.

The health of liberal democracy depends not just on its legal framework, but also on its citizens abiding by complementary and uncodified norms of behavior. For instance, democratic elections rely upon the norm that citizens will accept legitimate electoral outcomes even if their preferred candidates lose. While laws establish formal rules, liberal democracy ultimately relies on citizens internalizing norms of reasonableness to make those laws work as intended. If institutions discipline unreasonable citizens only in exceptional cases, this unintentionally incentivizes undesirable behavior. First, the effect of deterrence is reduced, which otherwise would inhibit unreasonable behavior. Second, reasonable citizens face greater difficulty in maintaining their habits of trust towards one another. Third, reasonable citizens may feel let down by the liberal order due to its inaction, further decreasing crucial institutional trust. Thus, abiding by the principle of clear and present danger provides insufficient institutional safeguards against unreasonableness, especially if the unreasonable have already grown powerful.

6.2.4 Judicial Interpretation

The final objection concerns specifically Nussbaum's point about containment by enshrining basic rights in the constitution. Besides the definitional issue that

---

[271] Ibid., p. xl.
[272] Ekeli also seems to imply this. See Ekeli, "Liberalism and Permissible Suppression", p. 179.

containment understood in this way is not clearly distinct from regular matters of justice, I argue that this approach faces an important judicial challenge.

First of all, containment by the judiciary has not always been effective in the last few decades. When one examines the cases of the past decades where liberal democracies devolved into autocracy, a common pattern emerges.[273] Usually, the critical step past the point of no return is when the authoritarian faction brings the constitutional court into its sphere of influence – this transforms the rule of law into the rule by law.[274] Modern authoritarian regimes use the judicial framework during their rise to power to erode and eventually dismantle liberal rights from within. In light of this, we can understand why unreasonable movements see control over the courts as essential.[275] First, to dismantle institutional constraints: controlling the courts removes the check by an independent judiciary and allows unreasonable movements to centralize power. Second, to sideline political dissent, for the unreasonable want to use the legal system to harass opposition figures under the guise of law. Third, to provide a veneer of legitimacy, as a pliant judiciary lends an air of legality to the illiberal actions of the unreasonable. In summary, the capture of the judiciary is in practice an effective and pivotal strategy for unreasonable movements to dismantle the liberal rights that are enshrined in the constitution.

Moreover, this strategy is feasible in the first place since judges have substantial discretionary space to interpret the meaning, scope, and application of rights, even if those rights are enshrined in the constitution and regarded as fundamental rights. This is because of several reasons. Firstly, constitutional language is often broad and open-ended, which allows for readings of the law that diverge from a reasonable account. Secondly, judges have flexibility in their methodological approaches, some of which are more compatible with unreasonable doctrines than others. Thirdly, constitutional rights often come into conflict, requiring a balancing of competing interests, and it is up to judges to determine which rights take precedence and balance them in a given case. Finally, ambiguities and gaps in the formulation of constitutional rights mean that interpretation is necessary to clarify the scope of the right in question. Hence, enshrining rights in a constitution does not effectively contain unreasonableness, since those rights are dependent upon judicial interpretation, and a judiciary influenced by unreasonable movements can subvert these rights under the guise of legal interpretation and application.

For these four reasons, I think that the reactive account of intervention as containment falls short. Although it successfully avoids marginalizing citizens, it does not succeed at pushing unreasonable political movements to the margins of society.

6.3 Preventive Containment

An alternative approach of containing the unreasonable by marginalization goes beyond the reactive account by stating that containment involves not just de-escalating current threats, but also preventing potential threats from actualizing. In this view, the liberal order ought to take measures to restrict the spread of unreasonable ideas and

---

[273] For instance, Venezuela under Hugo Chavez, Nicaragua under Daniel Ortega, Hungary under Viktor Orban, Poland under Andrejz Duda, Turkey under Recep Erdogan, and India under Narendra Modi.
[274] Ginsburg & Moustafa, "Rule By Law: The Politics of Courts in Authoritarian Regimes", p. 6.
[275] Wang, "Authoritarian Judiciary: How the Party-State Limits the Rule of Law", p. 50–85.

behavior, also when they do not pose an immediate threat. Quong has argued in favor of this model, but also immediately stipulated that preventive state intervention ought to abide by severe constraints:

> *"Any state action that contravenes this special duty [of civil non-interference] can therefore only be justifiable if it is supported by extremely compelling public reasons: reasons that are considered strong enough to outweigh the rights claim."*[276]

Ekeli formulated a similar criterion in the form of the 'significance principle':

> *"The more a particular restriction on the freedom to spread doctrines and ideas undermines the value of this liberty as an institutional means to promote free and well-informed discussion and deliberation, the stronger is the reason against this restriction."*[277]

Both criteria indicate that preventive action is only appropriate in a few cases, requires elaborate and convincing argumentation in support, and should only be implemented after a careful cost-benefit analysis that takes the practical context into account. Still, the core point remains that preventive containment can sometimes be a legitimate tool – even if it should always be used with reluctance and restraint. In particular, it is a proper measure when an unreasonable group starts participating actively in the liberal society and seeks political momentum, or when unreasonable citizens impede the capacity of other citizens to develop their moral powers.

The primary benefit to a proactive attitude in containing the unreasonable by the liberal order is a renewed and enduring appreciation for stability for the right reasons, which is of fundamental importance to the long-term prospering of liberal democracy. The liberal order has a duty towards reasonable citizens, who willingly support its rule, to keep the instruments of its demise out of the hands of those who wish to bring it about. In light of this, it is better to prevent than to cure. Most containment policies need quite some time to become effective – thus, preventive measures must be available to buy that time for other approaches to take over in the long run. I will briefly illustrate, using examples of Quong and Ekeli, what preventive containment could look like in practice, which rights are restricted, and in what way.

First, freedom of expression. This right is itself grounded in the reasonable ideal that all citizens ought to be regarded as free and equal persons. Therefore, its scope cannot cover unreasonable citizens disseminating dehumanizing messages, such as hate speech, for such expressions are diametrically opposed to the normative foundation of free speech. Thus, it is permissible for a liberal order to crack down on seditious speech or discriminatory rhetoric against marginalized ethnicities, religions, or social groups – even if this curtails the freedom of expression of unreasonable citizens. [278] For example, the liberal order might prevent escalation by shutting down broadcast organizations that contest core liberal values and are closely associated with unreasonable doctrines. Since the modern age has many powerful communication tools, unreasonable movements must be prevented from founding media spaces that serve as a breeding ground for unreasonable attitudes.

Second, freedom of education. This right is derived in political liberalism from the reasonable ideal that all citizens ought to be aided in cultivating their moral powers, so

---

[276] Quong, *Liberalism Without Perfection*, p. 310.
[277] Ekeli, "Liberalism and Permissible Suppression of Illiberal Ideas", p.179-180.
[278] Quong, *Liberalism Without Perfection*, p. 309-312.

that they may develop a capacity for justice. Therefore, education that inculcates in children an unreasonable attitude and illiberal values falls outside the scope of this right: forcing a comprehensive doctrine upon someone else contradicts accepting the burdens of judgment. Thus, the liberal order is justified in blocking unreasonable groups from establishing private schools as places of indoctrination. [279] For instance, the liberal order could shut down orthodox religious schools that display institutional sexism, denying young girls their status and capacities as free and equal citizens.

Third, freedom of association. This right is founded upon the reasonable ideal of society as a system of fair cooperation. Therefore, assemblies that are convened to promote their own dominance at the cost of the social norms of cooperation, undermine that ideal and are not covered in the scope of this right. Thus, the liberal order acts properly when it prohibits assemblies and associations convened for those purposes. [280] For example, the liberal order could deny unreasonable citizens to form political parties based on unreasonable doctrines, dissolve public rallies held with the explicit purpose of propagating unreasonable views, or regulate sources of funding and resources for growing unreasonable movements in liberal society.

6.4 Evaluating Preventive Intervention

Now that we have a better view of the benefits and application of preventive intervention, it is time to consider some of its complications. I will now discuss three challenges to preventive containment in particular.

6.4.1 Argument of Distrust

The first challenge, formulated by Ekeli, is the aforementioned 'argument of distrust'. Once the liberal order starts using its power to prevent the emergence of powerful unreasonable movements, this creates the possibility that this power will be abused later on. According to this view, a state with a proactive containment policy will slide down the slippery slope of prevention towards exploiting that power and wielding it against reasonable citizens. Indeed, it should be admitted that, as Quong writes: *"History is rife with examples of the abuse of political power by some to suppress the ideas of their political opponents."* [281] Governments have strong incentives to silence critics, and political liberalism pursues containment policies with the explicit goal of curtailing the challenges by the unreasonable. Since these policies are carried out by humans, that means that it will inevitably occur that preventive powers are mishandled – as is the case for every form of power. That caveat should not be dismissed.

However, Ekeli already points out that engaging in a slippery slope logic about the potential misuse of state power is itself a slippery slope. [282] When applied consistently, it would undermine any proposal that grants the state some regulatory power. Conversely, what if *not* adopting prevention as state practice sets a dangerous precedent? Perhaps, this will lead over time for the liberal democratic state to become paralyzed by anxiety about potential unintended negative consequences, and thus stop fulfilling its core tasks. All in all, this does not seem like a productive line of reasoning.

---

[279] Quong, *Liberalism Without Perfection*, p. 301-305.
[280] Ekeli, "Liberalism and Permissible Suppression", p. 183-184.
[281] Quong, *Liberalism Without Perfection*, p. 305.
[282] Ekeli, "Liberalism and Permissible Suppression", p. 188-189.

Rather, the real question is whether permitting the first step of preventive state action would greatly increase the probability of society having to suffer an oppressive regime. I think it will not. Although allowing preventive intervention would indeed slightly increase the chances of the liberal order sometimes treating its citizens unjustly (since every power will occasionally be misused, at the very least by accident), I argue that this increase is greatly outweighed by other factors. Firstly, implementing preventive policies will significantly diminish the closer threat. Not reasonable but unreasonable movements are simply the more likely perpetrators of instituting an oppressive regime. Secondly, preventive policies will put a halt to some ongoing oppressions in society. For instance, even if they currently abide by the liberal order as a modus vivendi, unreasonable doctrines that are sexist already act as an oppressive force to women within their own communities; preventive policies could somewhat reduce this oppression. Thirdly, liberal democracy has a historical record of developing firm checks and balances on power: it already possesses the appropriate tools to address misuse of preventive powers when it will happen. Hence, on aggregate, I think that preventive containment will likely decrease injustice. All in all, the argument of distrust does not succeed as an objection, although it is an important caveat.

6.4.2 Selective Restrictions

The second challenge, formulated by Sunstein, is that the preventive approach imposes restrictions selectively.[283] When the rights of citizens are being curtailed, there is additional cause for concern about possible injustices if the burden of these restrictions falls primarily on an identifiable out-group. In contrast, if liberty-infringing measures are supported by those citizens who are also subject to them, these measures are more likely to be justified. If restrictions apply broadly, then political incentives will act as a strong safeguard against the state abusing its preventive powers. Indeed, if preventive policies are disproportionally focused on identifiable minority groups, then intervention can quickly devolve into intrusion and injustice – especially if these groups have a longer history of being subject to state oppression. I think that Sunstein's warning is certainly warranted. Accordingly, the liberal order ought to ensure that the implementation of prevention measures faces a high burden proof, receives legal scrutiny, is routinely re-assessed, and is monitored by watchdog organizations.

Nevertheless, in my view, this challenge is not a fatal objection to preventive containment. Consider that a policy is discriminatory and unjust if it treats a certain group of citizens differently without a legitimate justification: the liberal order ought to abide by the egalitarian principle that similar cases are handled similarly and that a relevant distinction is necessary to justify differential treatment. Concerning the issue of curbing the rights of unreasonable citizens, there is indeed such a relevant and significant distinguishing factor. Politically active unreasonable citizens attempt to exploit their basic rights to undermine the liberal order that upholds those same rights for everybody, and ultimately aim to impose their own comprehensive doctrines on others. Differential treatment requires a relevant difference, and in this case, the intent to subvert democratic principles and use coercion to enforce their comprehensive doctrines constitutes such a distinction.

Furthermore, the prohibited grounds of discrimination represent personal characteristics that are innate, immutable, or fundamental to a person's identity and

---

[283] Sunstein, Laws of Fear, p. 205-209.

dignity.[284] In contrast, being unreasonable is a political attitude that can change, not an unalterable trait with deep intrinsic personal significance. The unreasonable do not constitute a clearly identifiable minority group similar to, for example, religious identity. Thus, even if the burdens of prevention are not equally borne by all citizens, this does not automatically render preventative policies discriminatory and unjust.[285] Though the challenge by Sunstein may help to navigate practical complications, it is not an argument against prevention as such.

6.4.3 Counterproductive Effects

The third challenge, formulated by Ekeli, is that restricting the rights of unreasonable citizens acts counterproductively.[286] The worry is that heavy-handed preventive measures can escalate conflicts with unreasonable groups, alienating them and hardening their resistance instead of defusing it through more measured approaches. Aggressive prevention could validate the narratives of unreasonable groups about state oppression, which aids their radicalization and recruitment efforts. Moreover, I think that an overreliance on preventive restrictions of rights would unduly divert resources away from addressing the root causes of rising unreasonableness.

Yet, while these risks ought to be acknowledged, I contend that there are several ways in which the potentially counterproductive effects of preventive intervention can be mitigated. Initially, by implementing preventive intervention only as a secondary option after community engagement has been tried first; for the nature of the preventive approach requires that it is tailored to its target doctrine. Later, by integrating preventive measures into a comprehensive strategy, combining them with initiatives that address the underlying factors that encourage unreasonableness.[287] Still, if these methods are unsuccessful, I think political liberalism ought to bite the bullet. Even if intervention causes resentment among some unreasonable citizens, non-intervention would cause resentment among many reasonable citizens who expect the liberal order to protect the shared ideal of society as a system of fair cooperation. Given that trade-off, the concerns of the reasonable take precedence.

6.5 Conclusion

In this sixth chapter, I have reviewed the case for the liberal order to achieve containment through marginalization: occasionally restricting certain rights of unreasonable citizens, typically concerning freedom of expression, education, and association. I have investigated the two main approaches to marginalizing the unreasonable, which both attempt to ward off unreasonable doctrines while avoiding unjust oppression of unreasonable citizens. The reactive approach, argued by Rawls and Sunstein, restricts itself to only responding to imminent escalation in cases of clear and present danger, and fears that more expansive policies lead too easily to abuse of power. However, I argued that the subsequent intervention will likely happen too late, undervalues the unity of society, provides perverse institutional incentives, and can be

---

[284] Gerards, J. H. (2007). The Grounds of Discrimination.
[285] Alternatively, one could emphasize that preventative measures, primarily targeting not specific citizens but doctrines, consist of broad policies applied across society. Thus, the burdens of prevention are actually also born by reasonable citizens who take care to stay that way.
[286] Ekeli, "Liberalism and Permissible Suppression", p. 190-191.
[287] Besides, for the reasons mentioned in §6.4.1, preventive action may, on aggregate, lead to decreased oppression and injustice.

circumvented by capturing the judiciary. The preventive approach, argued by Quong and Ekeli, expands upon the reactive account and additionally aims to prevent the emergence of unreasonable movements as a serious threat. I asserted that it is better to prevent than to cure - provided that the liberal order displays reluctance and restraint and institutes a series of checks and balances on this power. Still, although reactive and preventive policies form an indispensable part of an effective strategy for containing the unreasonable, these should be integrated into a comprehensive strategy and never serve as the default option to engage with the unreasonable.

This concludes Part II of this thesis project, over the course of which I have critically examined the primary strategies proposed within political philosophy to contain unreasonableness in a liberal society, focusing specifically on the approaches of conciliation, transformation, or marginalization. Through this critical review of currently available responses, I have endeavored to contribute towards a more comprehensive understanding of these potential strategies, the trade-offs between ideals of political liberalism, and the challenges that they face on normative or practical grounds. Thus, the stage has been set to consider how these challenges can be overcome and to explore in Part III what an adequate comprehensive strategy for containment may look like.

PART III: Resolving the Trilemma

In Part II, I have identified some approaches to containing the unreasonable as necessary, though not sufficient. In Part III, I propose additional ways to move towards a sufficient and coherent strategy for the successful containment of unreasonable doctrines and citizens, thus allowing political liberalism to overcome the trilemma of simple pluralism and answer the central research question of this thesis project:
"How should a liberal order navigating the fact of simple pluralism achieve the containment of unreasonableness while remaining normatively consistent with political liberalism?"
I argue that, in addition to those containment strategies already identified as promising in Part II, three extra strategies are needed. Individually, each addresses one challenge posed by the trilemma; collectively, they form a coherent three-pronged approach that could potentially fill the gap in adequately resolving the issue of unreasonableness.

The subsequent three chapters will each develop one part of the general resolution of the trilemma. In Chapter 7, I introduce a new typology of the unreasonable and argue that, with the right targeted strategy, several of these subgroups could become reasonable over time, thereby enhancing the *inclusivity* of the constituency of public justification. In Chapter 8, I discuss a public justification that provides unreasonable citizens with a plausible reason to accept their own coercion. Derived from the general preconditions of participation in public discourse, this public justification enhances the *legitimacy* of the liberal order. In Chapter 9, I advocate expanding the Rawlsian notion of reasonableness with an additional duty of vigilance. Also, I recommend an alternative approach in conversation with the unreasonable that does not focus on conjecture but epistemology. This will bolster those who are already reasonable, and thereby enhance the *stability* of the overlapping consensus.

Considered together, these three partial responses are envisioned not just as mutually compatible but as complementary, covering for each other's vulnerabilities. This is essential since, due to the inherent complexity of the trilemma of simple pluralism, a mere collection of measures will likely end up ineffective. Instead, it is required that all proposed strategies become integrated so that every dimension of the trilemma is addressed, but in such a way that the partial approaches do not undermine the efforts of the others. No single approach can be a panacea in containing the unreasonable. Rather, a successful containment policy must form one comprehensive whole out of mutually supportive partial strategies.

Before integrating these partial approaches in the conclusion of this thesis project, let us now first develop them separately.

Chapter 7: Inclusivity & Involvement

Thesis VII: That the category of unreasonableness is not homogeneous, but actually consists of four distinct subcategories (the anti-reasonable, the uncooperative, the unequitable, and the unaware) that require different approaches in order to become more reasonable – two of which could feasibly become reasonable over time by a strategy of involvement.

How can a liberal order practice its ideal of inclusivity towards all citizens when, given the fact of simple pluralism, the enduring and significant political presence of unreasonable citizens inhibits the formation of a compelling majority of reasonable citizens, which is what ensures stability for the right reasons? It has already been established that unreasonableness ought to be contained, but a comprehensive and sufficient strategy of containment can only be developed once the diverse group of citizens that are unreasonable has been conceptualized clearly. In this seventh chapter, I will explain first why I think that unreasonableness should not be theorized as a binary category. Next, I will develop four separate subcategories of the unreasonable – a distinction that I will also invoke in subsequent chapters. Then, I will elaborate on how involvement could likely persuade citizens from two subcategories in particular to become reasonable and contribute to a compelling majority of reasonable citizens. Ultimately, this alleviates the original inclusivity challenge posed by the trilemma.

7.1 Conceptualizing the Unreasonable

Originally, Rawls conceived the reasonable as a virtuous attitude consisting of two basic aspects: they abide by the criterion of reciprocity (cooperating on fair terms with other citizens, which they acknowledge as equal and free) and they accept the burdens of judgment (not imposing their own comprehensive doctrine upon others and expressing themselves during political deliberation in public reasons). In addition, I have argued being reasonable in the Rawlsian sense includes a third basic aspect: fulfilling the duty of prudence by the conscientious exercise of reason. Since missing any basic aspect renders one unreasonable according to Rawls, this results in a binary account: citizens are at a given moment either reasonable or they are unreasonable. [288]

It is understandable why Rawls did not provide a more detailed account of the unreasonable. In his view, the free institutions under a liberal order would over time generate the fact of reasonable pluralism; and unreasonableness would be of no political significance in the margins of society. Instead, Rawls *"lays out political ideals, principles, and standards as criteria of the reasonable"*[289], and in response to assertions that religious or philosophical truth ought to override the politically reasonable, Rawls states: *"We simply say that such a doctrine is politically unreasonable. Within political liberalism nothing more need be said."*[290] Concerning the unreasonable, he asserts that *"political liberalism does not engage those who think this way"*. Thus, Rawls does not further delve into subcategories of the unreasonable – that would be beside his point.[291]

---

[288] Their political attitude may change over time, and their degree of (un)reasonableness may differ; still, Rawls abides by this two-fold distinction.
[289] Rawls, *Political Liberalism*, p. 394-395.
[290] Ibid., p. 488.
[291] Ibid., p. 442.

However, it is of interest to this thesis project, which disagrees with Rawls' foundational assumption of the fact of reasonable pluralism. Instead, political liberalism ought to provide some account of how it ensures that a compelling majority of its citizens is reasonable; hence, unreasonable perspectives and ways become relevant due to their sheer ubiquity. The main risks of ignoring unreasonable perspectives are that it could allow them to fester and grow stronger without any attempt at containment, potentially destabilizing liberal institutions over time; additionally, that political liberalism appears intolerant itself, which could undermine political liberalism's moral authority and self-understanding. Whereas Rawls views any citizen who does not exhibit every basic aspect of reasonableness as unreasonable, I contend that it is theoretically interesting that citizens can fail to be reasonable in notably different ways. It would be sensible to develop a more complex understanding of the broader umbrella concept 'unreasonable' by dividing it into distinct subcategories, depending on which of the basic aspects of reasonableness is lacking.

I posit that four significant subcategories of unreasonable citizens are particularly relevant to Rawls and political liberalism. Three of these subcategories contain unreasonable citizens who are mostly reasonable, except for lacking one basic aspect of reasonableness. First, there are the 'uncooperative' citizens, who do not abide by the criterion of reciprocity. Second, there are the 'unequitable' citizens who do not accept the burdens of judgment. Third, there are the 'unaware' citizens who do not fulfill the duty of prudence. These three subcategories stand in contrast with the fourth, which contains 'anti-reasonable' citizens, who lack all basic aspects of reasonableness.[292] These four subcategories are displayed together in Table 1.

Table 1. Reasonableness and Four Types of Unreasonable Attitudes

	Abides by the criterion of reciprocity	Accepts the burdens of judgment	Fulfills the duty of prudence
Reasonable	Yes	Yes	Yes
'Uncooperative'	No	Yes	Yes
'Unequitable'[293]	Yes	No	Yes
'Unaware'	Yes	Yes	No
Anti-reasonable	No	No	No

Before we continue, several comments are in order concerning this typology. Firstly, there are eight possible configurations of the three basic aspects, yet I have chosen to leave out three political attitudes. Specifically, I excluded those subcategories where citizens display merely one of the three different basic aspects. These groups are more difficult to identify in practice, and they appear of secondary theoretical importance compared to the others. Analysis of these groups goes beyond the scope of this thesis project. Secondly, these attitudes occur on a spectrum and in different degrees. Citizens can move between categories during their lives, and the distinction may not always be readily apparent. Thirdly, this typology arises out of conceptual considerations and is

---

[292] See also §2.4
[293] I have chosen this less common spelling instead of 'inequitable' in order to preserve symmetry with the other subcategories and unreasonableness.

used for theoretical purposes; some philosophers have proposed other ways to further distinguish between types of unreasonable citizens.[294]

7.2 Four Subcategories of Unreasonableness

Now that the new subcategories have been set out, some specific examples are in order. To illustrate how forms of unreasonableness can differ in practice, I suggest per subcategory a political group and that appears to have adopted that particular political attitude. Additionally, to provide some contextual variety, I will also suggest as a wholly separate example a public intellectual who I would assign to that subcategory.

7.2.1 Uncooperative Citizens

First, let us consider *uncooperative* citizens, who accept the burdens of judgment and fulfill the duty of prudence, but are still not reasonable because they reject the criterion of reciprocity.

One example of a political group that could be classified as uncooperative is staunch libertarians. It cannot be disputed that this ideology has an analytical tradition that engages deeply with philosophical arguments and reasons conscientiously about their political beliefs. Furthermore, its valorization of liberty clearly indicates acceptance of the burdens of judgment. However, libertarians do not abide by the criterion of reciprocity when they refuse to support social policies that ensure fair terms of cooperation, such as welfare programs or progressive taxation.[295] Their commitment to absolute individual freedom can lead to an uncooperative attitude in ensuring that others have equal access to resources and opportunities. In this sense, libertarians can be unreasonable.

An uncooperative political theorist, in my view, is John Mearsheimer. As a professor of political science who publicly engages in debates on American foreign policy, he plausibly reasons conscientiously. Additionally, in his realist approach to international relations, ideologies are largely deemed irrelevant – so he appears to accept the burdens of judgment. Nevertheless, his advocacy of conducting politics almost solely based on power relations favors so-called Great Powers and dismisses the interests of minor nations; in practice, smaller states are denied the status of being free and equal.[296] Thus, Mearsheimer could be unreasonable in the uncooperative sense.

7.2.2 Unequitable Citizens

Second, let us examine *unequitable* citizens, who abide by the criterion of reciprocity and fulfill the duty of prudence, yet are not classified as reasonable due to their rejection of the burdens of judgment.

A political group that I would characterize as unequitable consists of committed Marxists. They surely reason conscientiously about their political and economic theories and often engage in activism. Moreover, their critique of exploitation under capitalism, their opposition to class-based privileges, and their advocacy for workers'

---

[294] I already discussed Kelly and McPherson in §4.2. In addition, Roberta Sala has proposed the subcategory of 'non-reasonable' citizens. However, I do not endorse her typology, as I find the concept hard to apply beyond the strict context of stability. Moreover, I think that Sala has provided insufficient normative and empirical reasons to deviate from Rawls. See: Sala, "Unreasonable People beyond Rawls", p. 253.
[295] Shapiro, "Is There a Libertarian Defense of the Welfare State?"
[296] Stallard, "What John Mearsheimer gets wrong about Ukraine."

rights reflects a fundamental belief in fair cooperation and the equality of all citizens. However, strict Marxists reject the burdens of judgment when they rule out the option of gradual economic reform and instead assert that the current capitalist system must be overthrown. This insistence does not recognize that reasonable disagreement exists about the best means to achieve social and economic justice.[297] In this case, strict Marxists are unreasonable in the unequitable sense.

One example of unequitable theorists, I think, is the New Atheism movement.[298] This collection of explicitly anti-theist intellectuals valorizes reasoning conscientiously and thus fulfills the duty of prudence. In addition, they have been vocal defenders of freedom of speech, arguing that all individuals should have the right to express their views freely regardless of their beliefs. This stance supports the reciprocal ideal of regarding all citizens as free and equal. However, New Atheist thinkers have repeatedly dismissed religious belief itself as fundamentally irrational or delusional,[299] and also have asserted that moderate religious believers provide cover for extremists by legitimizing faith itself.[300] Hence, they implicitly reject the burdens of judgment, as they do not fully respect the complexity and sincerity of religious convictions held by many other citizens. Thus, New Atheists are unreasonable in the unequitable sense.

7.2.3 Unaware Citizens

Third, we turn our attention to *unaware* citizens, who abide by the criterion of reciprocity and accept the burdens of judgment; however, they do not fulfill the duty of prudence and therefore cannot rightly be called reasonable in the full sense.

I argue that anti-vaccination advocates constitute an unaware political group.[301] Their stance is rooted in a concern for individual rights and the well-being of their children and communities. They emphasize the right of every individual to make personal health decisions, reflecting a belief in fair cooperation among all citizens. Additionally, many anti-vaccination advocates recognize the diversity of opinions on health and vaccination and stress the importance of open dialogue and asking critical questions. However, their position relies almost exclusively on anecdotal evidence, discredited studies, and misinformation that rejects the consensus among medical-scientific experts. Thus, anti-vaccination advocates are unreasonable in that they are unaware.

Among political theorists, I think that Noam Chomsky is a suitable example of an unaware public intellectual who can be unreasonable. Chomsky meets the criterion of reciprocity, as he emphasizes the importance of human rights and the fair treatment of all people, regardless of nationality. Furthermore, he acknowledges the complexity of international relations and engages with opposing perspectives as legitimate, thereby accepting the burdens of judgment. However, as Nikolas Kozloff writes: *"[Chomsky] has drawn the world's attention to the various misdeeds of the US and its proxies around the world, and for that he deserves credit. Yet, [he] has turned into something of an ideologue."*[302] Chomsky has been criticized, also by his political allies, for subjecting right-wing authoritarian regimes to significantly more rigorous scrutiny than communist

---

[297] Kleen, "Incrementalism vs. Revolution".
[298] Poole, "The Four Horsemen Review".
[299] Lash, "Where Does The God Delusion Come from?", p. 507–521.
[300] Harris, "The Virus of Religious Moderation".
[301] Blume, "Anti-vaccination Movements and Their Interpretations", p. 628–642.
[302] Kozloff, "Chomsky, Ali, and the Failure to Challenge the Authoritarian Left".

authoritarian regimes. The problem is in part that *"a kind of cult has developed around Noam Chomsky [...] which cannot believe they could ever be wrong, and produces ever more elaborate conspiracy theories to justify their mistakes."*[303] Despite his egalitarian principles and credentials, Chomsky has not displayed determination to overcome this personal bias or informational blindness, and I think is therefore unaware-unreasonable.

7.2.4 Anti-reasonable Citizens

Finally, let us discuss *anti-reasonable* citizens, who fulfill none of the three basic aspects and instead reject all three reasonable virtues: abiding by the criterion of reciprocity, accepting the burdens of judgment, and fulfilling the duty of prudence.

Among all subcategories, anti-reasonable groups are the most notorious; hence, many examples could be relevant. In §2.4, I already identified people with an RWA psychological profile as likely anti-reasonable citizens. Additionally, extremist hate groups and supporters of totalitarian regimes are clearly anti-reasonable. Indeed, all of these groups are highly unlikely to ever seek conciliation with the liberal order: they are actively opposed to reasonableness and see a reasonable attitude as a sign of weakness. However, given recent political tidings, I want to suggest another example of anti-reasonable groups here: conspiracy theory movements.[304] First, they reject reasoning conscientiously, as they renounce established facts and refrain from critical and evidence-based reasoning in favor of rationalizing elaborate and unfounded narratives. Second, they reject the criterion of reciprocity by promoting narratives that demonize perceived enemies or scapegoats while portraying themselves as the sole arbiters of truth. Third, they reject burdens of judgment by dismissing almost everybody who attempts to moderate their worldview as either a naïve simpleton or an agent of the conspiracy. Therefore, they oppose all basic aspects of being reasonable.

Arguably, the most prominent anti-reasonable political philosopher is Carl Schmitt.[305] Schmitt rejected the criterion of reciprocity by advocating for a hierarchical order where the sovereign authority imposes its will upon the populace without regard for equal treatment or personal freedom. Also, Schmitt rejected the burdens of reason by asserting that political decisions are ultimately based on subjective sovereign judgments rather than objective legal or moral principles. Moreover, Schmitt rejected the duty of prudence by denying that rational discourse and compromise could lead to political progress, and instead advocating that decisive action by the sovereign ought to resolve political conflicts. Since Schmitt's political theory is diametrically opposed to liberal democracy, he can rightly be described as an anti-reasonable thinker.

This concludes the section where the four new subcategories of unreasonable citizens are illustrated with examples. Together, these examples illustrate the significant differences between the subcategories of unreasonable citizens. There are several reasons why this four-fold distinction is important. First, this framework allows for a more nuanced analysis of various political groups and public figures. Breaking up a political coalition of unreasonable citizens is contingent upon being able to differentiate between its constituents. Second, it helps to clarify the motivations behind unreasonable behavior. For instance, understanding that some groups are

---

[303] Katerji, "The West's Leftist Intellectuals Who Traffic in Genocide Denial".
[304] Baurmann & Cohnitz, "Trust No-one? The (Social) Epistemological Consequences of Belief in Conspiracy Theories", p. 334-357.
[305] Vinx, "Carl Schmitt".

uncooperative because they reject the criterion of reciprocity helps us grasp the ideological underpinnings behind their stance. Third, it could foster more productive dialogue and enhance discussion from the perspective of the liberal order. Instead of engaging in broad and generalized arguments, the reasonable can now pinpoint specific aspects of unreasonableness and address them more effectively.

In short, if political liberalism must indeed convince some of the unreasonable to become reasonable, it must identify, examine, and adapt to the particular reasons for unreasonable dissent. What, then, seems to be the most promising approach per subcategory? I posit that marginalization is best suited for dealing with anti-reasonable citizens and that a mix of persuasion and transformation is appropriate for the other subcategories. I will discuss how to address the anti-reasonable in Chapter 8 and the unaware in Chapter 9; the rest of this chapter will focus on uncooperative and unequitable citizens.

7.3 Including the Uncooperative and Unequitable

How can the liberal order achieve the inclusion of uncooperative and unequitable citizens? In principle, these partially reasonable citizens are only one basic aspect shy of becoming reasonable and therefore are suitable candidates for targeted persuasion. Ideally, some specific strategy of engagement could convince a significant amount of uncooperative and unequitable citizens to become reasonable. However, the main challenge, I think, is that these unreasonable citizens must ultimately be convinced of assuming in their political attitude a basic aspect that they at present consciously oppose. Uncooperative and unequitable citizens are not unreasonable by accident or due to ignorance – they are unreasonable on purpose. In their view, being unreasonable is a superior moral-political attitude compared to being reasonable. Thus, the most straightforward approach will unlikely make a significant impact. For instance, reminding uncooperative citizens about the importance of social cooperation or unequitable citizens about the unfairness of imposing one's own doctrine upon others would miss the point; these reasonable ideals do not connect with an unreasonable worldview, and further inculcation will therefore be ineffective.

Additionally, the liberal order must avoid many other pitfalls. If it comes across as paternalistic, its efforts will breed resentment and perhaps turn counterproductive. If its approach is overly theoretical, it will likely run into the burdens of judgment and maybe alienate practice-oriented citizens. If its methods are too superficial, it will be ineffective – too far removed from the foundational liberal principles, and it will be disingenuous. Nevertheless, I think that three principles in particular will feature in any successful approach to ensure a compelling majority of reasonable citizens.

First, the liberal order ought to accept that its government has the permanent task of incentivizing unreasonable citizens toward reasonable behavior. Of course, liberal institutions and practices have a natural disciplinary effect over time. Still, by maintaining fair and respectful social systems that enable collective benefit, burden-sharing, and civic responsibilities, the liberal order could encourage additional support for the reasonable conception of justice. Specifically, the incentives must be arranged so that unreasonable citizens recognize that assimilation benefits them more than opposition to the liberal order. For this reason, community engagement should always be combined with concrete social connections, opportunities, and rewards.

Second, the liberal order ought to adjust its messaging when it attempts to persuade partially unreasonable citizens, framing the argument on *their* terms and appealing to *their* perspectives. Ultimately, these citizens can only be guided towards insights that resonate with their own moral compass: nobody but themselves can convince them of the merits of a fully reasonable attitude. Consider, for example, an uncooperative libertarian who is reasonable except for his rejection of the criterion of reciprocity. Then, a reasonable interlocutor should not frame the ideal of fair social cooperation in marked contrast with libertarian values, but as an essential means to uphold them.[306]

Third, the liberal order ought to acknowledge the role of sentiments in moral reasoning. Emotions like empathy, disgust, anger, and shame play a significant causal role in shaping people's moral-political judgments and motivating their actions.[307] When a liberal order engages in public reason, it should recognize this empirical reality rather than treat deliberation as a mostly rational process. Instead, a liberal order could leverage moral sentiments to inspire citizens towards ethical conduct that upholds its principles. In practice, a balance must be negotiated between recognizing the empirical reality and moral significance of emotions, while still maintaining space for reasoned principles and civil deliberation. Both reason and sentiment ought to be incorporated into the moral-political framework of the liberal order.

Having stipulated these three principles, I will now integrate them and develop an inclusive strategy of containment that combines persuasion and transformation.

7.4 Containment as Involvement

How should the liberal order engage with uncooperative and unequitable citizens? In essence, I argue that the inclusion of these two subcategories could be improved by an approach which I call *involvement*. Involvement is similar to and inspired by reasoning from conjecture, most notably since both feature a reasonable citizen[308] attempting to persuade an unreasonable interlocutor by suggesting insights they may not have had before. However, there are also some key differences. For instance, conjecture focuses primarily on doctrine, whereas involvement focuses primarily on attitude. Also, conjecture appeals to the power of reason, whereas involvement appeals to one's sense of justice. The following discussion provides a theoretical overview of the involvement strategy, setting out the general principles. I will address some potential criticisms in the next section.

Involvement starts from the premise that changing a moral-political fundamental attitude will inevitably provoke some cognitive dissonance during the process. This mental discomfort must be overcome somehow – or the persuasive effort will likely backfire and prove counterproductive. Thus, to avoid this scenario, involvement is only appropriate if the interlocutor meets three conditions. First, the interlocutor is already struggling with their political attitude, intuitions, or doctrine, and seeks to overcome the dissatisfaction caused by this cognitive dissonance. Thus, involvement occurs as a genuine effort to help, not as an imposition or challenge. Second, the interlocutor already has some sort of personal relationship with the reasonable

---

[306] For instance, by offering the reasoning that a commitment to genuine self-ownership for all necessitates conceiving society as a fair system of cooperation where everyone has a stake and no one is systematically excluded from substantive freedom due to unchosen circumstances.
[307] Hümmler & Schiesser, *Fact and Prejudice*, p. 51.
[308] This reasonable citizen could also be a community leader or government official.

'involver', who has listened attentively to the former's story and perspective. Thus, the interlocutor has been assured that they are principally regarded as a valued fellow citizen, and not reduced to their doctrinal beliefs as a dissident. Third, the interlocutor already fulfills the duty of prudence and values reasoning conscientiously in itself. This is why uncooperative and unequitable citizens are suitable targets for involvement whereas other unreasonable citizens are not: persuasion is only possible if revising views based on new insights is regarded as a strength instead of a weakness. Together, these three conditions serve to make the interlocutor likely to recognize that resolving their cognitive dissonance is compatible with becoming fully reasonable.

Assuming these preconditions, involvement starts with acknowledging that the unreasonable interlocutor is struggling on some moral-political issue and that some insight is missing before conceptual consistency can be achieved. For instance, their comprehensive doctrine clashes with their personal intuition, or they feel conflicting instincts and their doctrine provides insufficient guidance. Then, the involver establishes three corollaries: (1) that the interlocutor has a strong sense of justice – which is true, or they would not have felt bothered by this issue in the first place, (2) that the interlocutor values the duty of prudence – which is true, or they would already have rationalized their predicament away, and (3) that the interlocutor values one basic aspect of reasonableness (whether the criterion of reciprocity or the burdens of judgment) – which is true, or their position would be nearly indistinguishable from an anti-reasonable one. It will soon become clear why these elements provide the stepping stones towards reasonableness.

The core idea of involvement is that the missing basic aspect of reasonableness may be revealed using the other two basic aspects, which are present in this scenario. Thus, the involver may assert that abiding by the criterion of reciprocity and accepting the burdens of judgment are two sides of the same coin. After all, the willingness to propose fair cooperative terms arises from appreciating the need for reciprocal justification given the fact of simple pluralism. Conversely, accepting the burdens of judgment rules out imposing one's own comprehensive doctrine on other citizens, which in turn necessitates social cooperation on fair terms. Hence, these two basic aspects of reasonableness are complementary and mutually reinforcing, being intrinsically linked and representing different aspects of the same underlying virtue.

It is here that the corollaries come into play, in reverse order. If the involver has established (3), then the insight that the two basic aspects formulated by Rawls are extensions of each other should be mentally *conceivable* by the interlocutor. This insight does not have to be demonstrated as definitely true beyond all doubt; rather, the point is merely that adopting the missing basic aspect of reasonableness is plausible and a viable option. If the involver has established (2), then any additional cognitive dissonance that arises while mentally playing with this insight should be *manageable* for the interlocutor. Given that the open-minded interlocutor understands themselves as somebody who values intellectual growth, this self-perception will help them resolve their cognitive dissonance in a constructive way. If the involver has established (1), then this allows for the insight to *resonate* with the interlocutor. Validating the moral compass of the interlocutor will provide them a starting point from which to authentically develop their ethical reasoning within a reasonable framework.

The final step of involvement is that the involver helps the interlocutor to join a reasonable association that is both connected to their personal moral intuitions and also fosters the previously missing basic aspect. For instance, an association that exposes its members to citizens from a wild diversity of backgrounds fosters the criterion of reciprocity. Alternatively, an association that exposes its members to various comprehensive doctrines fosters the burdens of judgment. If the interlocutor indeed joins the reasonable association, this has three main benefits. First, becoming more reasonable can be a challenging process, and their new community can provide support while they work through this internal conflict. Second, social reinforcement makes the interlocutor more likely to internalize and consistently apply those norms of reasonableness in their own deliberation. Third, it provides the interlocutor with accountability to help sustain their commitment to reasonableness over time. In short, instead of aiming at a change in beliefs, involvement primarily focuses on changing behavior: encouraging the interlocutor to join a reasonable community and thereafter acquire over time a sense of belonging among reasonable citizens.

7.5 Rebutting Potential Criticisms

Having proposed this strategy of engaging with uncooperative and unequitable citizens, I will now defend involvement against three possible objections.

The first objection is that involvement engages too much in reasoning, and will therefore, like conjecture, run into the burdens of judgment. I want to confirm that, indeed, involvement should generally avoid reasoning as much as possible. But further, I will point out that involvement in fact entails very little reasoning. First, the three preliminary corollaries are less inferred propositions and more validations of aspects already present in the identity of the interlocutor. This process is not one of reasoning, but of clarification and identification. Second, the interconnection between the burdens of judgment and the criterion of reciprocity is postulated, not proved. The point of involvement is that this assertion could be true, not that it has been demonstrated as definitely true beyond all doubt. Unlike conjecture, involvement does not attempt to overcome the burdens of judgment – it merely proposes that it covers one more option. Thirdly, the recommendation to join a reasonable association is not a reasoned conclusion. Instead of an attempt to change the interlocutor's beliefs, it is merely a suggestion, an encouragement to temporarily experiment with changing their behavior given the lack of satisfying alternatives. Involvement does not directly engage with reasons or dogma – it can only succeed by appealing to the core intuitions and deep sentiments that underlie the relevant comprehensive doctrine.

Another potential objection is that the involver cannot sincerely link the basic aspects of reasonableness to the moral-political cognitive dissonance of the interlocutor, since they are often irrelevant to the issue at hand. Still, I maintain that the two basic aspects in question are relevant to virtually every moral-political dilemma. The criterion of reciprocity is central to any moral-political issue because it ensures that individuals' rights and perspectives are considered equally and no group is unfairly privileged over another. Also, many moral-political issues arise from conflicting interests and values, and the criterion of reciprocity, by requiring individuals to propose and accept fair terms of cooperation, is a crucial guideline for resolving such conflicts. Furthermore, the burdens of judgment acknowledge that reasonable people can disagree on moral and political issues due to the complexity of evidence, different life experiences, and

value systems. This basic aspect is vital for fostering an environment of tolerance and respect. Also, it promotes a culture of open-mindedness and ongoing dialogue. Hence, the involver will relieve the cognitive dissonance: either by introducing the burdens of judgment and taking off the pressure of having to resolve it, or by introducing the criterion of reciprocity and affirming the value of egalitarian aspirations and engaging in cooperation in good faith.

The final potential objection I discuss here is that the interlocutor might refuse to join a reasonable association. This concern is important because the success of the involvement strategy ultimately hinges on the interlocutor's willingness to engage with a community that embodies and promotes the missing aspect of reasonableness. Yet, the recommendation to join such an association is a natural next step for the interlocutor for several reasons. Firstly, because involvement emphasizes the importance of managing cognitive dissonance in the process of becoming reasonable, and joining a reasonable association does provide the interlocutor with a supportive environment where they can work through this mental discomfort. Secondly, because involvement presupposes that the interlocutor values conscientious reasoning, and a reasonable association validates the duty of prudence by providing a space where growth is encouraged. Thirdly, because deep-seated political attitudes can only be changed by habituation, and any reasonable insight obtained during a discussion needs to be reinforced through consistent practice and social engagement in order to truly take root. Exposure helps to solidify the abstract principles discussed during involvement into lived experiences, making the transition to reasonableness more tangible and less abstract.

For these reasons, I maintain that involvement is a promising strategy in helping a significant amount of uncooperative and unequitable citizens to become reasonable.

7.6 Conclusion

In this seventh chapter, I have argued why I think a strict binary of reasonable-unreasonable falls short of being a satisfying analysis. Hence, I have introduced a four-fold distinction of unreasonable citizens; subcategories that correspond to the lacking basic aspects of reasonableness (whether one specifically or all entirely). Moreover, I contend that each of these subcategories requires its own micro-targeted containment approach by the liberal in order to effectively persuade them towards greater reasonableness. I have provided three guiding principles for engaging with uncooperative and unequitable citizens. Finally, I have introduced the strategy of involvement as a suitable way to persuade and indirectly contain those two subcategories of unreasonable citizens. Since involvement can be hypothesized to be more effective than its alternative, conjecture, this approach alleviates the inclusivity challenge caused by the trilemma of pluralism – provided that a significant part of the unreasonable is indeed included and persuaded by the liberal order.

Now that I have considered uncooperative and unequitable citizens, who are still partially reasonable, I will now discuss the wholly unreasonable category of anti-reasonable citizens.

Chapter 8: Legitimacy & Intelligibility

Thesis VIII: That the liberal order can provide anti-reasonable citizens with adequate reason for excluding them from further deliberation on matters of legitimacy, on the grounds that meaningful membership in the constituency of public justification imposes certain epistemic norms that render citizens unintelligible if they pursue unreasonable proposals in that context.

How can a liberal order practice its ideal of legitimacy towards all citizens when unreasonable citizens persist in rejecting the reasonable public justifications that they are offered? This is particularly relevant concerning anti-reasonable citizens, who cannot realistically be expected to become reasonable at any point. In this eighth chapter, I argue in favor of a justification outside the reasonable framework which still gives unreasonable citizens adequate reason to accept their own coercion by the liberal order as legitimate. First, I determine what could be suitable grounds that are compatible with reasonableness but still can appeal to unreasonable citizens. Then, I will explore how intelligibility, necessary for the very process of public justification, inherently puts some epistemic constraints on participants in public deliberation. Next, I contend that this causes a dilemma for anti-reasonable citizens, as theorized by the accounts of Onora O'Neill and Julia Netter, that results in members of the constituency of public justification becoming unintelligible if they pursue unreasonable goals. Hence, unreasonable citizens cannot intelligibly advocate their own inclusion in the constituency of public justification. Thus, the liberal order is legitimate since it can offer an accessible justification to unreasonable citizens as to why their dismissal of public justifications is inconsequential to concerns of political justice. Finally, I will address some potential criticisms of the argument proposed in this chapter. If even anti-reasonable citizens can indeed be given adequate reason to accept that a liberal order is justified in providing only public justifications for coercive power, then this would alleviate the legitimacy challenge originally posed by the trilemma.

8.1 Grounds of Justification

It has been previously established that the legitimacy ideal in political liberalism requires that political power, especially concerning constitutional essentials, be exercised only in ways that all members of the constituency of public justification could reasonably be expected to endorse in light of their common human reason.[309] Thus, a liberal order becomes legitimate if it successfully provides justifications that are suitable for unanimous endorsement by its constituency of public justification. However, anti-reasonable citizens oppose the liberal order because they reject the public justifications offered for coercion. The liberal order cannot respond by offering non-public justifications, as this would favor some comprehensive doctrine and violate its independent character as a strictly political conception of justice. Neither can it assume that anti-reasonable citizens will eventually be persuaded—whether by conjecture, education, or other means—into becoming reasonable. Therefore, the only plausible way for the liberal order to remain legitimate is to exclude unreasonable citizens from its constituency of public justification and offer the unreasonable citizens

---

[309] See §3.3 and §4.2 of this thesis project.

in question an adequate reason in support—an argument with accessible premises that converges on a single conclusive insight if one reasons rationally. The goal of the argument is not to actually persuade a significant share of anti-reasonable citizens. Instead, it is meant to demonstrate to other, rational citizens why the liberal order can legitimately enforce boundaries on the deliberation of public justifications.

What normative grounds could adequately justify the exclusion of the unreasonable from the constituency of public justification? Since neither a reasonable framework nor an unreasonable comprehensive doctrine is suitable, the only available candidate seems to be public justification itself: to be identifiable as a distinct concept, it has certain inherent features. Thus, some requirements are necessarily imposed on citizens insofar as they can meaningfully act as members of its constituency. To be recognized and comprehended as part of the constituency of public reason, every participating citizen must meet certain basic epistemic standards that enable comprehensible communication. Falling short of these minimal standards means that the citizen in question is not engaged in a matter of public justification at all. The point is that if unreasonable citizens were to participate as members in the process where their community deliberates on justifications of coercive power, they would be subject to certain basic constraints for the descriptions in that conditional to make any sense at all. Ultimately, I will argue that these fundamental constraints are incompatible with pursuing unreasonable proposals, that they can be formulated in terms independent from reasonableness, and that unreasonableness is therefore a suitable criterion for exclusion from the constituency of public justification. But first, let us examine these minimal standards inherent to civic deliberation on public justifications more in detail.

The constituency of public justification consists of citizens who scrutinize public justifications of power to ensure fair mutual cooperation and do so by formulating reasons that are addressed to other citizens. Each citizen willingly proposes ideas that they believe all other members may potentially accept, and deliberate on ideas that other citizens have proposed. This process of public justification is contingent upon proposals that are based on reasons that are accessible to all other participants. Hence, a member of the constituency of public justification cannot in that role formulate proposals using jargon from their own comprehensive doctrine, and neither can they appeal to sources of reason that are not collectively accepted. Rather, a justification becomes *public* not merely by being communicated to others, but when other citizens acknowledge the connection between the proposal and its supporting reasons as rational and appealing. Of course, the constituency does not have to share the exact same reason for support; their personally preferred rational reasons may diverge as long as collective support is achieved on the proposal at hand. Still, proposing a political initiative involves every member being provided with supporting reasons that rationally demonstrate compatibility between each citizen's comprehensive doctrine and the proposal at hand. Together, this reveals that the process of public justification is contingent upon the basic requirement of intelligibility.

8.2 Intelligibility & Norms

In principle, public justifications in political deliberation must address each citizen, as the normative conception of society is derived from their individual perspectives. These justifications must also be intelligible to citizens, as one cannot consider a reason they do not understand. When citizens act intelligibly as participants in the

constituency of public justification, they appear rational in their reasoning and comprehensible in their assumptions. Intelligibility requires in particular that participants in communication share certain epistemic norms. If interlocutors lack these crucial shared epistemic norms, which emerge and evolve in a community of speakers over time, they cannot make themselves understood by the other party. These norms govern, for instance, what counts as attempts at communication, what qualifies as admissible evidence, what constitutes a relevant objection, and what indicates that someone's worldview is grounded in a shared reality.[310]

If one does not meet the conditions imposed by these norms within a community of speakers—such as the constituency of public justification in a liberal society—then they will not be regarded as acting intelligibly. For instance, if someone abides by private epistemic norms, they cannot participate in a public conversation, as their communicated reasons will be unintelligible to interlocutors at large. Particularly in a context centered around justification and respect for reasons, a person speaking irrationally for a prolonged period will be regarded as unintelligible, causing communication to break down. However, if all relevant parties have the sense that they understand or could understand each other, this warrants the assumption that all necessary core epistemic norms are present, allowing communication to continue.

Hence, it is essential that a citizen comes across as intelligible in the constituency of public justification when they engage with other members. This involves at least two norms merely by definition of 'public justification'. First, in support of their own proposals, a citizen must provide *reasons:* other citizens should be able to clearly recognize them as such, as opposed to irrational appeals or non-sequiturs. Second, a citizen must be prepared to deliberate these reasons with *other citizens:* all participants should be willing to engage in discussion of the reasons at hand, addressing each other as agents on an equal basis. Even if the citizen in question is unreasonable, acting as a member of the constituency of public justification necessarily means that one should act intelligibly in providing reasons and deliberately in addressing others. If unreasonable citizens were to defy these norms, this would result in their behavior being incomprehensible as having something to do with public justification at all.

Being regarded as an intelligible citizen who reasons rationally and treats others as agents involves not just speaking cohesively, but also a perceived general alignment between one's beliefs, intentions, and behavior. For instance, if a citizen proposes to prohibit smoking, they become unintelligible to the wider community of citizens if they make this proposal with a lit cigarette between their lips. This illustrates that, when it comes to intelligibility, coherency involves not only speech but also actions. Besides explicit verbal communication, some actions can also communicate a commitment to beliefs which could clarify or undermine the spoken message. In essence, when a citizen performs an action observed by others, this creates a premise that may be reinterpreted but cannot be denied if the account is to remain intelligible. Thus, the aforementioned epistemic norms constrain certain actions besides the act of speaking. Actions do not speak louder than words, but they may contradict them and render the agent unintelligible to their community.

---

[310] The relationship between intelligibility, politics, agency, and radicalism has been explored by Sina Kramer. See: Kramer, *Excluded Within: The (Un)Intelligibility of Radical Political Actors.*

8.3 The Unreasonable Dilemma

So far, I have established that the concept of the constituency of public justification imposes certain epistemic norms on citizens if their membership is to be meaningful. These norms include substantiating positions with reasons that are recognizable as rational and addressing these reasons to other citizens as agents. Additionally, for communication to be intelligible, some alignment between words and deeds is required. Now, let us consider anti-reasonable citizens who reject all three basic aspects of being reasonable: the criterion of reciprocity, the burdens of judgment, and the duty of prudence. This means they do not regard other citizens as free and equal, wish to cooperate under fair terms, refrain from imposing their own doctrine upon others, and attempt to reason conscientiously. If this subcategory of unreasonable citizens were to be included in the constituency of public justification, they would be called upon by other members to provide a rational justification for any unreasonable proposal—such as revoking some civic rights of a subgroup of citizens in violation of protected characteristics. I will now explain why this demand for rational justification causes a dilemma for unreasonable citizens, its two horns having been developed by Onora O'Neill and Julia Netter.[311] When anti-reasonable citizens target fellow members of the constituency of public justification in their proposals for disenfranchisement, they must either affirm or deny them being full-fledged agents. In either option, their position becomes self-contradictory, unintelligible, and untenable.

On the first horn, the unreasonable citizen who initiates a proposal disputing the basic rights of some fellow citizens will affirm that all citizens are agents in the full sense, including those in the targeted subgroup. Yet, if these citizens have done nothing blameworthy, there can be no rational reason to revoke their civic rights. Thus, if the unreasonable citizen persists in their proposal and seeks a rational supportive reason, they become committed to the idea that these citizens have committed some transgression and are sufficiently blameworthy that revoking their basic civic rights becomes an appropriate punishment. If this reason is accepted, the targeted subgroup of citizens would no longer be assigned equal moral concern. However, as Onora O'Neill has pointed out, for this line of reasoning to be rational, these targeted citizens must first be acknowledged as full agents, which is what fundamentally entitles them to equal moral concern.[312] Therefore, the unreasonable account is self-contradictory: it presupposes the agency of the targeted citizens, which is the inalienable characteristic that warrants their equal ethical standing, but it disputes that same characteristic by proposing to revoke the basic civic rights of fellow citizens. Thus, when an unreasonable citizen contradicts their initial commitment to the moral equality of all citizens, their actions as a member of the constituency of public justification render them intelligible, which disqualifies them from participating in further deliberation.

Nevertheless, public justification necessarily entails the procedural norm that members of the constituency not only provide rational reasons but are also willing to convince any other fellow member. Hence, for their status as members to remain meaningful, unreasonable citizens must engage in discussions of their proposals and supportive reasons with other citizens. This practice of justificatory discourse is contingent upon

---

[311] O'Neill, *Towards Justice and Virtue*, p. 97-102.
Netter, "Why Be Reasonable?", p. 98-101.
[312] O'Neill, *Towards Justice and Virtue*, p. 100.

all participants – crucially, including those in the targeted subgroup – being acknowledged as agents, capable of changing their minds and letting their beliefs influence their future actions. Therefore, once unreasonable citizens start addressing their reasons to those members facing disenfranchisement, a contradiction arises between their beliefs and behavior. The unreasonable citizens' proposal denies their interlocutors their equal moral status as agents, but the procedural norms require that the unreasonable appeal with their reasons to all other members of the constituency; engaging with the targeted citizens in rational dialogue grants the premise that the interlocutors are indeed agents. As Julia Netter phrases it: *"To engage in a justificatory discourse with a person can only be interpreted as a performative recognition of their agency."*[313] Thus, when an unreasonable citizen contradicts their initial commitment that the proposal does not target equal agents, their actions as a member of the constituency of public justification render them unintelligible, which disqualifies them from participating in further deliberation.

On either horn of the dilemma, unreasonable citizens cannot comply with the constraints imposed by the epistemic norms that are inherent to meaningfully engaging in public justification as a member of its constituency. Whether they fail to reason rationally or act intelligibly, the unreasonable are bound by unavoidable epistemic norms that they must either respect, thus abandoning their unreasonable agenda, or flout, rendering themselves unintelligible and politically irrelevant. The act of intelligible communication regarding public justifications entails minimal commitments to deliberative rationality and civic respect that make an anti-reasonable attitude untenable. Consequently, even fully unreasonable citizens have sufficient reason, beyond the reasonable framework of political liberalism and therefore accessible to them, to acknowledge that meaningful membership in the constituency of public justification is incompatible with hypothetically pursuing unreasonable proposals in that context. Therefore, unreasonableness is a suitable criterion for exclusion from this constituency and from making valid appeals to receive non-public justifications for the coercive use of political power by the liberal order.

8.4 Rebutting Potential Criticisms

I will now discuss three possible objections to the line of reasoning presented so far.

The first potential objection criticizes the assertion that unreasonable citizens become unintelligible due to self-contradiction in their reasoning. Imperfect rationality, such as committing a fallacy, may not necessarily transform an entire line of reasoning into incomprehensible gibberish; interlocutors may still convey meaningful ideas despite occasional lapses in rationality. To what extent does imperfect rationality cause communication to become unintelligible?

In response, I note that the argument is not committed to the premise that any instance of imperfect rationality causes complete unintelligibility. Rather, it is the degree of rational inconsistency combined with a particular context that renders an unreasonable citizen unintelligible – being not completely incomprehensible but rather unintelligible specifically as an acting member of the constituency of public justification. Hence, I apply this argument primarily to anti-reasonable citizens and not other unreasonable subcategories. Intelligibility arises when unreasonable citizens persistently engage in

---

[313] Netter, "Why Be Reasonable?", p. 99.

fundamental contradictions in reasoning or performative action specifically in a situation where all members of the constituency of public justification are engaged in deliberation. Here, pervasive inconsistencies disrupt the coherence of communication within the framework of public justification, where consistency and rationality are crucial for meaningful deliberation. The primary issue is not that unreasonable citizens are being irrational, but that they cannot possibly be arguing in good faith. Therefore, while imperfections in rationality may not always lead to unintelligibility, they do in the specific context of persistent self-contradiction within public justification.

The second potential objection is that the presented argument is hypocritical because it risks falling into the same pitfalls as the unreasonable attempts it seeks to counter. By advocating for the exclusion of unreasonable citizens from public justification, the argument itself may seemingly engage in the same behavior it condemns.

Despite a superficial resemblance, the argument is markedly different from the unreasonable proposals that it aims to prevent: in content, justification, and coherence. Firstly, whereas anti-reasonable citizens aim to deprive other citizens of basic civil rights, the exclusion argument still grants unreasonable citizens those same rights. It excludes them by necessity from deliberations in which these rights are developed, but that is a separate matter. Secondly, in contrast to the anti-reasonable, the argument does not discriminate based on a protected characteristic but merely an untenable political attitude. Unreasonable citizens are not permanently ostracized due to their identity; they could potentially change their views and be re-admitted into justificatory discourse. Thirdly, the argument maintains intelligibility within the framework of public justification: it formulates a rationally accessible reason and addresses unreasonable citizens to explain the grounds for their exclusion from the constituency, which affirms the status of unreasonable citizens as free and equal persons. Thus, the argument remains consistent with the epistemic norms that it posits, and it is sufficiently distinct from other unreasonable proposals.

The third potential objection is that unreasonable citizens could factually recognize other citizens as full agents without morally recognizing them as free equals. They may acknowledge others as independent and capable agents but nevertheless hold beliefs that lead them to deny the moral equality of others. How does accepting other citizens as factual agents commit unreasonable citizens to respect them as moral equals?

This question may have several possible answers. One is implied by O'Neill, who writes: *"When agents commit themselves to the assumption that there are certain others who are agents […], they cannot coherently deny these assumptions in working out the scope of ethical consideration to which they are committed."*[314] Applied to our case, this means that unreasonable citizens move by themselves from a factual towards a normative discussion when they reason that the citizens targeted by their proposal are somehow blameworthy, or when they address their reasons during deliberation to those same citizens. By either speech or action, they have already performed the acknowledgment of their interlocutor as a moral equal. In short, according to O'Neill, the argumentative ball is in the court of the unreasonable. Another answer is formulated by Netter, who writes that a person *"by virtue of her quality as an agent, which becomes evident to her upon reflection"* must necessarily *"claim some weight in other individuals' deliberations on how to*

---

[314] O'Neill, *Towards Justice and Virtue*, p. 100.

act towards herself", and demand *"to be granted moral concern by others."* However, since reasons are essentially public by nature, one *"cannot deny the legitimacy of the same claim by relevantly similar individuals, that is, other agents."*[315] It is for this reason that if one claims moral concern for oneself as an agent, one must also grant this same moral concern to others. For Netter, since the publicity of reasons generates a universalized moral claim, the recognition of others as agents therefore inherently carries the implication of their moral equality, as denying this would contradict one's own demand for moral consideration. Whether one prefers the account by O'Neill or Netter, it seems that this objection can be dealt with successfully.

8.5 Conclusion

In this eighth chapter, I have advocated that the liberal order can alleviate concerns about its legitimacy by offering an accessible justification to anti-reasonable citizens as to why they are excluded from further deliberation as members of the constituency of public justification. Since this justification is argued on grounds that are independent of the reasonable political framework – instead being derived from the epistemic norms inherent to public justification itself – even unreasonable citizens can be given an adequate reason to accept their own coercion by the liberal order as legitimate. Additionally, I have established why the process of public justification is contingent upon the intelligibility of its participants, which inherently puts some constraints on members of the constituency: they must substantiate their proposals with reasons that are recognizable as rational, and address these reasons to other citizens as agents. Additionally, some alignment between their words and deeds is required. Next, drawing upon the works of Onora O'Neill and Julia Netter, I explained why anti-reasonable citizens face a dilemma where in either scenario they become unintelligible as members of the constituency of public justification if they pursue unreasonable goals. Therefore, even unreasonable citizens have adequate reason to acknowledge that meaningful membership in the constituency of public justification is incompatible with hypothetically pursuing unreasonable proposals in that context. Because this demonstrates that being unreasonable is a suitable criterion for exclusion from this constituency, anti-reasonable citizens cannot challenge the legitimacy of the liberal order. Lastly, I offered rebuttals to three potential objections. In short, this chapter has illustrated that focusing on public intelligibility is a promising strategy for political liberalism to reinforce its claims to legitimacy since it draws not upon comprehensive doctrines but upon communication theory.

---

[315] Netter, "Why Be Reasonable?", p. 109.

This page has been intentionally left blank

Chapter 9: Stability & Vigilance

Thesis IX: That the liberal order should encourage reasonable citizens to perform the duty of vigilance, which entails safeguarding the social norms that enable civil political deliberation, engaging unaware citizens by asking about their epistemology, and cultivating a critical awareness of the benefits of liberal democracy.

How can a liberal society achieve its ideal of stability for the right reasons among a compelling majority of reasonable citizens when any overlapping consensus will continuously face the possibility of erosion by enduring unreasonable influences? I have already discussed that the liberal order has some limited options available to prevent unreasonable movements from becoming a threat to open institutions. [316] However, unlike inclusivity and legitimacy, which is mainly the responsibility of the liberal order to realize, the ideal of stability for the right reasons assigns a crucial role to the general citizenry. In this ninth and final chapter, I examine what reasonable citizens can do to prevent a stable overlapping consensus from backsliding into a mere modus vivendi, and in particular how they should engage with citizens who are unreasonable in the sense that they are unaware.

First, I elaborate on the difficulties in maintaining a compelling reasonable majority in a liberal society and the importance of Rawls' publicity condition. I contend that an approach concerning citizens who are unreasonable in the sense that they are unaware should draw upon several insights from the conversation technique of 'street epistemology' so that it may preserve the epistemic elements of reasonableness and the publicity condition of political liberalism. Next, I will argue that reasonableness should be expanded as a concept and develop the duty of vigilance, which entails that citizens should not only propose and accept but also proactively preserve fair terms of cooperation. Lastly, I will discuss several possible objections against this view. If reasonable citizens can support each other in maintaining a reasonable attitude, and if this strategy can make the unaware reconsider their epistemic approach, then this would alleviate the stability challenge originally posed by the trilemma.

9.1 Stability & the Publicity Condition

When stability denotes a modus vivendi, the main concern for the liberal order is to provide appropriate institutional checks and balances that can preserve the equilibrium of power in society. But when stability means stability for the right reasons, the essential role in safeguarding the stability of society is assigned to ordinary citizens. Whereas ensuring inclusivity and justification is primarily enabled by the liberal order, ensuring stability can only be done by the individual citizens who hopefully decide to willingly support the overlapping consensus. In other words, while the liberal order opens the door and shows it to its citizens, they are ultimately the only ones who can choose to walk through it.

In particular, I would like to highlight a two-folded reason why, instead of the state, citizens play the deciding role in maintaining stability in political liberalism. Firstly,

---

[316] See the discussion in §6.3 of this thesis project.

because the rule of law in society is contingent upon a fundamental resonance between enacted laws and citizens' informal understandings of those laws. [317] The crucial preservation of these norms as a social practice is not a legal but a civic-cultural matter, meaning that citizens carry individual responsibility in actively contributing to their continuation. Secondly, because liberal democracy is brought into practice by institutions whose functioning heavily depends upon citizen participation. As Rahel Jaeggi asserts, institutions are not static structures – rather, they are dynamic organizations of practices and attitudes that are deeply embedded in the lives of the general population.[318] A good institution does not only facilitate citizens in self-realization but also in identifying with that institution. An ongoing process of internal critique and renewal is essential to keeping institutions alive. In short, the stability of liberal democracy is intertwined with civil trust and participation in its institutions.

Thus, we may posit that citizens influence the gradual evolution of social norms and institutional trust in their everyday political interactions and attitudes. For instance, they affect what behavior is tolerated in political discussions or called out. Related to this is Rawls' *publicity condition,* which entails that all reasons and principles underlying the basic structure of society are publicly accessible and justifiable to all citizens. The publicity condition consists of three levels: 1) the political conception of justice regulates society in practice, 2) the general beliefs relevant to the principles of justice are widely endorsed by the citizenry, and 3) its full justification is presented and publicly available *"in the public culture, reflected in its system of law and political institutions, and in the main historical traditions of their interpretation."*[319] In short, when the publicity condition is satisfied, all citizens share a sense of what liberal democracy entails and what it requires in return.[320]

Fulfilling the publicity condition would bring many benefits: besides increasing political transparency, it helps citizens to genuinely support the principles of justice and bring about stability for the right reasons. However, it is less clear how the publicity condition is maintained over time: publicity does not just entail that the relevant information is available, but also that its content is absorbed by a critical mass of citizens. Then, what if the share of citizens that is familiar with the foundation and justification of the principles of justice declines, and this awareness disappears from public culture? Without the publicity condition, stability for the right reasons devolves into a modus vivendi as more and more citizens abide by the liberal order out of habit instead of principle. Rawls acknowledges this: *"[S]ome will not want to carry philosophical reflection about political life so far."*

Yet, he immediately concedes that *"certainly, no one is required to."*[321] Contra Rawls, I think that upholding the publicity condition over time requires a more assertive approach. I have established earlier that liberal democracy can easily be taken for granted due to the nature of its benefits.[322] But moreover, in the information age it is no longer sufficient for knowledge material to merely be available – instead, the relevant

---

[317] As already discussed in §6.2.3 in this thesis project.
[318] Jaeggi, "What is a (good) institution?", p. 1-13.
[319] Rawls, *Political Liberalism*, p. 66-67.
[320] Ibid., p. 71: *"To realize the full publicity condition is to realize a social world within which the ideal of citizenship can be learned and may elicit an effective desire to be that kind of person."*
[321] Ibid., p. 67.
[322] See §5.1 of this thesis project.

messages about the principles of justice need to 'cut through the noise' of the information overload that is ubiquitous in the modern era.[323] Conversely, when the political conception of justice is not publicly known but merely publicly available, this gives rise to the unreasonable subcategory of unaware citizens.

9.2 Unaware Citizens & Street Epistemology

Unaware citizens are considered unreasonable because even though they abide by the criterion of reciprocity and accept the burdens of judgment, they do not fulfill the duty of prudence and thus fail to comply with the epistemic elements of being reasonable. This attitude can manifest in several ways. For instance, they may be unwilling to overcome their personal prejudices or to address an informational gap on the issue at hand. Alternatively, they may be oblivious that they hold others to a more strict standard of rationality than themselves. Nevertheless, although unaware citizens do not reason conscientiously, they generally do not regard themselves as being hostile to liberal democracy as such. Rather than actively pursuing unreasonable goals, they instead reason from reasonable intentions by unreasonable means, which is what may cause them to inadvertently support unreasonable movements. This does not merely challenge the overlapping consensus, but it also makes it impossible for them to endorse stability for the right reasons. Since the responsibility to uphold this stability falls on reasonable citizens, how can they engage constructively with unaware citizens?

Attempts to persuade the unaware towards reasonableness face several pitfalls. Firstly, when somebody does not fulfill the duty of prudence, they are more prone to engaging in motivated reasoning.[324] Therefore, merely bringing certain information to the attention of the unaware is unlikely to help. Secondly, simply communicating one's own process of reasoning will probably also be insufficient. The unaware interlocutor cannot merely copy others' reasoning to become reasonable; instead, they must become capable of reasoning by themselves. Thirdly, it does not suffice to convince the unaware interlocutor on one particular topic: their stance on any specific issue is far less important than their underlying attitude concerning the epistemic elements of reasoning about moral-political matters. Recognizing these pitfalls, I contend that an effective approach may circumvent them by focusing on epistemology.

In developing a strategy to engage with unaware citizens, I take particular inspiration from 'street epistemology'.[325] This conversational technique aims to explore through respectful dialogue the reasons and methods behind the interlocutor's views, especially controversial beliefs. Its focus is not on any specific belief itself, but rather on how one arrives at that belief. In doing so, the "why" and "how" is prioritized over the "what". Ultimately, the goal is not necessarily to change someone's belief directly, but to encourage critical thinking about the reliability and justification of the process by which those beliefs are formed. Ideally, street epistemology helps people identify unreliable epistemic methods and attitudes, and move towards more reliable and evidence-based ways of forming beliefs.

Yet, street epistemology comes with several pitfalls of its own. One may become overly focused on getting the unaware interlocutor to question their beliefs, rather than also

---

[323] See: Marques & Batista, *Information And Communication Overload In The Digital Age*.
[324] See §2.3 of this thesis project.
[325] Hümmler & Schiesser, *Fact and Prejudice*, p. 55.

modeling good epistemology oneself – the technique risks becoming insincere if one does not keep an open mind themselves. Also, an insistence on asking probing questions may end up undermining the symmetry of the conversation and devolve into a one-sided critique. Instead, the conversation ought to build a dialogue where both participants are involved as open-minded equals.[326]

With these caveats in mind, I think that street epistemology can inspire attempts to make unaware citizens more reasonable in four specific ways. Firstly, by promoting epistemic responsibility: a strong emphasis ought to be placed on the responsibility of individuals to justify their beliefs rationally to foster a greater sense of epistemic duty and prudence. Secondly, by modeling good epistemic practices. If the reasonable interlocutor demonstrates intellectual humility, being open to new evidence, and conscientious reasoning, then this provides the unaware with an inspirational example of prudent reasoning. Thirdly, by fostering mutual understanding – for the goal should be creating a respectful and non-confrontational dialogue, which can help build trust and reduce resistance to new perspectives. Fourthly and crucially, by identifying and addressing epistemic gaps in knowledge or reasoning that cause the unaware to be classified as such. If reasonable citizens incorporate these elements into their interactions with unaware individuals, they can more successfully promote a conscientious and reflective approach to reasoning in the context of public moral-political deliberation.

9.3 Duty of Vigilance

So far, I have established that achieving and maintaining stability for the right reasons requires additional effort on the part of reasonable citizens if they want to successfully counteract the undermining effects of the ever-present unreasonable elements in society. I have also elaborated on why I think that a focus on epistemology would help to persuade unaware citizens toward reasonableness. Now, I will argue that the virtue of reasonableness should include an additional duty: the *duty of vigilance*. It describes the political attitude where a citizen does not only propose and abide by fair terms of cooperation, but also strives to safeguard, maintain, and bolster them. As an added moral responsibility, it addresses three needs. Firstly, the need to uphold social norms that underpin political liberalism and to provide community pushback if a citizen flaunts these standards. Secondly, the need to foster epistemic responsibility so that citizens with reasonable intentions are not led astray by erroneous reasoning. Thirdly, the need to cultivate awareness of those benefits of liberal democracy that otherwise would have been taken for granted. If the duty of vigilance becomes part of the public political culture of the citizenry, I contend that it can address these three needs and will consequently considerably improve the plausibility of stability as stability for the right reasons.

In general, the duty of vigilance consists of a moral duty in everyday political interactions to point certain things out: when somebody's behavior undermines the possibility for reasonable discussion, when somebody's justification seems unconvincing or confusing, or when everybody seems to take a fundamental benefit

---

[326] Another caveat is that the original developer, Peter Boghossian, deployed street epistemology exclusively to dissuade his interlocutors from adhering to religious comprehensive doctrines – see Boghossian, *Manual for Atheists*. This practice is in clear violation of the burdens of judgment and incompatible with political liberalism. I reject this application of the method.

for granted. As an ideal, it entails not merely improving our understanding of society and of others and of oneself, but also acting upon that understanding in such a way that these insights are practiced and preserved and passed on. The duty of vigilance is reactive, in the sense that it is always preceded by extensive listening, observation, and reflection; responsive, in the sense that the validity of what is pointed out depends on feedback from the community in that context; and receptive, in the sense that its practitioner is willing to consider that their warnings maybe apply the most to themselves. This duty cannot be fulfilled in the usual sense – rather, it is an ongoing commitment to actively and conscientiously engage in reasonable discourse.

In particular, there are three ways in which the duty of vigilance can contribute to the public political culture in a liberal society. Firstly, by giving participants in deliberation a way to hold each other accountable, calling out their acquaintances when their behavior falls short of the duty of civility or when their ideas are in tension with the basic aspects of unreasonableness. According to Jane Mansbridge's analysis of ordinary political exchanges, regular citizens are already wont to calling out people in their close social circle like this, often using rhetorical shorthands to respond to inappropriate statements.[327] The duty of vigilance is meant as an affirmation of these habits. The intervening comments are not necessarily reprimands but rather reminders: assuming reasonable intent whenever plausible, the interlocutor should still be notified when they come across in an unfortunate way. The central idea is that all citizens take on the joint role of moderator and contribute to the quality of their discussion.

Secondly, by fostering epistemic responsibility as reasonable citizens engage in respectful dialogue with their unreasonable peers while focusing on the core values and epistemological foundations of their political views. Instead of directly criticizing the interlocutor on some specific belief, relying on why-questions can both improve mutual understanding and also uncover the epistemic gaps in knowledge or reasoning that are the cause of an unaware political attitude. Translated to practice, the duty of vigilance encourages oneself and the interlocutor to seek out reliable information, question unfounded assumptions, and remain open to revising moral-political views based on new evidence. Thus, it could act as a catalyst for a public political culture where individuals are mindful of the epistemic responsibilities of reasonableness, which would help to prevent unaware citizens from being led astray by erroneous unreasonable arguments.

Third, by creating common ground through an increased awareness of the benefits of liberal democracy. A citizen who cultivates a vigilant attitude can, if appropriate, actively highlight those advantages provided by a liberal democratic system that are otherwise overlooked by citizens at large. Making the invisible benefits visible could foster a greater appreciation for these achievements. By emphasizing the tangible benefits that everyone is already enjoying under a liberal democratic system, vigilant citizens may bridge social divides and perhaps stimulate a sense of unity. Once these valuable outcomes are more widely recognized, a common ground emerges that may serve as a foundation for further productive discussions, where even many who are initially unreasonable may find some points that they appreciate about the liberal

---

[327] Mansbridge, "Everyday Talk in the Deliberative System", p. 211-238.
Mansbridge & Flaster, "The Cultural Politics of Everyday Discourse", p. 627–660.

order, thereby growing over time to support the overlapping consensus for reasons from within their own doctrine.

For all of these reasons, I advocate that political liberalism should encourage reasonable citizens to adopt the duty of vigilance in order to develop a more effective strategy for achieving stability for the right reasons.

9.4 Rebutting Potential Criticisms

In light of this proposed additional duty for the reasonable and approach to the unaware, I will now defend the duty of vigilance against three possible objections.

The first potential objection is that the duty of vigilance is redundant as the existing duty of civility already covers its content and introducing another duty would unnecessarily complicate the conceptual framework of political liberalism.

However, I think such an assertion is mistaken. The duty of civility covers three things explicitly: explaining oneself in terms of public reason, a willingness to listen to other citizens, and accommodating their views when appropriate.[328] This does not overlap with the duty of vigilance, which is about proactively maintaining the social norms of liberal democracy and safeguarding fair terms of cooperation. Whereas the duty of civility is about being respectful in political interactions oneself, the duty of vigilance is about actively promoting the conditions that make such interactions possible.

The second potential objection is that the duty of vigilance is too demanding. It questions the feasibility and fairness of imposing upon citizens the additional responsibility of vigilance in maintaining democratic social norms.

However, I contend that the additional tasks entailed by the duty of vigilance lie actually not that far beyond the current responsibilities of being reasonable and that the duty of vigilance is merely a natural extension of them. Firstly, there is reason to believe that quite a few citizens are in practice already calling each other out when reasonable social norms are flouted. Secondly, focusing on epistemology in conversations with unaware citizens may actually decrease one's burdens by making the conversating more constructive. Thirdly, raising awareness about the benefits of democracy can be integrated into everyday political interactions without much extra effort. Besides, if more citizens share the burdens of the duty of vigilance and make it a collective responsibility, the individual burden will become easier to bear.

Nevertheless, I think biting the bullet is also acceptable here. All citizens carry some responsibility for maintaining liberal democracy, and this is perhaps simply quite demanding. Expecting citizens to preserve and promote democratic values proactively does entail an additional burden that some may experience as challenging. Yet, if one assumes the fact of pluralism, these burdens are insignificant to the burdens that result from the alternative, which is overcoming pluralism through (the fact of) oppression.

The third potential objection is that the duty of vigilance is vulnerable to overreach and becoming intrusive. The concern is that encouraging citizens to actively engage with and correct the behaviors and beliefs of others might morph into a form of moral policing that infringes on personal autonomy.

---

[328] Rawls, *Political Liberalism*, p. 217.

In my view, this concern about potential overreach and intrusion can be a valid caveat in some contexts. However, I think it falls short as a critical objection against the duty of vigilance as such for several reasons. Firstly, I have framed the duty of vigilance as a supportive, respectful, and enlightening endeavor rather than a punitive one – it is primarily about promoting awareness and accountability. Secondly, the duty of vigilance operates within the virtuous attitude of reasonableness, which respects individual rights as it accepts the burdens of judgment and regards fellow citizens as free and equal. Thirdly, the duty of vigilance encourages individuals to uphold democratic principles not just in others but also in themselves. Moreover, the method of the duty of vigilance should always remain subordinate to the ultimate goal of fostering a culture of constructive public deliberation and mutual understanding.

9.5 Conclusion

In this ninth and final chapter, I have explored the pivotal role that reasonable citizens play in upholding the stability of a liberal society characterized by the right reasons. First, I have discussed the publicity condition and analyzed how citizens influence the gradual evolution of social norms and institutional trust through their everyday political interactions. Additionally, I have elaborated upon the subcategory of unreasonable citizens and why I think that the conversational technique of street epistemology can provide some insights into how to engage with them constructively and encourage the epistemic elements of reasonableness. Furthermore, I have argued in favor of expanding the concept of reasonableness with the additional duty of vigilance, which emphasizes the importance of not just proposing and honoring but also actively preserving fair terms of cooperation between citizens. Finally, I have addressed some potential objections against my suggested strategy. In short, I maintain that cultivating civil vigilance can improve the quality of public deliberation, encourage epistemic responsibility among unaware citizens, and reinforce stability for the right reasons. In this way, political liberalism can alleviate the stability challenge originally posed by the trilemma.

Over the course of Part III of this thesis project, I have argued several things. First, I introduced a new typology of the unreasonable that divides them into four subcategories. Furthermore, I have argued in favor of three new strategies that each engage with a different kind of unreasonable citizen: the approach of *involvement* concerning uncooperative and unequitable citizens, the *intelligibility*-justification concerning anti-reasonable citizens, and the duty of *vigilance* concerning unaware citizens. Additionally, I have developed in the three preceding chapters separate responses to each particular challenge (inclusivity, legitimacy, and stability) revealed by the trilemma of pluralism. Now it is time for the final conclusion and to integrate these insights into a comprehensive strategy that may achieve successful containment of unreasonable doctrines in a liberal society.

This page has been intentionally left blank

Conclusion

Throughout this thesis project, I have advanced a critical re-evaluation of Rawls' political theory. In particular, I have developed a philosophical analysis of how well political liberalism can contain unreasonable influences in a pluralistic society governed by a liberal order. Accordingly, I have focused on the central inquiry:

How should a liberal order navigating the fact of simple pluralism achieve the containment of unreasonableness while remaining normatively consistent with political liberalism?

In response, I contend that a liberal order should develop a multifaceted strategy that balances the core ideals of inclusivity, legitimacy, and stability. This requires a refined understanding of unreasonableness that distinguishes among different types of unreasonable citizens and tailors its approaches accordingly. Crucial steps include enhancing inclusivity by involving partially reasonable citizens, reinforcing legitimacy by excluding anti-reasonable citizens from public discourse, and bolstering stability by cultivating vigilance among reasonable citizens. When combined with existing approaches such as community engagement, civic education, and preventive intervention, the result is a comprehensive strategy that allows a liberal order to manage unreasonable influences without compromising its foundational principles.

I began by examining Rawls' concept of reasonableness and argued that its epistemic elements form a third basic aspect that I call the *duty of prudence.* Then, I argued that the recent state of moral psychological research suggests that Rawls should not postulate the fact of reasonable pluralism, but is restricted to the more modest fact of *simple pluralism* that features an enduring presence of unreasonable citizens. This creates tension between the three core ideals of political liberalism, which I modeled as a *trilemma* to indicate how seemingly separate challenges concerning unreasonable citizens are interconnected. In order to restore a cohesive harmony among its core principles, I advocate an approach of reflective equilibrium for political liberalism.

Next, I have provided a general overview of the current state of liberal political philosophy on the topic of unreasonableness and delved into the responses of conciliation, transformation, or marginalization. I found that *conciliation*, which attempts to placate unreasonable citizens, remains largely ineffective. Whenever *transformation* seems effective, it also struggles with some substantial normative objections. Lastly, I judged that *marginalization* of the unreasonable – e.g. restricting some of their basic rights – is necessary in emergencies but unsuitable as a primary approach. Overall, these current approaches are unable to overcome the trilemma.

Hence, I developed several insights that contribute to resolving it. First, I argued that the binary reasonable/unreasonable model is reductive, and instead introduced a typology that makes a fourfold distinction between subcategories of unreasonable citizens. Then, I advanced three separate responses to these distinct subcategories. Concerning partially reasonable citizens who reason conscientiously, I advocate a policy of *involvement.* When citizens are wholly unreasonable, I contend that concerns of *intelligibility* provide an adequate reason for their regrettable exclusion from deliberations about legitimacy. Finally, I propose that reasonable citizens should embrace a new moral responsibility in the *duty of vigilance* to bolster stability for the right reasons. Apart from contributing to the preservation of those social norms that

make civil discussion possible, this will help foster a greater sense of epistemic responsibility among unaware citizens. When considered together, these responses have the potential to alleviate each challenge posed by the trilemma, whether it concerns inclusivity, legitimacy, or stability. Now, the task is to integrate them with all other promising responses into a unified strategy for containment of the unreasonable.

Comprehensive containment of unreasonable doctrines necessitates a more nuanced understanding of unreasonable citizens. This new view is illustrated in Figure 1.

*Figure 1. Unreasonableness Reconceptualized*
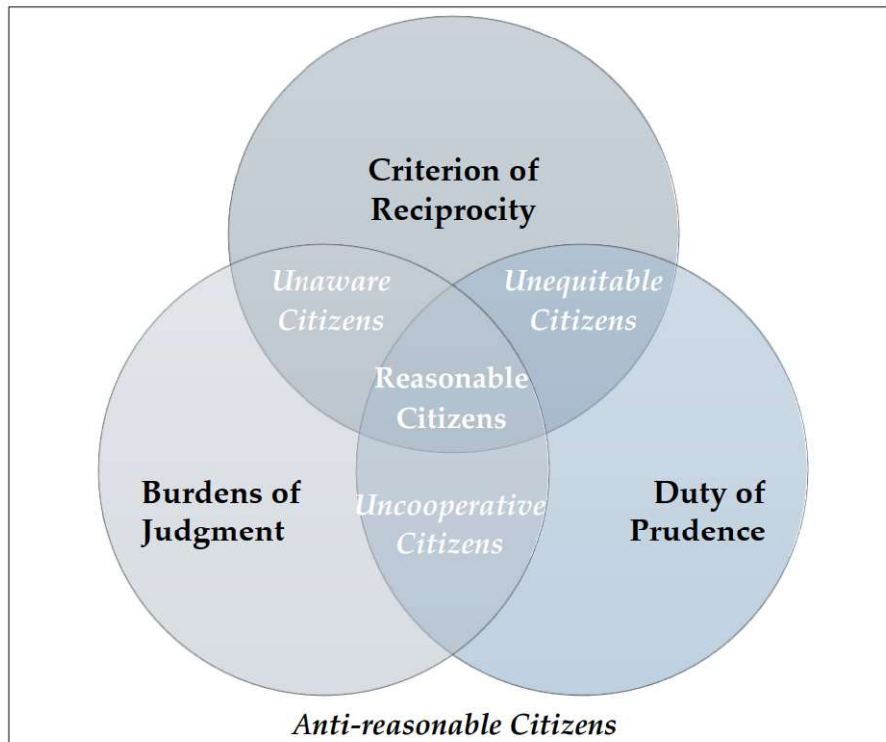
In this diagram, the blue circles express the three basic aspects of being reasonable. When they overlap completely, this sphere is the realm of reasonable citizens. When they overlap only partially, these spheres indicate the subcategories of unreasonable citizens that could over time grow to become willing supporters of the overlapping consensus. Conversely, anti-reasonable citizens are placed beyond the reasonable framework. They pose a constant challenge to reasonableness since their unavoidable presence acts as a continuous force in an outward direction, pulling the circles apart and adrift. As their political influence increases, this force grows and the spheres have less space in common, causing the space for reasonableness to shrink. Hence, the task for the liberal order and those who support it is to endeavor to make the circles converge as much as possible, progressing towards a compelling and enduring majority of reasonable citizens.

However, in consolidating its foundation of reasonable citizens, the liberal order faces both an empirical constraint in the form of simple pluralism and a normative constraint in its principles as expressed by political liberalism. Thus, the liberal order must navigate a trilemma which is illustrated in Figure 2.

Figure 2. Visualization of the Trilemma
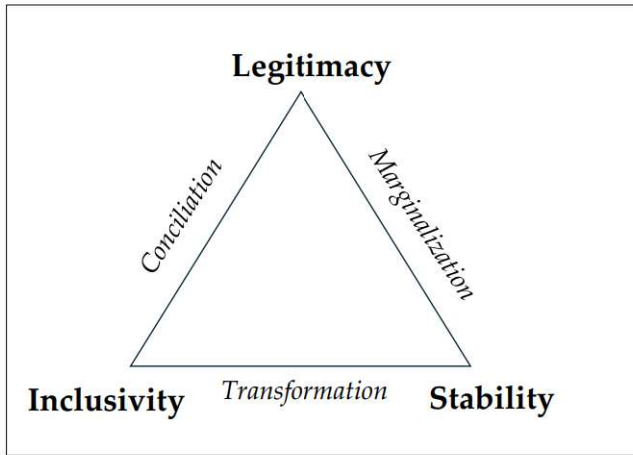
This figure illustrates the difficulties in realizing the three core ideals of political liberalism simultaneously. First, an approach of *conciliation* makes concessions to the unreasonable and incorporates them into the ranks of its constituency of public justification. However, since the unreasonable are generally persistent in their political attitude, stability for the right reasons becomes impossible – a mere modus vivendi is the best scenario. Second, an approach of *transformation* imposes extensive educational programs and nudging initiatives upon its citizens. However, since it often cannot provide adequate public reasons for these measures to its constituency of public justification, this approach would compromise legitimacy. An approach of *marginalization* restricts in exceptional circumstances some basic rights of the unreasonable, such as freedom of association. However, such policies create significant tension with the ideal of inclusivity and regarding all citizens as free and equal. Given these challenges, resolving the trilemma requires a comprehensive strategy where approaches are tailored to the specific context of that situation. When the space for reasonableness shrinks, implementing more severe measures becomes warranted on top of the policies that are already in place. This is illustrated in Figure 3.

Figure 3. Appropriate Containment Strategies As Sphere of Reasonableness Decreases
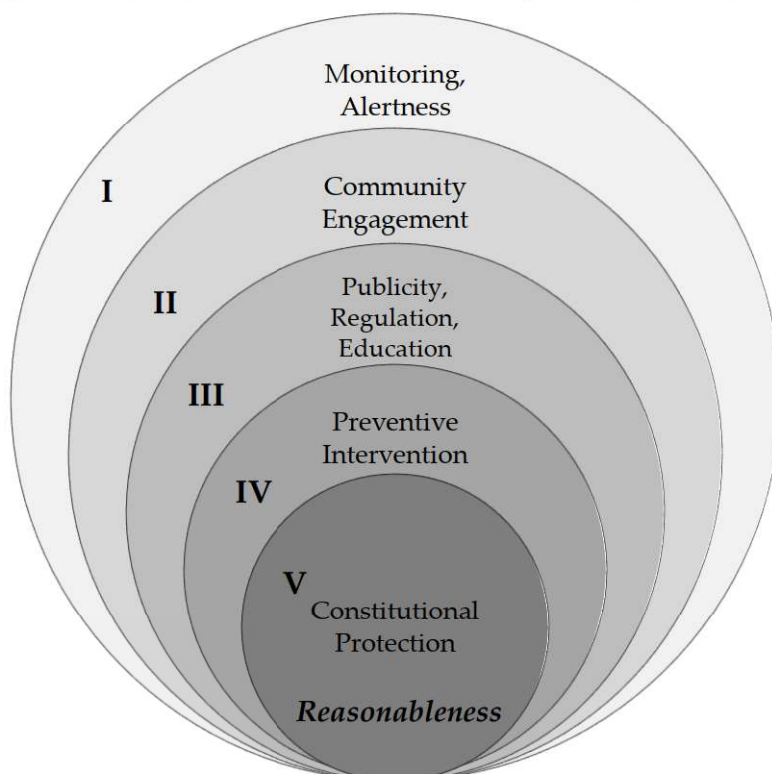
Since the level of unreasonableness in a liberal society can vary in degree between absent and mainstream, additional containment strategies should be implemented as the unreasonable threat grows. When unreasonable doctrines seem to have disappeared from society (see sphere I in Figure 3), the responsibilities of the liberal order still include alertness and monitoring, because unreasonableness can always resurface. When unreasonable influences are present in society but remain at its margins (sphere II), the liberal order should ensure that these groups in question are receiving the fundamental benefits of liberal democracy, especially social and economic opportunity. This replaces engagement via conjecture, which is best delegated to the leaders of the communities in question. When the unreasonable presence in society grows into a significant minority (III), more assertive measures come into play. Institutions should strive to uphold the publicity condition and also incorporate mechanisms for civil participation to garner support. If the rise of an unreasonable movement is tied to a social technology, strict regulation of that technology should be considered. Moreover, the liberal order should explore whether its education system can increase its emphasis on instilling in its students the fundamental values of freedom, equality, and fairness in a diverse teaching environment.

When the share of unreasonable citizens approaches a critical mass (IV), preventive intervention is warranted. Measures may include cracking down on hateful or seditious speech, dissolving public rallies of unreasonable movements, and restricting their funding – provided that these interventions act upon substantial proof and receive continuous legal scrutiny. If the unreasonable movement becomes widespread and passes the tipping point (V), the liberal order must prioritize protecting its constitution and judiciary. But regardless of the success of these defensive efforts, stability for the right reasons has disappeared and the liberal order must operate within a modus vivendi. In that case, political liberalism must begin anew, attempting to construct a more resilient conception of justice in that historical context. The goal of this roadmap is to avoid that outcome and guide the liberal order in navigating the challenges of the unreasonable.

Complementary to this range of measures is a duty of vigilance among ordinary reasonable citizens, who can engage with the unreasonable as follows. The starting assumption in public deliberation is that all participants are reasonable and operate in good faith. The first step when an interlocutor comes across as unreasonable is that reasonable citizens call this out in a friendly spirit, carrying the joint responsibility of moderating the public discussion. The second step is to show curiosity toward the perspective of the interlocutor and probe why they reason this way. This will clarify the political attitude of the interlocutor. If they are reasonable, the situation was merely a misunderstanding. If they are unaware, the epistemic gap in their reasoning can be identified. If they are uncooperative or unequitable, an approach of involvement is appropriate. If they are anti-reasonable, discussion in good faith becomes impossible and the interlocutor must regrettably be excluded from the deliberative process for the time being. Still, the reason for this exclusion should be communicated clearly: their unreasonableness renders them unintelligible in this context and community. The third step is to emphasize the benefits of liberal democracy, creating common ground.

Together, these general principles of a comprehensive strategy set out how a liberal order could contain unreasonableness and remain normatively consistent with political liberalism by tailoring inclusivity, clarifying legitimacy, and deputizing stability.

This thesis project has several limitations. First, while unreasonable citizens are an important factor in contemporary liberal democratic decline, they are not the sole factor. Other perspectives, such as institutional and socio-economic influences, should also be taken into account. Second, the efficacy of the proposed containment strategies remains unknown. Their full validation requires further empirical evidence. Third, political philosophy could benefit from additional interdisciplinary insights beyond psychology, such as those from law and communication studies, to enrich our understanding of unreasonableness and refine containment strategies.

There are also some remaining tensions within the framework of this thesis project. For instance, regarding education, political liberalism may not be sufficiently distinct from comprehensive liberalism. I am skeptical that this debate can be settled among reasonable interlocutors in the foreseeable future; therefore, I recommend that future containment strategies avoid emphasizing education as their main focus. Additionally, my proposals for involvement and vigilance are formulated in general terms that are unavoidably underdetermined. Clarifying these concepts requires an account that can also develop their practical implementation.

The findings of this thesis project have several significant implications for liberal political philosophy. Most importantly, Political liberalism should postulate simple instead of reasonable pluralism, reconceptualize reasonableness to include the duty of prudence, and adopt a multifaceted account of unreasonableness. There are also notable practical implications for political liberals. For example, attempts to engage with unreasonable citizens by focusing on doctrine are inadvisable. This also applies to initiatives that endeavor to cultivate liberal democratic attitudes in new generations by focusing on skills and knowledge instead of values and attitudes.

Finally, this thesis project points toward some promising areas for further philosophical inquiry. For instance, the relationship between social technology and liberal democracy is currently underexplored. Additionally, it remains unclear how the epistemic foundations of reasonableness intersect with democratic stability when public opinion diverges significantly from scientific consensus, as seen in discussions on climate change. Moreover, further research is warranted to determine conversational techniques that promote constructive dialogue with unreasonable citizens while remaining true to fundamental liberal democratic values.

In conclusion, while I acknowledge that enduring unreasonableness poses a serious challenge to liberal democracy, I hope that enacting the comprehensive strategy of containment set out in this thesis project increases the possibility that this challenge will be overcome and that reasonableness prevails. But, to give Rawls the final word:

"That is the hope; there can be no guarantee."[329]

---

[329] Rawls, *Political Liberalism*, p. 65.

Bibliography

Rawls

Rawls, John. "The Domain of the Political and Overlapping Consensus." *New York University Law Review* 64, no. 2 (1989): 233-255.

---, *A Theory of Justice: Revised Edition.* Cambridge: Harvard University Press, 1999.

---, *The Law of Peoples.* Cambridge: Harvard University Press, 2000.

---, *Justice as Fairness: A Restatement,* edited by Erin Kelly. Cambridge: Harvard University Press, 2001.

---, *Political Liberalism: Expanded Edition.* Columbia Classics in Philosophy. New York: Columbia University Press, 2005.

Introduction

Clark, Tom. "John Rawls Revisited: Politics Behind The Veil." *Prospect,* July 19, 2023. Link: https://www.prospectmagazine.co.uk/culture/62182/john-rawls-revisited-daniel-chandler-review.

Crothers, Lane, and Grace Burgener. "Insurrectionary Populism? Assessing the January 6 Attack on the U.S. Capitol." *Populism* 4, no. 2 (October 1, 2021): 129–45. Link: https://doi.org/10.1163/25888072-bja10025.

Douglas Rasmussen and Douglas Den Uyl. "Did Rawls Restore Political Philosophy?" *Law&Liberty,* April 14, 2021. Link: https://lawliberty.org/forum/rawls-and-the-restoration-of-political-philosophy/.

Erin Kelly and Lionel McPherson. "On Tolerating the Unreasonable." *The Journal of Political Philosophy* 9, no. 1 (2001): 38–55.

Forrester, Katrina. "The Future of Political Philosophy." *Boston Review,* September 17, 2019. Link: https://www.bostonreview.net/articles/katrina-forrester-future-political-philosophy/.

Freedman, Lawrence. "The Crisis of Liberalism and the Western Alliance." In *Survival: Trials of Liberalism,* 37–43. Routledge, 2022.

Gursozlu, Fuat. "Political Liberalism and the Fate of Unreasonable People." *Touro Law Review* 30 (2014): 35.

Levitsky, Steven, and Daniel Ziblatt. *How Democracies Die.* New York: Crown Publishing Group, 2018.

Quong, Jonathan. *Liberalism Without Perfection.* Oxford: Oxford University Press, 2011.

Robert Nozick. *Anarchy, State, and Utopia.* New York: Basic Books, 1974. Link: https://doi.org/10.1515/9780773585669-007.

Spinoza, Baruch. *Theologico-Political Treatise.* Edited by Martin Yaffe. First Edition. Newburyport: Focus Publishing, 2004.

Part I

Pogge, Thomas. *John Rawls: His Life and Theory of Justice.* Oxford: Oxford University Press, 2007.

Chapter 1

Huddy, Leonie, David O. Sears, and Jack S. Levy. "Introduction: Theoretical Foundations of Political Psychology." In *The Oxford Handbook of Political Psychology, 2nd Ed*, 1–19. New York, NY, US: Oxford University Press, 2013.

Legal Information Institute. "Colorable Claim." *Cornell Law School*. Accessed June 16, 2024. Link: https://www.law.cornell.edu/wex/colorable_claim.

Quist, Ryan M., and Miriam G. Resendez. "Social Dominance Threat: Examining Social Dominance Theory's Explanation of Prejudice as Legitimizing Myths." *Basic and Applied Social Psychology* 24, no. 4 (December 1, 2002): 287–93. Link: https://doi.org/10.1207/S15324834BASP2404_4.

Talisse, Robert B. *On Rawls: A Liberal Theory of Justice and Justification*. Wadsworth Philosophers Series. Belmont, CA: Wadsworth, Thomson Learning, 2001.

Valentini, Laura. "Ideal vs. Non-Ideal Theory: A Conceptual Map." *Philosophy Compass* 7, no. 9 (2012): 654–64. Link: https://doi.org/10.1111/j.1747-9991.2012.00500.x.

Wenar, Leif. "John Rawls." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2021. Metaphysics Research Lab, Stanford University, 2021. Link: https://plato.stanford.edu/archives/sum2021/entries/rawls/.

Chapter 2

Altemeyer, Robert. *The Authoritarians*. New Jersey: Cherry Hill Publishing, 2009. Available free online at https://theauthoritarians.org

Balliet, Daniel, Junhui Wu, and Carsten K. W. De Dreu. "Ingroup Favoritism in Cooperation: A Meta-Analysis." *Psychological Bulletin* 140, no. 6 (2014): 1556–81. https://doi.org/10.1037/a0037737.

Baumgartner Jennifer N., and G. Scott Morgan. "Mindfulness and Cognitive Depletion Shape the Relationship between Moral Conviction and Intolerance of Dissimilar Others." *Studia Psychologica* 61, no. 1 (2019): 31–41. https://doi.org/10.21909/sp.2019.01.770.

Bavel, Jay Van, and Dominic J. Packer. *The Power of Us: Harnessing Our Shared Identities for Personal and Collective Success*. London: Hachette UK, 2021.

Bavel, Jay Van, and William A. Cunningham. "Self-Categorization With a Novel Mixed-Race Group Moderates Automatic Social and Racial Biases." *Personality and Social Psychology Bulletin* 35, no. 3 (March 2009): 321–35. https://doi.org/10.1177/0146167208327743.

Brown, Rupert. "The Origins of the Minimal Group Paradigm." *History of Psychology* 23, no. 4 (November 2020): 371–82. https://doi.org/10.1037/hop0000164.

Buchanan, Allen. *Our Moral Fate: Evolution and the Escape from Tribalism*. Boston: MIT Press, 2020.

Cole Wright, Jennifer, Jerry Cullum, and Nicholas Schwab. "The Cognitive and Affective Dimensions of Moral Conviction: Implications for Attitudinal and Behavioral Measures of Interpersonal Tolerance." *Personality and Social Psychology Bulletin* 34, no. 11 (November 1, 2008): 1461–76. https://doi.org/10.1177/0146167208322557.

Ditto, Peter H., David A. Pizarro, and David Tannenbaum. "Chapter 10 Motivated Moral Reasoning." In *Psychology of Learning and Motivation*, 50:307–38. Elsevier, 2009. https://doi.org/10.1016/S0079-7421(08)00410-6.

Dunham, Yarrow. "Mere Membership." *Trends in Cognitive Sciences* 22, no. 9 (September 2018): 780–93. https://doi.org/10.1016/j.tics.2018.06.004.

Fischer, Ronald, and Crysta Derham. "Is In-Group Bias Culture-Dependent? A Meta-Analysis across 18 Societies." *SpringerPlus* 5, no. 1 (December 2016): 70. https://doi.org/10.1186/s40064-015-1663-6.

Hewstone, Miles, Mark Rubin, and Hazel Willis. "Intergroup Bias." *Annual Review of Psychology* 53, no. 1 (February 2002): 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109.

Irwin, John F., and Daniel L. Real. "Unconscious Influences on Judicial Decision-Making: The Illusion of Objectivity." *McGeorge Law Review* 42 (2010): 1.

Kahan, Dan M. "The Politically Motivated Reasoning Paradigm." SSRN Scholarly Paper. Rochester, NY, December 14, 2015. https://papers.ssrn.com/abstract=2703011.

Kunda, Ziva. "The Case for Motivated Reasoning." *Psychological Bulletin* 108, no. 3 (1990): 480–98. https://doi.org/10.1037/0033-2909.108.3.480.

Lodge, Milton, and Charles S. Taber. "The Rationalizing Voter: Unconscious Thought in Political Information Processing." *Social Science Research Network*, December 21, 2007. Available at https://ssrn.com/abstract=1077972 or http://dx.doi.org/10.2139/ssrn.1077972

Venaglia, Rachel and Maxwell, Laura. "How We Conducted Our International Study on Right-Wing Authoritarianism." *Morning Consult,* June 28, 2021. Link: https://pro.morningconsult.com/articles/right-wing-authoritarianism-international-study-methodology.

Mueller, Allison B., and Linda J. Skitka. "Liars, Damned Liars, and Zealots: The Effect of Moral Mandates on Transgressive Advocacy Acceptance." *Social Psychological and Personality Science* 9, no. 6 (August 2018): 711–18. https://doi.org/10.1177/1948550617720272.

Mullen, Brian, Rupert Brown, and Colleen Smith. "Ingroup Bias as a Function of Salience, Relevance, and Status: An Integration." *European Journal of Social Psychology* 22, no. 2 (March 1992): 103–22. https://doi.org/10.1002/ejsp.2420220202.

Nucci, Larry P. *Education in the Moral Domain*. Cambridge: Cambridge University Press, 2001.

Osborne, Danny, Thomas H. Costello, John Duckitt, and Chris G. Sibley. "The Psychological Causes and Societal Consequences of Authoritarianism." *Nature Reviews Psychology* 2, no. 4 (April 2023): 220–32. https://doi.org/10.1038/s44159-023-00161-4.

Otten, Sabine. "The Minimal Group Paradigm and Its Maximal Impact in Research on Social Categorization." *Current Opinion in Psychology* 11 (October 2016): 85–89. https://doi.org/10.1016/j.copsyc.2016.06.010.

Pronin, Emily, Daniel Y. Lin, and Lee Ross. "The Bias Blind Spot: Perceptions of Bias in Self Versus Others." *Personality and Social Psychology Bulletin* 28, no. 3 (March 2002): 369–81. https://doi.org/10.1177/0146167202286008.

Psychology Glossary. "Minimal Group Paradigm Definition." Alley Dog. Accessed June 16, 2024. Link: https://www.alleydog.com/glossary/definition.php?term=Minimal+Group+Paradigm.

Richerson, Peter J., and Robert Boyd. "The Evolution of Subjective Commitment to Groups: A Tribal Instincts Hypothesis." In *Evolution and the Capacity for Commitment*, 184–220. Volume III in the Russell Sage Foundation Series on Trust. New York, NY, US: Russell Sage Foundation, 2001.

Schmidt, Kathleen, and R. Grace Drake. "Minimal Group Procedures and Outcomes." *Collabra: Psychology* 9, no. 1 (December 21, 2023): 90187. https://doi.org/10.1525/collabra.90187.

Skitka, Linda J. "The Psychology of Moral Conviction." *Social and Personality Psychology Compass* 4, no. 4 (2010): 267–81. https://doi.org/10.1111/j.1751-9004.2010.00254.x.

Skitka, Linda J., and G. Scott Morgan. "The Social and Political Implications of Moral Conviction." *Political Psychology* 35, no. S1 (2014): 95–110. https://doi.org/10.1111/pops.12166.

Skitka, Linda J., G. Scott Morgan, and Daniel C. Wisneski. "Political Orientation and Moral Conviction: A Conservative Advantage or an Equal Opportunity Motivator of Political Engagement?" In *Social Psychology and Politics*, p. 57–74. Sydney Symposium of Social Psychology. New York: Psychology Press, 2015.

Westen, Drew, Pavel S. Blagov, Keith Harenski, Clint Kilts, and Stephan Hamann. "Neural Bases of Motivated Reasoning: An fMRI Study of Emotional Constraints on Partisan Political Judgment in the 2004 U.S. Presidential Election." *Journal of Cognitive Neuroscience* 18, no. 11 (2006): 1947–58. https://doi.org/10.1162/jocn.2006.18.11.1947.

Chapter 3

Bader, Veit, and Ewald R. Engelen. "Taking Pluralism Seriously: Arguing for an Institutional Turn in Political Philosophy." *Philosophy & Social Criticism* 29, no. 4 (July 1, 2003): 375–406. https://doi.org/10.1177/0191453703294002.

Barbara Walsh, Mary. "Private and Public Dilemmas: Rawls on the Family." *Polity* 44, no. 3 (July 1, 2012): 426–45. https://doi.org/10.1057/pol.2012.9.

Brennan, Jason. *Against Democracy*. Princeton: Princeton University Press, 2016. https://doi.org/10.1515/9781400882939.

Gunn, Paul. "Against Epistocracy." *Critical Review* 31, no. 1 (January 2, 2019): 26–82. https://doi.org/10.1080/08913811.2019.1609842.

Krick, Eva. "The Myth of Effective Veto Power under the Rule of Consensus. Dynamics and Democratic Legitimacy of Collective Decision-Making by 'Tacit Consent.'" *Négociations* 27, no. 1 (2017): 109–28. https://doi.org/10.3917/neg.027.0109.

Langvatn, Silje. "Legitimate, but Unjust; Just, but Illegitimate: Rawls on Political Legitimacy." *Philosophy and Social Criticism* 42 no. 2 (2015): 132-153. https://doi-org.proxy.library.uu.nl/10.1177/0191453715615386

Murphie, Andrew, and John Potts. *Culture and Technology*. London: Macmillan Education UK, 2003.

Myers, Todd. "Justifying Liberalism in Singapore." *American Institute for Economic Research.* Accessed June 16, 2024. https://www.aier.org/article/justifying-liberalism-in-singapore/.

Quong, Jonathan. *Liberalism Without Perfection*. Oxford: Oxford University Press, 2011.

Wright, Robert. *The Evolution Of God: The Origins of Our Beliefs*. New York: Hachette Book Group, 2010.

Part II

Daniel, Rhianwen. "The Liberal/Conservative Nationalism Divide: A Distinction without a Difference?" *Nations and Nationalism* 28, no. 2 (2022): 523–38. https://doi.org/10.1111/nana.12812.

Freeman, Samuel. *Rawls*. London: Routledge, 2007. https://doi.org/10.4324/9780203086605.

Hjerm, Mikael. "Defending Liberal Nationalism—At What Cost?" *Journal of Ethnic and Migration Studies* 30, no. 1 (January 1, 2004): 41–57. https://doi.org/10.1080/1369183032000170169.

Knops, Andrew. "Debate: Agonism as Deliberation – On Mouffe's Theory of Democracy." *Journal of Political Philosophy* 15, no. 1 (March 2007): 115–26. https://doi.org/10.1111/j.1467-9760.2007.00267.x.

Miller, David. *On Nationality*. Oxford Political Theory. Oxford: Oxford University Press, 1995.

Mouffe, Chantal. "Deliberative Democracy or Agonistic Pluralism?" *Social Research* 66, no. 3 (1999): 745–58.

Chapter 4

Ahmad, Hafal (Haval). "Youth De-Radicalization: A Canadian Framework." *Journal for Deradicalization*, no. 12 (September 14, 2017): 119–68.

Castaño-Pulgarín, Sergio Andrés, Natalia Suárez-Betancur, Luz Magnolia Tilano Vega, and Harvey Mauricio Herrera López. "Internet, Social Media, and Online Hate Speech. Systematic Review." *Aggression and Violent Behavior* 58 (May 1, 2021): 101608. https://doi.org/10.1016/j.avb.2021.101608.

Clayton, Matthew, and David Stevens. "When God Commands Disobedience: Political Liberalism and Unreasonable Religions." *Res Publica* 20, no. 1 (February 1, 2014): 65–84. https://doi.org/10.1007/s11158-013-9221-y.

Daniel Jacobsen, Mark. "Moral Dumbfounding and Moral Stupefaction." In *Oxford Studies in Normative Ethics: Volume 2*, edited by Mark Thomas, p. 289. Oxford: Oxford University Press, 2012.

Daniel, Rhianwen. "The Liberal/Conservative Nationalism Divide: A Distinction without a Difference?" *Nations and Nationalism* 28, no. 2 (2022): 523–38. https://doi.org/10.1111/nana.12812.

Doosje, Bertjan, Fathali Moghaddam, Arie Kruglanski, Arjan de Wolf, Liesbeth Mann, and Allard Feddes. "Terrorism, Radicalization and de-Radicalization." *Current Opinion in Psychology*, Intergroup Relations, 11 (October 1, 2016): 79–84. https://doi.org/10.1016/j.copsyc.2016.06.008.

Erin Kelly and Lionel McPherson. "On Tolerating the Unreasonable." *The Journal of Political Philosophy* 9, no. 1 (2001): 38–55.

Ferrara, Alessandro. *The Democratic Horizon: Hyperpluralism and the Renewal of Political Liberalism*. Cambridge: Cambridge University Press, 2014.

Freeman, Samuel. *Rawls*. London: Routledge, 2007. https://doi.org/10.4324/9780203086605.

Friedman, Marilyn Friedman, 'John Rawls and the Political Coercion of Unreasonable People', in *The Idea of a Political Liberalism: Essays on John Rawls*, edited by Victoria Davion and Clark Wolf. Oxford: Rowman & Littlefield, 2000.

Hinck, Robert S., Hayley Hawthorne, and Joshua Hawthorne. "Authoritarians Don't Deliberate: Cultivating Deliberation and Resisting Authoritarian Tools in an Age of Global Nationalism." *Journal of Deliberative Democracy* 14, no. 2 (December 14, 2018). https://doi.org/10.16997/jdd.310.

Hjerm, Mikael. "Defending Liberal Nationalism—at What Cost?" *Journal of Ethnic and Migration Studies* 30, no. 1 (January 1, 2004): 41–57. https://doi.org/10.1080/1369183032000170169.

Knops, Andrew. "Debate: Agonism as Deliberation – On Mouffe's Theory of Democracy." *Journal of Political Philosophy* 15, no. 1 (March 2007): 115–26. https://doi.org/10.1111/j.1467-9760.2007.00267.x.

March, Andrew F. *Islam and Liberal Citizenship: The Search for an Overlapping Consensus*. Oxford: Oxford University Press, 2011.

Miller, David. *On Nationality*. Oxford Political Theory. Oxford: Oxford University Press, 1995.

Mouffe, Chantal. "Deliberative Democracy or Agonistic Pluralism?" *Social Research* 66, no. 3 (1999): 745–58.

Napolitano, Giulia. "Conspiracy Theories and Evidential Self-Insulation." In *The Epistemology of Fake News*, edited by Sven Bernecker, Amy K. Flowerree, and Thomas Grundmann, p. 82–106. Oxford: Oxford University Press, 2021.

Preben Bertelsen. "Danish Preventive Measures and De-Radicalization Strategies: The Aarhus Model." *Panorama: Insights into Asian and European Affairs* 1, no. 241 (2015): 53.

Quong, Jonathan. *Liberalism Without Perfection*. Oxford: Oxford University Press, 2011.

Sara Canna, Jeff Day, and George Popp. "Common Characteristics of 'Successful' Deradicalization Programs of the Past." *Joint Staff DDGO*, February 1, 2020. https://apps.dtic.mil/sti/citations/AD1092721.

Schwartzman, Micah. "The Ethics of Reasoning from Conjecture*." *Journal of Moral Philosophy* 9, no. 4 (January 1, 2012): 521–44. https://doi.org/10.1163/174552412X628931.

Shabi, Rachel. "Deradicalising Britain: The Activists Turning Young Muslims Away from Extremism." *The Guardian*, March 18, 2016, sec. UK news.

https://www.theguardian.com/uk-news/2016/mar/18/deradicalising-britain-the-activists-turning-young-muslims-away-from-extremism.

Gabriele Badano and Alasia Nuti. "The Limits of Conjecture: Political Liberalism, Counter-Radicalisation and Unreasonable Religious Views." *Ethnicities* 20 no. 2 (2020): 293-311.

Walter, Nathan, Jonathan Cohen, R. Lance Holbert, and Yasmin Morag. "Fact-Checking: A Meta-Analysis of What Works and for Whom." *Political Communication* 37, no. 3 (May 3, 2020): 350–75. https://doi.org/10.1080/10584609.2019.1668894.

Chapter 5

Abramowitz, Alan I. *Disappearing Center: Engaged Citizens, Polarization, and American Democracy*. Illustrated edition. Yale: Yale University Press, 2011.

Anderson, Joel. "Scaffolding and Autonomy." In *The Routledge Handbook of Autonomy*, edited by Bel Colburn, p. 158-166. London: Routledge, Taylor & Francis Group, 2022.

Arocena, Felipe, Sebastián Sansone, and Nicolás Alvarez. "Technological Disruption and Democracy in the Twenty-First Century." *European Journal of Futures Research* 10, no. 1 (March 27, 2022): 3. https://doi.org/10.1186/s40309-022-00189-4.

Bader, Veit. *Secularism or Democracy?: Associational Governance of Religious Diversity*. Amsterdam University Press, 2007. https://doi.org/10.5117/9789053569993.

Baekgaard, Martin, Julian Christensen, Casper Mondrup Dahlmann, Asbjørn Mathiasen, and Niels Bjørn Grund Petersen. "The Role of Evidence in Politics: Motivated Reasoning and Persuasion among Politicians." *British Journal of Political Science* 49, no. 3 (July 2019): 1117–40. https://doi.org/10.1017/S0007123417000084.

Brennan, Jason. "Giving Epistocracy a Fair Hearing." *Inquiry* 65, no. 1 (January 2, 2022): 35–49. https://doi.org/10.1080/0020174X.2019.1663020.

Caitlin Drummond and Baruch Fischhoff. "Individuals with Greater Science Literacy and Education Have More Polarized Beliefs on Controversial Science Topics." *Psychological and Cognitive Sciences* 114, no. 36 (August 21, 2017): 87–92. https://doi-org.proxy.library.uu.nl/10.1073/pnas.170488211.

Callan, Eamonn. *Creating Citizens: Political Education and Liberal Democracy*. Oxford: Clarendon Press, 1997.

———. "Political Liberalism and Political Education." *The Review of Politics* 58, no. 1 (January 1996): 5–33. https://doi.org/10.1017/S0034670500051627.

Čavojová, Vladimíra, Jakub Šrol, and Magdalena Adamus. "My Point Is Valid, Yours Is Not: Myside Bias in Reasoning about Abortion." *Journal of Cognitive Psychology* 30, no. 7 (October 3, 2018): 656–69. https://doi.org/10.1080/20445911.2018.1518961.

Charles Marohn. *Strong Towns: A Bottom-Up Revolution to Rebuild American Prosperity*. New York: John Wiley & Sons, 2019.

Costa, M. Victoria. *Rawls, Citizenship, and Education*. New York: Routledge, 2010. https://doi.org/10.4324/9780203846629.

Dagger, Richard. *Civic Virtues: Rights, Citizenship, and Republican Liberalism*. Oxford: Oxford University Press, 1997.

Deng, Zongyi. "Bringing Knowledge Back In: Perspectives from Liberal Education." *Cambridge Journal of Education* 48, no. 3 (May 4, 2018): 335–51. https://doi.org/10.1080/0305764X.2017.1330874.

Diana Ionescu. "Lonely by Design: How Urban Planning Can Intensify Social Isolation." Planetizen, January 30, 2023. Link: https://www.planetizen.com/news/2023/01/121328-lonely-design-how-urban-planning-can-intensify-social-isolation.

Easterbrook, Gregg. *The Progress Paradox: How Life Gets Better While People Feel Worse*. 4th edition. New York: Random House Publishing Group, 2004.

Estlund, David. "Why Not Epistocracy?". In *Desire, Identity and Existence*, edited by Naomi Reshotko. Kelowna: Academic Printing and Publishing (2003), 53-69.

Fernández, Christian, and Mikael Sundström. "Citizenship Education and Liberalism: A State of the Debate Analysis 1990–2010." *Studies in Philosophy and Education* 30, no. 4 (July 1, 2011): 363–84. https://doi.org/10.1007/s11217-011-9237-8.

Gunn, Paul. "Against Epistocracy." *Critical Review* 31, no. 1 (January 2, 2019): 26–82. https://doi.org/10.1080/08913811.2019.1609842.

Gursozlu, Fuat. "Political Liberalism and the Fate of Unreasonable People." *Touro Law Review* 30 (2014): 35.

Harari, Yuval Noah. "Why Technology Favors Tyranny." *The Atlantic*, August 30, 2018. https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/.

Kahan, Dan M. "Ideology, Motivated Reasoning, and Cognitive Reflection." *Judgment and Decision Making* 8, no. 4 (July 2013): 407–24. https://doi.org/10.1017/S1930297500005271.

Kahan, Dan M., Ellen Peters, Erica Cantrell Dawson, and Paul Slovic. "Motivated Numeracy and Enlightened Self-Government." *Behavioural Public Policy* 1, no. 1 (May 2017): 54–86. https://doi.org/10.1017/bpp.2016.2.

Keith, Trevor, and Vicente del Rio. "New Urbanism, Automobile Dependency and Sense of Community: A Comparative Study of Two Residential Developments in California." *Focus* 1, no. 1 (April 1, 2004). https://doi.org/10.15368/focus.2004v1n1.11.

Klein, Ezra. "How Politics Makes Us Stupid." *Vox*, April 6, 2014. https://www.vox.com/2014/4/6/5556462/brain-dead-how-politics-makes-us-stupid.

Klein, Nicholas J., Kelcie Ralph, Calvin Thigpen, and Anne Brown. "Political Partisanship and Transportation Reform." *Journal of the American Planning Association* 88, no. 2 (April 3, 2022): 163–78. https://doi.org/10.1080/01944363.2021.1965495.

Lakoff, George. *Moral Politics: What Conservatives Know That Liberals Don't*. Chicago: University of Chicago Press, 1996.

Loren Lomasky. "Freedom and the Car." *Foundation for Economic Education*, December 1, 1997. Link: https://fee.org/articles/freedom-and-the-car/.

Love, Serena. "A Sense of Architecture in the Past: Exploring the Sensory Experience of Architecture in Archaeology." In *Elements of Architecture*. London: Routledge, 2016.

Macedo, Stephen. *Diversity and Distrust: Civic Education in a Multicultural Democracy*. Cambridge, Mass: Harvard University Press, 2000.

———. *Liberal Virtues: Citizenship, Virtue, and Community in Liberal*. Oxford: Oxford University Press, 1990.

National Safety Council. "Deaths by Transportation Mode," 2022. https://injuryfacts.nsc.org/home-and-community/safety-topics/deaths-by-transportation-mode/.

Neufeld, Blain. "Political Liberalism and Citizenship Education." *Philosophy Compass* 8, no. 9 (2013): 781–97. https://doi.org/10.1111/phc3.12064.

———. "Political Liberalism, Autonomy, and Education." In *The Palgrave Handbook of Citizenship and Education*, edited by Andrew Peterson, Garth Stahl, and Hannah Soong, 35–51. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-319-67828-3_12.

Ohl, Jessy J., Damien S. Pfister, Martin Nader, and Dana Griffin. "Lakoff's Theory of Moral Reasoning in Presidential Campaign Advertisements, 1952–2012." *Communication Studies* 64, no. 5 (November 1, 2013): 488–507. https://doi.org/10.1080/10510974.2013.832340.

Pfundmair, Michaela, Natasha R. Wood, Andrew Hales, and Eric D. Wesselmann. "How Social Exclusion Makes Radicalism Flourish: A Review of Empirical Evidence." *Journal of Social Issues* 80, no. 1 (2024): 341–59. https://doi.org/10.1111/josi.12520.

Philip Brickman. "Hedonic Relativism and Planning the Good Society." *Adaptation Level Theory*, 1971, 287-301.

Podschwadek, Frodo. *Educating the Reasonable: Political Liberalism and Public Education*. Berlin: Springer Nature, 2021.

Prior, Markus, Gaurav Sood, and Kabir Khanna. "You Cannot Be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions." *Quarterly Journal of Political Science* 10, no. 4 (2015): 489–518.

Reich, Rob. "Common Schooling and Educational Choice." In *A Companion to the Philosophy of Education*, edited by Randall Curren, 430–42. Hoboken: Blackwell, 2003.

Resseger, Matthew. "The Impact of Land Use Regulation on Racial Segregation: Evidence from Massachusetts Zoning Borders." *Mercatus Center*, October 11, 2022. https://econpapers.repec.org/paper/ajwwpaper/11220.htm.

Rosenberg, Shawn. "Democracy Devouring Itself: The Rise of the Incompetent Citizen and the Appeal of Right-Wing Populism," 2019. Available at: https://escholarship.org/uc/item/8806z01m.

Rozin, Paul, and Edward B. Royzman. "Negativity Bias, Negativity Dominance, and Contagion." *Personality and Social Psychology Review* 5, no. 4 (November 2001): 296–320. https://doi.org/10.1207/S15327957PSPR0504_2.

Sant, Edda. "Democratic Education: A Theoretical Review (2006–2017)." *Review of Educational Research* 89, no. 5 (October 1, 2019): 655–96. https://doi.org/10.3102/0034654319862493.

Scott, Robert. "The Meaning of Liberal Education." *On the Horizon* 22, no. 1 (January 1, 2014): 23–34. https://doi.org/10.1108/OTH-09-2013-0036.

Stulberg, Brad. "Loneliness Can Make You Vulnerable to Extremism." *TIME Magazine*, November 3, 2022. https://time.com/6223229/loneliness-vulnerable-extremist-views/.

Surridge, Paula. "Education and Liberalism: Pursuing the Link." *Oxford Review of Education* 42, no. 2 (March 3, 2016): 146–64. https://doi.org/10.1080/03054985.2016.1151408.

Van Der Ploeg, Piet. "Burgerschapsvorming Anders: Een Pleidooi Voor Zakelijk Onderwijs." *Pedagogiek* 35, no. 3 (December 1, 2015): 285–98. https://doi.org/10.5117/PED2015.3.PLOE.

Winner, Langdon. "Do Artifacts Have Politics?" *Daedalus* 109, no. 1 (1980): 121–36.

Winnie Hu. "'Hostile Architecture': How Public Spaces Keep the Public Out." *New York Times*, August 14, 2019. https://www.nytimes.com/2019/11/08/nyregion/hostile-architecture-nyc.html.

Chapter 6

Ekeli, Kristian Skagen. "Liberalism and Permissible Suppression of Illiberal Ideas." Inquiry 55, no. 2 (April 1, 2012): 171–93. https://doi.org/10.1080/0020174X.2012.661581.

Gerards, J. H., D. Schiek, L. Waddington, and M. Bell. *The Grounds of Discrimination*. London: Hart Publishing, 2007. https://hdl.handle.net/1887/13722.

Ginsburg, Tom, and Tamir Moustafa. *Rule by Law: The Politics of Courts in Authoritarian Regimes.* Cambridge: Cambridge University Press, 2008. https://summit.sfu.ca/item/15130.

Jhaver, Shagun, Christian Boylston, Diyi Yang, and Amy Bruckman. "Evaluating the Effectiveness of Deplatforming as a Moderation Strategy on Twitter." *Proceedings on Human-Computer Interaction 5*, no. 2 (2021): 381:1-381:30.

Nussbaum, Martha C. "Perfectionist Liberalism and Political Liberalism." *Philosophy and Public Affairs* 39, no. 1 (2011): 3–45. https://doi.org/10.1111/j.1088-4963.2011.01200.x.

Quong, Jonathan. *Liberalism Without Perfection*. Oxford: Oxford University Press, 2011.

Sallust. "Conspiracy of Catiline, Translated by John Selby Watson." *Perseus*. Accessed June 16, 2024. Link: https://www.perseus.tufts.edu/hopper/text?doc= Perseus%3Atext%3A1999.02.0124%3Achapter%3D52.

Sunstein, Cass. *Laws of Fear: Beyond the Precautionary Principle*. Cambridge: Cambridge University Press, 2005. https://doi.org/10.1017/CBO9780511790850.

Thomas, Daniel Robert, and Laila A. Wahedi. "Disrupting Hate: The Effect of Deplatforming Hate Organizations on Their Online Audience." *Proceedings of the National Academy of Sciences 120*, no. 24 (June 13, 2023): 120-144.

Wang, Yuhua. "Authoritarian Judiciary: How the Party-State Limits the Rule of Law." In *Tying the Autocrat's Hands*, 50–85. Cambridge: Cambridge University Press, 2015.

Part III

Chapter 7

Baurmann, Michael, and Daniel Cohnitz. "Trust No One?" In *The Epistemology of Fake News*, edited by Sven Bernecker, Amy K. Flowerree, and Thomas Grundmann. Oxford University Press, 2021.

Blume, Stuart. "Anti-Vaccination Movements and Their Interpretations." *Social Science & Medicine*, Patient Organisation Movements, 62, no. 3 (February 1, 2006): 628–642. https://doi.org/10.1016/j.socscimed.2005.06.020.

Daniel Shapiro. "Is There a Libertarian Defense of the Welfare State?" *Foundation for Economic Education.* Accessed June 13, 2024. Link: https://fee.org/articles/is-there-a-libertarian-defense-of-the-welfare-state/.

Hümmler, Holm, and Ulrike Schiesser. "Basic Strategies." In *Fact and Prejudice: How to Communicate with Esoterics, Fanatics and Conspiracy Believers*, edited by Holm Gero Hümmler and Ulrike Schiesser, 51–60. Berlin, Heidelberg: Springer, 2023. https://doi.org/10.1007/978-3-662-66032-4_4.

Katerji, Oz. "The West's Leftist 'Intellectuals' Who Traffic in Genocide Denial, from Srebrenica to Syria." *Haaretz*, November 24, 2017. Link: https://www.haaretz.com/opinion/2017-11-24/ty-article-opinion/the-wests-leftist-male-intellectuals-who-traffic-in-genocide-denial/0000017f-f346-d8a1-a5ff-f3cec4320000.

Kozloff, Nikolas. "Chomsky, Ali, and the Failure to Challenge the Authoritarian Left." *Al Jazeera*, September 22, 2012. Link: https://www.aljazeera.com/opinions/2012/9/22/chomsky-ali-and-the-failure-to-challenge-the-authoritarian-left.

Lash, Nicholas. "Where Does The God Delusion Come From?" *New Blackfriars* 88, no. 1017 (2007): 507–521. https://doi.org/10.1111/j.1741-2005.2007.00172.x.

Michael Kleen. "Incrementalism vs. Revolution," *Michael Kleen (Personal Blog)*, March 27, 2017. Link: https://michaelkleen.com/2017/03/27/incrementalism-vs-revolution/.

Poole, Steven. "The Four Horsemen Review - Whatever Happened to 'New Atheism'?" *The Guardian*, January 31, 2019. https://www.theguardian.com/books/2019/jan/31/four-horsemen-review-what-happened-to-new-atheism-dawkins-hitchens.

Sala, Roberta. "The Place of Unreasonable People beyond Rawls." *European Journal of Political Theory* 12, no. 3 (July 1, 2013): 253–70. https://doi.org/10.1177/1474885112465248.

Sam Harris. "The Virus of Religious Moderation." *Sam Harris (Personal Blog)*, March 19, 2005. Link: https://www.samharris.org/blog/the-virus-of-religious-moderation.

Stallard, Katie. "What John Mearsheimer Gets Wrong about Ukraine." *New Statesman*, October 5, 2023. https://www.newstatesman.com/world/europe/2023/10/john-mearsheimers-incorrect-views-on-everything.

Vinx, Lars. "Carl Schmitt." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2019. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2019/entries/schmitt/.

Chapter 8

Kramer, Sina. *Excluded Within: The (Un)Intelligibility of Radical Political Actors.* Oxford: Oxford University Press, 2017.

Netter, J. 2017. "Why Be Reasonable? Political Liberalism, Moral Pluralism, and Deep Disagreement." PhD thesis, University of Oxford. https://ora.ox.ac.uk/objects/uuid:26860521-a672-4ac5-bea9-b9e6497aff68.

O'Neill, Onora. *Towards Justice and Virtue: A Constructive Account of Practical Reasoning.* Cambridge: Cambridge University Press, 1996.

Chapter 9

Boghossian, Peter. *A Manual for Creating Atheists.* Durham: Pitchstone Publishing, 2014.

Figueiredo, Marques, Rui Pedro, and Batista Lopes Joao Carlos. *Information and Communication Overload in the Digital Age.* Pennsylvania: IGI Global, 2017.

Hümmler, Holm Gero, and Ulrike Schiesser. "Basic Strategies." In *Fact and Prejudice: How to Communicate with Esoterics, Fanatics and Conspiracy Believers*, edited by Holm Gero Hümmler and Ulrike Schiesser, 51–60. Berlin: Springer, 2023. https://doi.org/10.1007/978-3-662-66032-4_4.

Jaeggi, Rahel. "Was Ist Eine (Gute) Institution." In *Sozialphilosophie Und Kritik*, edited by Axel Honneth and Rainer Forst. Berlin: Suhrkamp, 2009. Link to English translation: https://www.philosophie.hu-berlin.de/de/lehrbereiche/jaeggi/mitarbeiter/jaeggi_rahel/RJWhat%20is%20a%20-good-%20institution.pdf

Mansbridge, Jane. "Everyday Talk in the Deliberative System". In *Deliberative Politics: Essays on Democracy and Disagreement*, ed. by Stephen Macedo, p. 211-238. Oxford: Oxford University Press, 1999.

Mansbridge, Jane and Flaster, Katherine. "The Cultural Politics of Everyday Discourse: The Case of the Male Chauvinist", *Critical Sociology*, 33 (2007): p. 627–660.

Plagiarism Form

Universiteit Utrecht

Faculty of Humanities
Version June 2024

Fraud and Plagiarism
Scientific integrity is the foundation of academic life. Utrecht University considers any form of scientific deception to be an extremely serious infraction. Utrecht University therefore expects every student to be aware of, and to abide by, the norms and values regarding scientific integrity.

The most important forms of deception that affect this integrity are fraud and plagiarism. Plagiarism is the copying of another person's work without proper acknowledgement, and it is a form of fraud. The following is a detailed explanation of what is considered to be fraud and plagiarism, with a few concrete examples. Please note that this is not a comprehensive list!

If fraud or plagiarism is detected, the study programme's Examination Committee may decide to impose sanctions. The most serious sanction that the committee can impose is to submit a request to the Executive Board of the University to expel the student from the study programme.

Plagiarism
Plagiarism is the copying of another person's documents, ideas or lines of thought and presenting it as one's own work. You must always accurately indicate from whom you obtained ideas and insights, and you must constantly be aware of the difference between citing, paraphrasing and plagiarising. Students and staff must be very careful in citing sources; this concerns not only printed sources, but also information obtained from the Internet.

The following issues will always be considered to be plagiarism:
- cutting and pasting text from digital sources, such as an encyclopaedia or digital periodicals, without quotation marks and footnotes;
- cutting and pasting text from the Internet without quotation marks and footnotes;
- copying printed materials, such as books, magazines or encyclopaedias, without quotation marks or footnotes;
- including a translation of one of the sources named above without quotation marks or footnotes;
- paraphrasing (parts of) the texts listed above without proper references: paraphrasing must be marked as such, by expressly mentioning the original author in the text or in a footnote, so that you do not give the impression that it is your own idea;
- copying sound, video or test materials from others without references, and presenting it as one's own work;
- submitting work done previously by the student without reference to the original paper, and presenting it as original work done in the context of the course, without the express permission of the course lecturer;
- copying the work of another student and presenting it as one's own work. If this is done with the consent of the other student, then he or she is also complicit in the plagiarism;
- when one of the authors of a group paper commits plagiarism, then the other co-authors are also complicit in plagiarism if they could or should have known that the person was committing plagiarism;
- submitting papers acquired from a commercial institution, such as an Internet site with summaries or papers, that were written by another person, whether or not that other person received payment for the work.

ChatGPT/Generative AI
You are not allowed to generate text, code, figures, images, etc. with Generative AI and present it as your own work. This is a form of fraud.

The rules also apply to rough drafts of papers or (parts of) theses sent to a lecturer for feedback, to the extent that submitting rough drafts for feedback is mentioned in the course handbook or the thesis regulations.

The Education and Examination Regulations (Article 5.14) describe the formal procedure in case of suspicion of fraud and/or plagiarism, and the sanctions that can be imposed.

Ignorance of these rules is not an excuse. Each individual is responsible for their own behaviour. Utrecht University assumes that each student or staff member knows what fraud and plagiarism entail. For its part, Utrecht University works to ensure that students are informed of the principles of scientific practice, which are taught as early as possible in the curriculum, and that students are informed of the institution's criteria for fraud and plagiarism, so that every student knows which norms they must abide by.

I hereby declare that I have read and understood the above.
Name: Marijn Eigenhuis
Student number: 5713935
Date and signature: June 16th, 2024   MLEigenhuis

Submit this form to your supervisor when you begin writing your Bachelor's final paper or your Master's thesis.

Failure to submit or sign this form does not mean that no sanctions can be imposed if it appears that plagiarism has been committed in the paper.

Appendix: List of Theses

I: That reasonableness as conceptualized by Rawls implies and ought to entail a third basic aspect in the form of the duty of prudence: a moral responsibility to exercise one's intellect conscientiously and to attempt to address personal prejudice, bias, fallacies, and informational blindness.

II: That recent moral psychological research indicates that ordinary citizens face significant cognitive difficulties in adopting the three basic aspects of reasonableness, which suggests that, even within the framework of ideal theory, it is more appropriate to postulate the fact of simple pluralism rather than reasonable pluralism.

III: That political liberalism aspires to the three core ideals of legitimacy, stability, and inclusivity, which cannot be fully realized simultaneously under the fact of simple pluralism, thus causing a trilemma.

IV: That a conciliatory approach to containment, as exemplified by a method of appeasement or conjecture, is an ineffective strategy that does not properly address the fundamental problem of unreasonableness.

V: That a transformational approach that incorporates institutional, educational, and technological options – although not sufficiently effective on its own – is an essential strategy to address the problem of unreasonable doctrines rejecting the political conception of justice.

VI: That a marginalizing approach, combining reactive and preventive intervention that restricts the rights of the unreasonable in exceptional cases, is necessary for successful containment; although, due to normative challenges, it should not function as the primary strategy.

VII: That the category of unreasonableness is not homogeneous, but actually consists of four distinct subcategories (the anti-reasonable, the uncooperative, the unequitable, and the unaware) that require different approaches in order to become more reasonable – two of which could feasibly become reasonable over time by a strategy of involvement.

VIII: That the liberal order can provide anti-reasonable citizens with adequate reason for excluding them from further deliberation on matters of legitimacy, on the grounds that meaningful membership in the constituency of public justification imposes certain epistemic norms that render citizens unintelligible if they pursue unreasonable proposals in that context.

IX: That the liberal order should encourage reasonable citizens to perform the duty of vigilance, which entails safeguarding the social norms that enable civil political deliberation, engaging unaware citizens by asking about their epistemology, and cultivating a critical awareness of the benefits of liberal democracy.