RECIPROCITY AND TRUST IN POLITICAL DELIBERATION

RECIPROCITY AND TRUST IN POLITICAL DELIBERATION: AN INVESTIGATION INTO THE NORMS OF DISCURSIVE CIVILITY

BY AYOMIDE AJIMOKO, B.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Arts

McMaster University © Copyright by Ayomide Ajimoko, December 2022

McMaster University MASTER OF ARTS (2022) Hamilton, Ontario (Philosophy)

TITLE: Reciprocity and Trust in Political Deliberation: An Investigation into the Norms of Discursive Civility AUTHOR: Ayomide Ajimoko, B.A. (McMaster University) SUPERVISOR: Professor Stefan Sciaraffa NUMBER OF PAGES: vi, 79

LAY ABSTRACT

In this thesis, I scrutinize possible explanations for why incivilities are widespread in contemporary political discourse. By appeal to a principle of reciprocity, my ultimate goal is to defend an explanation according to which citizen incivility in political discourse is *blameless*. According to the principle of reciprocity, citizens are not required to maintain civility if they have reason to believe that their interlocutors will not reciprocate civility in political discussion. Based on this principle, I argue that ordinary citizens in democratic societies across the US and industrialised west are often justified in being uncivil in political debate. For these citizens often have no reason to believe that their interlocutors will be civil. If the justificatory account of incivility is on the right track, then policies aimed at restoring civility in political discourse must be concerned to build citizen trust that others will reciprocate civility. Without this trust, citizens may not see themselves as having reason to be civil in political debate.

ABSTRACT

Much contemporary political discourse in the US and industrialized west is defective. According to a number of scholars, such as Scott Aikin and Robert Talisse, this defectiveness can plausibly be explicated in terms of a breakdown in political *civility*. In this thesis, I scrutinize possible explanations for why *incivilities* are widespread in political discourse. My goal is to defend an explanation according to which citizen incivility in political discourse is *blameless*. To do this, I appeal to a principle of reciprocity.

According to the principle of reciprocity, citizens are not required to maintain civility in political discourse if they have reason to believe that their interlocutors will not reciprocate civility. When applied to contemporary politics, this principle implies that ordinary citizens in democratic societies across the US and industrialised west are often justified in being uncivil in political debate. For these citizens often have no reason to believe that their interlocutors will be civil. If the reciprocity principle is right, then policies aimed at restoring civility in political discourse must be concerned to build citizen trust that others will reciprocate civility. Without this trust, citizens may not see themselves as having reason to be civil.

The thesis is divided into two chapters. In chapter 1, I defend the justificatory account of incivility against two competing accounts of political incivility—identity and group theories. These accounts pathologize political incivility as a kind of irrationality, but based on the argument from reciprocity, I argue that political incivilities are often rational and so justified. In chapter 2, I analyse the kind of trust that is necessary to build more civility in political discourse. In particular, I develop and defend a conception of *deliberative trust*, which is defined as the belief that one's interlocutor will reciprocally adhere to the norms of civility.

ACKNOWLEDGEMENTS

I want to express my sincerest gratitude to my supervisor, Stefan Sciaraffa, for his invaluable help throughout this process, especially during the COVID-19 pandemic. I particularly want to thank him for his patience and kind feedback whilst I reworked these ideas over multiple drafts. I also want to acknowledge my second reader, Violetta Igneski. Completing the master's in the midst of the pandemic was isolating and daunting, but I can point to distinct conversations with Violetta that kept me motivated. For this I am extremely grateful.

I would like to thank my family for their consistent love and support, and my partner Sarah and her family for their unfailing kindness and encouragement. Without you, I would not have completed this program. I am so thankful to share this achievement with you all.

Finally, I would like to acknowledge my friends. Thank you to Claire, Novan, Lois, Quinn, Leigh, Hasan, Bina, and Daniella for their support and inspiration. Thank you to my MA cohort, particularly Rachel, Corey, and Nicole for believing in me. I especially want to acknowledge Nicole for being a first-rate grad colleague and friend. Thank you for tolerating my anxieties and complaints about grad school and the pandemic. I will look back with fondness at our reading groups, impromptu Zoom calls, and discussions about reasons and normativity.

Thank you to everyone.

TABLE OF CONTENTS

INTRODUCTION.. 1

CHAPTER ONE: IS POLITICAL INCIVILITY EVER JUSTIFIED?...................... 9

1. INTRODUCTION 9
2. CIVILITY AND MUTUAL RESPECT 12
   2.1. Political Deliberation 12
   2.2. Civility in Political Discourse.. 14
   2.2.1. Respect for Political Equality 15
   2.2.2. Norms of Civility 17
   2.2.3. Political Incivility.. 19
3. IDENTITY & GROUP THEORIES OF POLITICAL INCIVILITY 20
   3.1. Identity Theories of Political Incivility 20
   3.2. Group Theories of Political Incivility.. 22
4. THE ARGUMENT FROM RECIPROCITY 24
   4.1. Stage One 24
   4.2. Stage Two.. 28
   4.3. Some Qualifications 31
5. THE UPSHOT 39
6. DELIBERATIVE TRUST AND DELIBERATIVE TRUSTWORTHINESS 40
7. CONCLUSION... 43

CHAPTER TWO: DELIBERATIVE TRUST OR SOCIAL TRUST? 45

1. INTRODUCTION 45
2. KEVIN VALLIER'S THEORY OF SOCIAL TRUST 49
   2.1. Preliminaries.. 49
   2.2. Vallier's Account of Social Trust 50
   2.2.1. The Domain of Social Trust 50
   2.2.2. The Focus of Social Trust: Adherence to Social-Moral Rules 51
   2.2.3. Moral Reasons, Social Trust, and Social Trustworthiness....................... 53
   2.2.4. How to Sustain Social Trust... 57
3. SOCIAL TRUST AND PUBLIC DISCOURSE 61
4. DELIBERATIVE TRUST, TRUSTWORTHINESS, AND PUBLIC JUSTIFICATION 64
   4.1. Deliberative Trust Must Track Deliberative Trustworthiness 65
   4.2. Deliberative Trust Depends on Public Justification 66
5. CONCLUSION... 69

CONCLUSION 71

BIBLIOGRAPHY 77

INTRODUCTION

Citizens in the US and across the industrialised west are undergoing a crisis of civility. When citizens discuss politics on social media, in the workplace, or at Thanksgiving dinners, they often insult and vilify their political opponents. Of course, incivility in public discourse is not new. One may go as far back as the European wars of religion to see the pattern of insults and vilifications between opposing social and ideological groups (Dees, 2022). But the contemporary scene brings its own set of concerns.

Today, Americans are more affectively polarised than ever. That is, they are more distrustful of and antipathetic towards each other (Iyengar et al., 2012). According to the American National Elections Studies (ANES), thermometer ratings of how partisans feel about opposing partisans and ideological groups have changed dramatically over the decades. An ANES survey tallies partisan affect using a scale of 0 to 100, where values between 0 and 50 represent cold feelings, and values between 51 and 100 warm feelings. In the US, the proportion of partisans who have cold feelings (feelings thermometer below 50) towards opposing-party partisans stayed below 10 percent between 1970 and 2000; however, between 2000 and 2016, that proportion doubled to 21 percent (Iyengar & Krupenkin, 2018, p. 203). Survey data from the Pew Research Centre shows that Americans have a hard time befriending counter-partisans too: 55 percent of Republicans and 64 percent of Democrats say they have "just a few" or "no" close friends from their opposing party (Pew Research Center, 2017).

The animus between partisans plays out in public discourse. In this context, civility is defined by acting in ways that manifests one's respect for others as political equals. This requires that, first, when faced with political disagreement, citizens should listen to each other to better understand each other's concerns; they should also be open to revising their own views in light

of available evidence. Second, since civility involves manifesting one's respect for others as equals, incivilities in political speech—such as insults, heckling, name-calls, and vilifications—seem to communicate disrespect for this equal status. Thus, citizens ought to refrain from such political incivilities during debate.

Given these standards of civility, contemporary political discourse falls disappointingly short. When citizens discuss politics on social media, in the workplace, at Thanksgiving dinners, at the barbershop and so on, they often insult and demean their political opponents. In directing such incivilities towards each other, they not only fail to listen and better understand each other, but also, they appear to convey a disregard for their opponents' status as political equals.

Hence, citizens, politicians, and academics alike call for *more civility* in political discourse. In a moving speech at the 2010 National Prayer Breakfast, Barack Obama makes a compelling plea for civility:

> "At times, it seems like we're unable to listen to one another; to have at once a serious and civil debate. And this erosion of civility in the public square sows division and distrust among our citizens…It makes politics an all-or-nothing sport, where one side is either always right or always wrong when, in reality, neither side has a monopoly on truth…Empowered by faith, consistently, prayerfully, we need to find our way back to civility. That begins with stepping out of our comfort zones in an effort to bridge divisions…Stretching out of our dogmas, our prescribed roles along the political spectrum, that can help us regain a sense of civility. Civility also requires relearning how to disagree without being disagreeable" (Barack Obama, Remarks by the President at the National Prayer Breakfast, 2010).

What might explain the defective state of contemporary political discourse? Are citizens blameworthy for engaging in political incivilities? In this thesis, I scrutinize possible reasons why incivilities are widespread in public discourse. My goal is to defend an explication of this phenomenon whereby citizen incivility in political deliberation is *blameless*. Support for this claim draws on a principle of reciprocity, according to which citizens are justified in being uncivil if they have reason to believe that their interlocutors will not reciprocate civility. This conclusion is important because it shows that citizens sometimes act rationally and so blamelessly when they choose to be uncivil in political debate. For, given the defective state of public discourse, citizens often fail to have reason to believe that their interlocutors will be civil. If citizens are justified for acting uncivilly in political discourse, then strategies aimed at restoring civility must account for this finding. More specifically, I argue, such strategies must involve ways of building citizen trust that others will reciprocate civility in political deliberation.

The thesis is divided into two chapters. In chapter 1, I defend the justificatory account of incivility against two competing accounts of political incivility. According to identity theories, political incivility arises from discomfort with political disagreement (Novaes, 2021; Taber & Lodge, 2006). On this view, when other citizens disagree with our political views, we take this to be a personal attack. This is because we take our political beliefs to be central to who we are. The discomfort with disagreement might then lead us to be uncivil—e.g., to name-call or bad mouth our political opponents.

According to group theories, in contrast, political incivilities can be analysed in terms of tribal impulses: psychological tendencies to favour one's group and to hold members of an opposing group in contempt (Huddy et al., 2015; Iyengar et al., 2012). Under this conception, partisans that identify with a given political party (e.g., Democrat or Republican) or ideological

group (e.g., Conservative or Liberal) develop tribal impulses to act in ways that favour their party while disdaining members of the opposing party. Consequently, tribal impulses drive citizen distrust, suspicion, and anger towards political opponents, which motivates them to insult, vilify, and demonize these opponents.

If political incivilities are grounded in personal discomfort with disagreement, or in tribalistic distrust of political opponents, then political incivilities are unjustified. On the identity framework, there is no reason to take disagreement personally, and we are not warranted in responding with hostility to such disagreements. With respect to group theories, tribalistic distrust embodies an unjustified reason to be uncivil in political discourse. Tribalistic distrust is caused by brute tendencies to hold one's political opponents in contempt, and this contempt for one's political opponents weakens the possibility of one taking their claims and concerns seriously; that is, it weakens one's commitment to treat them as a political equal. To willingly act on this tribalistic motive would be unjustified; therefore, political incivilities motivated by tribalistic distrust are unjustified.

Contra identity and group theories, I argue in chapter 1 that political incivilities are sometimes justified. If we appeal to a rational, and so justifiable, motivation for political incivility, then we can avoid the results predicted by identity and group models. This account of incivility appeals to an argument from reciprocity. According to the reciprocity argument, if one's interlocutor fails to reciprocate civility in political discussion, one has no reason to maintain civility. As Maxime Lepoutre explains, "plausibly, I will not commit to respectfully constraining the claims I make unless I feel that I can rely on my interlocutors also to fulfil a commitment to discursive civility" (Lepoutre, 2021, p. 134). Further, one might have reason to be uncivil given the uncivil behaviour of one's fellow discussant.

Scott Aikin and Robert Talisse generalise the argument from reciprocity to citizens in contemporary democratic societies that suffer from widespread political incivilities. On their view, an important property of the norms of civility is that they are collective norms: they apply to groups as a whole and not just individuals. Individuals in a collective acquire obligations to follow a collective norm because the collective norm applies to them as members of the group. Thus, individuals in such a collective have reason to follow a collective norm only on the condition that others do so too. If individuals in a given collective have reason to think that others will not follow a collective norm, then they have no reason to maintain adherence to that norm.

The idea, then, is that the norms of civility collectively apply to citizens as a whole, in virtue of their role as citizens in a democracy. Given this, the principle of reciprocity entails that citizens are required to be civil in political discourse so long as other citizens are civil too. It follows that if citizens have reason to think that the general public defects from civility, then they do not have reason to keep adhering to civility's requirements. As it turns out, contemporary political discourse is riddled with incivility. So, Aikin and Talisse conclude, "ordinary citizens often have adequate reason to assume that their opposition has pulled back from the commitment to civility, and so they also have adequate reason to hold that the requirements of civility are no longer in play in public life" (Aikin & Talisse, 2020, p. 98).

Importantly, what distinguishes Aikin and Talisse's explanation of political incivility is that it is also a justificatory account. For them, when reciprocity breaks down, citizens have no reason to keep maintaining civility: "once we are convinced that our opponents will not reciprocate civility in argument, we no longer have a rationale for upholding the standards that civility sets" (Aikin & Talisse, 2020, p. 98). But if this right, they argue, then citizens are

blameless for going against the standards that civility sets, especially given the defective state of contemporary discourse. In other words, they are permitted and so justified in being uncivil given the current state of discourse. To be sure, the permission to be uncivil does not imply that citizens can act however they wish in public discourse. For example, the permission to be uncivil might rule out cases of incivility like anger and outrage at one's political opponents. That said, at least a limited range of incivilities are permitted in virtue of the reciprocity argument.

The upshot of this account is that mechanisms that are geared towards building more civility in contemporary political discourse must account for civility's reciprocal character. Accordingly, such mechanisms must be concerned to build the trust that others will be civil in political discourse. In chapter 2, then, I outline and analyse a conception of trust that animates political deliberation amongst citizens. This analysis proceeds in two steps.

First, I posit that *deliberative trust* is the type of trust that citizens must exhibit in order to feel assured that their interlocutors in political debate will be civil. Deliberative trust is defined as the belief citizens hold about other citizens that each is a civil interlocutor. This entails that within a given interaction, one believes one's co-deliberator conscientiously adheres norms of civility, namely, 1) open-mindedness, and 2) avoiding bad faith behaviour, such as insults, vilifications, and so on.

I argue, however, that deliberative trust possesses a legitimacy condition according to which deliberative trust must be placed in actual deliberatively trustworthy citizens. *Deliberative trustworthiness* is defined as the virtue of being civil—i.e., of having the disposition to act in ways that the norms of civility require. To see the importance of the legitimacy condition on deliberative trust, consider A and B. Suppose B is a liar who is arrogant and dogmatic, then because B violates the standards of open-mindedness and avoiding bad faith behaviour, B is

uncivil and thus deliberatively *un*trustworthy. Further suppose that B hides this fact about himself well, and A is deceived into deliberatively trusting B when B is in fact deliberatively *un*trustworthy. The problem with this is that A's trust is misplaced. The danger of placing trust in untrustworthy citizens, as I shall argue in chapter 2, is that misplaced deliberative trust is unstable and might further deepen political incivility. For example, once A realises that her trust is misplaced in *un*trustworthy B, she is likely to resent B. This might lead her to develop a distrust of others in general and reinforce her feeling entitled to be hostile and disrespectful in discussions around politics. The result is that A's discovery of her misplaced trust might dispose her to even more incivility in her deliberative exchanges. Thus, it is crucial that deliberative trust is placed in deliberatively trustworthy citizens, given possible bad outcomes of misplaced trust.

The second step of the analysis is to investigate a requirement on the norms that anchor deliberative trust and trustworthiness. I argue that the norms of civility anchor deliberative trust and trustworthiness, and an important requirement on these norms is that they must be *publicly justified*. Drawing on Kevin Vallier's work on social trust, I explicate the strategy of sustaining trust around publicly justified norms of civility. Here, public justification plays a key role in securing the trust that others will reciprocate adherence to the norms of civility. According to the principle of public justification, a given coercive norm is justified if and only if each and every member of the public has a conclusive reason R to endorse that rule, law, policy, or action (Vallier, 2011, p. 262). If the norms of civility are publicly justified, this entails that citizens will follow them for their *own* reasons. This is crucial if these citizens are to be deliberatively trustworthy, for, as I shall argue in chapter 2, an important condition on the virtue of deliberative trustworthiness is that one sees oneself as having sufficient reason to follow civility's standards.

To close the thesis, I canvass two possible mechanisms to build deliberative trust and trustworthiness. First, I argue that by implementing integrative practices and policies, citizens can come to observe each other and socially mingle with each other (Anderson, 2010). The idea is that through integration citizens can be more open to being civil with each other in public discourse. Secondly, I argue that we can tap into the therapeutic aspects of trust, according to which trust has a bootstrapping effect: by trusting a person who is minimally trust responsive, one can unlock their potential to be even more trust responsive and thus trustworthy. On this view, even though the reciprocity argument permits us to be uncivil, one way we can build more civility is by choosing to be civil nonetheless. If we are civil despite having reason not to be, this might inspire our interlocutors to emulate our civility. In other words, by choosing to be civil, we may come to build the deliberative trustworthiness of others—i.e., we inspire the cultivation of the virtue of civility in them. And as they become deliberatively trustworthy, others can come to build more deliberative trust that citizens mostly reciprocate civility in political discourse. As a result, therapeutic trust might kickstart a bootstrapping effect whereby deliberative trust builds deliberative trustworthiness which in turn builds more deliberative trust that others will reciprocate civility.

Admittedly, these mechanisms for building trust and trustworthiness are not exhaustive. However, they provide us with an idea of how we might build deliberative trust and trustworthiness, and in turn build more civility in public discourse.

CHAPTER ONE: IS POLITICAL INCIVILITY EVER JUSTIFIED?

1. INTRODUCTION

Much contemporary political discourse in the US and industrialized west is defective. According to a number of scholars, this defectiveness can plausibly be explicated in terms of a breakdown in *civility*. In the context of democratic politics, civility is defined by acting in ways that manifests one's respect for others as political equals (Aikin & Talisse, 2020). This requires that, first, when faced with political disagreement, citizens should listen to each other to better understand each other's concerns, and also be open to revising their own views in light of evidence. Since civility involves manifesting one's respect for others as equals, incivilities in political speech—such as insults, heckling, name-calls, and vilifications—seem to communicate disrespect for this equal status. Thus, second, citizens ought to refrain from incivilities during political debate.

Given these standards of civility, contemporary political discourse falls disappointingly short. When citizens discuss politics on social media, in the workplace, or at Thanksgiving dinners, they often vilify and demean their political opponents. In directing such incivilities towards each other, they not only fail to listen and better understand each other, but also, they appear to convey a disregard for their opponents' status as political equals.

What might explain this defective state of contemporary political discourse? In this chapter, I defend an account of incivility against two competing accounts of political incivility. According to identity theories, political incivility arises as a result of discomfort with political disagreement (Novaes, 2021; Taber & Lodge, 2006). On this view, when other citizens disagree with our political views, we take this to be a personal attack. For our political beliefs make up our core values, so challenges to them seem like threats to our identities.

According to group theory, in contrast, political incivilities can be analysed in terms of tribal impulses: psychological tendencies to favour one's group and to hold members of an opposing group in contempt (Huddy et al., 2015; Iyengar et al., 2012). According to this conception, citizens that identify with a given political party (e.g., Democrats or Republican) or ideological group (e.g., Conservative or Liberal) develop tribal impulses to act in ways that favour that party or group (the "in-group") while disdaining members of the opposing party or group (the "out-group"). Consequently, tribal impulses drive distrust, suspicion, and anger towards political opponents. This tribalistic distrust then motivates citizens to insult, vilify, and demonize their opponents. In other words, tribalistic distrust drives political incivility.

On neither of these conceptions are political incivilities justified. On the identity framework, there is no reason to take disagreement personally, and even if we do, we are not warranted in responding with incivilities and hostilities. With respect to group theories, tribalistic distrust embodies an unjustified reason to be uncivil. Tribalistic distrust is caused by brute tendencies to disdain one's political opponents, and this disdain for one's political opponents weakens the possibility of one taking their claims and concerns seriously; that is, it weakens one's commitment to treat them as a political equal. This constitutes a distinct kind of wrong against one's political opponents since they are owed respect as political equals. To willingly act on this motive would be unjustified; therefore, political incivilities as motivated by tribalistic distrust are unjustified.

In this chapter, however, I repudiate the conclusion that political incivilities in contemporary democratic societies are unjustified. I defend an account of incivility that appeals to a reciprocity argument. According to the argument from reciprocity, citizens in political debate follow a principle of reciprocity whereby they will be civil only on the condition that

others are civil too. Following Scott Aikin and Robert Talisse, I argue that citizens have no reason to maintain civility in a political culture that disregards the norm. Thus, the argument from reciprocity supports the conclusion that citizens sometimes act rationally and justifiably by exhibiting uncivil behaviour in political debate.

The upshot is that strategies aimed at restoring civility in political discourse should account for the fact that citizens will not be civil unless they can trust that their interlocutors will be civil too. As such, strategies aimed at building more civility must be concerned to build citizen trust that their fellow citizens will reciprocate civility in political deliberation. To conclude the chapter, I outline an analysis of the conception of trust that is needed for citizens to engage civilly with one another: *deliberative trust*. Deliberative trust is defined as the belief citizens hold about other citizens that each is a civil interlocutor. This entails that within a given interaction, one believes one's co-deliberator conscientiously adheres norms of civility. *Deliberative trustworthiness* is defined as the virtue of being civil—i.e., of having the disposition to act in ways that the norms of civility require. I argue that deliberative trust possesses a legitimacy condition according to which deliberative trust must be placed in actual deliberatively trustworthy citizens.

To restore civility, then, we need to cultivate deliberative trust. But as I shall argue, to build citizen trust that civility is upheld, citizens need to observe that civility is indeed upheld. For citizens to see reciprocal civility, there must actually be deliberatively trustworthy citizens who uphold civility—i.e., citizens who are motivated to reciprocally uphold civility. But to get such trustworthy citizens, we need to develop deliberative trustworthiness in society. That is, we need to promote the virtue of being civil in political discourse. Therefore, building deliberative

trust is inextricably linked with cultivating deliberative trustworthiness. One must accompany the other.

The chapter is structured as follows. In section 2, I explicate the concept of civility and its requirements. In section 3, I describe how group and identity theories explicate the widespread phenomenon of political incivilities. Group and identity theories suggest that when citizens exhibit uncivil behaviour in political debate, they act in an unjustified manner. Section 4 rejects this suggestion. Here, I argue that on account of the argument of reciprocity, citizens act in a blameless or justified way when they engage in incivilities in political discourse. Sections 5, 6, and 7 explore the upshot of this justificatory argument, concluding with a sketch of the concepts of deliberative trust and deliberative trustworthiness.

2. CIVILITY AND MUTUAL RESPECT

In this section, I first provide an account of what political or democratic deliberation comprises. I then outline the concept of civility that I employ in the thesis.

2.1. Political Deliberation

Democratic deliberation, discourse, or debate (I use these terms interchangeably) describes the set of communicative activities in which citizens exchange considerations and rational arguments for and against a political proposal, such as a law mandating vaccines or a policy prohibiting abortions (Christiano, 1996; Cohen, 1997). Citizens care about proposals for a law or policy because these proposals specify the terms on which they can cooperate and interact with each other. Thus, their deliberations about these proposals are oriented towards the goal of agreement about how to live with each other, and short of agreement, their deliberations are oriented towards a better understanding of one another's preferences and interests (Aikin & Talisse, 2020, p. 18).

Political deliberation occurs in multiple forums across a democratic society. This way of thinking about political deliberation draws on the 'systemic approach' in deliberative democracy theory. On this picture, discourse between citizens in dispersed communicative forums combine to form a system of highly differentiated but interrelated parts: a *deliberative system* (Mansbridge et al., 2012, pp. 4–5). For instance, in their seminal introduction to the theory of deliberative systems in deliberative democracy, Jane Mansbridge, James Bohman, Simone Chambers, Thomas Christiano, Archon Fung, John Parkinson, Dennis F. Thompson and Mark E. Warren note:

> "A map of nodes in the deliberative system would reveal many nodes, with multiple forms of communication among them. Those nodes would include nation state bodies at different levels of government and with their different legislative houses, administrative agencies, the military, and the staffs of all of these; international bodies at different levels and their staffs; multinational corporations and local businesses; epistemic communities; foundations; political parties and factions within those parties; party campaigns and other partisan forums; religious bodies; schools; universities with their departments, fields, and disciplinary associations; unions, interest groups, voluntary associations and non-governmental organizations (NGOs) both ad hoc and long-standing; social movements with both their enclaves and their broader participation; the media including the internet, blogs, social media, interactive media, books, magazines, newspapers, film, and television; informal talk among politically active or less active individuals whether powerful or marginalized; and forms of subjugated and local knowledge that rarely surface for access by others without some opening in the deliberative system" (Mansbridge et al., 2012, p. 10).

According to this view, political discourse dispersed across multiple discursive contexts is linked systematically to yield democratic decisions. Thus, the burden of contributing to democratic decisions is not limited to formal arenas like the supreme court or legislature: political conversations in informal arenas like a coffee shop, partisan media, party caucus, and so on play a role too (Mansbridge et al., 2012, p. 2).

Following these 'systems' theorists, I will assume that as long as two or more citizens gather and discuss political proposals that affect them, they are engaging in democratic discourse or deliberation, for any such discussion is enmeshed in the larger deliberative system that ultimately yields democratic decisions. This assumption plausibly captures the kinds of discursive interactions we see between citizens in barbershops, Thanksgiving dinners and so on. The present discussion thus focuses on the norms of civility as they appear within myriad discursive contexts.

2.2. Civility in Political Discourse

Political civility is realized when parties are guided by and sufficiently conform to norms of civility in the context of political debate.[1] These norms require actions that realize respect for political equality. Thus, when A acts civilly towards B in political discourse (more on what this entails shortly), A's actions manifest her respect for B's political equality. To grasp what actions the norms of civility require, however, we need to inquire what respect for others as political equals involves. In what follows, I detail some important elements of respect for political equality in the context of political deliberation (section 2.2.1.). I then explain how following the norms of civility helps one express this respect (section 2.2.2.). This sets the stage for the discussion of political incivilities (section 2.2.3.).

[1] See Aikin and Talisse (2020, p. 17) for a discussion of this characterization of political civility.

2.2.1. Respect for Political Equality

In the context of political discourse, respect for one's interlocutor as a political equal requires that one places oneself under their deliberative influence.[2] As Elizabeth Anderson notes, "[a]ttentive listening to others' perspectives is an important way of expressing respect for others" (Anderson, 2022, p. 80). Jane Mansbridge et al. argue that "[openness] to being moved by the words of another is to respect the other as a source of reasons, claims, and perspectives" (Mansbridge et al., 2012, p. 11). And as Thomas Christiano states, one treats others with respect by "listening to what they have to say and by trying to take it into account or by rationally arguing against what they have to say" (Christiano, 2008, p. 199). This expresses respect for others since "it treats the other as a rational being with a perspective on value that is to be taken into account and that one can learn from" (Christiano, 2008, p. 199). Respect for political equality thus involves open-mindedness. Open-mindedness can be defined as a willingness to take up and seriously engage with political standpoints distinct from one's own (Baehr, 2011). It therefore implies a willingness to listen to one's interlocutor to better understand each other's concerns and a willingness to revise one's own views in the face of reasonable criticism. Moreover, if being open-minded involves surrendering one's arguments to reasonable criticism, this can only be successful if one is sincere and forthright with what one believes. Thus, open-mindedness also implies a willingness to be sincere and forthcoming with one's stances, reasons, and evidence.

A noteworthy feature of respect for political equality is that it is a collective or cooperative endeavor. On this view, respect for political equality helps realize the foundational

[2] Importantly, in addition to placing oneself under others' deliberative influence, respect for political equality requires that "citizens have equal votes, equal resources with which to negotiate with others and equal resources with which to participate in the process of discussion and debate over [policy proposals]" (Christiano, 2008, p. 229). For the present discussion I take these features of respect for equality for granted.

value of political equality. But this value can only be realized, i.e., honored or promoted, as a collective effort. This means that, within a given discursive forum, each discussant must contribute to the realization of political equality by treating each other with the respect that is due political equals. For example, suppose that A and B are debating abortion policy. If both A and B are open to taking each other's perspectives seriously, and (as I shall argue) if they refrain from insulting and harassing each other, then A and B realize the value of political equality: their joint effort to mutually respect each other as a political equal ensures that the value of political equality is realized between them. Suppose instead that A is forthcoming with her reasons and is open to listening to B, but B is rude and dogmatic. Here, mutual respect fails to obtain, and as such the value of political equality is not realized. For in this case, not all parties are treated as political equals, and political equality is realized only if all parties within the relevant forum enjoy respect from each other. Therefore, when B is rude and dogmatic, B fails to treat A with respect, and although A manifests her respect for B by being open-minded, political equality is not realized between them.

In the aforementioned, the collective setting in which respect for political equality is achieved involves only two individuals—A and B. However, respect for political equality can be much wider in scope, scaling up from dyadic interactions to group-based deliberative forums (e.g., a barbershop, family dinner, or citizen assembly), to the whole political community itself. In this, it follows that respect for political equality may fail to be realised by a political community as a whole, and yet it is secured within some of its parts. The reverse may also be true: respect for political equality might be realised by a political community as a whole, even though some of the discursive arenas that make up its constitutive parts fail to achieve mutual

respect.[3] Even so, whether the relevant collective setting for political equality is construed narrowly or broadly, participants in that setting contribute to realizing the value of political equality when they each endeavor to treat each other with respect—i.e., to take each other seriously as a source of claims and reasons.

2.2.2. Norms of Civility

Given that respect for political equality involves open-mindedness, it follows that close-mindedness, an unwillingness to listen to what others have to say and take it to account, undercuts one's respect for the political equality of others. Furthermore, behaviors embodying a close-minded attitude also undercut respect. Name-calling, insults, *ad hominem* attacks, demonization, and so on betray close-mindedness, so these behaviors undermine respect for political equality. Indeed, if A is prone to name-calling, browbeating, and hectoring B when they engage in political debate, A is less likely to take B seriously "as a source of reasons, claims, and perspectives" (Mansbridge et al., 2012, p. 11). B will plausibly feel slighted at A's rude remarks; he will believe that A is close-minded and not interested in taking his claims and concerns seriously. As a consequence, B might feel that A does not respect him as a political equal.

Civility is characterized by norms that specify distinctive ways one is to act if one wishes to express respect for political equality. These norms facilitate mutual respect for political equality, which in turn realizes the value of political equality. For example, for A to manifest her respect for B's equality, A may act in accordance with the norms of civility. Since respect

[3] Indeed, systems theorists note that for a deliberative system to secure values like mutual respect, its individual parts need not themselves secure mutual respect: "What might be considered low quality or undemocratic deliberation in an individual instance might from a systems perspective contribute to an overall healthy deliberation" (Mansbridge et al., 2012, p. 12). This line of thought is compelling, but as critics have pointed out, even though deliberative deficiencies in individual forums (such as a breakdown in mutual respect) might be accommodated by the deliberative system as a whole, these deficiencies might be harmful to citizens in these individual discursive forums. Hence, deliberative systems theorists need an account of the "appropriate normative criteria for determining when [deliberative deficiencies in individual forums] are legitimate" (Owen & Smith, 2015, p. 225).

involves openness to opposing views and perspectives, the norms of civility center around this openness. Thus, I propose the following two norms of civility:

Be open-minded: To be civil in political debate, one must be open-minded, meaning that one must be willing to take up and seriously engage with political standpoints distinct from one's own (Baehr, 2011). In other words, one must be disposed to listen to the reasons other discussants provide, reflect on one's own views, preferences, and interests on the basis of the reasons on offer, and revise one's views, preferences, and interests if reflection calls for it (Christiano, 1996, p. 117). As such, close-mindedness, the unwillingness to seriously consider opposing perspectives, is prohibited by the norm of open-mindedness.

Furthermore, open-mindedness requires that one endeavor to be sincere and forthright with the reasons and arguments that one takes to be compelling. To treat a fellow interlocutor as a political equal involves rendering one's arguments vulnerable to reasonable criticism, but this can only be successful if one is forthcoming with what one believes. Thus, the norm of open-mindedness also implies a requirement of sincerity in political debate.

Avoid bad faith behavior: The open-mindedness norm prohibits close-mindedness, but I also specify a related norm that requires citizens to avoid behaviors embodying a close-minded attitude. I refer to political speech and behavior expressing a close-minded attitude as "bad faith political speech and behavior." Bad faith political speech and behavior includes name-calling, hectoring, insulting, trolling, browbeating, discrediting, vilifying, and so on. Crucially, bad faith political speech and behavior does *not* include certain abrasive behaviors in discourse: for example, "heat and passion… rais[ing] [one's] voices, engag[ing] in sharp or biting rhetoric, and adopt[ing] an antagonist" but non-threatening posture toward others (Aikin & Talisse, 2020, p. 17). As Scott Aikin and Robert Talisse argue, civility does not require one to "maintain a posture

of calmness or politeness, or a pacifying and gentle tone of voice" (Aikin & Talisse, 2020, p. 17). Therefore, abrasive behaviors as those described above are consistent with an open-minded attitude, and so they do not violate the norm of avoiding bad faith behavior.

2.2.3. Political Incivility

Given the foregoing account of political civility, we can define the idea of political incivility. Political incivilities are defined as those attitudes, and behaviors embodying attitudes, that violate the norms of civility, i.e., open-mindedness and avoiding bad faith behavior. According to this view, close-mindedness in political discourse is an incivility because it is a violation of the open-mindedness norm. Similarly, bad faith political speech and behavior are political incivilities. As we have seen, such kinds of political speech embody a close-minded attitude: a person who engages in name-calling, discrediting, or trolling his political opponents displays an unwillingness to listen and take seriously what the latter have to say.

Importantly, political incivilities come in degrees. One degree of incivility might characterize a citizen that is close-minded but does not display it outwardly in bad faith behavior, but a higher degree could characterize another who is close-minded and name-calls and insults his political opponents. An even greater degree of political incivility may involve a citizen who not only is close-minded and engages in bad faith actions, but is also angry and hateful when he debates with his political opponents. In this paper, I develop and defend a principle of reciprocity according to which citizens are required to be civil in political debate on the condition that their interlocutors will reciprocate civility. Based on this principle, citizens are not required to maintain civility if they have reason to believe that their interlocutors will be *un*civil. That is, citizens are permitted to engage in incivilities if they have reason to believe that their interlocutors will be *un*civil. Despite this permission, however, reciprocity does not imply that

citizens have a "moral *carte blanche* to [act] however [they wish]" (Aikin & Talisse, 2020, p. 95). As such, not all instances of political incivilities are acceptable. As I shall argue, the permissible degrees of incivility are likely determined on a case-by-case basis (section 3.2.).

In what follows, I discuss possible explanations for the prevalence of political incivilities in contemporary political discourse. I first outline identity and group theories account of political incivilities. Then I discuss an account of political incivilities based on an argument from reciprocity.

3. IDENTITY & GROUP THEORIES OF POLITICAL INCIVILITY

3.1. Identity Theories of Political Incivility

According to identity theories, the driving forces behind political incivility are our (often irrational) responses to political disagreement. On this view, our political beliefs, values, and affiliations are a central part of who we are. For many of us these political beliefs and affiliations provide meaning and direction for how we navigate our lives as democratic citizens (Novaes, 2021). Often, our political views determine what we choose to do with our time and who we choose to associate with. Particularly for those who are actively engaged in politics—e.g., through voting, campaign building, volunteering, donating, etc.—political views and stances form a treasured aspect of their careers and personal lives.

Since our political beliefs are central to our personal identities, under identity theories, we are more likely to be guarded about them. Thus, any challenge to these beliefs and affiliations causes us to be uncomfortable (Novaes, 2021, p. 16). Indeed, we are more likely to consider political disagreement with our fellow citizens as personal attacks. This primes us to be combative, angry, and hostile when we discuss politics, for we fear that such a valuable aspect of our lives—our political beliefs and values—might be based in weak or false reasons.

Identity theories suggest, then, that instances of incivilities in political disagreement are unjustified. This is because for identity theories, political incivilities are an expression of discomfort with political disagreement. But even though political disagreement might be uncomfortable because it calls into question central aspects of who we are, this is not a good reason to be uncivil with of political opponents in debate. So, political incivilities cannot be justified on the identity theory model.

There are two problems with this view. First, identity theories assume that when we disagree combatively with our fellow citizens, our combativeness is a result of viewing their challenge as personal attacks. But if those positions we defend are not central to who we are, then the personal discomfort thesis is false. In contemporary discourse, it is conceivable that for many citizens their political beliefs play no central role in their personal identities. For example, many citizens debate politics as a fun intellectual exercise. For such persons, their incivilities might not stem from them feeling personally attacked.

The second, more pressing, issue with identity theories is that they are incomplete. In particular, they neglect cases of justified political incivility in which incivilities are directed at an interlocutor for violating a norm of civility. As I shall argue in the next section, the argument from reciprocity suggests that citizens have reason to follow the norms of political discourse on the condition that others do so too. If A is dogmatic and rude, thereby violating the norms of open-mindedness and avoiding bad faith behaviour, because of the reciprocity thesis B is justified in responding with sharp and combative remarks. Identity theories fail to capture this scenario, for they assume that all forms of combativeness in public discourse are a result of feeling attacked. Thus, identity theories are incomplete.

3.2. Group Theories of Political Incivility

According to group theories, political incivilities are a manifestation of tribal impulses. Tribal impulses are grounded in features of our social identities, and they embody brute psychological tendencies to favour persons who share a given social identity of one and to despise persons who do not share this identity. (Huddy et al., 2015; Iyengar et al., 2012). A social identity is defined as trait of a person which derives from that person's affiliation—and internalised sense of belonging—with a social group (Brewer, 1979). According to social identity theory, humans are social creatures, and for a variety of reasons we are disposed to create and form groups. The reasons for which we enter groups might be important, but they can also be trivial (Huddy et al., 2015, p. 3; Mason, 2018, p. 2). We form social groups to help meet our basic needs (e.g., food and shelter), and we enter social groups to realise intrinsic goods of love and friendship (Mason, 2018; Vallier, 2018, p. 17). But we also form groups on the basis of any shared characteristic, even the most trivial: for example, race, sports, religion, and one's political stances (Billig & Tajfel, 1973, p. 30). According to social identity theory, group membership—no matter its basis—gives us a sense of belonging, and this drives us to personally identify as a member of the group. As a consequence, the (social) identity of being a member of a group becomes central to who we are and what we do (Huddy et al., 2015, p. 3).

Social scientists argue, however, that the disposition to form groups faces limits. Once we have achieved cohesion with a certain group of people, we tend to form boundaries between those in this group (our "in-group") and those outside of it (our "out-group"). Thus, Marilynn Brewer argues that not only are humans driven by nature to form social groups: they are also driven to form *exclusive* social groups, separating the in-group from the out-group, often demonstrating hostility and disdain towards the latter (Brewer, 1979). Out-group animus, in

particular, is constituted by tribalistic distrust, suspicion, and sometimes hatred of opposing members of one's group. In the political context where resources for political power are scarce, we are so opposed to members of opposing groups that we take them to be inferior to us for no reason other than them being out-group members vying for the same resources and power (Achen & Bartels, 2016; Mason, 2018, p. 12).

If political incivilities are grounded in tribalistic distrust, then it appears that political incivilities are generally unjustified. For tribalistic distrust embodies an unjustified reason to be uncivil in political discourse. Tribalistic distrust is caused by brute tendencies to disdain one's political opponents, and this disdain for one's political opponents weakens the possibility of one taking their claims and concerns seriously; that is, it weakens one's commitment to treat them as a political equal. This constitutes a distinct kind of wrong against one's political opponents since they are owed respect as political equals. From this, it follows that tribalistic distrust seems to be the wrong kind of reason to be uncivil—for example, to be closeminded—with a member of an out-group, for example, a Conservative (or Liberal). To willingly act on this reason would be unjustified. Therefore, political incivilities as motivated by tribalistic distrust are unjustified.

This story is inadequate, however. It presupposes that the main reason why citizens are uncivil when they debate politics is because of underlying psychological drives: specifically, a brute drive to favour one's in-group while vilifying the out-group. But the problem with this approach is that it is silent on why certain group identities or affiliations—particularly, partisan identities—are more salient than others (for example, gender identity or sports team affiliation). Kevin Vallier nicely summarises this problem:

> [Group theories] tell an incomplete story about how our tribal psychologies are activated. Why are we retreating into new groups rather than

> sticking to our older, bigger group, namely the country as a whole? That is, why do we think we must stick to the red tribe or the blue tribe? Some institutional changes, such as an increase in publicly perceived racial and economic inequality, have helped activate tribal mindsets. But we need an explanation of why political identity has suffused American life so broadly in order to explain the degree of divergence in the United States (Vallier, 2020, p. 14).

Without an explanation of why partisan affiliation, as opposed to other group affiliations, is primed by our tribal impulses, it is not clear how much analytic traction group theories can offer for political incivility.

More worryingly, group theories suggest that political incivility is never justified. For political incivilities are grounded in tribalistic distrust, which constitutes a distinct kind of wrong against one's political opponents since they are owed respect as political equals. In what follows, however, I reject this conclusion. I defend an account of incivility that appeals to a reciprocity argument. This account both explains political incivility *and* justifies it. This account is important because, contrary to both identity and group theories, it demonstrates that citizens act rationally and justifiably by exhibiting uncivil behaviour in political debate.

4. THE ARGUMENT FROM RECIPROCITY

The argument from reciprocity comes in two stages. In the first stage, I explicate what the reciprocal character of civility involves. The second stage of the argument demonstrates how the reciprocal character of civility justifiably entitles citizens to political incivility.

4.1. Stage One

Aikin and Talisse suggest that a range of political incivilities can be justified by the reciprocal character of civility. This reciprocal character of civility can be characterised by a

principle of reciprocity according to which citizens in political debate are required to be civil as long as others are civil too. To grasp civility's feature of reciprocity, Aikin and Talisse appeal to another feature of civility: namely, that civility is a collective norm. A norm is collective when it applies to a group as a whole, not just individuals. For Aikin and Talisse, civility applies collectively in the sense that its standards govern citizens in a given democratic society as a group (Aikin & Talisse, 2020, p. 94). Under this picture, the demands of civility can only be met when citizens of a given democratic society generally follow the norms of civility. Admittedly, a given democratic society can still be civil even if a tiny minority of its citizens fails to abide by its civility standards, for example, by being rude and dogmatic in political speech. Even so, the point still stands that because the norm of civility governs citizens as a group, individual citizens are required to be civil because civility applies to them insofar as they are citizens.

Given that civility is a collective norm, it depends on most people generally following the norm. That is, its requirements can only be satisfied if a sufficient number of members of the collective follow the norm. As such, civility acquires its reciprocal character. The collective norm of civility is reciprocal in the sense that it is because it applies collectively—i.e., depends on most people generally adhering to the norm—that individual citizens have reason to follow it. As Aikin and Talisse put it, "[civility] prescribe[s] modes of conduct to us, collectively, in our role as citizens. Accordingly, individuals are required to [be civil] only when [civility is] embraced and generally practiced by the society at large" (Aikin & Talisse, 2020, p. 97).

To further illustrate this point that the collective application of civility grounds its reciprocal nature, consider Aikin and Talisse's analogy of the playground norm "keep your hands to yourself":

> "We teach children the policy "keep your hands to yourself" … But notice that the policy of keeping one's hands to oneself establishes a standard of conduct *for those on the playground*; more importantly, it is in virtue of its *collective application* that individuals are bound to comply with its requirements" (Aikin & Talisse, 2020, pp. 94–95; original emphasis)

For Aikin and Talisse, the norm "keep your hands to yourself" applies to a group, namely, the group of participants on the playground. Therefore, playground participants are required to follow this norm on the condition that others generally follow it too. That is to say, the norm "keep your hands to yourself" is a reciprocal norm. Indeed, to emphasise its reciprocal nature, Aikin and Talisse assert that the "keep your hands to yourself" norm is an abbreviation of the complex norm, "keep your hands to yourself on the condition that others are keeping their hands to themselves" (Aikin & Talisse, 2020, p. 95). By analogy, then, the norms of civility apply to citizens collectively, in their role as citizens. This collective applicability grounds their reciprocal character. Hence, individual citizens are required to be civil so long as others are civil too. In addition, we might notice the reciprocal character of the norms of civility by highlighting the fuller counterparts of the norms "be open-minded" ("be open-minded on the condition that others are open-minded") and "avoid bad faith behaviour" ("avoid bad faith behaviour on the condition that others avoid bad faith behaviour").

The explanation for reciprocity that appeals to collective applicability is on the right track, but Aikin and Talisse fail to provide a compelling argument for why civility is collectively applicable in the first place. On their view, civility is a public reciprocal virtue insofar as its standards apply collectively, and it contrasts with first-personal virtues like courage, which apply

to individuals and is not "contingent on the presence of other people"(Aikin & Talisse, 2020, p. 94). However, this distinction is not enough to vindicate the collective character of civility.

I argue that it is the collective nature of the value of respect for political equality that justifies civility's reciprocal character. As I argued above, respect for political equality is a collective achievement. This means that each participant to a democratic discussion must contribute to its realization. If a sufficient amount of people fails to contribute, then the collective endeavour to realise respect for political equality weakens, and accordingly the value will fail to be realised. I also argued that citizens contribute to the realisation of this value when they adhere to the norms of civility—i.e., when they are open-minded and avoid bad faith behaviour. On this view, the norms of civility apply collectively to discursive participants as whole because these norms are designed to secure the value of respect for political equality, which can only be collectively realised.

Therefore, because the value of respect for political equality can only be secured collectively, the norms of civility aimed at securing these collective values must gain obedience from everyone who participates. That is, since every participant has to contribute to the realisation of respect for political equality, and contribution to the collective realisation of this value is done by obeying the norms of civility, the norms of civility apply collectively to each participant.

Crucially, then, the reciprocity principle must hold if the value of respect for political equality will be realised. In other words, civility must be a reciprocal norm to realise respect for political equality. If two participants in a democratic debate follow the norms of open-mindedness and avoiding bad faith behaviour, then they can secure the collective achievements of respect for political equality. But if one of them defects, i.e., if one of them chooses to violate

the norms of civility, these collective realisation of respect for political equality will be lost, even if both the non-defector and defector might independently value these goods. Put another way, if one person chooses to defect, the non-defector's faithful adherence to the norms will not secure the value of respect for political equality, since a one-sided adherence to civility will fail to secure such values. Thus, unless each participant can trust that others will adhere to the norms of civility, they will have no reason to contribute to the realisation of these values.

4.2.Stage Two

The next stage of the argument is to show that the reciprocal character of civility entitles citizens in contemporary democratic societies to political incivilities. Because civility is a reciprocal norm, Aikin and Talisse argue, if one has reason to believe that one's interlocutor(s) in a deliberative interchange is or will be uncivil, then one has no reason to maintain civility with them. On this view, once B has good reason to believe that A will be uncivil, B is blameless for choosing not to maintain civility with A. Put another way, B is justified in being uncivil with A. Extending on their playground example, Aikin and Talisse remark:

> "[It] is in virtue of its *collective application* that individuals are bound to comply with ["keep your hands to yourself" norm] requirements. Consequently, when Adam violates the norm by grabbing Billy, and Billy retaliates, it would be absurd to criticize Billy for failing to keep his hands to himself. With Adam's violation, the collective norm is suspended, and in extricating himself, Billy does not himself *break* the rule. Indeed, Billy might nonetheless embody the virtue of being ungrabby; his action against Adam does not show otherwise"(Aikin & Talisse, 2020, pp. 94–95; original emphasis)

Here, because Adam has violated the norm of "keep your hands to yourself," Billy has no reason to keep following that norm. Crucially, because Adam violates the norm by grabbing

Billy, Billy has no reason to maintain the norm with Adam. Therefore, if Billy retaliates against Adam, "it would be absurd to criticize Billy for failing to keep his hands to himself" (Aikin & Talisse, 2020, p. 95). In a similar vein, the reciprocal character of civility vindicates political incivilities. Because one has no reason to maintain civility if one suffers incivility from another interlocutor, one seems permitted to retaliate in an uncivil manner in response. Recall A and B above. Suppose A upsets B by calling B names and insulting B. However, because civility is reciprocal in nature, and because B is not required to keep upholding civility with A, B is entitled (and not blameworthy) to respond sharply and rebuke A.

It therefore follows that citizens in contemporary democratic societies have reason to succumb to political incivilities within political discourse. The norm of civility applies to citizens as a whole, so if citizens have reason to think that the general public defects from civility, then citizens are not blameless for choosing to be uncivil in political discourse. As it turns out, however, given widespread incivility in contemporary political discourse, "ordinary citizens often have adequate reason to assume that their opposition has pulled back from the commitment to civility, and so they also have adequate reason to hold that the requirements of civility are no longer in play in public life" (Aikin & Talisse, 2020, p. 98). Thus, if the defective state of contemporary political discourse can de described in terms of widespread political incivility, then this defect not only has explication: citizens contributing to this defect are also justified.

Still, one might reject this conclusion because it seems to permit preemptive incivilities—i.e., incivilities that appear unprompted in the course of a give discussion. Even if it is true that the general public has defected from civility, this does not mean that in every political argument one finds themselves in, one ought to be preemptively *un*civil. To illustrate this charge, consider the following case:

> ***Vaccine***. In the workplace breakroom, Ben discusses his approval of government vaccine mandates with his co-workers. Ben thinks this is the best way to end the COVID-19 pandemic. Ava, Ben's co-worker, disagrees. She believes vaccines are a tool of social control. She gets irritable and calls proponents of vaccination, like Ben, "sheep." When colleagues try to carefully explain the problems with her views, Ava refuses to listen. Ben does not take the soft tactic like their friends. He does not mince words: he calls Ava's stances "stupid and dangerous."

In *Vaccine*, Ava and Ben direct incivilities at each other: Ava name-calls Ben by calling him "sheep," and Ben responds to Ava's remarks by calling her position "stupid and dangerous." According to the argument from reciprocity, Ava's preemptive incivility is rationally motivated and thus justified by her awareness of the fact that civility is not generally practiced by the members of her deliberative community. As such, Ava is justified in being *un*civil towards Ben from the start. And Ben is justified in being *un*civil in response.

In the next section, I specify a condition under which preemptive incivilities on account of the general public's defection from civility is justified. I argue that insofar as preemptive incivilities are a way to affirm one's status as a political equal, one is justified in being preemptively *un*civil in conditions of widespread defection from civility. That said, it is difficult to specify in advance the circumstances where preemptive incivilities affirm one's political equality. As a consequence, we might instead opt for a weaker conclusion according to which only Ben is justified in retaliating with incivilities to Ava's *un*civil behavior. The key point to note is that even if only the weak interpretation is more compelling, i.e., if only retaliatory incivilities (with some soon-to-be-specified qualifications) are permitted by the reciprocity

principle, we have still demonstrated that (retaliatory) incivilities can sometimes be rationally motivated, contra group and identity theories of political incivility.

In sum, the argument from reciprocity allows for the possibility that incivility is not always a pathology of irrationality. On this view, citizens are required to uphold civility insofar as others can be trusted to reciprocate civility, for only collective adherence to civility secures the kind of mutual respect that realizes the underlying value of political equality. In cases of non-reciprocity citizens are not required to maintain civility. Far from signaling irrationality, then, acting on one's positive reason to be *un*civil in cases of non-reciprocity is sometimes rational.

4.3. Some Qualifications

According to the reciprocity principle, citizens have no requirement to remain civil in instances of non-reciprocity. However, one might object that this does not imply having a reason to be *un*civil. It appears that the reciprocity argument supports the claim that, in instances of non-reciprocity, Ava and Ben each *lack reason to be civil*, but it does not seem to go as far as supporting the claim that, in those same instances, Ava and Ben each *have a reason to be uncivil*. As such, the reciprocity principle is vulnerable to an objection about scope: the principle does not seem to go as far as granting that one has positive reason to be *un*civil in political discourse.

To illustrate this charge, consider that Ava and Ben have a variety of options available besides incivility, even if they rightly judge that they have no reason to maintain civility in cases of non-reciprocity. For example, Ava and Ben might elect to remain civil in cases of non-reciprocity. In remaining civil in debate, they might employ a number of communicative strategies. For instance, drawing on recommendations from Elizabeth Anderson on communicating moral concern amidst political division, Ben and Ava might opt for sympathetic modes of communication, avoiding inflammatory speech and using more precise language to

capture what might be harmful or unfair about the COVID-19 vaccine mandates (Anderson, 2022, p. 76, 80).

Another possibility would be to disengage from certain kinds of political conversations. As Robert Talisse argues, taking a healthy distance from politics requires occasionally stepping away to remind oneself that one's political commitments are vulnerable to reasonable, if not decisive or formidable, criticism (though it does not entail suspending one's political judgement or avoiding politics altogether) (Talisse, 2021, p. 127). So, both Ava and Ben might choose to refrain from certain political discussions rather than remaining engaged in an *un*civil manner.

These foregoing options are acceptable in instances of non-reciprocity, but the question is whether incivility is at all permitted in addition. I contend that a *principle of resistance* implicit in the principle of reciprocity plays a role in permitting some instances of incivility. The key thought here is that in cases of non-reciprocity the principle of reciprocity implies "resisting the evil we actually receive" (Becker, 1986, p. 98).[4] To bring this principle of resistance into view, consider Lawrence Becker's distinctions between active or passive acceptance and active or passive rejection of evil received in cases of non-reciprocity (Becker, 1986, p. 97). Active acceptance implies returning good for evil received, and passive acceptance implies "making no protest at all" against evil received (Becker, 1986, p. 97). In contrast, active rejection involves fighting or struggling against an evil, and passive rejection involves merely refusing to submit to it. Importantly, active rejection does not imply repudiating evil *with evil* (Becker, 1986, pp. 96–97).[5]

[4] More accurately, Lawrence Becker advances that "given the 'responsive' nature of the reciprocity disposition generally, it seems appropriate to restrict the disposition to the aim of resisting the evil we actually receive" (Becker, 1986, p. 98).

[5] As such, based on the principle of resistance, not all incivilities are evil.

I assume that an evil is received in cases of non-reciprocity regarding the norms of civil discourse. This evil can be characterised by the message that is conveyed when one's interlocutor has (non-reciprocally) defected from political civility: namely, that one is not worth taking seriously as a political equal. This evil can be interpreted to arise from the general public: when the general public has defected from political civility, this sends the message to ordinary individual citizens that most citizens are *un*concerned to respect them as a political equal. On the other hand, this evil can also arise within smaller dyadic (or more group-based) deliberative exchanges, as when one's interlocutor is rude and dogmatic and so fails to respect one as a political equal.

If this assumption is right, then choosing to remain civil in conditions of non-reciprocity characterises a kind of active acceptance of this message. On the other hand, walking away from certain kinds of political conversations is passive rejection, but choosing to be *un*civil —for example, Ava calling Ben a "sheep," and Ben calling her views "stupid and dangerous"—constitutes active rejection of the message that one is not worth taking seriously as a political equal.

In the context of contemporary democratic politics, then, individuals are permitted and thus have a reason to engage in incivilities insofar as incivilities constitute *an active rejection of the message that they are not worth the respect due political equals*. Hence, in situations of non-reciprocity, individuals not only *lack the requirement to uphold civility*: they are also *permitted and so have reason to be uncivil*. Notice that this conclusion has both a weaker and stronger interpretation. On the stronger interpretation, the justificatory scope of the reciprocity argument encompasses preemptive incivilities, such as Ava being preemptively *un*civil with Ben. Here, the message that one is not a political equal is understood to come broadly from a society that fails to

heed the norms of civility. As such, Ava is justified in being *un*civil because she attempts to preemptively resist the message that she is of inferior status. In her mind, in order to stand up for her political equality, she needs to be *preemptively un*civil with Ben.

That said, one might plausibly object that preemptive incivility is misguided. Even if the general public defects from civility, this does not warrant Ava to be *uncivil* with Ben *in particular*. Therefore, we might opt for the weaker interpretation according to which only Ben is justified in being *un*civil, as his incivility is in response to Ava's preemptive incivility. Ava's initial incivility directly conveys to Ben the message that he is inferior, and in actively rejecting this message—for example, by calling her views "stupid and dangerous"—Ben affirms his own status as an equal who deserves to be taken seriously.

One qualification is in order. Whether both Ben and Ava, or only Ben, have reason to be *un*civil, having this reason does not imply that one has free rein over the range of incivilities at one's disposal (Aikin & Talisse, 2020, p. 95). So, the scope of permissible incivilities is limited. Admittedly, it is difficult to determine the scope of permissible incivilities *a priori*. As Lawrence Becker notes, "[what] counts as a justifiable act of resistance is…left to case by case considerations" (Becker, 1986, p. 98). Likewise, Aikin and Talisse argue that "it requires judgment and some context to determine where the line is between the acceptable and unacceptable" (Aikin & Talisse, 2020, p. 95). Moreover, what counts as an appropriate political incivility might differ from one cultural context to the other. For instance, in many parts of the US it is *un*civil to bring up controversial political discussions at the dinner table (Schaupp, 2011, p. 24). Accordingly, the permissibility of name-calling, hectoring, insulting, trolling, browbeating, discrediting, vilifying, and so on is subject to contextual considerations.

That said, one plausible *a priori* constraint on political incivilities draws on the value of political equality.[6] Under this view, only the degrees of incivilities that are consistent with asserting one's status as a political equal is justified. On this view, it may well turn out that preemptive incivilities are *un*justified: it is not clear that resisting the general public's message that one is inferior is a way of reasserting one's political equality. Here, the value of political equality might require that one be more judicious about the means by which one resists the general public's message.

Preemptive incivilities may be ruled out, but other controversial kinds of incivilities may be permitted. I shall defend one case in which this holds, namely, that case in which anger or outrage is a justified incivility in democratic discourse.

Political philosophers have compellingly argued for one such instance of justified outrage: resisting injustice and oppression. An important example is found in Malcolm X and the Civil Rights movement of the 1960s. Malcolm X was known for his anger-infused rhetoric which contrasted with the calmer, less abrasive tone of other Civil Rights movement leaders like Martin Luther King Jr. Consider the following excerpt from Malcolm X's *The Ballot or the Bullet* 1964 speech:

> If we don't do something real soon, I think you'll have to agree that we're going to be forced either to use the ballot or the bullet. It's one or the other in 1964. It isn't that time is running out—time has run out! 1964 threatens to be the most explosive year America has ever witnessed. The most explosive year. Why? It's also a political year. It's the year when all of the white politicians will be back

[6]Another plausible *a priori* constraint draws on principles of proportionality with respect to defensive harm. Principles of proportionality determine the appropriate level of defensive harm one can level at a person who is liable to receive this harm. For discussion on the morality and proportionality of defensive force, see Jonathan Quong (2020, especially ch. 4).

> in the so-called Negro community jiving you and me for some votes. The year when all of the white political crooks will be right back in your and my community with their false promises, building up our hopes for a let-down, with their trickery and their treachery, with their false promises which they don't intend to keep. As they nourish these dissatisfactions, it can only lead to one thing, an explosion (Malcolm X, 1964).

In this quote, we see Malcolm X's frustration with white politicians who visit black communities to campaign for support. According to Malcolm X, these politicians engage in "trickery" and "treachery," making "false promises," presumably about improving the conditions of black communities, but never fulfilling them. This leads him to think the only viable option to liberation is the bullet (although, to be sure, Malcolm X might not favour this tactic. Perhaps his anger leads him to view things in this way; that is what I contend needs justification).

We might plausibly interpret Malcolm X's claims as directed towards a systematic violation of the norms of civility by the white community as a whole. To grasp this claim, we need to imagine that systematic racism leads to systematic discounting of racially oppressed peoples' preferences and interests. Such systematic discounting shows up in the context of deliberation between the dominant group and the oppressed group. It arguably does so in either of two ways. First, systematic racism excludes blacks from the public sphere, and such exclusion makes it the case that whites and blacks will seldom interact in political discussion to exchange reasons and arguments. As a consequence, systematic racism causes whites to not pursue the norm of open-mindedness with blacks: racism excludes blacks from engagement in the public sphere, so even if whites are willing to listen respectfully to the reasons and arguments that

blacks offer, they might not be able to hear these reasons in the first place because of systematic exclusion.

Secondly, systematic racism leads white communities to violate the norms of civility directly and intentionally. Due to their racism, they might think that the reasons and concerns of blacks are not worth listening to. That is, systematic racism might lead whites to think that blacks are not worthy of respect and so are not worth taking seriously as equally authoritative sources of reasons. Thus, with respect to the norm of open-mindedness, systematic racism might cause a significant portion of whites to be unwilling to listen and engage with the reasons and arguments of black people. With respect to avoiding bad faith behaviour, systematic racism might lead whites to be explicitly racist with blacks, since they genuinely view them to be less than a political equal.

Therefore, we might plausibly take Malcolm X's claims as responding to a violation of the norms "be open-minded" and "avoid bad faith behaviour" by the white community as a whole, on account of their systematic racism. Given the reciprocity argument, Malcolm X is entitled to incivilities. The question, then, is: Is Malcolm X's outrage on account of white incivilities justified?

Many oppose the angry character of Malcolm X's speeches, especially given their summons to violence. One could possibly argue that the anger and outrage expressed by Malcolm X is wholly unjustified: angry, violent rhetoric cannot be condoned, even if it is in response to severe injustice. However, there are two reasons to resist this claim.

First, a call for more civility—or, more precisely, less anger-filled incivility—in this context might amount to silencing an oppressed person from rejecting their oppression. This is especially harmful in cases where marginalised persons lack the concepts and vocabulary to

express their concerns, and such rage-induced rhetoric might be one useful and appropriate substitute. As Maxime Lepoutre argues, "the pool of conceptual resources a society has for conferring meaning on activities and experiences ends up being especially adapted to dominant groups' experiences" (Lepoutre, 2021, p. 23) As a result, members of disadvantaged groups might "struggle to make sense of and to articulate their group- specific experiences" (Lepoutre, 2021, p. 23). And even if members of disadvantaged groups have the conceptual tools to articulate their problems, calls for more civility might hinder them from "appealing to concepts that are needed adequately to express important considerations or concerns" (Lepoutre, 2021, p. 23). If this is right, then Malcolm X's rageful rhetoric might be an appropriate empowering tool in combatting racial injustice.

Yet, it is not exactly clear that rageful rhetoric is an empowering tool against injustice. As Martha Nussbaum argues, for instance, anger is counter-productive in repudiating injustice. One reason, among many, is that anger "increases the other party's anxiety and self- defensiveness," and thereby "does nothing to move matters forward" (Nussbaum, 2016, p. 230). However, as Maxime Lepoutre argues, this objection overlooks a distinctive epistemic value of anger. According to Lepoutre, anger-infused rhetoric aimed at repudiating injustice can, in the right conditions, be a useful epistemic tool in shedding light on injustice and holding persons accountable. In particular, "anger can help [one's discussants] imaginatively experience what it is like to be in the speaker's shoes, how the world feels or appears from where they stand. Put differently, anger-infused narratives can enable the audience to empathize more fully with the speaker" (Lepoutre, 2021, p. 61). As Lepoutre maintains, "Malcolm X's rage foregrounds important properties of the injustice black Americans encounter." For example, Lepoutre argues that Malcolm X's *The Ballot or the Bullet* speech highlights the systemic and inescapable nature

of American racism, something that the average white person might not have gotten insight about were it not for Malcolm X's anger-filled speech. Moreover, Lepoutre argues, Malcolm X's anger challenges dominant ideas of what counts as reasonable and civil behaviour in democratic politics. He notes that "[by] highlighting patterns of exclusion, Malcolm X's anger helps make rational sense of the apparently unreasonable attitude which consists in embracing violent action, even when the odds of success are unfavourable. If reforming the American system from within is hopeless, and if one's situation is no better than jail, then violence aimed at putting one's opponents in their place may seem the best option" (Lepoutre, 2021, p. 69).

For these reasons, Malcolm X's anger-filled incivility seems to be justified. In the first place, his is a response to antecedent deliberative violations from the white community. The response is intense outrage, so we might be inclined to claim that such outrage is never justified. However, in the second place, given the epistemic value involved in highlighting racial injustice, Malcolm X's response seems fitting and appropriate.

Again, this is a defence of just one instantiation of anger. It might turn out that other instances of anger are wholly unjustified. Yet, nuanced instances of anger, such as an incivility that is aimed at rejecting injustice, might be justified by the argument from reciprocity.

5. THE UPSHOT

If the argument from reciprocity holds, then we have a compelling argument for the claim that citizens act rationally by exhibiting uncivil behaviour in political debate. This conclusion repudiates identity and group models of political incivility according to which citizens are unjustified in being uncivil as their incivility is grounded in either personal discomfort or tribalistic distrust, which constitutes a wrong. The argument from reciprocity provides us with a

plausible alternative motive behind incivility that not only explains widespread political incivility but also justifies it.

The upshot is that if citizens are blameless for the defective state of political discourse, then strategies to ameliorate this state must account for this finding. In particular, strategies must account for the argument from reciprocity, according to which citizens will not be civil unless they can trust that other citizens will uphold civility like them. Thus, if a reform strategy is aimed at building more civility, it must also be concerned to build more trust that others will reciprocate civility.

To build citizen trust, however, citizens need to observe that civility is indeed upheld. But for citizens to observe reciprocal civility, there must be trustworthy citizens who publicly uphold civility—i.e., citizens who are motivated to reciprocally uphold civility in their discursive interactions. Yet, to get such trustworthy citizens, we need to develop trustworthiness in society. That is, we need to promote the virtue of being civil in political discourse so that other citizens can trust that civility is upheld (and will be reciprocated) by members of the public. As such, building trust and trustworthiness are inextricably linked. To build more civility in political discourse, we need to build citizen trust that others will reciprocate civility, but to build this trust we need to build this civility. One mechanism must accompany the other.

To conclude this chapter, I outline an analysis of the type of trust that is needed for citizens to engage civilly with one another. I then explain why developing trustworthiness—i.e., developing the virtue in citizens to be civil—is important for developing trust.

6. DELIBERATIVE TRUST AND DELIBERATIVE TRUSTWORTHINESS

I contend that *deliberative trust* is the type of trust that citizens must exhibit in order to feel assured that their interlocutors in political debate will be civil. Deliberative trust is defined

as the belief citizens hold about other citizens that each is (or cares and tries to be) a civil interlocutor. This entails that within a given interaction, one believes one's co-deliberator conscientiously adheres norms of civility, including be open-minded and avoid bad faith behaviour, such as insults, vilifications, and so on.

There are two key features to note about deliberative trust. First, deliberative trust consists of the belief that other citizens are civil in political debate. This belief leads to positive attitudes and feelings towards one's political opponents (cf. Govier, 1992). These feelings imply the absence of dislike towards an interlocutor with an opposing political viewpoint. They instead seem to communicate respect. Furthermore, these positive attitudes that are constitutive of deliberative trust help to maintain and enrich deliberative interactions. Thus, if A deliberatively trusts B, she is likely to be respectful and kind towards B. This means that she treats B with goodwill: she refrains from vilifying and launching invectives at B and is charitable when they exchange reasons. A's deliberative trust might also manifest in open-mindedness towards B's views on a range of political issues. Open-mindedness involves a willingness to take up and seriously engage with political standpoints distinct from one's own (Baehr, 2011). Since A thinks B is a civil interlocutor, she might be more willing to see things from B's perspective on other issues that they might not at first agree on. This might enlighten her to her own bias and limited point of view. And, more importantly, her deliberative trust might sustain a civil deliberative exchange between her and B, where each feels like the other conveys respect in the conversation.

Understood this way, deliberative trust that others will be civil, that is, follow the norms of be open-minded and avoid bad faith behaviour, actually leads one to be civil themselves. In other words, deliberative trust grounds positive feelings that encourage both open-mindedness

and a desire to be kind and treat one's interlocutor with good will and respect. Since open-mindedness and avoiding bad faith behaviour (by treating one's interlocutor with good will and respect) are constitutive of civility on this conception, deliberative trust leads to more civility.

Secondly, deliberative trust possesses a legitimacy condition, according to which, for deliberative trust to be valid, it must be placed in actual deliberatively trustworthy citizens. *Deliberative trustworthiness* is defined as the virtue of being civil, i.e., a sincere and conscientious follower of the deliberative norms of open-mindedness and avoiding bad faith behaviours. The conditions in which A legitimately deliberatively trusts B must be those where B is actually deliberatively trustworthy: that is, instances where B cares about civility norms and tries to observe them as much as possible. If A is tricked into deliberatively trusting B when B is in fact deliberatively *un*trustworthy—for example, B is insincere in his arguments and frequently lies and disinforms others—then A's trust is misplaced. The danger of placing trust in untrustworthy citizens, as we shall see in chapter 2, is that deliberative trust must be grounded in accurate beliefs about one's co-discussant, or else it is unstable and might further deepen political incivility. For example, once A realises that her trust is misplaced in *un*trustworthy B, she is likely to feel betrayed and resentful towards B. These feelings of betrayal and resentment might then lead her to be uncivil with B, to insult and vilify him. Although she might be justified in doing this, A might go on to develop a distrust of others in general and reinforce her feeling entitled to be hostile and disrespectful in discussions around politics. The result is that A's discovery of her misplaced trust might dispose her to even more incivility in her deliberative exchanges. So, it is crucial that deliberative trust actually tracks deliberative trustworthiness, for misplaced trust might exacerbate rather than mitigate political incivility.

In this, it follows that to ameliorate the problems associated with political incivility, what primarily matters is the cultivation of deliberative trustworthiness, i.e., the virtue of caring and following discursive norms. In chapter 2 I elaborate more on this point, but for now suffice it to say that for empirical strategies geared at building more trust, such strategies must prioritise developing deliberative trustworthiness. This is because for deliberative trust to be legitimate, it must be grounded or placed in actually trustworthy agents. With the cultivation of deliberative trustworthiness, citizens develop the (deliberative) trusting belief that other citizens care about being civil in political discourse. Accordingly, they will come to see themselves as having reason to be civil in political discourse.

7. CONCLUSION

Much contemporary political discourse in the US and industrialized west is riddled with political incivilities. Political incivilities are defined as those attitudes and behaviours embodying attitudes that convey a disregard for the equal political status of democratic citizens. When citizens discuss politics on social media, in the workplace, or at Thanksgiving dinners, Political incivilities are those disrespectful forms of political speech such as name-calling, insulting, trolling, demonising, vilifying, discrediting.

In this chapter, I defended an account of political incivility that appeals to an argument from reciprocity. This account both explains *and* justifies political incivility. According to the argument from reciprocity, civility is a reciprocal norm, so citizens in political debate follow a principle of reciprocity with respect to the requirements of civility. This principle of reciprocity states that citizens have reason to be civil only on the condition that others are civil too. As Maxime Lepoutre explains, "plausibly, I will not commit to respectfully constraining the claims I

make unless I feel that I can rely on my interlocutors also to fulfil a commitment to discursive civility" (Lepoutre, 2021, p. 134).

But if this thesis about reciprocity is right, then citizens are *blameless* for going against the standards that civility sets, especially given the defective state of contemporary discourse. In other words, they are justified in being uncivil given the current state of prevailing uncivil discourse. This directly contrasts with identity and group theories of political incivilities, which explains political incivilities in terms of personal discomfort or intertribal prejudices. By investigating an account of political incivilities based in the reciprocal structure of civility, we have the basis for a more persuasive explanation of political incivilities: in particular, this account is persuasive because it supports the controversial claim, contra identity and group theories of political incivilities, that citizens act rationally by exhibiting uncivil behaviour in political debate.

The upshot is that if citizens are blameless for rampant incivilities in political discourse, then strategies to fix this defect must account for this finding. First, strategies aimed at building civility must specifically focus on cultivating trust that others will reciprocate civility. Second, such strategies must also work to encourage the motive in citizens to be civil in discourse so that other citizens can trust that civility is generally embraced (and will be reciprocated) by members of the public. In other words, strategies must not only focus on building citizen trust: they must also focus on developing citizen trustworthiness. In chapter 2, I further analyse these concepts of deliberative trust and trustworthiness.

CHAPTER TWO: DELIBERATIVE TRUST OR SOCIAL TRUST?

1. INTRODUCTION

In chapter 1, I defended an account of political incivility that appealed to an argument from reciprocity. This account both explains *and* justifies political incivility. According to the argument from reciprocity, civility is a reciprocal norm. So, citizens in political debate follow a principle of reciprocity with respect to the requirements of civility, whereby citizens have reason to be civil only on the condition that others are civil too. As Maxime Lepoutre explains, "plausibly, I will not commit to respectfully constraining the claims I make unless I feel that I can rely on my interlocutors also to fulfil a commitment to discursive civility" (Lepoutre, 2021, p. 134). If this thesis about reciprocity is right, then in a culture where political incivility prevails, citizens of that culture are *blameless* for going against the standards that civility sets. It follows that citizens in contemporary democratic societies are often justified in being uncivil given the current state of rampant incivility in public deliberation.

If citizens are justified in being uncivil, tactics aimed at fixing the problem of widespread incivility must account for this finding. More specifically, strategies aimed at building civility must focus on cultivating *deliberative trust* that others will reciprocate civility. In this vein, I further elaborate on the concepts of deliberative trust and trustworthiness in this chapter. Deliberative trust is defined as the belief citizens hold about other citizens that each is a civil interlocutor. This entails that within a given interaction, one believes one's co-deliberator conscientiously adheres norms of civility, including open-mindedness and avoiding bad faith behaviour, such as insults, vilifications, and so on. Deliberative trustworthiness is defined as the virtue of being civil, i.e., a conscientious follower of the deliberative norms of open-mindedness and avoiding bad faith behaviours.

Kevin Vallier has recently defended a contrasting theory of trust. In two books, *Must Politics be War* and *Trust in a Polarized Age*, Vallier suggests that a cultivation of *social trust* is an antidote to widespread political incivility. Vallier defines social trust as the belief, held by each member of society, that other members of society will follow publicly recognized social-moral rules for moral reasons. Vallier's focuses on social-moral rules such as "do not physically assault others," and "keep your promises," and the general rules that constitute liberal rights practices (e.g., free speech, free association, and so on). On this conception, social trustworthiness is understood as the virtue of adhering to social-moral rules because one has sufficient moral reason to do so (Vallier, 2020, p. 23). According to Vallier, high social trust leads to less extreme political opinions, secures consensus for public policy, results in more patience to let social reforms "work themselves out before complaining," and makes it easier for people to adjust to demographic changes (Vallier, 2020, p. 69).

One important feature of Vallier's theory is that the social-moral rules that form the basis of social trust and trustworthiness must be publicly justified. In a publicly justified social order, a given coercive social norm, rule, law, policy, or action is justified if and only if each and every member of the public has a conclusive reason R to endorse that norm, rule, law, policy, or action (Vallier, 2011, p. 262). If the general network of social-moral norms in a given social order is publicly justified, this implies that members of the public have sufficient reason to endorse and thus follow the norms that animate that order. Thus, a crucial consequence of public justification is that, in a publicly justified order, each citizen has reason to be socially trustworthy because they can see themselves as having sufficient reason to follow social moral rules. Recall that social trustworthiness is understood as the virtue of adhering to social-moral rules because one has sufficient moral reason to do so (Vallier, 2020, p. 23). If this right, then in a public justified

societal order, where all generally endorse and abide by the social norms of that order, each citizen will see themselves as having reason to be trustworthy.

My aim in this chapter is to qualify and extend to the sphere of public discourse Vallier's theory of social trust and trustworthiness. In doing so, I further explicate the concepts of deliberative trust and trustworthiness. I argue that, to address the problem of political incivility, we need a kind of trust that is more distinctive than social trust, but that also rests on public justification. As it stands, Vallier's theory of social trust and trustworthiness is too thin to capture the contours of public deliberation amongst citizens. The account is anchored around the norms that animate liberal rights practices, such as "do not physically assault others," and "keep your promises," but these norms do not directly apply to the context of political deliberation between citizens. As a result, it is not clear how social trust, understood as grounded in thin publicly justified liberal rights norms, can be the kind of trust that is pertinent to restoring civility in public discourse. Therefore, I argue that we need a substantive account of the deliberative norms—i.e., norms that regulate political discussion—that anchor political discourse and that also depend on public justification. These deliberative norms are the norms of civility, and they ground a *deliberative* kind of trust and trustworthiness.

Even though the present account of deliberative trust differs from Vallier's, there are still important similarities. In particular, a key feature that the theory of deliberative trust and trustworthiness borrows from Vallier's theory is *public justification*. Public justification plays a key element in securing the trust that others will reciprocate adherence to the norm of civility. Recall that, according to the principle of public justification, a given coercive norm is justified if and only if each and every member of the public has a conclusive reason R to endorse that rule, law, policy, or action (Vallier, 2011, p. 262). So, if the norms of civility are publicly justified,

then this means that citizens will follow the norms of civility for their own reasons. This is crucial if these citizens are to be deliberatively trustworthy, for an important condition on the virtue of deliberative trustworthiness is that one sees themselves as having sufficient reason to follow the norms of civility.

I conclude the chapter by defending a consideration that empirical strategies geared at building more trust and civility must bear in mind. I argue that such strategies must prioritise developing deliberative trustworthiness, i.e., the virtue of caring and following discursive norms, in addition to cultivating deliberative trust. As we shall see, for deliberative trust to be legitimate, it must be grounded or placed in actually trustworthy agents. With the cultivation of deliberative trustworthiness, citizens develop the (deliberative) trusting belief that other citizens care about being civil in political discourse. Accordingly, they will come to see themselves as having reason to be civil in political discourse.

In section 2, I provide an overview of Kevin Vallier's theory of social trust and social trustworthiness. Section 3 applies this theory of social trust to the context of political discourse between citizens. In this section, I argue that Vallier's theory of social trust and trustworthiness as grounded in liberal rights social-moral norms is too thin a basis of trust that one's interlocutors will reciprocally follow the norms of civility. In section 4, I then discuss the distinct kind of trust and trustworthiness that is relevant for discursive reciprocity: deliberative trust and trustworthiness. I also discuss how public justification is a crucial requirement on deliberative trust and trustworthiness if they are to succeed in resolving the issue of incivility in political deliberation. Finally, I discuss the importance of basing deliberative trust in deliberative trustworthiness.

2. KEVIN VALLIER'S THEORY OF SOCIAL TRUST

2.1. Preliminaries

Before I canvass Kevin Vallier's theory of social trust, it is important to address the kinds of issues that his theory is aimed at addressing. The central cause for concern for Vallier is the problem of partisan divergence. Partisan divergence—or what is also known as 'political polarisation'—denotes the constellation of phenomena relating to the increasing political distance between citizens in the industrialized West, especially in the US. Phenomena involved in partisan divergence include issue- and affect-based polarisation as well as political sorting (Vallier, 2020, p. 4). Issue-based polarisation means that citizens increasingly hold diverging policy attitudes and opinions (McCarty, 2019, p. 9; Vallier, 2020, p. 3). Affective polarisation means that citizens increasingly dislike and distrust each other (Iyengar, 2021; Iyengar et al., 2012; Iyengar & Krupenkin, 2018; Vallier, 2020, p. 3). Political sorting is understood as the increasing association of distinct social identities with partisan identity (Mason, 2018, p. 18; McCarty, 2019, p. 15). For example, we find that Republican partisans in the US tend to be increasingly conservative (ideological identity), white (racial identity) and evangelical (religious identity) (Mason, 2018, p. 18).

Given that Vallier's concern is partisan divergence in general, his theory of social trust is likely not intended to single out political incivilities as the main issue of concern. *A fortiori*, his theory is probably not geared to specifically address political incivilities *within* public discourse. To avoid discussing a strawman, then, the present study reads Vallier's theory of social trust as a potential solution (likely unintended by Vallier) to the question of incivilities pervasive in democratic discourse. Thus, I develop my account of deliberative trust and trustworthiness based on this reading of Vallier.

2.2.Vallier's Account of Social Trust

2.2.1. The Domain of Social Trust

Vallier begins his account by focusing on *social cooperation* as an important site of social trust and trustworthiness. A social cooperative scheme is defined as a scheme in which members of a collective C rely on each other to achieve, through cooperation, both individual and collective goals, commitments, or projects. For example, if A needs to build a house but lacks the expertise, she can rely on a building contractor, B, to finish this job for her. According to Vallier, social-moral rules are a crucial component of social cooperation between the members of a collective C. A social-moral rule is defined as "a publicly recognized prescription for a group of persons to engage, or to be allowed to engage, in certain lines of conduct according to the relevant context" (Vallier, 2018, p. 32). Social-moral rules are important because they facilitate and regulate social cooperation between the members of C. In addition, they are backed by empirical expectations that persons generally comply *and* normative expectations that persons ought to comply (Vallier, 2018, p. 34). So, persons are expected to comply as a matter of statistical prediction (empirical expectations). And if a person refuses to comply with a social-moral rule, then they can be held accountable (because of normative expectations).

Vallier's account of social-moral rules begins at the point of social cooperation, but it does not end there. Importantly, social-moral rules constrain other kinds of actions persons can take, so social-moral rules that guide cooperation must go beyond dividing social roles of labour and establishing rules pertinent to one's roles in the cooperative endeavour. Thus, persons are required to "be honest, not to harm others without cause, not to kill the innocent, to keep promises, to show gratitude, to aid the impoverished," and so forth (Vallier, 2020, p. 24). More generally, the constraints prescribed by social-moral rules stem from a society's *rights practices*.

These rights practices constrain how we treat each other in our everyday interactions, both inside and outside cooperative schemes. For example, A's right to physical security requires that I refrain from hitting him as he walks across from me on the street. The right to assistance is enjoyed by everyone, so if I see a drowning child in a pond I have a duty to jump in and save them. In addition to the practices implicated above, other examples of rights practices include free speech rights and property holding rights (Vallier, 2020, p. 22). Free speech allows individuals to speak their minds without fearing coercive sanction from other citizens and from the state, and property rights protect individual private property from coercive interference from others and from the state. We can therefore see Vallier's account as encompassing both cooperative *and* non-cooperative interactions between members of society.

2.2.2. The Focus of Social Trust: Adherence to Social-Moral Rules

As mentioned, if a person violates a social-moral rule—for example, if I hit A in the face as he walks across from me—they can be held accountable for doing so. The practice of holding others accountable is implied by the meaning of the term "social-moral." This term is meant to capture the fact that these rules are 1) social in the sense of prescribing (permitting, requiring, or prohibiting) a set of actions for a set of people, and 2) moral insofar as they are categorical (i.e., we follow them because they are right and not just because they are socially beneficial) and are central to accountability practices, such as blame or punishment. Importantly, social-moral rules justify the reactive attitudes of indignation and resentment we have when others violate these rules. Thus, when A cooperates with B to achieve the goal of building a house, she relies on social-moral rules, compliance with which facilitates the myriad tasks associated with house-building, which A trusts B to implement (Vallier, 2018, p. 28). But if B violates one of these rules—say, B performs a shoddy job due to laziness and incompetence—then A can rightly

blame and resent him for his actions. She can simply reference, as grounding her castigations, the social-moral rule that requires B to execute his part of the plan correctly.

Given this background idea of social-moral rules, Vallier goes on to define the concept of social trust. On Vallier's view, when A believes that B will help her achieve her goals by following social-moral rules, A is said to socially trust B. More precisely, Vallier defines social trust as follows:

> ***Social Trust***: The belief, held by each member of society about society in general, that other members of society follow social-moral rules for moral reasons (Vallier, 2020, p. 24).

Under this conception, for A to trust B (whether in a cooperative endeavour with B or in a different context), A must not only believe that B is willing and able to abide social-moral norms: she must also believe that B is motivated to do so because of moral reasons. This raises a question of how to understand social trustworthiness. Vallier defines social trustworthiness as follows:

> ***Social Trustworthiness***: The virtue of adhering to social-moral rules because one has sufficient moral reason to do so (Vallier, 2020, p. 23).

According to this definition, trustworthy members of society care about moral rules, and their concern for these rules arises out of the moral reasons that drive them to care. One crucial feature of Vallier's theory of social trust and trustworthiness is that social trust must be rational. For Vallier, social trust is rational when it is formed on the basis of evidence of trustworthiness. As a consequence, the validity or legitimacy of social trust depends on whether it is well-placed, i.e., placed in trustworthy citizens.

For Vallier, social trust must have this rational, evidence-based ground because trust cannot survive awareness of the fact that it is based in poor evidence "and so will be less stable than trust based in good evidence" (Vallier, 2020, p. 25). Moreover, if a person's trust is misplaced—i.e., placed in an *un*trustworthy citizen and grounded in false beliefs about the latter's moral motivations—this leaves them vulnerable to deceit and domination. Thus, if persons are trustworthy in the sense that they abide by moral rules for moral reasons, then citizens can be able to *trust them for the right reasons* (Vallier, 2020, p. 23). On this conception, the right reasons that motivate social trust are grounded in the belief that others abide social norms because of their moral reasons. So, citizens' trust must be rationally based in the actual trustworthiness of other citizens.

2.2.3. Moral Reasons, Social Trust, and Social Trustworthiness

But what are moral reasons? For any given member of society, Vallier defines a moral reason to be a consideration that motivates compliance with social-moral rules (Vallier, 2018, p. 41, 2020, p. 24). On this conception, a moral reason moves one to care about following a given social-moral norm. Vallier holds an internalist view of moral reasons according to which the description of an agent's moral reasons must partially appeal to that agent's beliefs, desires, values, or commitments. As such, moral reasons are in a sense psychologically accessible to an agent, so that when the agent is asked to explain or justify their actions, they can site the moral reasons they have.

Two further features distinguish moral reasons. First, moral reasons, grounded in one's beliefs and values, "must be seen on reflection as morally appropriate motives" (Vallier, 2020, p. 26). For example, B might be moved by personal ideals, goodwill, duty, religion, or a promise to follow moral rules (Vallier, 2020, p. 24). In contrast, inappropriate motives would involve

motivation by selfishness or ill-will. Moral motivations are diverse, says Vallier, because they reflect the evaluative and viewpoint diversity of members of society, a key assumption held by political philosophers about the result of free expression in liberal societies (Rawls, 2005). However, despite evaluative diversity, Vallier notes that certain improper motivations, such as those grounded in mere self-interest (particularly when this does not invoke moral reasons) or bad will (for example, abiding moral rules with the aim of manipulating or cheating others), are still ruled out.

Secondly, moral reasons must "ground our practice of accountability" (Vallier, 2020, p. 26). When we hold violators of social-moral norms to account, we not only invoke the existence of the rules, but also, we appeal to the reasons that the violators have to not defy social-moral rules. We can, for instance, criticise a pacifist who violently injures others by appealing to her beliefs and values as reasons that speak against that action.

Importantly, in order to hold persons like the pacificist accountable, we must be able to decipher what their reasons are; we must be able to perceive, that is, the moral considerations that motivate them. Thus, for moral reasons to play a role in accountability practices, they must also satisfy a publicity condition. As Vallier puts it:

> "[O]ther members of the trusted person's society [must be able to see] the [trusted person's moral] reasons as reasons for her according to her own evaluative standards. That is, people [must be able to see] on reflection that she has the reason, even if they reject the reason for themselves" (Vallier, 2020, p. 26).

Therefore, in order to hold them accountable to rights practices, we must be able to decipher what their reasons and motivations are. This, however, raises an important concern about how we can detect people's motivations. Vallier is sensitive to this problem. He remarks:

"The trouble is that moral rule compliance does not usually signal why people comply, since people's reasons for action are frequently obscure" (Vallier, 2020, p. 82).

To resolve this problem, Vallier appeals to the fact that we can often discern a person's motivations, particularly in situations where morality and self-interest conflict. One critical note is that self-interest sometimes coincides with morality: it can be in one's self-interest to act morally. However, this is often not the case. Complying with moral rules is usually somewhat costly, says Vallier, so self-interest often motivates rule violation (Vallier, 2020, p. 83). Relatedly, the fact that rule compliance can assail self-interest confirms that actual rule compliance can be inspired by moral reasons. If in some instance, however, one thinks rule following is induced because of a fear of sanction, then moral reasons are likely not efficacious (Vallier, 2020, p. 83, note 162).

The test, then, is whether in a given instance a person is moved by self-interest or by moral reasons. Vallier offers the case of John who loses his wallet, which contains a large sum of money. Reba finds the wallet and returns it with the money intact, perhaps, say, out of goodwill. According to Vallier, John can discern that Reba complied with a social-moral rule, namely, returning a lost wallet with the belongings intact. John can also see that Reba does this for a moral reason, since it was in her self-interest to keep the money, and she would not have been punished if she did not return the wallet. So long as Reba passes this test, John rationally trusts her: as John sees it, Reba publicly complies with a moral rule for a moral reason. On the other hand, if John learns that Reba found the wallet but chose not to return it out of self-interest, she cannot be trusted (Vallier, 2020, p. 83).

Another test case Vallier highlights concerns *violating* a moral rule for moral reasons where self-interested reasons favour rule compliance, or *following* a moral rule for self-interested

reasons where moral reasons disfavour compliance. Vallier argues that in each scenario following the moral reason rationally grounds an observer's trust:

> Suppose Reba is (and John knows she is) a member of a pacifist religion, but there is a military draft, and the social norm is to comply with it. He then observes Reba openly refusing conscription and being arrested. This may increase his trust in her. Conversely, [suppose] John knows that Reba is a member of a religion that forbids her from complying with a common social practice, such as saying the Pledge of Allegiance, but he sees that she says it anyway. He may thereafter trust her less. (Vallier, 2020, pp. 83–84).

The key point here is that acting on moral reasons is the basis of trust and trustworthiness. In the first scenario, the moral rule requires that Reba enter the draft; in the second, the moral rule requires saying the Pledge of Allegiance. In both cases, Reba's moral reasons, stemming from her deeply held values, favour violating the moral rule—even though it might have been in her self-interest to fight for her country or say the Pledge of Allegiance. Since John believes (because he observes) Reba acts for a moral reason, John is said to trust her, under Vallier's account.

We can thus summarize Vallier's theory of social trust into three parts. To begin, social trust is defined as the belief which each member of a society holds that other members of society will generally follow social-moral rules for moral reasons. And social trustworthiness is understood as the virtue of adhering to social-moral rules for moral reasons.

So, the first component of Vallier's theory is that social trust and social trustworthiness occurs in many contexts in which members of society interact with each other. These contexts include social cooperation to achieve a certain goal and non-cooperative interactions such as

when I refrain from punching the next man I see on the street. The second component is that social trust is focused on the rule-following behaviour of other citizens; in particular, social-moral rule-following that is motivated by moral considerations, such as goodwill, duty, or religion. Finally, social trust must be well-placed: the belief that others are trustworthy must actually be informed by the truth that others are trustworthy. Put another way, the epistemic reasons for which A socially trusts B must be grounded in the social trustworthiness of B. As Vallier asserts, social trust must include a true belief that "people ordinarily have adequate moral motivation to follow moral rules" (Vallier, 2020, p. 50).

2.2.4. How to Sustain Social Trust

According to Vallier, only a society with a "broad range of liberal rights practices" can sustain social trust. But there are two jointly sufficient criteria a liberal rights regime must meet to sustain trust. First, its social-moral rules must be publicly justified. To understand why this requirement is important, we must take a short detour to inquire what it means exactly. According to a number of public reason liberal theorists, Vallier included, the requirement of public justification finds rationale in a liberty principle:

> ***Liberty Principle***: Liberty is the norm, and unjustified coercion is always *pro tanto* wrong (Vallier, 2011, p. 262).

The liberty principle states that all kinds of *unjustified* social coercion or pressure is *wrong*. So, it applies not only to a society's laws and public policies, but also to the social-moral rules that purport to regulate cooperative and non-cooperative actions between citizens and pressure them to act in certain ways. We encountered above the claim that liberal democratic societies host citizens with diverse moral worldviews, which support diverse moral reasons. These worldviews highlight distinct and often incompatible conceptions of value stemming from

conflicting religious, moral, and philosophical convictions. Thus, in order for coercion by (social-moral) rules and laws in a liberal society to be just or right, they must be publicly justified for every member of a society: that is, each member must have a conclusive reason to endorse and adhere to those rules and laws. This leads to the public justification principle:

> ***Public Justification Principle***: A coercive rule, law, policy, or action is justified if and only if each and every member of the public has a conclusive reason R to endorse that rule, law, policy, or action (Vallier, 2011, p. 262).

Public justification in Vallier's account means that each citizen endorses her society's scheme of liberal rules and rights practices based on *her own* conception of value, personal ideals, or religious beliefs—which ground her own moral reasons. Vallier's theory of social trust focuses on social-moral rules. On this image, the conclusive reason R that a given citizen might have to follow her society's social-moral rules must flow from her beliefs and values, which implies that different citizens might have different reasons for following their coercive laws institutions.[7]

It is because of his conception of social trust that Vallier takes the public justification of social moral-rules to be important. Recall that Vallier takes social trust to be the belief that one's fellow citizens follow social-moral rules for moral reasons, which stem from often incompatible moral values with one's own. But a publicly justified regime just is that liberal social order where social-moral rules are endorsed and adhered to for diverse moral reasons. Therefore, for citizens to develop trust that others follow moral rules for moral reasons, she must know that her

[7] Vallier terms his account of public justification a *convergence* account. This view contrasts with *consensus* accounts of public justification, according to which each citizen must appeal to the *same* reason R. Examples of consensus theorists include John Rawls (Rawls, 2005), Joshua Cohen (Cohen, 1997), and Jonathan Quong (Quong, 2011). Notoriously, consensus theorists often exclude religious reasons, and more generally reasons that are not shared by all citizens, as candidates for R. This has been met with the charge that consensus accounts of public justification are needlessly exclusionary (Vallier, 2011).

society's social-moral rules are publicly justified (Vallier, 2020, p. 26). If she knows that her society's social-moral rules are publicly justified—i.e., if she knows that other citizens endorse and follow moral rules for moral reasons—then she can maintain the trusting belief that is constitutive of social trust. Public justification thus forms one part of the conditions for a liberal regime to sustain social trust.

But how can one know that one's social-moral order is publicly justified? That is, how can one know that other citizens endorse and follow moral rules for moral reasons? This leads to the second of the (jointly sufficient) criteria that a liberal rights regime must satisfy to maintain social trust: namely, exercising- and complying with- social-moral rules and rights practices must be publicly observable. That is, social trustworthiness, the virtue of being a follower of social-moral rules for moral reasons stemming from duty, goodwill, religion, and so on, must be publicly seen.

To grasp this claim, consider that in a public justified societal order, where all generally endorse and abide by the social norms of that order, each citizen will see themselves as having reason to be trustworthy because each member of the public has conclusive reason to endorse social-moral norms. This reason motivates compliance with these norms. That is, this reason motivates social trustworthiness, understood as the virtue of abiding by social-moral rules because one has sufficient moral reason to do so As a consequence, when citizens publicly conform to social-moral rules that are justified by their own lights, such as "protecting and exercising liberal rights," other citizens can observe that they not only endorse social-moral rules but also are motivated by the right kinds of reasons to follow these rules (Vallier, 2020, p. 22). For Vallier, "observable acts of protecting and exercising liberal rights… establish the empirical expectations required for social trust" (Vallier, 2020, p. 22, p.50). This is because when citizens

see that their fellow citizens conform to social-moral rules for moral reasons, they form the trusting belief that citizens generally follow those for moral reasons. Accordingly, through observance of social trustworthiness, they can come to know that their social order is publicly justified.

Recall Vallier's example of Reba returning John's wallet. Reba's action satisfies the social-moral norm of "returning a lost wallet to the owner with the belongings intact." John can suspect that this norm is publicly justified, but he would be more confident in his trust if he can see this norm actually complied with for a moral reason. Reba's behaviour might thus play a crucial role for John in forming this belief. *Ex hypothesi*, Reba is conflicted between self-interested reasons calling for her to keep the wallet versus moral reasons (say, goodwill) enjoining her to return the wallet. Given that she submits to her moral reasons and returns the wallet, Reba adheres to the social-moral norm for moral reasons. John's observance of this return wallet might then create and solidify his social trust.

These two conditions—public justification and public social trustworthiness—are necessary for a liberal society to sustain social trust. Social trust involves trusting others to follow social-moral norms for moral reasons. This implies trusting that others have sufficient reason to abide by social-moral norms. For Vallier, only a publicly justified order can satisfy the criteria in which each member of the public has a conclusive reason to follow social-moral rules. But this is not enough to sustain social trust. Even if each member of the public has a conclusive reason to follow social-moral rules, we might still question how we can know that this condition holds in the society. Thus, compliance with social-moral norms for moral reasons must be publicly seen. Compliance with social-moral norms for moral reasons characterises the virtue of

social trustworthiness. Therefore, social trustworthiness must be publicly seen in order to sustain trust.

With this, we have the beginnings of a strategy to address the problem of widespread political incivility: if we can create, maintain, and promote publicly justified liberal institutions, and promote public compliance with the norms of these institutions, we can create and sustain social trust and trustworthiness. On this reading of Vallier's theory, social trust, which provides an antidote to uncivil politics, depends on publicly justified liberal orders in which social trustworthiness is publicly observable: for such orders, argues Vallier, are the best tools on offer to ground the type of trust that is necessary to target political incivility.

3. SOCIAL TRUST AND PUBLIC DISCOURSE

Although Vallier's account of social trust offers us with the beginnings of an account of nature of the trust that is necessary to target political incivility, I argue that this account needs some modification to succeed on this score. As it stands, Vallier's account of social-moral norms is too thin to capture the distinct trusting relationship within political discourse, namely, a trust in others to reciprocate adherence to deliberative norms such as civility. Deliberative norms instead provide a substantive basis for the reciprocal trust that is relevant for discursive settings.

To understand this charge, we need to investigate Vallier's claim that social-moral norms centre around a society's *rights practices*. Vallier defines a social-moral rule as "a publicly recognized prescription for a group of persons to engage, or to be allowed to engage, in certain lines of conduct according to the relevant context" (Vallier, 2018, p. 32) Social-moral rules are important because they facilitate and regulate social cooperation. Notice that, for Vallier, social-moral rules constrain actions persons can take, so social-moral rules go beyond dividing social roles of labour and establishing rules pertinent to one's roles in the cooperative endeavour. In

particular, social-moral norms derive from a society's *rights practices*, defined as liberal rights that constrain how we treat each other in our everyday interactions. Thus, persons are required to "be honest, not to harm others without cause, not to kill the innocent, to keep promises, to show gratitude, to aid the impoverished," and to respect liberal rights, such as the right to "physical security," free speech rights and property rights .

Since Vallier's account of social-moral rules is centred mainly around liberal rights practices, this account of norms is too thin to anchor the trust and trustworthiness that characterises political discourse among citizens. The central kind of trust in political discourse concerns whether citizens *trust that others will reciprocate following norms of political discourse*. This suggests that the relevant social-moral norms are deliberative, and thus might be more complex than a rights-based social-moral norm like "do not physically assault others" or "do not to kill the innocent."

To further elucidate this point, consider the following case of *Vaccine* from chapter 1:

> ***Vaccine***. In the workplace breakroom, Ben discusses his approval of government vaccine mandates with his co-workers. Ben thinks this is the best way to end the COVID-19 pandemic. Ava, Ben's co-worker, disagrees. She believes vaccines are a government tool of social control. She gets irritable and calls proponents of vaccination, like Ben, "sheep." When mutual friends try to carefully explain the problems with her views, she refuses to budge. Ben does not take the soft tactic like their friends. He does not mince words: he calls Ava stupid and dangerous for harbouring "anti-vaxxer" beliefs. Friends accuse Ben of being uncivil. Others pick sides, and debate soon becomes bitter and aggressive.

It is conceivable in this case that Ben and Ava still follow the liberal rights practices that characterise their interaction in *Vaccine*. For example, Ben does not mince his words, but he might not deny Ava's right to speak her mind freely. So, Ben and Ava might still respect each other's right to free speech, among other liberal rights. And yet, Ben and Ava might still distrust each other to reciprocate following deliberative norms, such as civility, even though they can trust each other to respect each other's fundamental liberal rights.

The point is not that there is *no* social-moral norm qua liberal right that is violated in *Vaccine*, which violation must be held accountable. Rather, the point is that even if all the social-moral norms qua liberal rights are complied with, and everyone knows that they are complied with, the trust that is necessary to sustain Ben and Ava's deliberative interchange might still be missing. This is because the trust that is necessary to sustain their interchange centres on reciprocal *deliberative* norms, which are not straightforwardly abstract liberal rights. With deliberative norms, Ben needs to trust Ava to be a faithful follower of the deliberative norms that structure their interchange, and Ava needs to trust Ben in turn.

To get the right kind of trust, then, we need to go beyond a conception of social-moral norms thinly construed in terms of liberal rights practices. We need a substantive conception of the social-moral norms that anchor political deliberation between citizens. One central social-moral norm for political discourse, I have argued, is the norm of civility. Recall from chapter 1 that civility can be defined by two norms:

Be open-minded: Here, one must be open to the reasons other discussants provide, reflect on one's own views, preferences, and interests on the basis of the reasons on offer, and revise one's views, preferences, and interests if reflection calls for it (Christiano, 1996, p. 117). Open-

mindedness also involves "rationally arguing against what [others] have to say" if what they say fails to convince us.

Avoid bad faith behaviour: On this image, civility proscribes closemindedness or dogmatism, the unwillingness to be open to the reasons and arguments of others; it also prohibits bad faith behaviours like name-calling, browbeating, hectoring, insulting, vilifications and so on.

These norms of civility represent substantive norms of political discourse. Unlike social-moral norms pertaining to rights practices, the norms of civility are social-moral norms that animate political discussions between citizens specifically.

With the norms of civility, we have a substantive account of the social-moral norms that anchor the trust citizens have that others will reciprocate adherence. In *Vaccine*, for example, Ben needs to trust Ava to be a faithful follower of the norms of civility, and Ava needs to trust Ben to be civil in turn. In other words, Ben needs to deliberatively trust Ava to reciprocate civility, and also be deliberatively trustworthy himself. The same goes for Ava.

4. DELIBERATIVE TRUST, TRUSTWORTHINESS, AND PUBLIC JUSTIFICATION

Deliberative trust and deliberative trustworthiness are a highly distinctive kind of trust suited to public deliberation amongst citizens. Their distinctiveness from Vallier-type social trust and social trustworthiness comes from their dependence on the norms of civility. Deliberative trust is defined as the belief citizens hold about other citizens that each is (or cares and tries to be) a civil interlocutor provided that others are too. This entails that within a given interaction, one believes one's co-deliberator conscientiously reciprocates the norms of civility, including open-mindedness and avoiding bad faith behaviour, such as insults, vilifications, and so on.

In contrast, d*eliberative trustworthiness* is defined as the virtue of being civil—i.e., of having the disposition to act in ways that the norms of civility require. For example, a citizen A who endeavours in political discussions to be open-minded and to refrain from bad faith actions is civil: thus, they are also deliberatively trustworthy. In other words, their being civil makes them worthy of the deliberative trust of others.

Although deliberative trust and deliberative trustworthiness are a highly distinctive from Vallier-type social trust and social trustworthiness, I argue that they embody some of the central features of Vallier's account.

4.1. Deliberative Trust Must Track Deliberative Trustworthiness

The first feature, which we have already encountered in chapter 1, is that deliberative trust must track deliberative trustworthiness to be legitimate, in the same way social trust must track social trustworthiness to be legitimate.[8] Recall that for Vallier social trust is only valid when it is placed in socially trustworthy citizens. On his view, social trust—the belief that other citizens follow social-moral rules for moral reasons—must be rational and based in evidence of trustworthy rule-following behaviour of others. ground because trust evidence e a true belief that "people ordinarily have adequate moral motivation to follow moral rules" (Vallier, 2020, p. 50). That is, social trust must include the belief that others are indeed trustworthy. If social trust lacks this basis, then it leaves those who trust vulnerable to being deceived and harmed by the untrustworthy trusted. Moreover, such misplaced trust cannot last through time compared to well-placed trust, since misplaced trust cannot survive awareness of the fact it is based in poor evidence.

[8] Chapter 1, section 6.

In a similar vein, deliberative trust is only valid when it is placed in deliberatively trustworthy citizens. That is, A's trust that B will reciprocate civility in debate must be based in evidence that B will indeed reciprocate civility. Drawing on Vallier's theory of social trust, we might conclude a few things about the danger of placing deliberative trust in deliberatively untrustworthy citizens. First, misplaced deliberative trust is too unstable to last since it is not based in evidence of trustworthiness. Secondly, the interlocutor who misplaced their deliberative trust falls prey to the harms of deceit and manipulation of deliberatively untrustworthy citizens.

Finally, misplaced deliberative trust might further deepen political incivility when it is discovered. For example, suppose A misplaces her deliberative trust in *un*trustworthy B, who is an arrogant and dogmatic interlocutor but who hides his vices through lies. Once A realises B's true nature, and realises that her trust has been wrongfully placed in B, she is likely to feel betrayed and resentful towards B. These feelings of betrayal and resentment might then lead her to be uncivil with B, to insult and vilify him. Although she might be justified in doing this to B (especially given the argument from reciprocity in chapter 1), A might develop a distrust of others in general and feel entitled to hostility and incivilities in political discussions. So, the danger of misplaced trust is that not only is it unstable and leaves the one who trusts in a vulnerable position: misplaced trust that others will reciprocate civility may worsen political incivilities.

4.2. Deliberative Trust Depends on Public Justification

The second feature shared by Vallier-type social trust and social trustworthiness and deliberative trust and trustworthiness is public justification. Recall the public justification principle:

> ***Public Justification Principle***: A coercive rule, law, policy, or action is justified if and only if each and every member of the public has a conclusive reason R to endorse that rule, law, policy, or action (Vallier, 2011, p. 262).

In Vallier's theory of social trust and social trustworthiness, public justification applies mainly to a society's social-moral norms. Here, each citizen endorses her society's network of social-moral rules and rights practices based on *her own* conception of value, personal ideals, or religious beliefs—which ground her own moral reasons. On this image, the conclusive reason R that a given citizen might have to follow her society's social-moral rules must be grounded in her own beliefs and values, which suggests that different citizens might have different reasons for following their society's network of social-moral rules and rights practices.

The importance of public justification arises for Vallier when he discusses how to build and sustain social trust. Public justification occurs when members of a given public endorse a given social-moral rule for their own moral reasons, and Vallier defines social trust as the *belief* that one's fellow citizens follow social-moral rules for moral reasons. Putting these claims together, then, social trust is defined as the belief that the social-moral rules that animate one's society are publicly justified—i.e., are endorsed (and thus followed) by each member of the public for their own reasons. As a consequence, social trust crucially depends on public justification.

Likewise, public justification plays a central role for deliberative trust. Deliberative is defined as the belief that others will reciprocate civility public discourse. To form this trusting belief, one must believe that each member of the public has sufficient motivation to follow the norms of civility in political discussion. But the claim that each member of the public has sufficient motivation to follow the norms of civility in political discussion just is the claim that

the norms of civility are publicly justified. So, deliberative trust consists in the belief that the norms the norms of civility are publicly justified. If this is right, then public justification plays a crucial role for sustaining deliberative trust, for in order to form the belief that others will reciprocate civility due to their moral reasons, one must believe that others endorse the norms of civility for their own moral reasons; that is, one must believe that the norms of civility are publicly justified.

As with social trust, one might plausibly ask about deliberative trust: how can one know that the norms of civility are publicly justified? That is, how can one know that other citizens endorse and follow the prescriptions of civility for moral reasons? The answer for this, as with Vallier-type social trust, is that abiding by the norms of civility must be publicly observable. The reasoning for this is similar to the reasoning with Vallier-type social trust.

If the norms of civility are publicly justified, then each citizen will see themselves as having reason to reciprocate being civil, for they have a conclusive reason to do so, which reason stems from their beliefs and values. Because this reason motivates citizens to be civil, and because deliberative trustworthiness consists in reciprocating the norms of civility since one has sufficient moral reason to do so, if the norms of civility are publicly justified, then each citizen will see themselves as having reason to be deliberatively trustworthy.

It is because of public observation of deliberative trustworthiness that citizens know that the norms of civility are publicly justified. In other words, when citizens reciprocate civility in public debate, other citizens will observe two things about them: first, that these citizens follow the norms of civility, and second, that they reciprocally follow these norms because they endorse them on the basis of moral reasons. Observation of these two things is observation of deliberative trustworthiness. So, through observance of deliberative trustworthiness, citizens can come to

know that civility is publicly justified. And because they know that civility is publicly justified, citizens harbour deliberative trust, which is the kind of trust that is relevant for public discourse.

5. CONCLUSION

In this chapter, I argued that to address the problem of political incivility we need a kind of trust that is more distinctive than Vallier-type social trust and trustworthiness but that also rests on public justification. This is because Vallier's theory of social-moral norms is too thin to capture the distinct nature of trust within public deliberation amongst citizens. In contrast, I advanced that we need a substantive account of the deliberative norms that anchor political discourse and that also depend on public justification. The norms I focused on were the norms of civility. With civility, we have an account of the kind of trust that is relevant for discourse, namely, *deliberative trust* that others will reciprocate civility in discourse. Deliberative trust is defined as the belief that others will reciprocate civility public discourse. Like social trust, deliberative trust rests on public justification. This entails that the norms of civility on which deliberative trust rests must be publicly justified: each citizen must see themselves as having sufficient reason to be civil.

I argued that for citizens to develop the belief that others will reciprocate civility, they must observe actual civil behaviour from others. On this conception, being reciprocally civil on account of one's moral reasons means that one is *deliberatively trustworthy*. As such, for citizens to develop deliberative trust, i.e., the belief that others will reciprocate civility, they must publicly observe *deliberative trustworthiness* in others, i.e., civil behaviour grounded in moral motives.

Notice that based on this, empirical strategies aimed at addressing the problem of widespread political incivility by trying to cultivate more civility must account for the reciprocity

argument, according to which citizens will not be civil unless they can trust that other citizens will uphold civility like them. However, as I have argued in chapter 1, to build more civility in political discourse, we need to build citizen trust that others will reciprocate civility, but to build this trust we need to build this civility. These two mechanisms thus dovetail together.

That said, the fact the deliberative trust must track deliberative trustworthiness (section 4.1) suggests that building deliberative trustworthiness is more fundamental than building deliberative trust. There are two main reasons to think that deliberative trustworthiness is more fundamental. The first draws on the preceding observation that in order to form deliberative trust at all, citizens must be able to publicly confirm deliberative trustworthiness as a virtue held by other citizens. Here, if citizens cannot publicly observe deliberatively trustworthy behaviour, they will not form the trust that others will reciprocate civility in public debate, and this might make them feel entitled to incivilities. However, if citizens can publicly observe such trustworthy behaviour, then they will form the trust that others in society reciprocate civility in political debate. This will motivate citizens to also be civil in political debate.

The second reason is that deliberative trustworthiness is a legitimacy condition on deliberative trust. For deliberative trust to be valid or legitimate, it must be placed in citizens who are deliberatively trustworthy. Misplaced deliberative trust is too unstable to last; it leaves the one who trusts vulnerable; and it might exacerbate political incivilities when discovered. Therefore, it is important that strategies aimed at building deliberative trust must be concerned that this trust is based in deliberative trustworthiness. Indeed, it seems that building deliberative trustworthiness should take priority here so as to ensure that, wherever it is cultivated, deliberative trust tracks deliberative trustworthiness.

CONCLUSION

The purpose of this thesis has been two-fold. First, I defended the claim that citizens are mostly blameless for the problem of widespread incivilities in political discourse. According to the argument from reciprocity, citizens have reason to act civil only if they can trust that others will reciprocate civility. Given that contemporary political discourse across the US and industrialised west is riddled with political incivility, citizens often do not have reason to be civil in political discourse, for they cannot trust that others will reciprocate civility. As a result, citizens act rationally when they choose to be uncivil in political debate. To be sure, this does not give them a "moral carte blanche" to act as they please in political debate. Yet, a limited range of political incivilities seems to be justified.

Secondly, I analysed the nature of the trust that others will reciprocate civility in political debate—i.e., a deliberative form of trust. Deliberative trust must be well-grounded, meaning that it must be placed in deliberatively trustworthy citizens, where deliberative trustworthiness is defined as the virtue of following the rules of civility because one takes oneself to have reason to do so. I advanced that for empirical strategies looking to resolve the problem of political incivility through building deliberative trust, they must prioritise deliberative trustworthiness. For legitimate deliberative trust crucially depends on deliberative trustworthiness.

To conclude this thesis, I canvass two strategies to help build deliberative trust and trustworthiness.

Social Integration: One way to build civic trust and trustworthiness is by implementing integrative practices and policies (Anderson, 2010). Given how divisive politics is today, citizens have become politically sorted. Political sorting is understood as the increasing association of distinct social identities with partisan identity (Mason, 2018). For example, we find that

Republican partisans in the US tend to be increasingly conservative (ideological identity), white (racial identity) and evangelical (religious identity), but Democrat partisans tend to be liberal, racially diverse, and irreligious (Iyengar, 2021; Mason, 2018). In this, it is unsurprising that liberals only mingle socially with other liberals, and conservatives with other conservatives. Social media algorithms worsen the situation because people typically follow news sources and political pundits that align with their views. This results in widespread echo chambers. According to C. Thi Nguyen, echo chambers are epistemic communities organized around core beliefs adherence to which 1) determines membership in the echo chamber, and 2) motivates "a disparity in trust" between members and non-members (Nguyen, 2020, p. 146). On this view, adherence to the core beliefs of an echo chamber motivate a kind of distrust of non-members. This distrust is usually expressed as epistemic discrediting, where members of the echo chamber systematically regard non-members as dishonest, epistemically malicious, or a poor source of epistemic resources (Nguyen, 2020). Put another way, the distrust expressed as epistemic discrediting involves viewing non-members of one's echo chamber as offenders of norms of political epistemic exchange—as *un*trustworthy co-deliberants.[9]

Integrative policies that promote intermingling across party divides may therefore allow citizens to cultivate deliberative trust and trustworthiness. With respect to deliberative trustworthiness, for example, modifying social media algorithms to diversify the viewpoints in one's feed might help citizens encounter sincere challenges to their own political viewpoints. This would help citizens become more open-minded insofar as they regularly engage rival

[9] We might also contrast virtuous echo chambers with vicious ones. An echo chamber centered around political extremism—say, white supremacy—is a vicious echo chamber. But one centered around gender oppression activism is comparatively virtuous. This is not to deny that echo chambers in themselves are pernicious. To the extent that echo chambers promote systematic discounting of types of evidence or type of non-member there is *pro tanto* reason to think that all echo chambers are epistemically pernicious.

viewpoints and are encouraged to revise their stances given countervailing evidence. Another possibility is opposing-party pundits debating each other civilly on TV and social media. If citizens see institutional actors (especially those they trust) engage political opponents with an open mind and with good will, then citizens might be motivated to emulate such behaviour.

With respect to deliberative trust, recall from chapter 2 that one important way of building trust that others reciprocate civility is by making deliberative trustworthiness public. This involves ensuring that compliance with civility on the basis of one's moral reasons—i.e., reasons based in moral considerations like goodwill, religion, or moral duty—is seen by others. Integrative policies are particularly useful for this. In giving citizens the opportunity to socially mingle with each other—by transforming social media algorithms, for example—they will have ample opportunity to see others act trustworthily, which might create and sustain the trusting belief that others indeed reciprocate civility in political debate.

Therapeutic Deliberative Trust. Therapeutic trust denotes a unique feature of trust according to which trust has a bootstrapping effect: by trusting a person who is minimally trust responsive, one can unlock their potential to be even more trust responsive and thus trustworthy (Frost-Arnold, 2014; Holton, 1994; McGeer, 2008). Richard Holton presents the case of a shopkeeper who hires an ex-convict tried for petty theft (Holton, 1994). The shopkeeper entrusts the till into the employee's care. According to Holton, her trusting act involves the *participant stance*, a stance we take towards agents, i.e., beings that can act for reasons and be held responsible to those reasons. For Holton, the distinguishing feature of trust and the participant stance is liability to betrayal and resentment, so that the shopkeeper is liable to feel betrayed and resentful if the employee steals. Importantly, however, by taking the participant stance towards the ex-convict, the shopkeeper welcomes the ex-convict back to the moral community: she says

to him that he is worthy to be treated as a person who can act for reasons. In doing this, the therapeutic argument goes, the trust that the ex-convict will not steal will infect the ex-convict with a desire not to steal. As Phillip Pettit and Victoria McGeer, extending Holton's argument, contend, the manifest reliance of the shopkeeper on the ex-convict empowers the ex-convict insofar as it provides him a reason to fulfil the shopkeeper's trust (McGeer & Pettit, 2017, p. 15).

Something like therapeutic trust might apply to political discourse too. On this conception, we might attempt to tap into the bootstrapping, therapeutic effect of trust to develop trustworthiness. Consider this modified case of *Vaccine* from chapter 2:

> ***Vaccine***. In the workplace breakroom, Ben discusses his approval of government vaccine mandates with his co-workers. Ben thinks this is the best way to end the COVID-19 pandemic. Ava, Ben's co-worker, disagrees. She believes vaccines are a government tool of social control. She gets irritable and calls proponents of vaccination, like Ben, "sheep." When mutual friends try to carefully explain the problems with her views, she refuses to budge.

Given this, however, Ben can respond in a variety of ways. For example:

a) Ben can try to convince Ava of his position despite her hurtful behaviour
b) Ben can decide to walk away from the conversation
c) Ben can respond in a sharp and abrasive manner; for instance, by calling Ava's position "irrational and dangerous."
d) Ben can name-call and vilify Ava by calling her an "anti-vaxxer"
e) Ben can respond in anger and outrage towards Ava: he could raise his voice and pound his fist while rebuking Ava for her stance.

Recall that following Lawrence Becker, we can classify these possible responses to Ava's bad faith behaviour by two cross cutting distinctions: active or passive acceptance, and active or passive rejection (Becker, 1986, p. 97). Active acceptance implies returning good for evil received, and passive acceptance implies "making no protest at all" against evil received. In contrast, active rejection involves fighting or struggling against an evil, and passive rejection involves merely refusing to submit to it. On this view, options *a*—trying to convince Ava despite his hurt feelings—characterises a kind of active acceptance of Ava's behaviour. Option *b* characterises passive rejection: by disengaging from the conversation, Ben refuses to submit to Ava's incivility as it communicates disrespect. Options *c*, *d*, and *e* portray, to varying degrees, active rejection of Ava's remarks: by calling her an anti-vaxxer, by responding sharply, or by responding more intensely with anger or outrage, Ben actively tries to repudiate Ava's bad faith behaviour.

The aim of this thesis was to vindicate incivilities like options *b*, *c*, *d*, and in special circumstances, *e*. However, option *a* is a perfectly viable, civil response. Indeed, option *a* is the *therapeutic approach* to building more deliberative trustworthiness. In option *a*, rather than responding in an uncivil manner to Ava's closemindedness and name-calling (e.g., calling Ben and their colleagues "sheep"), Ben instead stays in the conversation and tries to win Ava over. In other words, Ben still deliberatively trusts Ava to reciprocate civility. By continuing to engage with her open-mindedly, and refraining from insulting her (e.g., by calling her an "anti-vaxxer"), Ben treats Ava as someone worth taking seriously a person and as a political equal—i.e., as a person who can act for reasons and who is as an authoritative source of reasons and claims. As a consequence, Ava might be moved by Ben's good-natured gesture and become more civil herself—become more open to the charges of others and less likely to engage in bad faith

behaviour like name-calling and demonisations. Thus, Ben’s therapeutic deliberative trust, in opting to convince Ava of his position despite her hurtful behaviour, might motivate Ava to become more deliberatively trustworthy herself.

The result of tapping into this therapeutic feature of deliberative trust is a bootstrapping or looping effect: as one deliberatively trusts in a therapeutic way, one might motivate others to reciprocate civility, for they view one’s therapeutic trust as a sign of goodwill and as a reason for them to reciprocate civility. In turn, this results in publicly observable deliberatively trustworthy behaviour, which grounds the deliberative trust of others, which, when therapeutic, encourages deliberative trustworthiness, and so on. Admittedly, this strategy is vulnerable to many obstacles. For instance, some citizens might be stubborn and wholly unresponsive to therapeutic deliberative trust, in which case other strategies to build their trustworthiness would be more appropriate. That said, the therapeutic approach might be more useful for interpersonal settings in which participants are responsive to each other and are motivated to be civil if another citizen places therapeutic deliberative trust in them.

BIBLIOGRAPHY

Achen, C. H., & Bartels, L. M. (2016). *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton University Press.

Aikin, S. F., & Talisse, R. B. (2020). *Political Argument in a Polarized Age: Reason and Democratic Life*. Polity Press.

Anderson, E. (2010). *The Imperative of Integration*. Princeton University Press.

Baehr, J. (2011). The Structure of Open-mindedness. *Canadian Journal of Philosophy*, *41*(2), 191–213.

Becker, L. C. (1986). *Reciprocity*. Routledge.

Billig, M., & Tajfel, H. (1973). Social Categorization and Similarity in Intergroup Behaviour. *European Journal of Social Psychology*, *3*(1), 27–51.

Brewer, M. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*(2), 307–324.

Calhoun, C. (2000). The Virtue of Civility. *Philosophy & Public Affairs*, *29*(3), 251–275.

Christiano, T. (1996). *The Rule of the Many: Fundamental Issues in Democratic Theory*. Westview Press.

Christiano, T. (2008). *The Constitution of Equality Democratic Authority and its Limits*. Oxford University Press.

Cohen, J. (1997). Deliberation and Democratic Legitimacy. In J. Bohman & W. Rehg (Eds.), *Deliberative Democracy: Essays on Reason and Politics* (pp. 67–91). MIT Press.

Frost-Arnold, K. (2014). The Cognitive Attitude of Rational Trust. *Synthese*, *191*, 1957–1974.

Govier, T. (1992). Trust, Distrust, and Feminist Theory. *Hypatia*, *7*, 16–33.

Holton, R. (1994). Deciding to Trust, Coming to Believe. *Australasian Journal of Philosophy*,

72(1), 63–76.

Huddy, L., Mason, L., & Lene Aarøe. (2015). Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity. *American Political Science Review*, *109*(1), 1–17.

Iyengar, S. (2021). The Polarization of American Politics. In J. de Ridder & M. Hannon (Eds.), *The Routledge Handbook of Political Epistemology* (pp. 90–100). Routledge.

Iyengar, S., & Krupenkin, M. (2018). The Strengthening of Partisan Affect. *Political Psychology*, *39*, 201–218.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not Ideology: A Social Identity Perspective on Polarization. *The Public Opinion Quarterly*, *76*(3), 405–431.

Lepoutre, M. (2021). *Democratic Speech in Divided Times*. Oxford University Press.

Mansbridge, J., Bohman, J., Chambers, S., Christiano, T., Fung, A., Parkinson, J., Thompson, D., & Warren, M. E. (2012). A systemic approach to deliberative democracy. In J. Parkinson & J. Mansbridge (Eds.), *Deliberative Systems: Deliberative Democracy at the Large Scale* (pp. 1–26). Cambridge University Press.

Mason, L. (2018). *Uncivil Agreement: How Politics Became Our Identity*. The University of Chicago Press.

McCarty, N. (2019). *Polarization: What Everyone Needs to Know*. Oxford University Press.

McGeer, V. (2008). Trust, Hope and Empowerment. *Australasian Journal of Philosophy*, *86*(2), 237–254.

McGeer, V., & Pettit, P. (2017). The Empowering Theory of Trust. In P. Faulkner & T. Simpson (Eds.), *The Philosophy of Trust* (pp. 14–34). Oxford University Press.

Nguyen, C. T. (2020). Echo Chambers and Epistemic Bubbles. *Episteme*, *17*(2), 141–161.

https://doi.org/10.1017/epi.2018.32

Novaes, C. D. (2021). Reassessing Different Conceptions of Argumentation. In A. Tanesini & M. P. Lynch (Eds.), *Polarisation, Arrogance, and Dogmatism: Philosophical Perspectives* (pp. 7–24). Routledge.

Nussbaum, M. (2016). *Anger and Forgiveness*. Oxford University Press.

Pew Research Center. (2017). *The Partisan Divide on Political Values Grows Even Wider*. https://www.pewresearch.org/politics/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/

Rawls, J. (2005). *Political Liberalism: Expanded Edition*. Columbia University Press.

Schaupp, K. (2011). Epistemic Peers and Civil Disagreement. In D. Mower & W. L. Robison (Eds.), *Civility in Politics and Education* (1st ed., pp. 23–43). Routledge.

Taber, C. S., & Lodge, M. (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science*, *50*(3), 755–769.

Vallier, K. (2011). Convergence and Consensus in Public Reason. *Public Affairs Quarterly*, *25*(4), 261–279.

Vallier, K. (2018). *Must Politics Be War? Restoring Our Trust in the Open Society*. Oxford University Press.

Vallier, K. (2020). *Trust in a Polarized Age*. Oxford University Press.