

[This paper is due to be published in *Arrogance and Polarization in Debate*, edited by  
Alessandra Tanesini and Michael Lynch]

## **THE POLARIZATION TOOLKIT**

Quassim Cassam

University of Warwick

*Abstract:* This paper outlines and defends a ‘toolkit approach’ to political polarization. This approach sees polarization as a tool or strategy employed by political actors as a means of achieving their political objectives. Understood as a condition, polarization often results from the effective employment of the tools in the polarization toolkit. These tools include: myth-making, stereotyping, polarizing speech, propaganda, conspiracy theories and ‘othering’. Various explanations of the effectiveness of these tools are explored. The toolkit approach to polarization is contrasted with one that sees ‘moral empathy gaps’ as the key to polarization and recommends greater empathetic understanding as the antidote. Questions are raised about this antidote and the Moral Foundations Theory that underpins it. The latter is criticised for over-generalizing from the American experience. It is suggested that depolarization requires the use of a ‘depolarization toolkit’. Some of technique of depolarization are analogous to those of polarization. Others are very different. The reason that it is easier to polarize than to depolarize is that polarizers have the best tools at their disposal.

1

Political polarization has been defined as ‘a condition where political officials and ordinary citizens are so deeply divided that there is no basis for compromise or even productive communication among them’ (Aikins & Talisse 2018). This condition is widely assumed to be politically dysfunctional and perhaps also harmful in other ways.<sup>1</sup> If this assumption is correct then it is important to understand the causes and mechanisms of political polarization. In order to devise effective depolarization strategies, it is necessary to understand why and how this type of polarization occurs in the first place.

1

Many accounts of political polarization have been proposed. One idea is that this type of polarization is the result of, or exacerbated by, *moral empathy gaps*.<sup>2</sup> These are gaps in our ability to empathize with moral reactions that differ from our own. In politics, moral empathy gaps are said to result in a lack of mutual understanding and a tendency attribute partisan differences to the malevolence of one's political adversaries.<sup>3</sup> Accordingly, polarization can be countered by promoting greater empathetic understanding. If we can achieve this then, as Hannon argues, 'we might continue to disagree, but at least we would understand each other better; and this may help us to work together. In contrast, misunderstanding can lead to cynicism and contempt for others, which is part of what causes polarization' (2019: 9).

This line of thinking is, though seductive, problematic in a number of respects. In the first place, there are questions about the extent to which political polarization can really be accounted for by moral empathy gaps or lack of understanding. There are many forms of political polarization that, at least on the face of it, have little to do with such factors. Where moral empathy gaps exist, they may be an effect rather than a root cause of polarization. Once the mechanisms of polarization are understood it becomes evident that empathetic understanding is unlikely to be an effective antidote. Furthermore, partisan differences *are* sometimes due to malevolence. If one side is correct in seeing the other as malevolent then it is entitled to reject demands for understanding or compromise. It is sometimes neither possible nor appropriate to empathize with one's political opponents.

Underlying these concerns are questions about the correct analysis of polarization. One view is that political polarization is a *state* or *condition* defined by the ideological distance between two polarized groups or parties. On a different interpretation, polarization is a *process*:

We therefore define polarization as a process whereby the normal multiplicity of differences in a society increasingly align along a single dimension, cross-cutting

differences become instead reinforcing, and people increasingly perceive politics and society in terms of “Us” and “Them” (McCoy, Rahman and Somer 2018: 18).

As well as defining polarization as a process rather than a condition, this account does not view ideological polarization as the only, or even the primary, form of polarization. For example, it is possible for a political system to be polarized along ethnic lines even in the absence of any fundamental ideological differences between the two groups or their political representatives.

What triggers the process of polarization? One possibility is that it is triggered by the actions of political actors. This suggests that polarization is a political *strategy* or *tool* that is knowingly and deliberately employed by political actors as a means of achieving their own political ends.<sup>4</sup> These ends typically include consolidating their support and weakening their opponents.<sup>5</sup> Understood in this way, polarization need not be pernicious but it often is. It can lead to authoritarianism, intolerance and disagreements over basic facts. It is arguable that Donald J. Trump’s victory in the 2016 American Presidential election was at least partly the result of the effective use of a polarization strategy.<sup>6</sup> Those who deplore this strategy would be advised to study its workings and devise an effective response. Since political polarization is a *political* phenomenon, it calls for a *political* response.

The discussion below will proceed as follows: part 2 will explore the suggestion that political polarization is exacerbated by moral empathy gaps. This analysis relies on Moral Foundations Theory (MFT), the approach to the moral foundations of politics expounded by Jonathan Haidt in his book *The Righteous Mind: Why Good People are Divided by Politics and Religion*. The focus in part 2 will be on the significant limitations of MFT and accounts of polarization based on this theory. Although these accounts are not without merit, they also display a certain degree of naivety about the politics of polarization. One manifestation of

this naivety is the suggestion that polarization can be mitigated by closing moral empathy gaps.

Part 3 will explore the idea that polarization is a political strategy. There is no conflict between thinking of polarization as a condition, a process and a strategy. The strategy triggers the process that results in the condition. The key is therefore to understand the mechanics and enabling conditions for the effective implementation of the polarization strategy. Causing and deepening political polarization is something that (some) political operators *do*, and it is important to understand how they do it. It is helpful to think in terms of a *polarization toolkit* or a *polarization playbook*. This consists of a set of strategies or tricks of the trade that those intent on causing polarization employ for their own ends. This toolkit includes myth-making, stereotyping, polarizing speech, propaganda, conspiracy theories and ‘othering’, defined as ‘the attribution of relative inferiority and/or radical alienness to some other/ out-group’ (Brons 2015: 83). The dehumanizing and demonizing of an out-group by an in-group is the most pernicious form of othering but an effective tool in the polarization toolkit.

Victims of dehumanization cannot be expected to respond to their predicament by cultivating an empathetic understanding of the motives of their persecutors. These motives are, in any case, often obvious. It would be more to the point to require in-group members to cultivate an empathetic understanding of the out-group but anti-out-group propaganda may prove an insuperable obstacle to such an understanding. An effective political response to political polarization will need to include a strategy for countering the range of polarization strategies. This challenge will be the focus of part 4. One issue is whether it is feasible to counter polarization by employing a *depolarization toolkit*. If it makes sense to think in these terms then a further question concerns the extent of the overlap between the depolarization toolkit and the polarization toolkit.

Moral Foundations Theory says that ‘there are (at least) six psychological systems that comprise the world’s many moral matrices’ (Haidt 2012: 211): care/ harm, liberty/ oppression, fairness/ cheating, loyalty/ betrayal, authority/ subversion, and sanctity/ degradation.<sup>7</sup> These systems are the moral analogue of taste receptors. MFT explains cultural variations in morality by noting that there are cultural and historical variations in the sorts of things that trigger a particular system. Furthermore, it is also possible to understand in these terms the dispute between liberals (in the American sense) and conservatives. The liberal left builds its moral matrix on three of the six foundations (the first three) but prioritizes the care/ harm foundation: ‘the most sacred liberal value is caring for victims of oppression’ (Haidt 2012: 345). In contrast, conservatism is claimed by MFT to rest more or less equally on all six moral foundations.<sup>8</sup>

How does this help to explain the worsening polarization of American politics? Ditto and Koleva argue that the American “culture war” between liberals and conservatives makes perfect sense when viewed as constructed upon their differing moral sensitivities.<sup>9</sup> Political attitudes are shaped by what people feel and, just as importantly, what they do not feel. People on the left feel particular empathy for victims of oppression but lack the visceral distaste of conservatives for “sacrilegious” acts like flag burning. While conservatives are not indifferent to victims of oppression ‘they do not “feel victims’ pain” with a liberal’s intensity’ (2011: 332). Partisan conflict is fueled by such gaps in our ability to empathize with moral reactions that differ from our own. When we fail to appreciate the visceral responses that motivate another person’s moral concerns we tend to ‘attribute partisan differences, not to differing moral sensitivities, but to more accessible social-cognitive constructs such as intellectual deficiency or malevolent intention’ (2011: 332). Overcoming moral empathy gaps is hard. However, ‘a hard-won empathy for the moral intuitions of our political adversaries

could lead to more benign (and perhaps more productive) interpretations of their character, motivations, and policy preferences' (2011: 332).

Empathy can be understood as emotional or as cognitive.<sup>10</sup> One can understand another person's feelings without sharing them. Empathy in this cognitive sense is bloodless. Emotional empathy, as the label suggests, engages the empathizer's emotions. In the case of feelings like pain, the difference between cognitive and emotional empathy is between merely understanding that another person is in pain and "feeling" their pain. More generally, emotional empathy is 'the act of coming to experience the world as you think someone else does' (Bloom 2016: 16). In these terms, moral empathy gaps are emotional empathy gaps, that is to say, limitations in the ability of liberals and conservatives to *feel* each other's moral intuitions. The assumption is that cognitive empathy is not enough to overcome polarization or provide for a genuine understanding of opposing political perspectives.

A striking feature of MFT is that it tends to take conservative and liberal accounts of the moral foundations of their politics at face value. The factors by which they *say* they are influenced in their thinking about contentious political issues are taken to be the ones by which they *are* influenced. For example, after describing the left's emphasis on the value of fairness and its understanding of fairness as equality, Haidt adds:

On the right, the Tea Party movement is also very concerned about fairness. They see Democrats as "socialists" who take money from hardworking Americans and give it to lazy people (including those who receive welfare or unemployment benefits) and to illegal immigrants (in the form of free health care and education) (2012: 160).

Fairness, for the right, means proportionality, the principle that 'people should be rewarded in proportion to what they contribute, even if that guarantees unequal outcomes' (2012: 161). Yet the extreme social and economic inequalities that the Tea Party movement is prepared to

countenance far outstrip anything that could plausibly be accounted for by differential contributions.

A related point that Haidt's discussion overlooks is the extent to which expressions like 'hardworking Americans' and 'lazy people' are highly loaded. Those who benefit from welfare are (incorrectly) assumed to be predominantly black, and 'lazy people' is used on the right as a coded racial epithet.<sup>11</sup> This is consistent with research that shows that the attitudes of whites towards welfare are strongly influenced by the extent to which they see blacks as lazy. Politicians on the right assume that, as Martin Gilens puts it, by engaging with issues like welfare and crime they can 'exploit whites' racial animosity and resentment while diminishing the appearance of race baiting' (1996: 593). It is naïve to talk about the moral foundations of conservatism in the U.S without mentioning the right's attitude towards race and the use of the terminology of fairness to stoke racial resentment.

Another example of MFT's unquestioning attitude is provided by Haidt's remarks about the American culture wars in relation to biomedical issues. Conservative opposition to abortion is traced to its commitment to the sanctity of life. Yet those who oppose abortion on this basis include many who are strongly in favour of capital punishment and have no qualms about the sanctity of life in this connection. There may be ways of reconciling conservative attitudes towards abortion and capital punishment. On the other hand, it is also possible that in reality the 'sanctity foundation' plays a less significant role in conservative thinking than Haidt supposes. In this domain myths are all too easily taken for political reality.

Even if MFT is taken at face value there are questions about its implied analysis of polarization. Many accounts of polarization comment on its symmetry: *neither* side "gets" where the other side is coming from, and the moral empathy gaps described by Ditto and Koleva affect both sides of the political divide. Why would this be if, as Haidt claims, the moral foundations of conservatism include the moral foundations of liberalism? If 'there is no

foundation used by the left that is not also used by the right' (Haidt 2012: 333) then one would expect the left to have difficulty empathizing with the right's gut moral reactions but not vice versa. Given this supposed conservative advantage it is hard to understand why the hostility between left and right is as mutual as is commonly supposed or, for that matter, why there is *more* political polarization now than before.

The most serious objection to accounts of polarization inspired by MFT is that they fall into the trap of over-generalizing from the American experience. Polarization in other parts of the world is very different from polarization in the United States, and measures of polarization aren't necessarily measures of ideological distance. This is one lesson of a 2019 special issue of the *Annals of the American Academy of Political and Social Science* on pernicious polarization.<sup>12</sup> The eleven country case studies of polarized politics reveal some commonalities but also striking differences. To give just one example, a study of polarization in Bangladesh describes it as involving 'competing definitions of national identity, rather than divergent economic ideologies or class-based social cleavages' (Rahman 2019: 173). The fundamental division is between one bloc that ties national identity to religion and another that ties it to ethnicity and the use of the Bengali language. It isn't obvious that these and other instances of polarization are best understood in terms of MFT. Different conceptions of national identity need not have different moral foundations in anything like Haidt's sense.

Even in the US context there are key aspects of political polarization that are missed by analyses in terms of moral foundations. In the US, as in Bangladesh, the foundations that are at the root of some of the most bitter political divisions are the foundations of the nation. Polarizing actors build on pre-existing socio-political cleavages, and these include 'formative rifts':

We refer to long-standing and deep-cutting divisions that either emerged or could not be resolved during the formation of nation-states, or sometimes during fundamental re-formations of states such as during transitions from communism to capitalism, or authoritarian to democratic regimes, as *formative rifts* (Somer and McCoy 2019: 8).

Examples includes rifts about the basis of national identity in Bangladesh and ‘the legacy of unequal citizenship rights that were conferred upon African Americans, Native Americans, and women during the foundation of the United States’ (Somer and McCoy 2019: 15). The effects of these rifts are still being felt. Regardless of whether liberals and conservatives have different moral sensibilities, they clearly have quite different perceptions of the history of their nation and what it means to be an American. This formative rift has in turn enabled the exploitation of racial animosity and resentment as tools of polarization. The emphasis on moral foundations misses this and many other important aspects of polarization, both in the United States and elsewhere.

### 3

Reflecting on the limitations of the Moral Foundations approach to polarization brings into focus the need for a more realistic understanding of the mechanics of polarization. One way to do this is to imagine oneself in the shoes of a political actor who calculates that their best hope of achieving their political objectives is to sow discord between different groups and deepen any pre-existing divisions. By purporting to represent the interests of the larger of the two groups they might hope to increase and solidify their public support. To make the scenario more realistic, it can be stipulated that moral scruples don’t figure in the polarizer’s thinking. The only question that interests them is: what works? This can be read as a question about the polarizer’s toolkit, about the various means by which it is possible to increase levels of polarization if that is one’s political objective.

This practical approach to polarization echoes Jason Stanley's perceptive analysis of fascism.<sup>13</sup> 'Fascism' is Stanley's label for 'ultranationalism of some variety (ethnic, religious, cultural), with the nation represented in the person of an authoritarian leader who speaks on its behalf' (2018: xiv). Fascism and polarization are closely related, given that 'the most telling symptom of fascist politics is division. It aims to separate a population into an "us" and a "them"' (Stanley 2018: xvi). Not all polarizers are fascists but all fascists are polarizers. Stanley identifies what he regards as the core 'tactics', 'strategies' or 'techniques' of fascist politics.<sup>14</sup> Most, if not all, of these tactics, strategies and techniques belong in the polarization toolkit.

Consider the significance of myth-making. One form of myth-making consists in the invention and invocation of a mythic past to justify existing divisions and hierarchies.<sup>15</sup> This is the focus of Stanley's account but myth-making can take many other forms. The circulation of negative stereotypes about the inherent nature of an out-group is a form of myth-making that is more concerned with the present than the past. Examples include the description of members of a racial group as lazy or of immigrants as criminals. These are not just false generalizations. They are also what have been called 'generics', and this is key to their myth-making role. Generics are generalizations that omit quantifiers like 'some' and 'all'.<sup>16</sup> It is one thing to say that all or most or some Fs are G. It is another to say 'Fs are G'. In Haslanger's words:

In choosing a generic, it appears that one is saying *of a kind of thing*, specified in the statement, that its members are, or are disposed to be G (or to G) *by virtue of being of that kind*. The speaker conveys that being G is somehow rooted in what it is to be F: G-ing is what Fs do (or are disposed to do) by virtue of being F. This locates the source of the Gness in being (an) F (2012: 457).

Since it is plainly not in the nature of immigrants to be criminals, a generic like ‘immigrants are criminals’ is a form of myth-making. It propagates and perpetuates a myth or legend that serves to deepen divisions between immigrants and non-immigrants and cause the latter to be hostile, or more hostile, toward the former. Myth-making in the form of generics often has pride of place in the polarizer’s toolkit.

Generics are a form of polarizing speech but polarizing speech need not take the form of generics. Polarizing speech can be defined as any speech that, intentionally or otherwise, creates or deepens polarization.<sup>17</sup> When an in-group leader uses a speech to vilify or ridicule an out-group that is polarizing speech. Nothing as comparatively subtle as generics needs to be involved. Another potent form of polarizing speech consists in the articulation of a specific grievance against an out-group. The real or imagined grievances to which polarizing speech draws attention might be political, economic or cultural. As noted by McCoy and Somer, ‘one of the most notable characteristics of pernicious polarization is the Manichean, moralizing character of political discourse. Leaders and supporters alike describe their own and opposing political groups in black and white terms as good and evil’ (2019: 244). There is also the giving of derogatory nicknames (“crooked Hilary”) to opponents and critics and the categorizing of entire sections of the community with pejoratives. This form of polarizing speech has become a staple of political discourse in the United States and elsewhere.

Myth-making, stereotyping and polarizing speech are forms of propaganda. Stanley defines political propaganda as ‘the employment of a political ideal against itself’ (2015: xiii). To put it another way, ‘political propaganda uses the language of virtuous ideals to unite people behind otherwise objectionable ends’ (Stanley 2018: 24).<sup>18</sup> The Confederacy’s use of the concept of liberty to justify slavery was an example of propaganda in this sense but there are forms of polarizing propaganda that are not propaganda in Stanley’s sense. Consider, for example, the government of Myanmar’s insistence on describing the Rohingya as insurgents,

and illegal immigrants. The Rohingya, a Muslim minority, have been present in Myanmar for generations. Anti-Rohingya discourse in Myanmar promotes the idea that the Rohingya are not indigenous to Myanmar and pose a threat to the Buddhist majority. This is propaganda pure and simple, even though it does not use a political ideal against itself.

How, then, should the notion of propaganda be understood? For present purposes, propaganda, or at any rate political propaganda, is best understood as the deliberate attempt to alter, reinforce or otherwise affect a people's political views or behaviour by manipulating their emotions.<sup>19</sup> Myth-making, stereotyping and polarizing speech are forms of propaganda to the extent that they are particular ways of manipulating people's emotions for political ends. As Jason Brennan notes, 'propaganda can be used for good or for bad' (2017: 36), and its underlying message can be true or false.<sup>20</sup> The description of the Rohingya as illegal immigrants was a form of what might be called *negative* propaganda.<sup>21</sup> It was propaganda based on a false claim, its fundamental objective was to represent the Rohingya as the Other and thereby to distinguish them from other ethnic groups in Myanmar, and the means which it did this was to manipulate the emotions of the Buddhist majority.

The othering of an ethnic out-group – the representation of it as an alien or inferior Other - is often a precursor to violence directed against it. As Wade notes, the othering of the Rohingya in Myanmar stripped them of 'the qualities that normally inhibit the use of violence against a fellow human being' (2017: 97).<sup>22</sup> Indeed, according to Wade, at least some of the Buddhists who participated in extreme anti-Rohingya violence in 2012 saw the Rohingya as sub-human. The 1982 Citizenship Act had already made the Rohingya stateless by stipulating that to be a citizen of the country one had to be on a list of Myanmar's 135 "national races". The fact that the Rohingya were not listed was a clear indication of their status as the Other. As a result of their exclusion from the 2014 census, 'the Rohingya were pushed even further away, their statelessness reaffirmed' (2017: 218). The fomenting of ethnonationalism and

othering of out-groups is in the polarization playbook, and far too many governments have read the book.

Another chapter in the polarization playbook is devoted to conspiracy theories. A conspiracy theory is not just a theory about a conspiracy. Conspiracy theories are forms of political propaganda that ‘function to denigrate and delegitimize their targets, by connecting them, mainly symbolically, to problematic acts’ (Stanley 2018: 58).<sup>23</sup> There are few more effective ways to polarize a community than to accuse minorities of involvement in nefarious conspiracies to undermine the majority. The Nazis who promoted anti-Semitic conspiracy theories understood only too well that the scapegoating of an out-group deepens in-group/out-group polarization and makes it easier for the in-group to regard the out-group as the ‘enemy within’. Conspiracy theories are forms of myth-making that stereotype and other targeted out-groups. Rather than working in isolation they use several of the tools in the polarizing toolkit to build a polarizing narrative.

It’s one thing to identify the various tools and strategies that polarizers can and do use to achieve their polarizing objectives. It’s another to work out why these tools and strategies are as effective as they are.<sup>24</sup> Many communities are vulnerable to polarization but how is this vulnerability to be accounted for? Are there varying *degrees* of vulnerability? If so, what explains these variations? These are questions to which there are no quick or easy answers, but several possibilities suggest themselves. One is that what make us vulnerable to polarization are our *epistemic vices*.<sup>25</sup> A different approach says that polarization ‘activates the latent tendencies in the population toward ethnocentrism – generally favorable views toward the in-group and unfavorable stereotyping of the out-group’ (McCoy and Somer 2019: 244). Although not all polarization is along ethnic or national lines, it is nevertheless striking how often polarization is bound up with ethnocentrism or ethnonationalism. If we can explain

the prevalence of these deformations, we will have gone a long way towards explaining some of the most pernicious and destructive forms of polarization.

Epistemic vices are character traits, attitudes or ways of thinking that get in the way of knowledge or understanding.<sup>26</sup> Examples of such vices include prejudice, wishful thinking and closed-mindedness. Anti-out-group prejudice obviously plays a major role in polarization and so does closed-mindedness, an unwillingness to listen or give serious consideration to out-group perspectives. In the case of Myanmar, anti-Rohingya prejudice is underpinned by the closed-minded dismissal of any historical evidence that points to a longstanding Rohingya presence in the country. Such evidence is rejected out a wish to see the Rohingya as illegal aliens. As well as wishful thinking, another thinking vice that polarizers exploit is Manichean thinking. As noted above, polarizing discourse is Manichean, but Manichean *discourse* is underpinned by Manichean *thinking*, a tendency to think in black and white terms and to overlook complexity and nuance. As well as being epistemically vicious, this way of thinking makes one vulnerable to anti-out-group propaganda.

It is unlikely that the tendencies toward ethnocentrism and ethnonationalism that are so expertly exploited by polarizers can be adequately explained in terms of generic epistemic vices. An interesting alternative explanation is offered by Arjun Appadurai, who connects ethnonationalism and polarization with the idea of the nation state:

[T]here is a fundamental, and dangerous, idea behind the very idea of the modern nation-state, the idea of a national “ethnos”. No modern nation, however benign its political system and however eloquent its public voices may be about the virtues of tolerance, multiculturalism, and inclusion, is free of the idea that its national sovereignty is built on some sort of ethnic genius (2006: 3).<sup>27</sup>

What Appadurai interprets as an inherent ethnicist tendency in the very idea of the nation state is especially toxic in the context of globalization, where the speed with which people

and resources move across national boundaries produces new forms of uncertainty in social life.<sup>28</sup> In the context of rapid migration and refugee movement one form of such uncertainty is expressed by the question: ‘how many of “them” are there now among us?’ (Appadurai 2006: 5). It is this uncertainty that is exploited by the politics of them and us. What it promises is certainty about who “we” are and who “they” are.

There are also psychological factors that enable polarization and increase the potency of the tools in the polarization toolkit. These psychological factors include out-group bias, ‘the tendency to favor members of one’s own community and discriminate against outsiders’ (Livingstone Smith 2011: 49). In its most extreme form, this bias can lead to dehumanization, the perception of the Other as less than human.<sup>29</sup> As well as being a psychological enabling condition for group polarization more generally, out-group bias also explains the success of specific polarization techniques. Aldous Huxley described nationalist propaganda in its most extreme form as aiming to do only one thing, to ‘persuade one set of people that another set of people are not really human and that it is therefore legitimate to rob, swindle, bully, and even murder them’.<sup>30</sup> This type of propaganda depends for its effectiveness on a tendency to favor members of one’s own community and discriminate against outsiders

This is a far from comprehensive survey of polarization strategies and the factors that enable polarization. No doubt there are many items in the polarization toolkit that have not been listed. No doubt there are many enabling conditions of polarization that have not been mentioned, and the ones that have been mentioned are far from uncontentious. Nevertheless, it is important to see that polarization is something that can be brought about or deepened by using a range of techniques, and that there are specific background conditions that account for the effectiveness of these techniques. Without an understanding of these conditions and these techniques there is little of hope of developing realistic and viable depolarization

strategies. What might such strategies be, and what are their chances of success? These are the questions to which it is now necessary to turn our attention.

4

Theorists who see polarization as resulting from a lack of mutual understanding also tend to regard greater empathy as the key to countering polarization. In order to understand why other people have political opinions that are different from, and opposed to, our own we need to ‘empathize with their thinking’ (Hannon 2019: 8).<sup>31</sup> Empathy helps us to understand each other and thereby reduces polarization. Empathetic understanding of an adversary’s point of view also makes it less tempting to attribute their stance to malevolence. Accordingly, if one is concerned about political polarization and interested in putting together a depolarization toolkit then it seems that practical measures to promote empathetic understanding in political debate should certainly be part of one’s toolkit. The practical challenge is to specify the nature of such measures.

One question about this line of thinking is whether it exaggerates the extent to which polarization is the result of misunderstanding. The depth of some divisions is due to real and irreconcilable differences between people and communities. Amos Oz makes this point in relation to the Israeli/ Palestinian conflict:

Well, first I have bad news for you: some conflicts are real, they are much worse than a misunderstanding. And then I have some sensational news for you: there is no essential misunderstanding between Palestinian Arab and Israeli Jew. The Palestinians want the land they call Palestine. They have very strong reasons to want it. The Israeli Jews want exactly the same land for the same reasons, which makes for a perfect understanding between the two parties, and for a terrible tragedy (2012: 8).

What is needed is not better understanding but compromise, as distinct from capitulation. The enemy of compromise is fanaticism and that is why the problem is so intractable. Fanaticism,

Oz insists, is ‘an ever-present component of human nature; an evil gene, if you like (2012: 46).<sup>32</sup>

Another question about the idea that empathy is an antidote to polarization concerns the identity of those for whom greater empathy is called for. Consider a scenario in which an in-group and an out-group are deeply polarized as a result of the skillful implementation of a deliberate polarization strategy by a group of polarizers. It might be true that the latter would be more likely to abandon their polarization strategy if they properly empathized with victims of polarization. However, calls for empathy are likely to fall on deaf ears on the reasonable assumption that the polarizers have malevolent motives and stand to gain politically from polarization. There is equally little to be said for attempts to empathize with the polarizers. Empathizing with them is highly unlikely to have any impact on their activities. Furthermore, one might be reluctant on ethical grounds to empathize with malevolent acts or malevolent motives.

The remaining possibility is that in the envisaged scenario the empathy that offers the best hope of a reduction in polarization is in-group empathy for the out-group or out-group empathy for the in-group. Even if Oz is right that there is no fundamental *misunderstanding* between in-group and out-group, this leaves open the possibility that greater mutual empathy will result in greater mutual *sympathy* and in this way counteract polarization. Sympathy, in the sense of pity or sorrow for someone else’s predicament, reduces contempt and hostility, and that has to be a good thing. On the other hand, it is questionable whether it is realistic or reasonable to expect a persecuted out-group, such as the Myanmar Rohingya, to empathize or sympathize with their in-group persecutors. Furthermore, as long as the polarizing strategies described above are still in operation there is little chance of calls for greater empathy having any effect. All of this points to the need for a political response to polarization. For present purposes, such a response will consist in the employment of political tools or methods to

combat polarization. These will include strategies to counter polarization strategies. Only by countering the latter is there any hope of countering polarization.

Since polarizers use a toolkit to cause or exacerbate polarization a natural thought is that anti-polarizers need their own *depolarization toolkit*. What might the contents of such a toolkit be, and to what extent do they overlap with the contents of the polarization toolkit? On the second of these questions, it is clear that there are some tools in the polarization toolkit that it would be highly inappropriate to use for the purposes of depolarization. Depolarization cannot be achieved by othering or the promotion of conspiracy theories. These techniques are part of the problem rather than part of the solution. It is certainly relevant for depolarizers to question the motives of polarizers but questioning the motives of conspiracy theorists doesn't make one a conspiracy theorist. Equally, the othering of an out-group by an in-group should not be countered by the othering of the in-group by the out-group. Such an approach is likely to increase rather than decrease polarization.

A less straightforward question concerns the role of propaganda in depolarization. If anti-out-group propaganda is an effective technique for deepening polarization might it not be feasible to counter polarization by using a different type of propaganda? As noted above, propaganda can be used for good ends, and its underlying message can be true. Propaganda that has both of these features is propaganda in the positive sense.<sup>33</sup> In the present context, the relevant form of positive propaganda is what might be called *humanizing propaganda*. The othering of out-groups is one of the most effective tools in the polarization toolkit and the ultimate objective of othering is dehumanization. This is where humanizing propaganda comes into its own as a depolarization technique. Humanizing propaganda uses a variety of techniques to induce the in-group to see out-group members, such as the Rohingya, as fully human and deserving of equal concern and respect.<sup>34</sup> The sense in which it is still propaganda is that it works by manipulating people's emotions.

Among the techniques of humanizing propaganda is rhetoric. Rhetoric is ‘the art that enables us to present the truth with eloquence’ (Skinner 1996: 2). This characterization of rhetoric, which derives from Cicero, implies that propaganda need not be false. After all, the out-groups that humanizing propaganda induces in-groups to regard as fully human *are* fully human. This is also where empathy plays a role. The empathy that matters for depolarization isn’t empathy with the political demands of an out-group but the more basic empathy that is needed to see out-group members as fully human. Rhetoric wouldn’t be required if facts and evidence were sufficient to change people’s minds. The need for rhetoric ‘stems from the fact that reason lacks any inherent capacity to persuade us of the truths it brings to light’ (Skinner 1996: 2). Depolarization is not just a matter of changing people’s *opinions* but also of changing their *attitudes*. Bringing about such a change means engaging with their emotions, and this is what political propaganda does. Accordingly, humanizing propaganda and rhetoric both belong in the depolarization toolkit, and they are both political responses to political polarization.

However potent one’s depolarizing rhetoric there are limits to what it can achieve. If two groups are deeply polarized there is every possibility that neither side will be susceptible to depolarizing rhetoric. Such rhetoric might be seen by both sides as bogus and is liable in any case to be drowned out by the polarizing rhetoric to which they are still exposed. As well as contending with polarizing rhetoric, depolarizing rhetoric will also come up against the structural factors that support polarization. If ethnonationalism is built into very idea of the nation state then that would be one such factor. Such is the potency of ethnonationalism and ethnocentrism that mere rhetoric is unlikely to be an effective means of countering them. They are powerful forces that polarizers are much better at exploiting than depolarizers are at combatting.

Another tool in the depolarization toolkit is intergroup contact. According to what has been dubbed the ‘contact hypothesis’, face-to-face contact between members of opposing groups lessens prejudice and improves intergroup relations.<sup>35</sup> By the same token, it alleviates polarization. However, empirical research in this area indicates that ‘contact is not a panacea’ (Paolini et al. 2018: 12). Unsuccessful or negative interactions can deepen polarization, and one of the effects of polarization is to lead people to avoid contact with members of opposing groups. Much social interaction today is online, where there is a marked tendency only to interact with like-minded people. Other voices are not heard in the ‘epistemic bubbles’ we inhabit online. There is also evidence that exposure to opposing views on social media can actually *increase* political polarization.<sup>36</sup>

It is hard not to conclude that the polarization toolkit is, in terms of its effectiveness, far superior to the depolarization toolkit. This is a depressing thought in human terms, but it makes sense of historical and political reality. It is inescapable that it is easier to polarize than to depolarize. The question is why that should be, and the toolkit approach answers that question: polarizers have all the best tools. The point of this pessimistic conclusion is not to induce despair but to instill a sense of reality in philosophical discourse about polarization. Everything that can be done to counter polarization should be done but it is also important to be realistic about the prospects of depolarization. Homilies about the importance of civility and mutual understanding are not enough.

Given the limitations of the depolarization toolkit a different approach is needed, one that focuses as much on limiting the damage done by political polarization as on combating polarization. Damage limitation in this context will mean designing our political institutions so that they minimize the negative impacts of polarization. Relevant design features should include institutional checks and balances that compel political actors to compromise, whether they like it or not. The protection of minorities is another important objective. Much as one

would like to contribute to the design of such safeguards it is doubtful that philosophy has much to say in this area.

## References

- Abramowitz, A. & McCoy, J. (2019) 'United States: Racial Resentment, Negative Partisanship, and Polarization in Trump's America', *Annals of the American Academy of Political and Social Science* 681 (1): 137-156.
- Aikin, S. & Talisse, R. (2018) 'Our Polarization Problem', *3 Quarks Daily* 18 December 2018: <https://www.3quarksdaily.com/3quarksdaily/2018/12/our-polarization-problem.html>.
- Allport, G. (1954) *The Nature of Prejudice* (Cambridge, MA: Addison-Wesley).
- Appadurai, A. (2006) *Fear of Small Numbers: An Essay on the Geography of Anger* (London and Durham: Duke University Press).
- Bail, C. et al. (2018) 'Exposure to opposing views on social media can increase political polarization', *Proceedings of the National Academy of Sciences* 115 (37): 9216-9221.
- Brennan, J. (2017) 'Propaganda about Propaganda', *Critical Review* 29 (1): 34-48.
- Brons, L. (2015) 'Othering, an Analysis', *Transcience* 6 (1): 69-90.
- Cassam, Q. (2007) *The Possibility of Knowledge* (Oxford: Oxford University Press).
- Cassam, Q. (2019a) *Vices of the Mind: From the Intellectual to the Political* (Oxford: Oxford University Press).
- Cassam, Q. (2019b) *Conspiracy Theories* (Cambridge: Polity Press).
- Bloom, P. (2016) *Against Empathy: The Case for Rational Compassion* (London: Vintage Books).
- Ditto, P. & Koleva, S. (2011) 'Moral Empathy Gaps and the American Culture War', *Emotion Review* 3 (3): 331-332.
- Feinberg, M. & Willer, R. (2015) 'From Gulf to Bridge: When Do Moral Arguments Facilitate Political Influence?', *Personality and Social Psychology Bulletin* 41 (12): 1665-1681.

- Gilens, M. (1996) ‘“Race Coding” and White Opposition to Welfare’, *American Political Science Review* 90 (3): 593-604.
- Haidt, J. (2012) *The Righteous Mind: Why Good People are Divided by Politics and Religion* (London: Penguin Books).
- Hannon, M. (2019) ‘Empathetic Understanding and Deliberative Democracy’, *Philosophy and Phenomenological Research* (version of record online 22 July 2019).
- Haslanger, S. (2012) ‘Ideology, Generics, and Common Ground’, in S. Haslanger *Resisting Reality: Social Construction and Social Critique* (Oxford: Oxford University Press): 446-477.
- Khoo, J. (2017) ‘Code Words in Political Discourse’, *Philosophical Topics* 45 (2): 33-64.
- LeBas, A. (2019) ‘Can Polarization be Positive? Conflict and Institutional Development in Africa’, *American Behavioral Scientist* 62 (1): 59-74.
- Leslie, S-J. (2008) ‘Generics: Cognition and Acquisition’, *Philosophical Review* 117 (1): 1-47.
- Livingstone Smith, D. (2011) *Less Than Human: Why We Demean, Enslave, and Exterminate Others* (New York: St. Martin’s Press).
- Margalit, A. (2009), *On Compromise and Rotten Compromises* (Princeton: Princeton University Press).
- McCoy, J. (2018) ‘Extreme political polarization weakens democracy - can the US avoid that fate?’, *The Conversation* 31 October 2018.
- McCoy, J., Rahman, T. & Somer, M. (2018) ‘Polarization and the Global Crisis of Democracy: Common Patterns, Dynamics, and Pernicious Consequences for Democratic Politics’, *American Behavioral Scientist* 62 (1): 16-42.

- McCoy, J. & Somer, M. (2019) 'Towards a Theory of Pernicious Polarization and How It Harms Democracies: Comparative Evidence and Possible Remedies', *Annals of the American Academy of Political and Social Science* 681 (1): 234-271.
- Nagel, T. (2012) 'The Taste for Being Moral', *New York Review of Books* 6 December 2012.
- Oz, A. (2012) *How to Cure a Fanatic* (London: Vintage Books).
- Paolini, S. et al. (2018) 'Seeking and avoiding intergroup contact: Future frontiers of research on building social integration', *Social and Personality Psychology Compass* 12 (12): 1-19.
- Pettigrew, T. & Tropp, L. (2006), 'A Meta-Analytic Test of Intergroup Contact Theory', *Journal of Personality and Social Psychology* 90 (5): 751-83.
- Rahman, T. (2019) 'Party System Institutionalization and Pernicious Polarization in Bangladesh', *Annals of the American Academy of Political and Social Science* 681 (1): 173-192.
- Skinner, Q. (1996) *Reason and Rhetoric in the Philosophy of Hobbes* (Cambridge: Cambridge University Press).
- Somer, M. & McCoy, J. (2019) 'Transformations through Polarizations and Global Threats to Democracy', *Annals of the American Academy of Political and Social Science* 681 (1): 8-22.
- Stanley, J. (2015) *How Propaganda Works* (Princeton: Princeton University Press).
- Stanley, J. (2018) *How Fascism Works: The Politics of Us and Them* (New York: Random House).
- Tirrell, L. (2012), 'Genocidal Language Games', in I. Maitra and M. McGowan (eds.) *Speech and Harm: Controversies over Free Speech* (Oxford: Oxford University Press): 174-221.
- Wade, F. (2017) *Myanmar's Enemy Within: Buddhist Violence and the Making of the Muslim 'Other'* (London: Zed Books Ltd.).

---

<sup>1</sup> Political polarization is widely assumed to be dysfunctional and to weaken democracy. See, for example, McCoy 2018. Where political polarization prevails, ‘the common ground among politically opposed falls out, resulting in political deadlock’ (Aikins & Talisse 2018). Polarization makes tolerance ‘increasingly costly and tenuous for individuals and political actors across opposite sides of the polarization. Electorates lose confidence in public institutions and normative support for democracy may decline’ (McCoy, Rahman & Somer 2018: 18). On the other hand, ‘polarization, difference, and a certain dose of antagonistic competition are part of the democratic game and can even have democratizing consequences at times’ (Somer & McCoy 2019: 10). On the role of polarization in democratization and institution-building, see LeBas 2018. Polarization needn’t be pernicious but often is. Pernicious polarization causes epistemic as well as political harms. These harms include the tendency of each side to claim a monopoly on truth and wisdom.

<sup>2</sup> Ditto and Koleva 2011.

<sup>3</sup> This is one of the key insights of Ditto & Koleva 2011 and Haidt 2012. See also Feinberg and Willer 2015.

<sup>4</sup> For the idea of polarization as a tool or strategy see Somer & McCoy 2019: 13.

<sup>5</sup> Somer & McCoy 2019: 13.

<sup>6</sup> Compare the analysis in Abramowitz & McCoy 2019. As they see it, ‘Trump masterfully articulated and reinforced the existing divides in the electorate, but did not create them’ (2019: 139).

<sup>7</sup> For further discussion see Nagel 2012.

<sup>8</sup> Haidt describes this as the ‘conservative advantage’ (2012: 180).

<sup>9</sup> Ditto & Koleva 2011.

<sup>10</sup> See Bloom 2016: 17.

<sup>11</sup> Stanley 2015: 123. For further discussion see Khoo 2017.

---

<sup>12</sup> *Annals of the American Academy of Political and Social Science* 681 (1).

<sup>13</sup> It also echoes the insightful account of polarization in Somer & McCoy 2019. As they note, ‘more often than not polarization is a political strategy that [political] actors employ to achieve wide-scale political ends’ (2019: 13). In the same passage Somer & McCoy describe polarization as a ‘tool to consolidate supporters and weaken opponents’.

<sup>14</sup> See the Introduction to Stanley 2018.

<sup>15</sup> As described by Jason Stanley in chapter 1 of Stanley 2018.

<sup>16</sup> For more on generics see Leslie 2008 and Haslanger 2012.

<sup>17</sup> ‘Polarizing speech articulates or even suggests a grievance, stoking fears, anxieties, and resentments that then become expressed as hostility, bias, and eventually enmity. By choosing the cleavage or grievance to highlight, political elites drive the polarization in important ways’ (McCoy & Somer 2019: 240). Note the assumption that polarization driven by political elites.

<sup>18</sup> These two characterizations of propaganda are not obviously equivalent. See Brennan 2017 for further discussion of Stanley’s views.

<sup>19</sup> For the idea that propaganda works by manipulating people’s emotions see Brennan 2017: 36.

<sup>20</sup> This is also Stanley’s view. See below, footnote 36, for further discussion of the sense in which the underlying message of a piece of propaganda can be ‘true’.

<sup>21</sup> For the possibility of positive propaganda see the discussion below.

<sup>22</sup> In Livingstone Smith’s terminology, the Rohingya were victims of dehumanization. He defines this as ‘the act of conceiving of people as subhuman creatures rather than as human beings’ (2011: 26). Dehumanization overrides inhibitions against killing and ‘has the specific function of unleashing aggression in war’ (Livingstone Smith 2011: 71). See, also, Tirrell 2012 on ‘genocidal language games’.

---

<sup>23</sup> The idea that conspiracy theories are forms of propaganda is explained and defended in Cassam 2019b.

<sup>24</sup> The intuitive distinction here is between the *means* of achieving something and the background *enabling conditions* for achieving that thing by those means. This distinction is explained in chapter 1 of Cassam 2007.

<sup>25</sup> Epistemic vices are explained in Cassam 2019a.

<sup>26</sup> This is the ‘Obstructionist’ view of epistemic vices defended in Cassam 2019a.

<sup>27</sup> As Alessandra Tanesini has noted in private correspondence, what Appadurai says here is not true of all nation states. It is not true of immigrant states like Australia and Canada.

<sup>28</sup> Appadurai 2006: 5.

<sup>29</sup> Livingstone Smith 2011: 2.

<sup>30</sup> This is from a speech by Huxley given in 1936 and quoted in Livingstone Smith 2011: 21.

<sup>31</sup> And, if Haidt is right, with their gut reactions.

<sup>32</sup> On compromise, see Margalit 2009.

<sup>33</sup> Allied propaganda against the Nazis might be regarded as positive propaganda. The Nazis really were awful and that is one sense in which the underlying message of the propaganda was ‘true’. However, Brennan notes that some of this propaganda not only attempted to manipulate people’s emotions, it also ‘tried to instil fear and racist paranoia’ (2017: 36). In this sense, it was not wholly positive. A more general question is whether it can ever be a wholly positive thing to try to change people’s minds by manipulating their emotions.

<sup>34</sup> Stanley argues that ‘there is a kind of propaganda that is politically necessary to overcome fundamental obstacles to the realization of democratic ideals’ (2015: 109-110). He calls this type of propaganda ‘civic rhetoric’ (2015: 115). An example is ‘appeal to emotion, of the sort evoked by art, in the services of the message that Blacks deserve equal respect as humans and citizens’ (Stanley 2015: 110).

---

<sup>35</sup> The contact hypothesis was proposed by Gordon Allport in Allport 1954. Allport's hypothesis is supported by a 2006 meta-analysis of 515 studies which found that intergroup contact typically reduces intergroup prejudice. See Pettigrew & Tropp 2006.

<sup>36</sup> Bail et al. 2018.