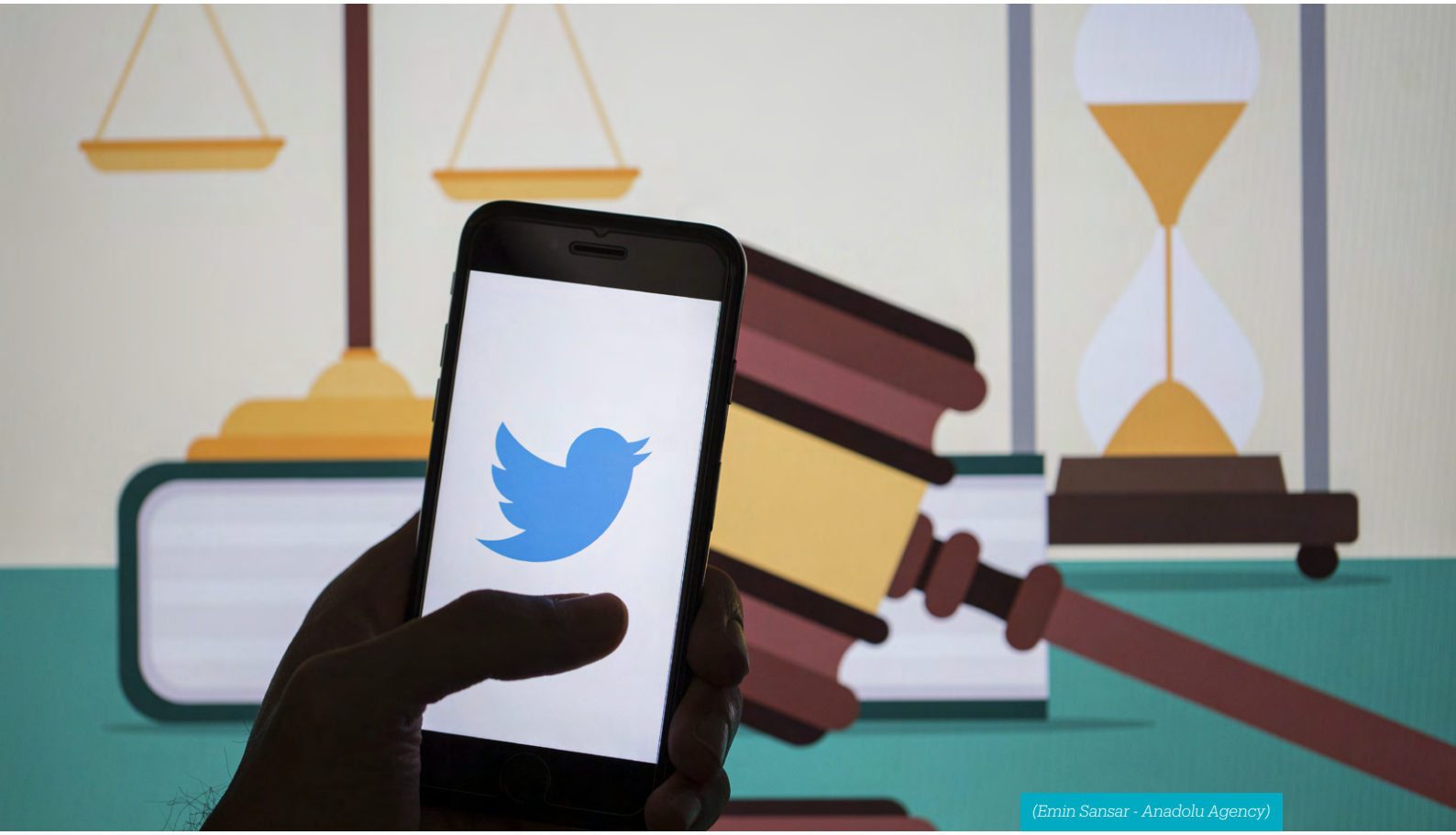
TRT WORLD research centre

Democracy Between Social Media, Political Polarisation, and Disinformation

Dr Serkan Birgel



(Emin Sansar - Anadolu Agency)

Coming off the back of the recent Twitter takedowns in Turkey, this document briefly revisits the nature of the relationship between social media, political polarisation, and disinformation, in the context of this nexus for public policy and democracy. The literature on the topic is vast, but the paper will still attempt to introduce this relationship and explore the aforementioned themes.

The opinions expressed in this policy outlook represent the views of the author(s) and do not necessarily reflect the views of the TRT World Research Centre.

Introduction

This document briefly revisits the nature of the relationship between social media, political polarisation, and disinformation, in the context of this nexus for public policy and democracy. The literature on the topic is vast, but the paper will still attempt to introduce this relationship and explore the aforementioned themes. The paper comes off the back of a recent Twitter takedown of social media bots (automated accounts that engage in social media propaganda tools, derived from the word 'robot'), in Turkey; an event that was criticised as ineffective and selectively anti-AK Party in design. The aforementioned development is juxtaposed with other recent developments that have commanded attention as of late, so as to probe further the relationship between social media, political polarisation, and disinformation. There is bound to be some level of disagreement in the realm of competitive politics, but polarisation is taken to be the stalwart division of a society such that points of commonality, or policies for the common good, become ever more difficult to find. All of these elements go on to affect public debate, policy, and, ultimately democracy. The author takes as a given that all forms of disinformation and polarising politics (across all political colours and stripes) should be assuaged, with a view to counteract polarisation and to ultimately improve the civility of contemporary political debate. Finally, though the fact that social media bots are used does not necessarily nullify the content of the message advanced, the use of bots to advance otherwise cogent or legitimate points of view may still tarnish the content at hand (source credibility effects).

Social Media and Politics: New Promises and Old Consternations

The advent of social media did at one stage raise euphoria at the possibility of a 'golden age of global democratisation' (cited in Tucker *et al.* 2018), where idealistic expectations of a revolutionary role were had for what was once considered a new medium, begotten from the rise of the internet. No doubt numerous instances of good have been achieved through the use of social media, oftentimes used as a conduit to raise public and government awareness of people in need; as a way to disseminate crucial information to the public for the greater good; and a platform to build and engage with communities, amongst other benefits too numerous to cite. Social media bots have been used for some good as well, though they are by-enlarge infamous for malignant political attacks. The human condition, however, remains the same, and perennial issues between mass media and politics readily persist. For Thomas Rid, we now live in an age of disinformation, one with a long and fascinating history that in essence aims to ebb away 'liberal epistemic orders' by placing ideology ahead of objectivity. With the onset of the new fact of life that is social media, a new parlance has been found for age-old pitfalls, perhaps the most common being that of 'fake news' (lies), alongside various categories of disinformation (sophistry, misleading narratives), all of which play a role in the dissemination of propaganda. Both the content of mass media and effects of mass media upon the ebb and flow of politics (especially elections), political rationality, and human behaviour (including those of elites and the 'masses'), are well-established domains of thought that have rightly been the subject of widespread scrutiny. The scholar John Street has, for example, described the relationship between politics and mass media as a relationship of power at its heart, although power in the abstract should probably not be seen as the be-all and end-all of the affair between media and politics, as well as that of media and democracy.

With social media comes greater digital activism, and how this burgeoning field of activism interacts with established systems of governance is an important area of enquiry. The interplay between policy and mass media is not simply one-way; a disposition that enhances the agency and/or stakes attributed to media narratives. With a fear of stating the obvious, access to greater volumes of information through social media does not necessarily entail an enhanced democracy. Nor does, according to one study from Boxell *et al.* (2017), internet access and social media usage necessarily increase levels of polarisation. Nor in fact is fake news, disinformation, and one-sided narratives, limited to anonymised or automated social media actors. Further still, not all disseminators of what are regarded as extreme views insincere in their views or motivated only by electoral engineering, perhaps driven by ignorance or spite at wider socio-economic and cultural concerns. This, in turn, raises the question of who is a bot or not, and where the line is drawn by social media regulators, and how deeply issues today are understood in light of their political context. In any case, the content of 'traditional' media actors should always continue to be assessed in their own right for ideological and methodological bias. To take a step back even further, the notion that there is an immediate or absolute positive relationship between media and democracy is one that can indeed be problematised. More so, mass media is not the only institution that promotes democracy, and the role of other domains of life such culture and religion, education, and political and economic systems, and the differences amongst them, must also be considered in tandem with mass media institutions. The decline of trust and confidence in the latter may be both a cause and consequence of the traction that 'fake news' attains. Global corporations that run social media platforms are now a subject of policy debate in an effort to make such institutions more democratically accountable where weaknesses exist; Facebook and the Cambridge Analytica scandal being a very high-profile case to date that has helped catalyse the debate at hand. Rather than

apolitical, platforms such as Twitter are openly political, going so far as to flag and restrict the U.S. President's tweets. Beyond comparative issues of access and the speed at which information may be disseminated, social media can be such a platform that the sheer volume of information is skewed to one side of a debate. Writing for the Harvard Kennedy School's Shorenstein Centre on Media, Politics, and Public Policy, Wilson and Starbird (2020) find in their study that on Twitter, content that challenges the activities of the White Helmets group in Syria outweighs content that supports them. Another example can be found in Nimmo *et al.* (2020), who note how a coordinated series of social media operations emanating from Russia conducted a 6-year campaign that promoted more extreme, binary, and one-sided narratives, favourable to dominant Russian narratives.

In more general terms, Tucker *et al.*'s (2018) recent meta-review of the academic literature concerning the interplay between social media, political polarisation, and disinformation; a nexus that has been of acute concern in the contemporary age. There is a reciprocal relationship between these three elements, where social media functions as the arena within which political polarisation and disinformation ensure. Within this realm exist: bots, trolls (antagonistic humans raring for controversy), and even 'cyborgs' (human-run accounts that periodically deploy bots). Added to these definitions is that of 'burners' which are single-use accounts abandoned after one post. Bot detection is said to be difficult, partly owing to the sheer scale bots function, the ease at which they can be generated in en mass, and where the link between any malicious activity and specific actors becomes especially difficult, all of which and more is expounded upon further by Gorwa and Guilbeault (2018). The authors here provide a more detailed typology of the various types of malicious activity related to social media, and advocate for an enhanced understanding and classification system for automated bots that would then aid effective policy action. The authors go on to reason that even the most advanced bot detection methods hinge on human judgement; a facet that is perhaps the final or ultimate source of analysis and interpretation across several different domains, but one that is problematised further by the fact that public API's of providers such as Twitter (which researchers rely on), may not, in fact, provide a fair or transparent account of the nature of a platform's content. Still, to briefly give a sense of the scale of bot prevalence only (the definition of which is said to still be vague and ill-defined), though Twitter itself has voiced certain caveats to the use of tools such as 'Botometer' and 'Bot Sentinel' that purport to identify who is and who is not a bot, a 6-week Pew Research Centre study conducted in 2017 found that 2/3's of tweeted links were shared by suspected bot accounts identified with such tools. On closer inspection, the numbers can be staggering. 58 million accounts were suspended in the final months of 2017 alone. Dr. Onur Varol (soon to be an Associate Prof.), who developed the social bot system of Botometer, found that across a period of three months from October 2015 onwards, 9-15% of active Twitter accounts exhibited bot-like behaviours/were bots; a percentage that seems to conform with the take-down figures of late 2017. Not all bots engage in political messaging, but just how those that do go on to drive political discourse (especially a clear assessment of electoral impact), alongside what is described as 'slippery slope' of who exactly is or is not a bot as discussed further here in the New York Times. There is also the role of advertising to consider, and whether or not headlines and news dissemination becomes geared towards the dissemination of short-term controversy and scandal, in order to attract viewers. This then makes the reader geared toward an event rather than the context from which events emanate from. The scholar Jacques Ellul (1967, 56-57) raised awareness of a "stupefying lack of continuity" and "the most superficial level of politics". Where attention becomes the commodity, Ellul asserts that a lack of continuity is created; "[...] my attention, attracted today by Turkey, will be absorbed tomorrow by a financial crisis in New York, and the day after tomorrow by parachutists in Sumatra"...whereon the consumer reacts to the news, rather than developing a long-term view and understanding of political context.

Disenchantment, Polarisation, and Elections

Rather than drive enhanced political engagement, experiences on social media can have the effect of leaving individuals disenchanted with political debate, thereon leading to political disengagement. Here, a regular flow of controversy and counter-controversy takes the place of cogent, long-term, and better-informed political debate. Cited in Tucker et al. 2018, two studies, one by Barberá and Rivero (2015) and the other by Preotiuc-Pietro et al. (2017) lay credence to the hypothesis that users of Twitter with the more extreme ideological interpretations share disproportionately more content than moderate users. The paper also notes that for Twitter, content that is "outrageous and counterintuitive is more likely to be shared", where the motivation to both debate or provoke both aids and abets the sharing of news that is inaccurate or exaggerated. There is also the role of 'endorsement heuristics' (the inclination to take as a given content that has been endorsed by others, without scrutinising the material for yourself). What begins to emerge here is that beyond the sheer prevalence of political polarisation and misinformation, the immature manner in which debates are readily had with the use of incendiary language, selective cynicism, general incivility, and immaturity, are also drivers of said disenchantment and disengagement, but one that helps to harden existing divisions (the backlash or boomerang effects). The interplay between pathology and the cogency of points of politics should be regarded as unavoidable, for one cannot escape the

emotional drivers that are fundamental to the human condition. But what polarisation may do is depend far too much on pathological division and appeals to emotion, such that a fair assessment of political ideology becomes difficult to have. More so on polarisation, and with all the caveats of research design, Bail *et al.* (2018) found that "Republicans who followed a liberal Twitter bot became substantially more conservative posttreatment. Democrats exhibited slight increases in liberal attitudes after following a conservative Twitter bot, although these effects are not statistically significant". Polarisation would have existed in some form before cleavages are exploited by highly antagonistic, one-sided narratives. All of this, in turn, can be detrimental to public policy and democracy, as elites may absorb such divisions rather than quelling divisions. The story, however, is never entirely doom and gloom, and social media in principle does offer an easy opportunity to challenge stereotypes and create positive interactions, deliberation, and political compromise, and a decrease in polarisation in general. Free speech is encouraged but, as always, rights do come with responsibilities. Inevitably, social media providers do have established regulations of what is permissible and impermissible which, in an abstract and detached manner, may seem questionable. In practice, however, limitations of free speech (in terms of say, hate speech), are understandable and welcome. Regulation is a process that can be fraught with difficulties and seemingly perplexing decisions. In June 2020, dozens of Tunisian, Syrian, and Palestinian journalists and activists had their Facebook accounts deactivated. Account holders claim their profiles were miscategorised for having links to terrorism.

Tucker *et al.* (2018, 9) state that political conversations "occur most often between people with close personal ties – spouses, close friends, and relatives". This statement may not entirely be accurate. Rather, it can be reasonably stated that social media expands political conversations beyond far close relationships upon an audience of potentially millions of strangers across the world. Rather than a private affair then, politics becomes ever more public, with one possible detriment being that otherwise banal facts of life or interactions are now 'politicised'. Such politicisation may widen a gulf between actual and virtual, where the consumer of social media finds themselves part of a debate that is far-removed from their locale, which may then in turn aid disenchantment with politics further as an effective force for policy change in one's immediate or near neighbourhood. The communities formed online may not resonate with or come at the expense of communities formed locally. The dissonance can also be seen when bots are used to artificially enhance a political agent's popularity, which may then go on to have detrimental impacts in the long-run. Returning to the general sphere of social media, encountering political views different to one's own may not, in fact, be a one-way gilded path to greater political deliberation, compromise, and the reduction of polarisation. Instead, the opposite readily occurs, where opposing viewpoints are reinforced. A lot of this has to do with the biases that all people live with, but also in the manner in which political debates are conducted themselves. Though by no means an endemic phenomenon, political polarisation in Turkey, for example, is said to be "long-standing and structural" by Kırdemir (2020) in a report for EDAM but enjoys nonetheless regular social media use for political communication "representing a population that is almost equally distributed into different opinion groups".

The prevalence and impact of partisan actors remains a fruitful area of investigation, where there seems to be more consensus and understanding on the former (prevalence), rather than the latter (impact). As noted in Ferrara (2017), who charted a disinformation campaign related to the French Presidential elections of the same year, coordinated attacks on political candidates through the medium of social media have been dated back to 2010. The problem is partly methodological, as with the measurement of bot prevalence. Though there may very well be meaningful impacts, the measurement of the exact effects of social media (with extension to other forms of media), in also hard to discern, though the field of study on the matter progresses in this regard. "Measuring the actual impact of trolling and online influence campaigns is probably impossible" (Starbird, *cited in* Rid). Still, within this space is the potential to exaggerate impact where online metrics devoid of context or a critical assessment of the data can exaggerate the effects of disinformation which can then create an opportunity "for more, and more convincing, disinformation about disinformation" (Rid 2020, 382). On elections per se, a lot of the discussion of the subject matter has been galvanised by on-goings related to the 2016 U.S. Presidential election, where Facebook, in particular, was used to disseminate fake news. One can imagine how during times of election, misinformation and hyper-partisanship may be promoted by both domestic and international actors. Allcot and Gentzkow (2017) demonstrate how the production of fake news (aided by the drive for advertising revenue), was widely shared and heavily tilted in favour of the Donald Trump campaign. But how the sharing and generation of such content actually goes on to influence a voter's decision at the ballot box is difficult to identify. The authors also found that social media was also not the dominant source of election-related news, and that confirmation bias was alive and well. Without nuance, there remains a tendency to draw a clear-cut linear causality, especially those opposed to Trump, and to reason, something along the lines that fake news wholly won Trump the election, or Russian interference wholly won Trump the election.

The idea of social media regulation is also an unmissable part of the discussion here, one that deserves further attention but will only be briefly introduced here. What

seems to be the case is that the mere mention of social media regulation raises the well-known lines of argumentation that revolve around notions of free speech. The topic is, of course, a serious one, and highly emotive. What may sometimes be amiss in discussions thereof is that social media is already regulated in some way shape or form by the various platform providers themselves. In more abstract terms, 'free speech' is also liable to be taken in a way that may renege on the consequences of speech or the responsibilities that come with rights such as that of free speech. As one small anecdote only, the reader will have noticed for example the recent stance Twitter has taken against the U.S. President. There have been further curtailments of accounts deemed to be in some way shape or form counter to the values and/or mission social media platforms aspire to. The justification of such curtailments should always be up for discussion, but the point to be taken is that regulation does exist from the offset. Here an article from the BBC begins to describe the various measures governments around the world have adopted or are considering to adopt in light of the burgeoning political forum that social media represents, with all its potential pitfalls as discussed thus far throughout this review. The European Union, for example, seems particularly keen to counter-act extremist/terror-related content. Turkey will share the same concern. One can, unfortunately, foresee points of dissonance given any contemporary prejudices, but if there is to be regulation it can and should be done in a participatory manner that one hopes will make the case for necessary regulation more amenable. Freedom House for example advocates for incremental change rather than sweeping overhaul, transparency, and civil society participation. The method of regulation should not affect the overall need for it.

June 2020 Turkish Twitter Takedowns

All bots and related sources of disinformation should be countered and removed (in perpetuity), in order to combat polarisation, aid the health of public policy and debate, and thereon improve democracy. Social media bots are also a global phenomenon, in no way endemic to Turkey or any single party. In this light, the news of what has been described as an ineffective, counter-productive, and biased suspension by Twitter of 7,340 accounts in Turkey (as described later), that were created between 2008-2020, is a recent and pertinent case study in the discussion of how to approach the nexus between political polarisation, disinformation, and social media. Twitter had shared with the Stanford Internet Observatory (hereafter SIO), the meta-data of accounts described as 'state-linked' that were suspended from three countries: China, Russia, and Turkey. Assessments of the report are on-going, though there are enough secondary resources at the moment to mount an initial critique, one that can be juxtaposed with closer inspections of the accounts themselves that TRT World Research Centre understands will be published in the near future.

Firstly, that Turkey is assessed alongside China and Russia will be a source of an initial concern that grows as the report is read further. From SIO's Cyber Policy Centre, Grossman *et al*. (2020) noted inter alia that the Twitter takedown included, (and hence can be deduced as non-exclusively): "centrally managed compromised accounts that were used for AK Party cheerleading", accounts critical of opposition parties, but also accounts that lobbied the AK Party government for policy reform. Many of the accounts in question were also said to be compromised by hackers, and reactive to popular content without the user's knowledge. As per Twitter, 'centrally-managed' is the key phrase that raises the suspicion that the accounts were controlled by the government, though the phrase is not well-defined, nor is there any degree or organised activity defined. On this, what the authors do concede – a concession amiss in some critical receptions of the report – is that "[…] *the exact relationship between these accounts and the AK Party is unclear…*", with the existence of pro-AK Party accounts that lobbied the AK Party for policy reform problematising a clear-cut conclusion. The report nevertheless went on to surmise as to why certain domains and hashtags were used to lobby the AK Party, but the manner in which it did so pushed the reader to the single conclusion on offer. The existence of 'centralised' pro-AK Party accounts critical of the AK Party was hypothesised to have, "*by respectfully addressing the government and asking for a regulation change", been a "strategy of creating legitimacy for upcoming reforms*".

Secondly, disinformation is also a cross-party phenomenon in Turkey. The limited targeting of pro-AKP activity can be contrasted with the scope and tone of what the report identifies only as the 'Gülen movement'. Here, one cannot but help note the report's momentary take on the use of 'Bylock' – the encrypted messaging app used by members of the long-standing Fetullah Gülen cult known also as the Fetullah Gülen Terrorist Organisation (FETÖ). The designation is a result of the group's coup attempt in 2016 and Janus-faced political activities targeting Turkish democracy. The report asserts that "*the government deems having ByLock on one's phone as tantamount to proof of being a part of the Gülen movement, and thus proof of being a terrorist*". Before addressing this point, exactly why the report felt the need to momentarily assert a pro-FETÖ viewpoint in an exclusively anti-AK Party report is odd, where a topic as expansive as that FETÖ is treated in a sentence or two without any further context or critique. It is pertinent to be cognisant of the fact that recognition of FETÖ's nefarious side is cross-party common sense in Turkey, and not simply an 'Erdoğan', government, or AK Party issue. In tone and content, the statement is dismissive and off-hand, one that implies that the debate on FETÖ and the use of ByLock is over and done to the detriment of the

AK Party. Firstly, though it is certainly understandable that the SIC wishes not to take sides so as to avert implicating itself in Turkish domestic politics, by omission and lack of reflexivity on such an inordinate point of controversy, the report does not manage to escape the realm of Turkish domestic politics at all. Secondly, considering the weight of evidence that demonstrates the role of the Gülen cult in the 2016 coup attempt, it is odd to describe the group simply as a 'movement' or to belittle the seriousness of how the group should be approached, or to disparage the very key role that the use of ByLock played in the organisation's communication strategy. To deem all users of ByLock as terrorists is excessive, but how true that dictum actually holds, and how far that statement vindicates the 'Gülen movement' and by extension detracts from the seriousness of which the Gülen cult is to be approached, is questionable.

The third point is that unsurprisingly, but perhaps also unavoidably, social media users particularly hostile to the AK Party made rapid use of the publication, demonstrating the value of the material for online propaganda. Latent political polarisation or hyper-partisan politics was not quelled. Instead, the material quickly became ammunition for an already-polarised political debate. However, as suggested immediately following publication, if it is indeed the case that automated social media accounts (bots) solely associated with the AK Party were considered, whilst conversely bots favourable to other political parties were not considered, then the report itself will have inadvertently veered into the category online propaganda. Online propaganda packages potentially factual information in such a way so as to disparage the opposing viewpoint only, which may in turn aid political polarisation, and then ultimately have a detrimental impact on the quality of public policy and democracy. In an article for the New York Times the director of the SIO, Alex Stamos, echoes that sentiment with the quote that: "disinformation about disinformation is still disinformation, and it is harmful to the overall fight". Associate Professor Akın Ünver of Kadir Has University, Turkey, provided insightful comments on the matter. In a social media thread, Prof. Ünver expressed that the network removed had been inactive for months, that it was not a "very central cluster" among a wider network that has not itself registered less activity, being described as 'dormant'. Crucially, non-, and anti-government disinformation networks that do exist in Turkey and for Prof. Ünver are more active in periods of political contest, were not removed. There is cause for concern that a peripheral, dormant, and stagnant network associated with the AK Party only was removed, conversely re-igniting activity against what might be seen as foreign interference and clear anti-AK Party bias which renders for Prof. Akın Ünver elsewhere, Twitter as a thoroughly political actor. In the end, more remains to be done to chart the prevalence and impact of social media bots on the Turkish political scene, as with the realm of politics in general. But the select targeting of bots allied with one particular party may do more harm than good, where instead bots of all colours and stripes should be assuaged, the civility of public discourse should be encouraged, and the reasons for bias (domestic or international) made transparent and debatable.