The Routledge Companion to Media Disinformation and Populism

Howard Tumber, Silvio Waisbord

The evolution of computational propaganda

Publication details

Dariya Tsyrenzhapova, Samuel C. Woolley

Published online on: 24 Mar 2021


Routledge

Accessed on: 12 Jul 2023

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
THE EVOLUTION OF COMPUTATIONAL PROPAGANDA

Theories, debates, and innovation of the Russian model

Dariya Tsyrenzhapova and Samuel C. Woolley

Acknowledgements

For their help and feedback, the authors would like to acknowledge colleagues in the School of Journalism and Media and the Center for Media Engagement (CME), both at the University of Texas at Austin. For their generous support of the Propaganda Research Team at CME, we would like to thank the Beneficial Technology team at Omidyar Network, the Open Society Foundations-US, and the Knight Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these funders. Please direct correspondence to the School of Journalism and Media, Belo Center for Media, 300 W Dean Keeton St, Austin, TX 78712.

Introduction

Russia’s digital interference during the 2016 US presidential election and the UK Brexit referendum helped catalyse broadscale international debates about the global spread of political manipulation online (Mueller 2019; U.S. Senate Foreign Relations Committee 2018). During the Arab Spring protests, Facebook and Twitter were celebrated for their power to mobilise crowds for social movements (Tufekci and Wilson 2012). The recent outbreak of dis/misinformation and inorganic political communication campaigns has, however, led many to argue that social media are now vessels for ‘new and powerful forms of disguised propaganda’ (Farkas 2018, 6), as well as channels for ‘organized and coordinated’ public opinion manipulation (Bradshaw and Howard 2017, 4).

But computational propaganda campaigns on social media are not simply powerful and organised; they are massive, far-reaching, and international. The people who wage them make use of legitimate advertising tools on sites like Facebook as well as illicit tactics, including fake profiles and social media bots – pieces of software algorithmically programmed to spread propaganda messages (Woolley and Howard 2016). Russian efforts during the 2016 US election underscore these features.
During the investigation of US special counsel Robert S. Mueller III, Twitter identified nearly 4,000 accounts associated with Russia’s Internet Research Agency (IRA) and over 50,000 automated accounts linked to the Russian government (Mueller 2019). All told, in the ten weeks prior to the 2016 US election, 1.4 million people directly engaged with these tweets by acts of quoting, liking, replying, or following (Twitter Public Policy 2018). In a similar effort, Facebook disabled 5.8 million fake accounts. This included purging 120 IRA-linked pages, which posted 80,000 pieces of content between June 2015 and August 2017, according to Facebook general counsel Colin Stretch. Facebook also discovered that the IRA had disseminated more than 3,500 ads through a number of groups: ‘Stop All Immigrants’, ‘Black Matters’, ‘LGBT United’, ‘United Muslims of America’, etc. These ad campaigns cost the IRA as little as $100,000 (Mueller 2019), only a tiny fraction of its $125 million monthly budget in the run-up to the US election, (United States of America v Internet Research Agency LLC 2018) and reached almost one in three Americans, or 126 million US-based users (U.S. Senate Committee on Crime and Terrorism 2017).

The Russian propaganda machine has been successful in harnessing social media as platforms for stimulating social and political unrest – particularly in pushing polarisation, apathy, and disbelief both online (Bastos and Farkas 2019; Bessi and Ferrara 2016) and offline (Allbright 2017). The government in Moscow and many other powerful political actors well understand Manuel Castells’s premise surrounding information warfare in a network society: ‘torturing bodies is less effective than shaping minds’ (Castells 2007, 238).

In this chapter we focus on the global diffusion of the Russian computational propaganda model in order to highlight broader changes in the use of social media in attempts to deceptively alter the flow of information during high-stakes political events. Drawing on literature from propaganda studies, reflexive control theory (RC), and information diffusion, we conceptualise how false news messages, built with the intention of igniting the flame of social disagreement, can be harnessed in order to mobilise social movements globally.

There is a clear and continued need for more effects-driven research in computational propaganda. We point to several pieces of work that provide theoretical insight into how behavioural changes may occur as a result of an audience’s exposure to disinformation online. As Bernhardt, Krasa and Polborn (2008) note, the residual effects of this manipulation can include affective polarisation and media bias, which lead to electoral mistakes. In response to targeted criticisms about the lack of demonstrable effects of computational propaganda on behavior at the voting booth (Metaxas, Mustafaraj and Gayo-Avello 2011), we include a preliminary discussion on the difficulties of measuring first-order (direct) socio-political effects in the era of digital dis/misinformation. We join other scholars in calling for more computational propaganda and digital dis/misinformation research exploring the measurement of second-order (indirect) behavioural changes.

A historical overview of propaganda

The rise of organised social media campaigns in recent years led scholars to advocate for the continued relevance of propaganda studies (Jack 2019; Woolley and Howard 2016), which may inform the current debates on the propagating scale of information warfare. Classical scholarship defines propaganda as a form of deliberately organised manipulation of public opinion. It primarily manifests a small group of institutional actors exerting power over a larger populace, striving to ‘activate people’, ‘arouse questions’, and ‘lead to critical reactions’ (Hemánus 1974, 215).

In manipulating public opinion, propagandists aim to convert suggestions into ‘strong active beliefs’, triggered and reinforced through past memories (Moscovici 1993, 73). Through the
manipulation of symbols’ (Hemánus 1974, 215), propaganda shapes people’s thoughts, beliefs, and actions by creating ‘pictures in our heads’ or igniting stereotypes (Lippmann 1922). Propaganda escalates emotions of ‘hope or fear’, which lead audiences to internalise information without scrutiny and deliberation (Lee 1952, 62). If the persuasion is successful, Ellul (1973, 166) contends that the suggested notions seamlessly transform into ‘prejudices and beliefs, as well as objective justifications’, representing new notions of reality. Therefore, classic scholarship suggests that propaganda does not just manipulate the mind; it can also cause other forms of distress as it alienates those who spread and consume it – subtly and slowly making them give up beliefs and integrity to obey someone else (Ellul 1973).

Media (and therefore, uses of propaganda) have changed rapidly over the past century since Lasswell’s (1927) pioneering scholarship in the fields of political communication and media studies. As of 2019, 43 percent of US adults consumed news via Facebook while 12 percent relied on Twitter (Pew Research Center 2019). The networked architecture of social media commands the information flow, algorithmically warranting certain messages wider exposure and exhibiting potential for new and hyper-specific types of audience engagement (Van Dijck 2012). Jack (2019) accounts for the uncertainty of how today’s active audiences, imbued with the power of social sharing, may inadvertently be contributing to a wider spread of dis/misinformation online. By introducing a framework of ‘wicked content’, Jack (2019) calls for treating propaganda as ‘a sensitizing concept’, which also includes fake news (448). She suggests that the complementary concept will retain room for ‘ambiguities of meaning, identity, and motivation, along with unintentional amplification and inadvertent legitimization’ of problematic digital content circulating in an ever-more-complex media ecosystem (449).

Computational propaganda: a global overview

Computational propaganda employs automated and algorithmic methods to spread and amplify messages on social media coupled with the overt propaganda tactics of ideological control and manipulation (Woolley and Howard 2018). Unlike early information systems for propaganda, discussed by pioneer researchers (Ellul 1973; Lee 1952), the current form operates via ‘the assemblage of social media platforms, autonomous agents, and big data’ (Woolley and Howard 2016, 4886). It can rely upon political bots or automated computer code built to mimic human users and push out targeted political messages. Content that begins with bots is often immediately spread across a wide network of high-volume users, including politicians, influencers, and pundits (Woolley and Guilbeault 2017).

Computational propaganda is an international phenomenon. Russian global digital machinations well display this. Brexit was ‘a petri dish’ used to test and gear up for digital information campaigns during the 2016 US election (Mayer 2019). There is a growing body of research detailing the ways in which computational propaganda was strategically leveraged across several government types: by oppressive political regimes, including in China (Yang 2017), Hong Kong (Zhong, Meyers Lee and Wu 2019), Iran (FireEye Intelligence 2018), Saudi Arabia (Benner et al. 2018), Venezuela (Forelle et al. 2015) and by embattled democracies such as Brazil (Arnaudo 2017), Turkey (Akin Unver 2019), and the Philippines (Bengali and Halper 2019).

Political groups in the United States and other Western democracies, many of which face issues of growing polarisation and illiberalism, have employed similar digital tactics in efforts to shape public opinion in their favour. Various studies on Twitter traffic and the 2016 US election indicate that bots have accounted for anywhere from 8 to 14 percent of active accounts (at a
given time), and they were responsible for generating from nearly one-fifth to one-third of the overall Twitter traffic surrounding the event (Shao et al. 2018; Bessi and Ferrara 2016). In Italy, bots have been used to deliberately target influencers during political events by enticing them to share misleading information (Stella, Ferrara and De Domenico 2018). Respected media outlets in Western democracies are also targets of digital disinformation campaigns, many amplified by bots. Lukito et al. (2018) found that a total of 116 stories from 32 major news outlets, including The Washington Post, National Public Radio, USA Today, The New York Times, and other major outlets embedded at least one IRA tweet as a source, often showcasing an account expressing strong partisan beliefs. Bots, largely, spread and amplify messages from low-credibility sources (Shao et al. 2018), and their conversations rarely leave the confines of social media platforms; thus, being quoted by the mainstream media could be attributed to second-order indirect effects, influencing public opinion.

Consequences of computational propaganda

The first documented instance of using social bots and ‘astroturf’ campaigning in political communication on social media dates back to the 2010 US midterm elections (Ratkiewicz et al. 2011). However, researchers are still working to determine concrete societal and electoral effects of political bot communication more narrowly and dis/misinformation and propaganda on social media more broadly. To date, scholarship studying fake news and bot-driven content has devoted considerable attention to highlighting the role automated social media accounts play in disseminating low-credibility content, including false news reports, hoaxes, clickbait headlines, etc. (Shao et al. 2018; Bastos and Farkas 2019). But knowing that many people saw or shared this sort of problematic political content, as Lazer et al. (2018) note, is not the same as knowing if digital disinformation affected users’ behavior.

New work in computational propaganda suggests that bots were successful in amplifying negative and inflammatory stories and sentiments during the 2017–2018 anti-government protests in Iran (Thieltges et al. 2018) and the 2017 Catalan referendum in Spain (Stella, Ferrara and De Domenico 2018). Researchers found that this amplification then resulted in increased polarisation amongst dissenting voices. Political polarisation, according to Bernhardt, Krasa, and Polborn (2008), ultimately causes important information to be lost due to media bias. This, in turn, can lead to voters lacking quality information when they go to the polls; thus, making electoral mistakes becomes more likely.

Scholarship exploring audience engagement with fake news during the 2016 US election suggests that the anticipated effects are ‘less than you think’ (Guess, Nagler and Tucker 2019) because ‘persuasion is really hard’ (Little 2018, 50). When evaluating the success of the computational propaganda campaigns during elections, it is important to note that effects may not always be obvious or linear. Second-order effects may be easier to track than first-order ones, indirect ones more viable then those that are direct. Importantly, many bots and other groups working to amplify disinformation do not work to communicate directly with people on social media – they are constructed to trick trending algorithms and journalists into re-curating content to users via existing ‘trusted’ sources (Woolley and Guilbeault 2017).

While there has been a consistent effort to study the quantitative breadth of computational propaganda campaigns (Stukal et al. 2019; Ruck et al. 2019; Spangher et al. 2020; Bessi and Ferrara 2016; Shao et al. 2018; Badawy, Ferrara and Lerman 2018), much research still has to be done to examine the persuasive qualities of the groups who generate such campaigns and the consumer demand for those messages.
Theorising the diffusion of propaganda over social media

The rise of social media as crucial tools for information sharing has disrupted the traditional pathways information used to follow while travelling to its intended audience. This inherently changes the way public opinion is formed today. Recent studies detailing news-sharing practices emphasise Twitter’s ‘crowd phenomenon’ (Kim, Newth and Christen 2014), which makes it a conducive news diffusion space, facilitating immediate and wide audience exposure to information. As Hummel and Huntress (1952) note, this is exactly what propaganda aims for: if it does not reach its intended audience by making people ‘listen, or read, or watch’, it ultimately fails (51).

For the past decade, communications scholars have made a concerted effort to study what makes messages go viral on Twitter (Stieglitz and Dang-Xuan, 2013; Hansen et al. 2011), but until recently, there has not been a holistic approach to explaining why emotionally charged and politically polarised messages diffuse at a higher rate and scale in digital arenas. The mediated skewed diffusion of issues information theory (MSDII), however, works to bridge this gap (McEwan, Carpenter and Hopke 2018).

MSDII provides an alternative theoretical perspective to well-known theories concerning echo chambers (Jamieson and Cappella 2008) and filter bubbles (Pariser 2011). Using exposure as a key metric to determining message diffusion, MSDII suggests that predisposed views (or personal ego involvement with an issue), message quality (the shorter it is, the more factual, less biased and of stronger argument quality it is perceived), and user’s social network ecosystem (the larger the following, the more likely it is to encounter opposing views) each contribute to wider information sharing on social media (McEwan, Carpenter and Hopke 2018).

Ego involvement, or individuals’ determination to express views dear to them while also seeking out and sharing information that supports personal beliefs, is conceptually similar to individuals’ decisions to practice selective exposure in response to like-minded information and selective avoidance when encountering ideologically conflicting views (Stroud 2011). This practice solidifies a ‘web of belief’, as discussed by Quine and Ullian (1978), which allows people to reject or dismiss information based on how well it fits with their social group conventions. MSDII contends, despite the fact that a social media news feed provides opportunities for users to receive ‘attitude-consistent and attitude-inconsistent messages’, ego involvement with an issue still can make it difficult for the networked communities of strong-tied users to ‘accurately access’ the quality of divergent arguments regardless of positive message attributes (McEwan, Carpenter and Hopke 2018, 2). With this in mind, mere exposure to like-minded information only strengthens ideological beliefs and attitudes and has little to do with actually changing them. Increased presence on social media, combined with connections to like-minded individuals, solidifies one’s views and perceptions, turning the cognitive process into a ‘cyclical feedback loop, (McEwan, Carpenter and Hopke 2018, 9), intensified over time through repeated exposure and engagement.

Crucially, MSDII needs further empirical testing of its key propositions. It does, however, provide new insights into how social media, given their technological affordances and embedded network ties, facilitate and extend the reach of computational propaganda once it hits a like-minded audience. It holds exciting potential to contribute to our understanding of why those exposed to computational propaganda could be susceptible to ‘biased argument processing’ (McEwan, Carpenter and Hopke 2018, 4) that contributes to ‘affective and behavioral divide’ across party lines – a phenomenon Iyengar et al. (2019) define as ‘affective polarization’ (134).
The progression of the Russian model

The Russian government puts great emphasis on directing information operations – an umbrella term for propaganda efforts that today include computational propaganda. Their use of online manipulation strategies highlights the broader international phenomenon of computational propaganda. They primarily see information operations generally as ‘a decisive tool’ for maintaining social control rather than just ‘a supporting element of state power’ (Allen and Moore 2018, 61). This explains their government-sponsored attempts to mold public opinion through a diverse network of agents and channels, tracking and targeting audiences at home (Yapparova 2019) as well as abroad (MacFarquhar 2020). In scholarly debates, Russia’s propaganda efforts have been interpreted through the lens of the reflexive control theory (RC) (Thomas 2004; Bjola 2018; Till 2020; Kowalewski 2017). Developed in the wake of the Cold War by the Russian military, this elaborate form of propaganda is still seen by scholars as ‘an empirical black box’ (Bjola 2018, 23).

Russian attempts at reflexive control played out prominently during the 2016 US election. The IRA leveraged an army of roughly 400 human curators who used bots and other digital tools in disseminating content (Chen 2015). Till (2020) described this group as ‘fully employed “agents of influence”’ used in an opaque digital form of Russian ‘statecraft’ (7). As a part of a minimum daily quota, each ‘agent’ was expected to post comments on 50 articles, run six Facebook accounts, post three posts on each Facebook account, hold two discussions on Facebook, and manage ten Twitter accounts, populating those with at least 50 tweets a day (Singer and Brooking 2018).

Conceptually, the goal behind RC is to control the reflex of the opponent by finding ‘the weak links’ (for example, racial tensions) and further exploiting them to ‘sharpen ideological polarisation, maximize political disunity and weaken democratic institutions’ (Bjola 2018, 22). Walker and Ludwig (2017) call this form of deliberate manipulation ‘sharp power’, which acts like ‘the tip of [the] dagger’, aiming to ‘pierce, penetrate, or perforate’ the media system of targeted countries. Bjola (2018) states that manipulation is accomplished through ‘cognitive mapping’ and ‘micro-targeting’ of the intended audiences (22).

The rise of advanced computation and big data has opened up a wide range of possibilities for novel reflexive control tactics, used by Russia and a variety of other powerful political actors. Specifically, these innovations allow for new means of ‘exploiting moral, psychological . . . as well as personal characteristics’ in the form of ‘biographical data, habits, and psychological deficiencies’ to acquire ‘the highest degree of reflex’, which maximises the chances of defeating the opponent by covertly influencing the other’s perception of reality (Thomas 2004, 241–242).

The sheer size of the global digital audience, with 3.8 billion users now present on social media (Kemp 2020), simplifies the task of orchestrating computational propaganda campaigns: the audience is vast, and exposure to a message is immediate. Moreover, automated technology used on Twitter and other social media sites is increasingly cheap and easy to obtain. An investigation from The Daily Beast found that a social bot account could be purchased online for less than five cents (Cox 2017). The fact that the Russian government inundated important political conversations with bots to tilt public opinion in its favour exemplifies a shrewd strategy of solving an essential dilemma between either selectively censoring problematic content or blocking Facebook and Twitter within Russia entirely (Sanovich 2019). Given the present power of social media in the global information infrastructure, neither of these measures would have been possible without causing a public outcry. In launching computational propaganda via reflexive control, Russia made a power move internally and externally by turning an old worry into a boon.
Early news coverage surrounding the 2020 US election suggests that the Russian propaganda machine made efforts to gear up for a more sophisticated and abstruse ‘reflexive control’ operation aimed at influencing the general public. A CNN investigation run in collaboration with researchers at Clemson University, Facebook, and Twitter found a network of more than 200 accounts stirring up racial tensions amongst the Black Lives Matter community, drawing thousands of shares and reactions. Masquerading as US-based users, these accounts turned out to be run by trolls based in Ghana and Nigeria (Ward et al. 2020). Facebook later deemed these accounts to be perpetuating ‘coordinated inauthentic behavior’ and linked their activities to the Russian IRA (Gleicher 2020).

Russian attempts to control US public opinion aren’t just occurring over social media. Another arm of the Russian government propaganda – Radio Sputnik (formerly the Voice of Russia and RIA Novosti) recently opened for business on US soil. Radio Sputnik now plays at 104.7 FM in Kansas City, Missouri. Acting through a broker in Florida, the Russian government prepaid the Missouri-based Alpine Broadcasting Corporation $324,000 for three years’ worth of news programming at an hourly rate of $49.27 (MacFarquhar 2020).

Conclusion
Computational propaganda, as exemplified by the Russian model, rapidly evolves to satisfy current demands and a changing informational environment. During the 2016 US election, Russia enjoyed modest levels of success in leveraging digital dis/misinformation campaigns in order to encourage digital polarisation and offline protest. They, and other governments around the world, have leveraged computational propaganda in efforts to exert reflexive control over their opposition – internally and abroad. They and others will continue to deploy bots and other mechanisms of computational propaganda in attempts to manipulate public opinion.

Classic literature from media and propaganda studies, paired with more recent work on computational propaganda, information operations, and political communication helps in formulating theoretical understandings of the current and ever-changing state of digital political manipulation. Reflexive control theory and the recent developments in theories of news diffusion offer a comprehensive explanation for why computational propaganda from Russia and other political actors works to spread and take hold.

There is a marked need for concerted and continued empirical work on computational propaganda in order to discern the behavioral effects – and cultural impact – of this relatively new problem. These efforts should be compared and combined with broader efforts on disinformation and polarisation, issues closely tied to research on computational propaganda. While first-order behavioral effects may prove difficult to track, second-order ones may be more possible to discern. Computational propaganda has been established as a phenomenon; now the scholarly community needs to work to figure out what this phenomenon does to peoples’ behaviour.

References
Akin Unver, H 2019, ‘Russia has won information war in Turkey’, Foreign Diplomacy, 21 April. Available at: https://foreignpolicy.com/2019/04/21/russia-has-won-the-information-war-in-turkey-rt-sputnik-putin-erdogan-disinformation/


Cox, J 2017, ‘I bought a Russian bot army for under $100’, Daily Beast, 13 September. Available at: www.thedailybeast.com/i-bought-a-russian-bot-army-for-under-dollaro100


The evolution of computational propaganda


Kemp, S 2020, ‘Digital 2020: 3.8 billion people use social media’, *WeAreSocial*, 30 January. Available at: https://wearesocial.com/blog/2020/01/digital-2020-3-8-billion-people-use-social-media


Little, AT 2018, ‘Fake news, propaganda, and lies can be pervasive even if they aren’t persuasive’, *Critique*, 11(1), pp. 21–34.


Yang, Y 2017, ‘China’s Communist Party raises army of nationalist trolls’, Financial Times, 29 December. Available at: www.ft.com/content/9ef9592-c2bd-11e7-97e2-916d4bca0da

Yapparova, L 2019, ‘The information nation: A Kremlin-managed research center is selling services that can deanonymize anyone in Russia’, Meduza, 20 September. Available at: https://meduza.io/en/feature/2019/09/20/the-information-nation?fbclid=IwAR0FhXlp36V9HlaBxD6zxtqeInkh1nEmbWkjmRnX0TTP1bUgUZ6ri1BEBovY