































organizations (Ellul 1965; Jowett and O’Donnell 2012). On social media, however, this investigation strategy is challenged. Determining ‘who gains the most’ is difficult, when propaganda can potentially derive from a small partisan group, a large-scale organization, or even a single individual seeking to further an agenda or simply provoking others by *trolling* (Phillips 2012). This raises new epistemological challenges (Schou and Farkas 2016): How can we investigate disguised sources and intentions on social media? How does the credibility of a social media profile or page increase through its likes, shares, and comments? How can users become more critical of information streams produced by aggressive and hateful posts as well as antagonistic reactions? And how can we assess the magnitude and significance of disguised propaganda, when fake profiles can reach thousands of users within days, before social media companies delete them? These questions require urgent scholarly attention.

In the context of vertical propaganda, large-scale organizations take advantage of the decentralized structure of social media by orchestrating far-reaching campaigns that are nonetheless difficult to identify as such. Two vital components in this regard are the use of so-called *troll armies* and *social bots* for social media astroturfing (Benedictus 2016). On social media, astroturfing encompasses the orchestration of user profiles by an organization, such as a government agency or private corporation, to simulate public support or opposition towards a particular topic. This form of disguised manipulation can serve as both propaganda of agitation and integration, as organizations seek to consolidate power through attacks on perceived opponents as well through the manufacturing of widespread support. Astroturfing can also rely on both obfuscated and impersonated sources, as organizations might pay users to post content from their own social media accounts or through networks of fictitious profiles. In practice, these modalities are often interconnected. In China and Russia, government agencies have orchestrated large-scale troll armies, in which people are paid to promote government agendas through social media profiles (Tong and Lei 2013, Benedictus 2016). In China, this has been coined the ‘50-cents party’, as users were rumored to receive 50 cents for each social media post they create in support of the government (Aro 2016).

Large-scale organizations engage in social media astroturfing for a number of reasons. This form of disguised propaganda can potentially have widespread influence on public

opinion and be an effective tool to silence critics through aggressive campaigns (Aro 2016). Astroturfing might also serve to divert public attention from contemporary crises by flooding social media with non-related content. This method, which relies on a “strategic distraction from collective action, grievances, or general negativity” (King et al. 2017), has been used extensively in China, where the government is estimated to orchestrate 448 million social media posts per year (King et al. 2017). This content is first and foremost produced by human laborers (King et al. 2017), yet astroturfing can also rely on social bots.

Social bots are user profiles controlled by software that algorithmically produce and disseminate content. During the 2016 U.S. presidential elections, a large-scale study estimates that close to 20 percent of all content on Twitter concerning the elections was produced by social bots (Bessi and Ferrara 2016). Bots were found on both sides of the political spectrum, although a majority supported the Republican candidate, Donald Trump (Kollanyi et al. 2016). A key application of bots was to disseminate conspiracy theories and disinformation, popularly referred to as ‘fake news’ (Shao et al. 2017). As of the time of writing, the U.S. Congress is investigating whether some of this activity was orchestrated by Russian agencies (Wakabayashi and Shane 2017), but despite the potential political implications of these activities, it remains difficult to identify the sources behind automated propaganda:

Concluding, it is important to stress that... it is impossible to determine who operates such bots. State- and non-state actors, local and foreign governments, political parties, private organizations, and even single individuals with adequate resources... [could] deploy armies of social bots and affect the directions of online political conversation.

(Bessi and Ferrara 2016)

As with all disguised propaganda on social media, platforms’ decentralized mode of content proliferation and potential anonymity provided for propagandists complicate both epistemological boundaries and empirical investigations of vertical and horizontal propaganda. This raises serious and urgent questions of accountability, transparency,



and contestation of disguised propaganda – for scholars, users, policy makers, law enforcement, and (perhaps most importantly) the corporations owning these platforms. So far, social media companies have largely placed the responsibility for countering these phenomena on users. However, this strategy is not a viable solution to contemporary or future democratic consequences posed by disguised propaganda.

### **Countering disguised propaganda on social media**

The vast popularity of social media platforms makes it difficult for companies, such as Facebook and Twitter, to identify and moderate problematic content. As a solution to this challenge, companies construct their policy enforcement principles around user engagement, often deploying commercial content moderators solely when users flag content for violations of company policies (Roberts 2016). Reimagining and reengineering this division of labor is a difficult endeavor:

The huge numbers of members that popular social media sites boast and the vast volume of content these members post make it impossible for the staff of the host companies to pro-actively monitor and edit the contents. As we've seen, the only way content guidelines - in particular, those related to hate speech - can be applied is through the active engagement of real people... Almost inevitable, this task falls mainly on the users of the social media sites.

(Foxman and Wolf 2013, 106)

Identifying hate speech and disguised propaganda has to rely on human judgment, as algorithms cannot adequately analyze cultural contexts of each post (at least not yet). Due to the ubiquity of social media platforms, such human judgment has to derive from users. Following this argument, an encouraging solution to disguised propaganda could seem to be the formation of citizen groups, actively fighting propaganda by reporting fake pages and profiles to social media companies. Promising as this initiative may be, users can only superficially counteract disguised propaganda under current conditions (Farkas and Neumayer 2017).

There are many challenges involved in building alternative spaces to fight disguised propaganda on social media platforms. New forms of digital editing tools make it increasingly difficult to determine, if pictures and videos are manipulated. The decentralized structure of social media platforms makes it difficult to find and contest propaganda before it potentially reaches a wide audience. The biggest challenge, however, is the way in which social media companies place the responsibility for countering propaganda on their users, yet only provide limited and opaque opportunities for them to act. As a result, tactics to manipulate users become increasingly sophisticated, while collective resistance cannot.

The idea of empowerment of crowds acting and creating together has been present in early discourses about social media. Tim O'Reilly coined the term Web 2.0 with one key component being the “wisdom of the crowds” (O'Reilly 2005). For social media companies, crowdsourcing became an effective marketing discourse, in which they present their platforms as spaces of participation, decentralization, spontaneous interaction, and lack of hierarchy – ideas hijacked from the radical left (Žižek 2009). In the case of fighting disguised propaganda, these ideas about social media shift the responsibility to the users. They, however, have to navigate limitations of architectures and policies provided by social media corporations. Instead of empowering activists, “power has partly shifted to the technological mechanisms and algorithmic selections operated by large social media corporations” (Poell and van Dijck 2015, 534)

On Facebook, users are provided only with a ‘report’ button to notify the company of content violations (Farkas and Neumayer 2017). How Facebook processes these reports remains highly opaque. Consequently, users cannot know how or on what grounds Facebook takes action. Even *if* Facebook deletes a profile or page, the creators can typically remain anonymous and continue their work. This makes it incredibly difficult for users or authorities to hold anyone accountable. These challenges also complicate the work of journalists or researchers trying to study the implications of disguised propaganda, as a page or profile might reach thousands of users within days and then disappear without notice. To limit the potential contemporary and future threat of disguised propaganda, users should be able to identify, mobilize, organize, and collectively resist

manipulation much more effectively. Although anonymity can be beneficial for democratic discussion in many ways, it is problematic for counter-action that creators of disguised propaganda can stay completely anonymous and avoid any consequences. For this to change, social media corporations need to be held accountable for countering propaganda on their platforms. In the current situation, crowdsourced user actions mainly seem to serve as a diversion from corporate responsibility and questions of accountability.

## **Conclusion**

Disguised propaganda has undergone a series of profound changes alongside technological developments throughout the 20<sup>th</sup> and 21<sup>st</sup> century: from impersonated propaganda in early 1900s newspapers (e.g. *The Protocols of the Elders of Zion*), to clandestine radio during the Second World War, all the way up until present-day social bots and troll armies on social media. Alongside this significant evolution, scholarly contributions, such as analytical and epistemological frameworks, are continuously challenged. As this chapter has made apparent, digital media platforms complicate the fundamental notion of disguised propaganda *de facto* deriving from large-scale organizations. Additionally, digital media challenge existing conceptual boundaries, such as Ellul's (1965) conceptualization of vertical and horizontal propaganda. These technical developments, however, do by no means render these profound conceptual works redundant. Contrarily, revisiting concepts of propaganda studies (such as Ellul 1965); Sproule 1994; and Herman and Chomsky 1988) enables us to explore how disguised propaganda changes in digital and social media, but also to outline their continuity across different media technologies. More scholarly engagement with disguised propaganda on social media is necessary to further develop concepts at the intersection of internet research and propaganda studies. Research in this field should expand methodological, analytical, conceptual, and epistemological frameworks, but also support resistance against disguised propaganda that produces hatred and racism. Scholars should not only strive to understand the development of propaganda, but also challenge and contest manipulation and deception in contemporary and future online spaces.

## **References:**

- Aro J (2016) The cyberspace war: propaganda and trolling as warfare tools. *Eur View* 15:121–132. doi: 10.1007/s12290-016-0395-5
- Auerbach J, Castronovo R (2013) Introduction: Thirteen Propositions about Propaganda. In: Auerbach J, Castronovo R (eds) *The Oxford Handbook of Propaganda Studies*. Oxford University Press, Oxford, pp 1–16
- Baym NK (2015) Social Media and the Struggle for Society. *Soc Media + Soc* 1:205630511558047. doi: 10.1177/2056305115580477
- Becker H (1949) The Nature and Consequences of Black Propaganda. *Am Sociol Assoc* 14:221–235.
- Benedictus L (2016) Invasion of the troll armies: from Russian Trump supporters to Turkish state stooges. *Guard*.
- Bessi A, Ferrara E (2016) Social bots distort the 2016 us presidential election online discussion.
- boyd danah m., Ellison NB (2007) Social Network Sites: Definition, History, and Scholarship. *J Comput Commun* 13:210–230. doi: 10.1111/j.1083-6101.2007.00393.x
- Castells M (2013) *Communication Power*, Second Edi. Oxford University Press, Oxford
- Chaykowski K (2017) Mark Zuckerberg: 2 Billion Users Means Facebook’s “Responsibility Is Expanding.” *Forbes*
- Daniels J (2009a) Cloaked websites: propaganda, cyber-racism and epistemology in the digital era. *New Media Soc* 11:659–683. doi: 10.1177/1461444809105345
- Daniels J (2009b) *Cyber Racism: White Supremacy Online and the New Attack on Civil Rights*. Rowman & Littlefield Publishers Inc., New York
- Daniels J (2014) From Crisis Pregnancy Centers to Teenbreaks.com: Anti-abortion Activism’s Use of Cloaked Websites. In: *Cyberactivism on the Participatory Web*. pp 140–154

- Donath JS (1998) Identity and deception in the virtual community. In: Smith MA, Kollock P (eds) *Communities in Cyberspace*. Routledge, London, pp 22–58
- Ellison NB, Boyd DM (2013) Sociality Through Social Network Sites. In: Dutton WH (ed) *The Oxford Handbook of Internet Studies*. Oxford University Press, Oxford, pp 151–172
- Ellul J (1965) *Propaganda: The Formation of Men’s Attitudes*. Vintage Books, New York
- Farkas J, Neumayer C (2017) “Stop fake hate profiles on Facebook”: Challenges for crowdsourced activism on social media.
- Farkas J, Schou J, Neumayer C (2017) *Cloaked Facebook Pages: Exploring Fake Islamist Propaganda in Social Media*. *New Media Soc*. doi: <https://doi.org/10.1177/1461444817707759>
- Floridi L (1996) Brave.Net.World: the Internet as a disinformation superhighway? *Electron Libr* 14:509–514.
- Foot KA, Schneider SM (2002) Online Action in Campaign 2000: An Exploratory Analysis of the U.S. Political Web Sphere. *J Broadcast Electron Media* 46:222–244. doi: 10.1207/s15506878jobem4602\_4
- Foxman AH, Wolf C (2013) *Viral Hate*. Palgrave Macmillan, New York
- Good KD (2013) From scrapbook to Facebook: A history of personal media assemblage and archives. *New Media Soc* 15:557–573. doi: 10.1177/1461444812458432
- Hancock JT (2012) *Digital deception: Why, when and how people lie online*. Oxford University Press, Oxford
- Herman ES, Chomsky N (1988) *Manufacturing Consent: The Political Economy of the Mass Media*. Pantheon Books, New York
- Herman ES, Chomsky N (2008) *Manufacturing Consent: The Political Economy of the Mass Media*. The Bodley Head, London

- Jowett GS, O'Donnell V (2012) Propaganda and Persuasion. SAGE Publications, Los Angeles
- King G, Pan J, Robert ME (2017) How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument. GkingHarvard.edu. doi: 10.1017/S0003055417000144
- Kollanyi B, Howard PN, Woolley SC (2016) Bots and Automation over Twitter during the U.S. Election.
- Leiser M (2016) AstroTurfing, “CyberTurfing” and other online persuasion campaigns. *Eur J Law Technol* 7:1–27.
- Linebarger PMA (2010) Psychological Warfare. Coachwhip Publications, Darke County, Ohio
- Lippmann W (1946) Public opinion (Vol. 1). Transaction Publishers, New Brunswick and London
- Milan S (2013) Social Movements and Their Technologies: Wiring Social Change. Palgrave Macmillan, New York
- O'Reilly T (2005) What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>. Accessed 6 Oct 2017
- Phillips W (2012) The House That Fox Built: Anonymous, Spectacle, and Cycles of Amplification. *Telev New Media* 14:494–509. doi: 10.1177/1527476412452799
- Piper P (2001) Better Read That Again: Web Hoaxes and Misinformation.
- Poell T, van Dijck J (2015) Social Media and Activist Communication. In: Atton C (ed) *The Routledge companion to alternative and community media*. Routledge, London, pp 527–537
- Roberts ST (2016) Commercial Content Moderation: Digital Laborers' Dirty Work. In: Noble SU, Tynes B (eds) *The Intersectional Internet: Race, Sex, Class and Culture Online*. Peter Lang, New York, pp 147–160

- SantaVicca EF (1994) The Internet as reference and research tool: a model for educators. *Ref Libr* 19:225–236.
- Schou J, Farkas J (2016) Algorithms, Interfaces, and the Circulation of Information: Interrogating the Epistemological Challenges of Facebook. *KOME - An Int J Pure Commun Inq* 4:36–49. doi: 10.17646/KOME.2016.13
- Shao C, Ciampaglia GL, Varol O, Jul SI (2017) The spread of fake news by social bots. In: *ArXiv*. <https://arxiv.org/abs/1707.07592v2>. Accessed 5 Oct 2017
- Soley LC, Nichols JS (1987) *Clandestine Radio Broadcasting: A Study of Revolutionary and Counterrevolutionary Electronic Communication*. Praeger, New York
- Sproule MJ (1994) *Channels of Propaganda*. EDINFO Press, Bloomington
- Tate M, Alexander J (1996) Teaching critical evaluation skills for World Wide Web resources. *Comput Libr* 16:49–54.
- Thomson K (2011) White Supremacist Site MartinLutherKing.org Marks 12th Anniversary. *Huffingt. Post*
- Tong Y, Lei S (2013) War of position and microblogging in China. *J Contemp China* 22:292–311. doi: 10.1080/10670564.2012.734084
- van Dijck J (2013) *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press, Oxford
- Wakabayashi D, Shane S (2017) Twitter, With Accounts Linked to Russia, to Face Congress Over Role in Election. In: *New York Times*. <https://www.nytimes.com/2017/09/27/technology/twitter-russia-election.html>.
- Winseck D (2008) Information Operations ‘Blowback’: Communication, Propaganda and Surveillance in the Global War on Terrorism. *Int Commun Gaz* 70:419–441. doi: 10.1177/1748048508096141
- Zhang J, Carpenter D, Ko M (2013) Online astroturfing: A theoretical perspective. *19th Am Conf Inf Syst AMCIS 2013* 4:2559–2565.

Žižek S (2009) Violence: Six sideways reflections. Profile Books, London

Pre-print