Understanding Online Information Operations: Development of an Influence Network for Scientific Inquiry Testing Environment (INSITE)

Courtney Crooks
Advanced Concepts Laboratory
Georgia Tech Research Institute
Atlanta, Georgia
courtney.crooks@gtri.gatech.edu

Tom McNeil
Advanced Concepts Laboratory
Georgia Tech Research Institute
Atlanta, Georgia
tom.mcneil@gtri.gatech.edu

Ben Sawyer
School of Modeling, Simulation & Training
Univerity of Central Florida
Orlando, Florida
sawyer@ucf.edu

Matthew Canham
School of Modeling, Simulation & Training
Univerity of Central Florida
Orlando, Florida
mcanham@isst.ucf.edu

David Muchlinski
School of International Affairs
Georgia Institute of Technology
Atlanta, Georgia
david.muchlinski@inta.gatech.edu

***Abstract*— Influence operations that promote propaganda, disinformation, and the propagation of social hysteria represent an existential threat to the United States. Effective countermeasures must be developed that can respond in near real-time and anticipate future adversarial actions. One of the most significant hurdles to developing effective countermeasures is the lack of a complex and dynamic testing environment that provides adequate assessment of algorithms and automated detection tools. Through the integration of social sciences with applied mathematics, dynamic, multi-factorial phenomena such as social influence and response behavior within complex social systems may be investigated with scientific rigor. The proposed capability will fulfill a critical need for developing a social-centric model to understand and assess complex influence factors and design of social engineering counter-measures that promote national security interests. To develop such a model, the research community also needs an accessible social media platform that is controlled by researchers, for researchers, and can be used to test new ideas in a realistic setting. A researcher-controlled platform will not only provide unprecedented access to data but will also allow researchers to test mitigation intervention strategies that would be impossible to implement in existing social media platforms.**

Keywords—social engineering, influence operations, information operations, national security, computational social science

I. Influence Operations, an Existential Threat to Democracy

Since the 1990s national security policy advisors have been warning about the threat posed by social engineering and influence operations conducted over online social networks. The ability for near-peer state actors, non-state terrorist groups, and international criminal organizations to project power domestically represents a new threat vector that the U.S. is ill-equipped to counter [4]. International extremist movements rely on influence operations both to recruit new members and to further propagate messaging [58]. Social media allows extremist groups to project power through domestic homegrown violent extremists in a way that was not possible a few years ago [37]. While traditional social engineering has focused on influencing psychological and social aspects of a target population, the emergence of Cyber-Physical Systems (CPS) and subsequent interlinkage with Cyber-Social Systems (CSS) has significantly widened the attack surface for malicious actors. For example, a recent study demonstrated the feasibility of disrupting the power grid and causing a blackout by sending consumers false coupons that incentivized increased electricity usage during peak hours [62]. These interconnected systems increase infrastructure complexity and add to the potential for disruption.

Compounding these current challenges, automated agents in the form of "bots" are becoming increasingly sophisticated and relying to a greater degree on machine learning and neural network technologies to become more "human-like" in their actions, and more capable of passing human intelligence tests such as CAPCHA systems [60]. These bots may act in an amplifying or dampening roles to promote or diminish desired messages [25]. Bots may take on false personas as "sock puppets" and may be partially human controlled as "cyborgs" [32].

National security leaders currently lack sufficient analytics-based situational awareness about adversarial influence operations to make optimal decisions in complex social systems scenarios (e.g., political, socioeconomic, public safety

and health infrastructure). Increasing reliance on automation by threat actors needs to be countered with proactively defensive capabilities. One of the most formidable challenges to developing such capabilities is the lack of an environment to assess influence processes and test proactive defenses in.

II. A New Approach to Understanding Online Information Operations

Most research focused on online behavior is reliant on commercial social media platforms, which researchers cannot control but are only able to access through application program interfaces (API) or publicly available datasets [44]. The limitations of relying on these existing platforms are primarily ethical concerns, lack of terrestrial correlates, and the inability to manipulate the platform to test interventions.

Because researchers are limited to the data that they can obtain from social media platforms it is often extremely difficult to impossible to collect and analyze demographic information from research subjects. In cases where this data has been collected, serious breaches of privacy have occurred, as in the Cambridge Analytica's use of this data to micro-target users with political advertisements [73]. This means that research derived from social media behavior will be limited in the degree of psychological insights that might be derived. Another limitation of this data is the lack of insight into corollary terrestrial behaviors. For example, being Facebook friends with someone has a different meaning than being friends with someone "in real life". Lacking access to research participants on online social media platforms limits researcher access to this information and insight.

In order to effectively counter influence operations in online social networks, scientists must develop methods to not only automatically detect and respond to such attacks but to also test intervention and mitigation tools, techniques, and strategies. This capability is nearly impossible to implement in a commercial social media platform that must remain financially viable. A non-commercial online social network that is explicitly designed for research purposes is the only way to effectively test interventions without breaching the ethics associated with informed consent, or risking irreversible commercial damage and platform viability. The development of automated counter-tactics, techniques, and procedures (TTPs) will require an environment that these can be tested in.

III. Exploration of Social Influence through Computational Social Science

Further research is needed to discover explanations for how organized influence operations spread messaging through social network structures, regardless of whether the messaging is pro-social, inadvertently inaccurate, or propaganda. It is possible that content may be identified and interpreted through multi-disciplinary perspectives to gain a deeper understanding of meaning, emotionality, and perhaps intent of specific social group behaviors and dialogue [49]. Analytic tools such as text analyzers and machine learning algorithms, can provide automated analysis capability to assess patterns, themes, and psychologically-based content in language samples, such as those extracted from semi-structured interviews and social media posts, or large disparate data sets, such as those extracted from social media conversations. Keeping humans in the loop as an integral part of the methodological platform can further add contextual richness to patterns discovered by the automated algorithm.

Through the integration of social sciences with applied mathematics (i.e., computational social science), dynamic, multi-factorial phenomena such as social response behavior within complex social systems may be investigated with scientific rigor. An adaptive, generalizable, and scalable model that incorporates complex social psychoanalytic processes and cyber-technical systems factors will enable the development of robust modeling and simulation concepts [19]. The proposed capability will fulfill a critical need for developing a social-centric model to understand and assess complex influence factors and design of social engineering counter-measures that promote national security interests.

It is necessary to bring together diverse perspectives for such research. Methodological developments across computer science and applied mathematics now allow for the automated collection, analysis, and understanding of massive amounts of data previously considered to be unmanageable [34]. Yet these data remain interpretable only within theories carefully constructed and tested by rigorous social analysis [68]. Digital data created by users of social networks provides historical snapshots of political collective action. Such data have been analyzed to understand and predict political instability [61], radicalization and violence [13], terrorist recruitment [49], and electoral violence [48]. Previous analyses show that it is possible to discover psychological profiles of social media users from the language they use on these platforms, and that such profiles may be useful in predicting which type of user is likely to be more or less susceptible to influence operations [51]. Leveraging the massive volume of data humans use to imbue events, messages, and communities with meaning on social media, can uncover the meaning-making process which leads social media users to engage in diverse forms of political action [75].

Prior research has demonstrated that sound social science theoretical foundations, coupled with computational and applied mathematics such as machine learning, provide complimentary methodological means by which to analyze text-intensive data and large data sets produced by various forms of media, to formulate dynamic, data-driven models of complex social phenomena that are difficult to represent by other means [7, 40, 14, 66]. Through network modeling and data-driven dynamic visualizations, researchers may create artificial society simulations to study social change in response to impactful societal events [63].

IV. Multi-theoretical Bases of Social Influence

A. Social Psychoanalytic Explanations of Influence

Social psychoanalytic literature discusses common themes related to how the formation of belief systems, individual, and social identity, lead to reinforcement of those beliefs through observable behaviors in order to avoid conflict with conscious and unconscious identity architectures (for a comprehensive

discussion of social psychoanalytic perspectives, see [43]). Reference [70] summarized several theoretical underpinnings within the social and behavioral sciences that contribute to understanding social engineering campaigns. What remains to be fully understood, is under what conditions can identity and belief structures be manipulated, altered, or become dependent on, factors present in the experienced environment.

Theories of moral development help explain potential motivational forces that contribute to pro-social and radicalized behaviors in a social system. Reference [45], in his historic discussion on obedience to authority, discussed societal factors that encourage detachment from moral consequences and foster blind obedience to authority figures. Since Milgram's research, others have explored how similar social factors contribute to terroristic or radicalized behaviors. Conversely, reference [57]'s review of violent social mobilizations discussed evidence for evolutionary psychological factors that assist most people in avoiding large group manipulation, radicalization, and mobilization. Reference [57] argued that the individuals most likely to be manipulated by demagogue leaders and disinformation are in some way already predisposed toward engaging in conflict, and that it is the coordination of these individuals into one large group capable of violent action that is the key element that must be achieved for mobilization to occur.

Psychoanalytic theory discusses demoralization as a contributing factor for susceptibility to conversion and influence by social groups, such as those that could be considered radicalized or cults [13, 5]. Demoralization in this context, is described as failure to meet expectations, inability to cope with a pressing problem, and sense of being powerless, hopeless, helpless, and isolated [28]. Demoralization is thought to be a common characteristic of crisis or social breakdown and encourages a state in which the demoralized person may be more susceptible to suggestion by others. Although the demoralized person may also experience a heightened suspicion of others, this may paradoxically result in a higher likelihood to engage in help-seeking behavior and a greater likelihood to trust those providing help [28]. "Such demoralization increases a person's susceptibility to emotionally charged methods of influence that arouse hope by offering detailed guides to behavior based on an inclusive, infallible assumptive world [28, p85]". The concept of moral vindications, which according to reference [42] are formed and reinforced through experiential learning, also contribute to the understanding of social breakdown, help-seeking motivation, and susceptibility to emotional influence [42, 21].

Shared group identity, social support structure, and influential leader factors are discussed throughout the open literature as key features associated with ideologies and national identities potentially associated with propensity to radicalize [e.g., 24, 71, 72, 69]. For example, reference [69] leverages psychoanalytic concepts to propose a continuum of belief structures and behavior describing the path toward radicalization. That is, on one end of the continuum is a committed social activist; whereas in the center, a fanatic idealist that is not yet aggressively behaving; and finally, on the opposite end is an aggressive martyr turning against himself or others. Reference [69] also draws attention to past trauma and shame/humiliation as key historical experiences that could provide common links and common formative factors among radicalized individuals. In line with reference [69]'s ideas, is reference [74]'s discussion of social psychological processes contributing to the psychology of terrorism. They investigate three contributing factors to evolving terroristic motivations: individual need to engage in political violence, ideological narrative justifying political violence, and social network that promotes radicalization through some means.

B. Cognitive and Trait-Based Explanations of Influence

Reference [16] research bridged moral development and cognitive processes, suggesting that both have influence on decision-making; specifically, difficult decisions such those eliciting an approach-avoidance reaction. Their findings argued that systemic, strategic, and tactical motivations must each be considered in order to understand what underlying factors may have influenced decisions, behaviors, and outcomes of approach-avoidance scenarios. Another contribution of cognitive research toward understanding influence, is from the literature on cognitive bias in information processing [39]. Reference [39] suggests that predispositions to perceiving and processing information in a particular way are based in part on beliefs, prior experience, and bottlenecks in attentional capabilities. Advances in neuroscience research suggest that specific areas of the brain may become activated when we are presented with emotionally charged stimuli. Presumably, this contributes to emotional reactivity followed by emotionally-driven observable behaviors. Thus, bad actors can manipulate by influencing our underlying cognitive processes and outward behaviors, if they can manage to produce and deliver effective emotional stimuli to their target audience [38].

Attempts to investigate the impact of social media usage on socio-behavioral constructs such as cognitive, relational, moral, and personality functioning are increasing rapidly [e.g., 64, 22, 3]. Research in the open literature has investigated personality cues emerging from social networking system (SNS) data by leveraging trait theory such as Goldberg's factor model (aka, the Big Five) [17, 33]. In other research, reference [66] found that people with typically undesirable personality features associated with the Dark Triad (i.e., psychopathy, narcissism, and Machiavellianism) [54] were more likely to arrive at utilitarian versus deontological moral decisions, and less likely to perceive meaning in life.

Reference [20] investigated relationships among social networking preferences and personality features of university students, presumed to have been raised in an environment where digital information and associated technologies factor prominently, as represented on the Digital Natives Scale [67]. Among a battery of 13 questionnaires administered, were the Short Dark Triad (SD3) [56], Varieties of Sadistic Tendencies (VAST) [54], and Spheres of Control (SOC) scale to represent cognitive style [55]. Findings from this survey suggested that Machiavellianism was positively correlated to thriving on instant gratification, while psychopathy was negatively correlated with growing up with technology. VAST scores for Direct Sadism, and Vicarious Sadism, were all negatively correlated to growing up with technology, but positively

correlated to reliance on graphics. Growing up with technology was also positively correlated to Spheres of Control (SOC) overall score, personal control, and interpersonal control. Several future studies could be developed from these findings to continue research on relationships between growing up in a technological environment and empathy.

Research exploring vulnerability to Remote Online Social Engineering (ROSE) attacks connects cognitive influence with trait explanations, indicating that some users are more susceptible to social engineering than others [9, 8, 2, 46, 65, 52]. ROSE attacks are recognized as a serious, current societal threat particularly in the cyberspace domain [41]. One potential factor is that certain techniques may be cognitively irresistible when tailored to victims' specific personality profiles. Research within the domain of tailored marketing supports this perspective with some individuals being differentially susceptible to certain influence tactics [50]. Ongoing empirical research suggests that individuals with certain traits may also be more vulnerable to specific messaging in ROSE attacks [9, 65].

C. Social Network Explanations of Influence

Being social animals, human beings are influenced by their perceived peer group [11], and those individuals they identify closely with [12]. Because of this, the structure of our social networks can have a tremendous impact the evolution of our moral orientation and whether we judge circumstances positively or negatively [30]. An ongoing debate rages between network structuralist and social relationalists about whether social network structure or social relationships have more influence over an individual's attitudes, behaviors, and cognitions [29]. A structuralist perspective considers factors such as the number of connections that have a certain property, and how this might influence an individual's own behavior. For example, having a number of obese friends means that one's own likelihood of being obese increases [10].

In contrast, a network relationist perspective considers the degree of influence that a given connection has on an individual. For example, an older sibling may have greater influence on an individual than a cousin might, even though these are both first degree connections from a social network structural perspective [29]. These considerations are crucial to understand within an IO context because they suggest different approaches to delivering counter-messaging to an adversarial campaign. The first approach emphasizes the number and strength of connections, while the latter emphasizes the degree and valence of influence that different individuals present.

Unfortunately, little quantitative research exists linking violent political behavior to online influence operations [15]. Social media offers prospective radicals an opportunity to develop social ties and find validation through others, thus providing the critical element of social interaction at relatively low cost [36]. The make-up of these online communities, and how processes of radicalization cascade across and/or within them, is poorly understood. Radical groups use social media to spread their violent messages and to reach global audiences [77]. These websites, by linking ideologically affiliated individuals, enable extremists to forge a sense of shared identity, even if it is only virtual [59], which may facilitate violent action on the part of some members. Closeness of individual ties to such groups is positively correlated with the production of violent or hateful messages in online communities [18].

V. Understanding the Evolution of Social Extremism

In the current information environment, social media communities, or other online targets such as video gaming and other technology hobbyist groups, may be targeted by extremist groups to facilitate actions that encourage destructive public activity, such as spreading propaganda, influencing social factors, and recruiting new members [26, 1]. The role of information operations on cultural ideology, the effects on behaviors of exposed online communities, and the risks that potential extreme social behaviors pose to our national security, need to be considered in light of our current understanding of destabilizing social responses such as radicalization and moral panic. Social engineering seeks to influence social attitudes and behaviors on a population scale by exploiting trust or fear; and the human being is demonstrably vulnerable to this influence [47, 76]. Malicious actors may use disinformation tactics to encourage moral panic, distrust in public institutions, and irresponsible or harmful social behaviors [53, 23]. Further, poor operational security (OPSEC) practices play a significant role in exposing service members to potential attacks by online threats. A report by the Network Contagion Research Institute (NCRI) suggested that military communities may be especially susceptible to social engineering attacks by insurgent groups attempting to promote radicalization of military members through socially networked memes [31].

Exposure to violent radicalization themes online may also influence online opinions and offline behaviors toward the promoted violent content [35]. Malicious actors are aware of this opportunity to exert their influence, and actively seek out potentially vulnerable service members. We also argue that it is important to remember how social engineering campaigns can foster a spectrum of behaviors; the end goals of such a campaign could range from societal destabilization to constructive pro-social activism. For example, reference [70] summarized socio-psychological concepts contributing to the promotion of healthy public behaviors during the COVID-19 pandemic. Reference [70]'s paper illustrates why we need to study not only how radicalized behavior emerges and leads to the potential for destructive activities, but also how we are resilient to and can combat these destructive responses.

VI. Proposed Research Platform

We submit that developing a research-centric social network environment that can be tailored, semi-contained, and observed as a naturalistic community, is key to empirically exploring and understanding the complex interplay of these factors to counter adversarial influence operations. Through this naturalistic social network environment, the research community will have the capability to explore critical research questions, discover and trace influence vectors, and test interventions to mitigate propagation of targeted information campaigns across this platform.

The development of such an Influence Network for Scientific Inquiry Testing Environment (INSITE) should meet the following minimal requirements. INSITE participation would be exclusive to those actively signed up as participants in an IRB-approved study at a given institution. First, studies using the INSITE should have the ability to provide participants with informed consent [27], and the ability to opt out of the study at any time as well as have their data removed. Second, the INSITE should be localized to the institution (or institutions) involved in the research. Third, the INSITE needs to have the ability to manipulate the network structure and platform content. This will allow for the testing of interventions and automated detection techniques. Fourth, the INSITE should be compatible with both iOS and Android devices. Finally, the code base should be made openly and freely available to the greater research community. Hosting this code on a Git Hub or equivalent platform will allow for both widespread access but also community involvement in improving the platform and developing additional capabilities.

We propose that such a platform will offer several advantages to the research community to build a better understanding of how influence behaves and propagates across social media and social networks. Research dedicated INSITE would address ethical concerns of conducting research on a commercial platform by explicitly informing research volunteers about the nature of the research study [27], the INSITE research platform, and nature of data collection. The development of an INSITE will create the ability for researchers to investigate the influence of "deep" psychological factors, such as personality and other persistent individual traits, on the diffusion and acceptance of messages across the environment. In this environment, researchers will have the ability to follow terrestrial correlates to online behaviors. Currently, an assumption of most online influence studies is that online behaviors correspond to behavior in real life (IRL); however, as previously stated a "Facebook friend" is not necessarily equivalent to having a friend IRL. The INSITE will allow researchers to track research subjects over time through both online behavior and external psychometric assessments to gain insights about the evolution of both platform-level behaviors and individual psychological development. This will provide tremendous opportunity to observe the influence that memes and other online phenomenon are having on the attitudes of the individual. The INSITE will allow researchers to manipulate platform dynamics, empirically derive and test potential disinformation countermeasures, and develop real-time automated social engineering detection and mitigation techniques.

Current practices of "de-platforming" or "shadow banning" of social media personalities is extremely controversial and may violate free speech protections. Furthermore, these actions are not known to be an effective means of countering disinformation and may also have the opposite effect of increasing saliency of disinformation. For example, one de-platformed individual was recently nominated to the Republican Party ticket for a Florida Congressional seat. It is currently difficult or impossible to directly compare the results of one study exploring online influence on Facebook, with another study of influence on Twitter. The development of an INSITE will allow for better comparison across research studies. A researcher-controlled platform will not only provide unprecedented access to data but will also allow researchers to test disinformation countermeasure strategies that would be impossible to implement in existing social media platforms.

REFERENCES

[1] Alarid, M. (2016). Recruitment and Radicalization: The Role of Social Media and New Technology. Prism: The Journal of Complex Operations, xxx.

[2] Albladi, S. M., & Weir, G. R. (2018). User characteristics that influence judgment of social engineering attacks in social networks. Human-centric Computing and Information Sciences, 8(1), 5.

[3] Andreassen, C., Pallesen, S., & Griffiths, M. (2017). The relationship between excessive online social networking, narcissism, and self-esteem: Findings from a large national survey. Addictive Behaviours, 64, 287-293.

[4] Arquilla, J., & Ronfeldt, D. (2001). Networks and netwars: The future of terror, crime, and militancy. Rand Corporation.

[5] Barak, A., & Suler, J. (2008). Reflections on the Psychology and Social Science of Cyberspace. In A. Barak, Psychological aspects of cyberspace: Theory, research, applications (pp. 1-12). Cambridge, UK: Cambridge University Press.

[6] Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. Cognition, 121, 154-161.

[7] Burger, A., Oz, T., Kennedy, W. G., & Crooks, A. T. (2019). Computational Social Science of Disasters: Opportunities and Challenges. Future Internet, 11(103), 1-31. doi:10.3390/fi11050103

[8] Canham, M., Constantino, M., Hudson, I., Fiore, S. M., Caulkins, B., & Reinerman-Jones, L. (2019). The enduring mystery of repeat clickers. Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019). USENIX Advanced Computing Systems Association.

[9] Canham, M., Posey, M. C., Strickland, D., & Constantino, M. (unpublished). Phishing for Long Tails: Examining Organizational Repeat Clickers and Protective Stewards. Sage Special Issue on Organizational Cybersecurity.

[10] Christakis, N. A., & Fowler, J. H. (2009). Connected: The surprising power of our social networks and how they shape our lives. Little, Brown Spark.

[11] Cialdini, R. (2016). Pre-suasion: A revolutionary way to influence and persuade. . Simon and Schuster.

[12] Cialdini, R. B., & Cialdini, R. B. (2007). Influence: The psychology of persuasion (Vol. 55). New York: New York: Collins.

[13] Cohen, S. J. (2019). The unconscious in terror: An overview of psychoanalytic contributions to the psychology of terrorism and violent radicalization. Int J Appl Psychoanal Studies, 216-228.

[14] Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., . . . D. Helbing, D. (2012). Manifesto of computational social science. The European Physical Journal: Special Topics, 214, 325–346.

[15] Conway, M. (2017). Determining the role of the internet in violent extremism and terrorism: Six suggestions for progressing research. Studies in Conflict & Terrorism, 40(1), 77-98.

[16] Cornwell, J. F., & Higgins, E. T. (2015). Approach and avoidance in moral psychology: Evidence for three distinct motivational levels. Personality and Individual Differences, 86, 139-149.

[17] Correa, T., Bachmann, I., Hinsley, A. W., & de Zúñiga, H. G. (2013). Personality and Social Media Use. Organizations and Social Networking: Utilizing Social Media to Engage Consumers, 21, 41-61. doi:doi:10.4018/978-1-4666-4026-9.ch003

[18] Costello, M., & Hawdon, J. (2018). Who are the online extremists among us? Sociodemographic characteristics, social networking, and online experiences of those who produce online hate materials. Violence and Gender, 5(1), 55-60.

[19] Crooks, C. L. (unpublished). Influence and Vulnerability in the Information Environment: Implications for Cyber-Enabled Information Operations and National Security. In M. (. Kosal, Innovate for Future Threats: Disruptive Innovation Efforts and Uses of the Technology Environment by State and Non-State Actors.

[20] Crooks, C. L., Martin, S. M., Fang, S., Feinberg, R., & Narvilkar, M. (2020). Social Networking Behavior and Psycho-Social Factors of the Digital Native (Unpublished Manuscript).

[21] Dahmer, H. (1993). Psychoanalytic social research. Free Associations, 3(4), 490-99.

[22] De Choudhury, M., & Kiciman, E. (2017). The language of social support in social media and its effect on suicidal ideation risk. ICWSM, 32-41.

[23] Del Vicario, M., Bessi, A., Zollo, F., Petronic, F., Scala, A., Caldarelli, G., . . . Quattrociocchi, W. (2016). The spreading of misinformation online. PNAS, 113(3), 554-559. doi:doi/10.1073/pnas.1517441113

[24] Drury, J. (2020). Recent developments in the psychology of crowds and collective behaviour. Current Opinion in Psychology, 35, 12-16. Retrieved from https://doi.org/10.1016/j.copsyc.2020.02.005

[25] Dubois, E., & McKelvey, F. (2019). Political Bots: Disrupting Canada's Democracy. Canadian Journal of Communication, 44(2).

[26] Ferrera, E. (2015). Manipulation and abuse on social media. SIGWEB Newsletter.

[27] Flick, C. (n.d.). Informed consent and the Facebook emotional manipulation study. Research Ethics, 12(1), 14-28.

[28] Frank, J. D., & Frank, J. B. (1993). Persuasion & Healing, 3rd edition. MD: Baltimore: Johns Hopkins University Press.

[29] Friedkin, N. E. (2006). A structural theory of social influence (Vol. 13). Cambridge University Press.

[30] Friedkin, N. E., Proskurnikov, A. V., & Bullo, F. (2019). Positive contagion and the macrostructures of generalized balance. Network Science, 7(4), 445-458.

[31] Goldenberg, A., & Finkelstein, J. (2020). Cyber Swarming, Memetic Warfare and Viral Insurgency: How Domestic Militants Organize on Memes to Incite Violent Insurrection and Terror Against Government and Law Enforcement. Network Contagion Research Institute.

[32] Gorwa, R., & Guilbeault, D. (2020). Unpacking the social media bot: A typology to guide research and policy. Policy & Internet, 12(2), 225-248.

[33] Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in onlien social networks: Self-reported Facebook-related behaviors and observable profile information. Cyberpsychology, Behavior, and Social Networking, 14(9), 483-488.

[34] Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political analysis, 21(3), 267-297.

[35] Hassan, G., Brouillette-Alarie, S., Alava, S., Frau-Meigs, D., Lavoie, L., Fetiu, A., . . . Sieckelinck, S. (2018). Exposure to Extremist Online Content Could Lead to Violent Radicalization: A Systematic Review of Empirical Evidence. International Journal of Developmental Science, 12(1-2), 71-88 .

[36] Helfstein, S. (2012). Edges of radicalization: Individuals, networks and ideas in violent extremism. Combating Terrorism Center at West Point.

[37] Herridge, C. (2012). The Next Wave: On the Hunt for Al Qaeda's American Recruits. Crown Forum.

[38] Heslen, J. J. (2020). Neurocognitive hacking: A new capability in cyber conflict? Politics and the Life Sciences, 39(1), 87-100.

[39] Hills, T. T. (2018). The dark side of information proliferation. Perspectives on Psychological Science, 1-8. Retrieved from https://journals.sagepub.com/doi/10.1177/1745691618803647

[40] Keuschnigg, M., Lovsjo, N., & Hedstrom, P. (2018). Analytical sociology and computational social science. J Comput Soc Sc, 1, 3-14. Retrieved from https://doi.org/10.1007/s42001-017-0006-5

[41] Krombholz, K., Heidelinde, H., Huber, M., & Weippl, E. (2015). Advanced social engineering attacks. Journal of Information Security and Applications, 22, 113-122.

[42] Kumar, V. (2017). Moral vindications. Cognition, 167, 124-134. Retrieved from http://dx.doi.org/10.1016/j.cognition.2017.05.005

[43] Layton, L. (2020). Toward a Social Psychoanalysis: Culture, Character, and Normative Unconscious Processes. New York, NY: Routledge.

[44] Meel, P., & Vishwakarma, D. K. (2019). Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. Expert Systems with Applications, 112986.

[45] Milgram, S. (2009). Obedience to Authority. New York: NY: Harper Perennial Modern Thought.

[46] Moody, G. D., Galletta, D. F., & Dunn, B. K. (2017). Which phish get caught? An exploratory study of individuals′ susceptibility to phishing. European Journal of Information Systems,, 26(6), 564-584.

[47] Mouton, F., Leenen, L., & Venter, H. S. (2016). Social engineering attack examples, templates and scenarios. Computers & Security, 59, 186-209.

[48] Muchlinski, D., Yang, X., Birch, S., Macdonald, C., & Ounis, I. (unpublished). We need to go deeper: Measuring electoral violence using convolutional neural networks and social media.

[49] O'Halloran, K. L., Tan, S., Wignell, P., Bateman, J. A., Pham, D. S., Grossman, M., & Moere, A. V. (2019). Interpreting text and image relations in violent extremist discourse: A mixed methods approach for big data analytics. Terrorism and Political Violence, 31(3), 454-474.

[50] Oyibo, K., Orji, R., & Vassileva, J. (2017). Investigation of the Influence of Personality Traits on Cialdini's Persuasive Strategies, 13. 13.

[51] Panicheva, P., Ledovaya, Y., & Bogolyubova, O. (2016). Lexical, morphological and semantic correlates of the dark triad personality traits in Russian facebook texts. (pp. 1-8). Artificial intelligence and natural language conference (AINL), IEEE .

[52] Pattinson, M., Jerram, C., Parsons, K., McCormac, A., & Butavicius, M. (2012). Why do some people manage phishing e-mails better than others? Information Management & Computer Security, 20(1), 18-28.

[53] Paul, C., & Matthews, M. (2016). The Russian "Firehose of Falsehood" Propaganda Model. RAND.

[54] Paulhus, D. L., & Jones, D. N. (2015). Measures of dark personalities. In Measures of Personality and Social Psychological Constructs (pp. 562-594). Elsevier. doi:http://dx.doi.org/10.1016/B978-0-12-386915-9.00020-6

[55] Paulhus, D. L., & Van Selst, M. (1990). The spheres of control scale: 10 yr of research. Personality & Individual Differences, 11(1), 1029-1036.

[56] Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. Journal of Research in Personality, 36, 556-563. doi:10.1016/S0092-6566(02)00505-6

[57] Peterson, M. B. (2020). The evolutionary psychology of mass mobilization: how disinformation and demagogues coordinate rather than manipulate. Current Opinion in Psychology, 35, 71-75. Retrieved from https://doi.org/10.1016/j.copsyc.2020.02.003

[58] Petraeus, D., & Amos, J. (2006). US Army/Marine counterinsurgency field manual.

[59] Post, J. M. (2005). When hatred is bred in the bone: Psycho-cultural foundations of contemporary terrorism. Political Psychology, 26(4), 615-636.

[60] Priyadarshini, I., & Cotton, C. (2019, October). Internet Memes: A Novel Approach to Distinguish Humans and Bots for Authentication. In Proceedings of the Future Technologies Conference, (pp. 204-222).

[61] Ramakrishnan, N., Butler, P., Muthiah, S., Self, N., Khandpur, R., Saraf, P., & Kuhlman, C. (2014, August). 'Beating the news' with EMBERS: forecasting civil unrest using open source indicators. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, 1799-1808.

[62] Raman, G., AlShebli, B., Waniek, M., Rahwan, T., & Peng, J. C. (2020). How weaponizing disinformation can bring down a city's power grid. PloS one, 15(8), e0236517.

[63] Sawyer, R. K. (2004). Social explanation and computational simulation. Philosophical Explanations, 7(3), 219-231.

[64] Singh, S., Farley, S. D., & Donahue, J. J. (2018). Grandiosity on display: Social media behaviors and dimensions of narcissism. Personality and Individual Differences, 134, 308-313.

[65] Sudzina, F., & Pavlicek, A. (2017). Propensity to click on suspicious links: Impact of gender, of age, and of personality traits. BLED.

[66] Taylor, G. R. (1977). Prediction and social change: The need for a basis in theory . Futures, 404-414.

[67] Teo, T. (2013). An initial development and validation of a Digital Natives Assessment. Computer & Education, 67, 51–57.

[68] Titiunik, R. (2015). Can big data solve the fundamental problem of causal inference? PS: Political Science & Politics, 48(1), 75-79.

[69] Twemlow, S. W. (2005). Extreme Religious Fundamentalism and Violence: Some Psychoanalytic and Psychopolitical Thoughts. The International Journal of Psychoanalysis, 86(4), 957-973. doi:10.1516/EY30-319C-5DLJ-VCL3

[70] Van Bavel, J. J. et al. (2020). Using social and behavioural science to support COVID-19 pandemic response. Nature: Human Behavior. Retrieved from https://doi.org/10.1038/s41562-020-0884-z

[71] Volkan, V. D. (2009). Large group identity, international relations, and psychoanalysis. International Forum of Psychoanalysis, , 18(4), 206-213.

[72] Volkan, V. D., & Fowler, J. C. (2009). Large-group Narcissism and Political Leaders with Narcissistic Personality Organization. Psychiatric Annals, 39(4), 214-223.

[73] Ward, K. (2018). Social networks, the 2016 US presidential election, and Kantian ethics: applying the categorical imperative to Cambridge Analytica's behavioral microtargeting. Journal of media ethics, 33(3), 133-148.

[74] Webber, D., & Kruglanski, A. W. (2018). The social psychological makings of a terrorist. Current Opinion in Psychology, 19, 131-134.

[75] Wignell, P., Tan, S., O'Halloran, K. L., & Lange, R. (2017). A mixed methods empirical examination of changes in emphasis and style in the extremist magazines Dabiq and Rumiyah. Perspectives on Terrorism, 11(2), 2-20.

[76] Workman, M. (2007). Gaining access with social engineering: An empirical study of the threat. Information Systems Security, 16(6), 315-331. doi:10.1080/10658980701788165

[77] Zhou, Y., Reid, E., Qin, J., Chen, H., & Lai, G. (2005). US domestic extremist groups on the Web: link and content analysis. IEEE intelligent systems, 20(5), 44-51.