Cyberpsychology, Behavior, and Social Networking
Volume 21, Number 2, 2018
© Mary Ann Liebert, Inc.
DOI: 10.1089/cyber.2016.0706

A Game Theoretical Approach to Hacktivism: Is Attack Likelihood a Product of Risks and Payoffs?

Jessica E. Bodford, PhD, and Virginia S.Y. Kwan, PhD

Abstract

The current study examines hacktivism (i.e., hacking to convey a moral, ethical, or social justice message) through a general game theoretic framework—that is, as a product of costs and benefits. Given the inherent risk of carrying out a hacktivist attack (e.g., legal action, imprisonment), it would be rational for the user to weigh these risks against perceived benefits of carrying out the attack. As such, we examined computer science students' estimations of risks, payoffs, and attack likelihood through a game theoretic design. Furthermore, this study aims at constructing a descriptive profile of potential hacktivists, exploring two predicted covariates of attack decision making, namely, peer prevalence of hacking and sex differences. Contrary to expectations, results suggest that participants' estimations of attack likelihood stemmed solely from expected payoffs, rather than subjective risks. Peer prevalence significantly predicted increased payoffs and attack likelihood, suggesting an underlying descriptive norm in social networks. Notably, we observed no sex differences in the decision to attack, nor in the factors predicting attack likelihood. Implications for policymakers and the understanding and prevention of hacktivism are discussed, as are the possible ramifications of widely communicated payoffs over potential risks in hacking communities.

Keywords: hacking, hacktivism, decision making, descriptive norms, game theory

Introduction

A *cyber-attack* is defined as an attempt to damage, destroy, or gain illegal access to a computer network or system. Within the broader family of cyber-attacks falls *hacking*, or the act of gaining access to—and control over—third-party computer systems.[1] Although *hacktivism*, which forms the focus of this work, shares many of the methods observed in hacking, its name evinces a clear distinction that sets it apart from other forms of cyber-attacks: As Jordan and Taylor[2] aptly state, ''hacktivism is activism gone electric'' (p. 3)—it is the emergence of protest in cyberspace to make a statement that is often morally, politically, ethically, or legally driven,[3] estimated to incur costs up to $114 billion each year to governments and corporations.[4] To be sure, such attacks should not be carried out lightly: In the past, apprehended members of major hacktivist groups had faced criminal charges, including fines up to $183,000 and prison sentences of a maximum of 10 years.[5,6] Thus, before taking part in a hacktivist attack, it is imperative that, in line with game theory, the member weigh both end goals (hereafter referred to as *payoffs*) and consequences if caught (*risks*).[7,8]

Game theory states that within a zero-sum game between decision makers with limited possible moves, it is possible to calculate the strategies that would guarantee the optimal payoffs for each player given the other player's actions. In its most general sense, this framework holds that rational decision makers should carefully weigh all known consequences of their decisions before acting. In the present study, game theory can be considered a strategy by which an individual chooses an action (here, to attack or not to attack) based on available information (risks and payoffs involved). Because neither risks nor payoffs are typically known before taking part in an attack, they are both subjective in nature—the hacktivist must estimate the probability and severity of both risks and payoffs, and base their final decision of attack on these estimations.

Hacktivism and Anonymous

Since the late 1990s, hacktivism has evolved into a distinct, publicly recognized form of protest, with hacktivists commonly portrayed as watchdogs that regularly perform valued functions for society—namely, the enforcement of rights to freedom of opinion, expression, and information.[9,10]

These principles have been closely tied with hacktivist activity, such that in the face of social injustice or restrictions of basic human rights, hacktivists are likely to rise to action. Due to the dispersed and largely unidentifiable nature of hacktivist groups (e.g., Anonymous), such actions are

Department of Psychology, Arizona State University, Tempe, Arizona.

primarily coordinated through online forums that guarantee full anonymity to users, including the exclusion of user IP addresses to all but system administrators (e.g., 4chan).[11] In the wake of a triggering event, one or more hacktivists would post a call to action on 4chan with details of the trigger events precipitating the attack, as well as details regarding how members should join in this attack.[12] Responses to these calls may range from dozens to tens of thousands of members depending on the trigger event and necessary actions.

Research Question and Hypothesis

Therefore, we seek to explore the following question: To what degree do hackers' perceptions of risks and payoffs guide their determined likelihood of carrying out a hacktivist attack? Following game theory, one would expect participants to form mental ratios of the risks and payoffs inherent in a given hacktivism scenario, and that this ratio will help guide their overall perceived likelihood of attack—that is, participants would consider an attack more likely as risks decrease and payoffs increase.

Furthermore, it is both timely and important to explore whether sex differences and prevalence of peer hacking behaviors predict these perceptions. In past research on social activism, social activists tend to be young, politically active, White, highly educated, and male[13]—qualities that have also described many well-known Anonymous members in the past.[14,15] Beyond their youth, there exists an implicit assumption that a notable majority of hacktivists are male; however, little research has addressed whether sex differences exist in attack likelihood. Regarding peer hacking behaviors, research on *descriptive norms* suggests that the degree to which known others act on a particular behavior influences one's own willingness to engage in that behavior.[16] Applied to the present study, if several friends had engaged in hacking in the past, an individual may be more likely to consider hacking to be an ordinarily approved behavior and a rewarding experience, and to increase his or her perception of attack likelihood.

Method

Participants

Undergraduate students who enrolled at a large, public university in the United States were invited to participate in an online study. Participants were current majors in computer science, computer science engineering, informatics, information systems, or a related discipline to ensure sufficient technological knowledge to (a) fit the demographic build of those we might expect to use 4chan and similar Anonymous communication channels, (b) recognize and consider partaking in Anonymous calls to action, and (c) consider themselves prospective "Anons" (i.e., members). Four hundred seventy-four students (43.9 percent female) participated in the study, displaying a mean age of 22.60 ($SD = 3.36$) and a modal age of 19. Overall, 63.8 percent of our sample identified as White, and 17.1 percent identified as Asian or Middle-Eastern.

Materials

Participants read a scenario that aimed at inducing them into the frame of mind of a student who happens to come across an Anonymous call to action. They then answered a series of questions regarding perceived risks, payoffs, overall attack likelihood, peer prevalence, and demographic variables.

Guided visualization scenario. Each participant viewed a short, guided visualization scenario asking him or her to adopt the mindset of an imagined third-person target who was similar in age range, university affiliation and major, and gender. Decades of psychological research have supported guided visualization as a strategy that increases the vividness and clarity of participants' mental imagery through text-based prompts.[17,18] We asked participants to consider the actions of an imagined other to avoid the possibility that participants would withhold information or opinions due to conflicting interests, morals, or values (Fig. 1B).

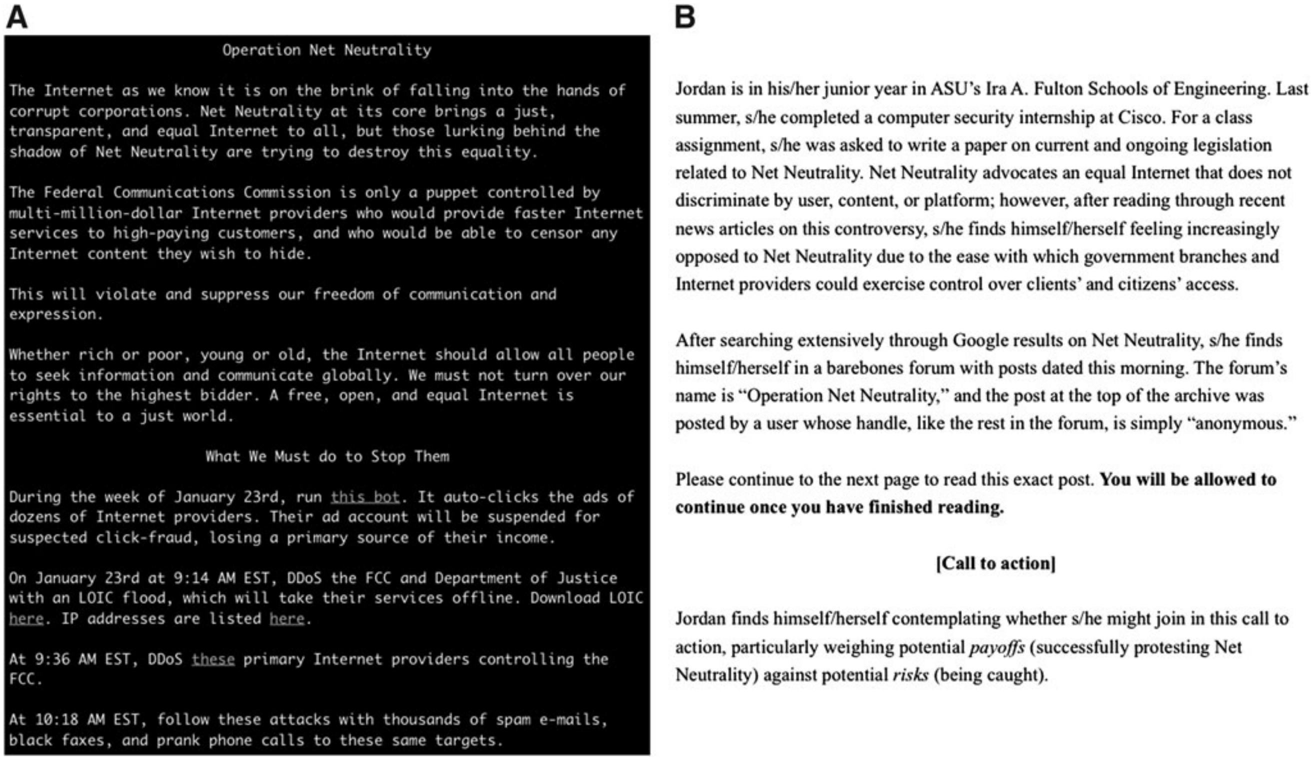
Call to action. Because past hacktivist activity had primarily revolved around issues of social injustice,[14,19,20] we composed a call to action stemming from a trigger event that was socially unjust and also relevant at the time of data collection—namely, Net Neutrality. Our guided visualization scenario outlined a situation in which, while Googling up-to-date news stories on the unfolding controversy, "Jordan" stumbles across a page containing a call to action that directly opposes Net Neutrality. "Jordan" acted as our third-party target name due to its gender neutrality, with pronouns altered to convey either a male or female subject depending on the participant's own gender. The language and format (e.g., font, wording; see Fig. 1A) of our call to action closely followed that of an Anonymous-released video available through YouTube.[21]

Risks, payoffs, and likelihood. After undergoing this guided visualization, participants were asked to report subjective risks, separated into probability of being caught and severity of punishment if caught; and subjective payoffs, comprising probability of succeeding and magnitude of payoffs if successful. We, therefore, examined whether:

$$P(\mathrm{A}) = f(P[\mathrm{R_S}] \times M[\mathrm{R_S}]) + f(P[\mathrm{P_S}] \times M[\mathrm{P_S}]) + f(\mathrm{R_S P_S}) \tag{1}$$

in which the probability of attack likelihood ($P[\mathrm{A}]$) is predicted by (a) the probability of a success ($P[\mathrm{P_S}]$) or failure ($P[\mathrm{R_S}]$) occurring, and (b) the magnitude or severity of the benefits ($M[\mathrm{P_S}]$) or detriments ($M[\mathrm{R_S}]$) posed. These items were further delineated into the risks and payoffs posed to Jordan and the broader group of hacktivists to capture self- and collective-focused estimations. We then presented a single item to capture perceived likelihood of attack, worded in the third-person: *How likely do you think Jordan is to participate in this call to action*?

Our independent variables of interest were grouped into three overarching categories: risk, payoff, and attack likelihood. Risk variables comprised (a) subjective risk, (b) likelihood of being caught, and (c) severity of punishment if caught; payoff variables included (a) size of payoffs, (b) likelihood of success, benefit (c) to personal pride, (d) to others, (e) through self-challenge, and ability to (f) boost status, (g) fight Net Neutrality, and (h) raise public awareness against Net Neutrality. Our three risk variables were highly correlated (average $r[470] = 0.518$, $0.415 \leq r \leq 0.694$, $p < 0.001$), as were our eight payoff variables (average $r[447] = 0.425$, $0.225 \leq r \leq 0.663$, $p < 0.001$),



Operation Net Neutrality

The Internet as we know it is on the brink of falling into the hands of corrupt corporations. Net Neutrality at its core brings a just, transparent, and equal Internet to all, but those lurking behind the shadow of Net Neutrality are trying to destroy this equality.

The Federal Communications Commission is only a puppet controlled by multi-million-dollar Internet providers who would provide faster Internet services to high-paying customers, and who would be able to censor any Internet content they wish to hide.

This will violate and suppress our freedom of communication and expression.

Whether rich or poor, young or old, the Internet should allow all people to seek information and communicate globally. We must not turn over our rights to the highest bidder. A free, open, and equal Internet is essential to a just world.

What We Must do to Stop Them

During the week of January 23rd, run this bot. It auto-clicks the ads of dozens of Internet providers. Their ad account will be suspended for suspected click-fraud, losing a primary source of their income.

On January 23rd at 9:14 AM EST, DDoS the FCC and Department of Justice with an LOIC flood, which will take their services offline. Download LOIC here. IP addresses are listed here.

At 9:36 AM EST, DDoS these primary Internet providers controlling the FCC.

At 10:18 AM EST, follow these attacks with thousands of spam e-mails, black faxes, and prank phone calls to these same targets.

Jordan is in his/her junior year in ASU's Ira A. Fulton Schools of Engineering. Last summer, s/he completed a computer security internship at Cisco. For a class assignment, s/he was asked to write a paper on current and ongoing legislation related to Net Neutrality. Net Neutrality advocates an equal Internet that does not discriminate by user, content, or platform; however, after reading through recent news articles on this controversy, s/he finds himself/herself feeling increasingly opposed to Net Neutrality due to the ease with which government branches and Internet providers could exercise control over clients' and citizens' access.

After searching extensively through Google results on Net Neutrality, s/he finds himself/herself in a barebones forum with posts dated this morning. The forum's name is "Operation Net Neutrality," and the post at the top of the archive was posted by a user whose handle, like the rest in the forum, is simply "anonymous."

Please continue to the next page to read this exact post. **You will be allowed to continue once you have finished reading.**

[Call to action]

Jordan finds himself/herself contemplating whether s/he might join in this call to action, particularly weighing potential *payoffs* (successfully protesting Net Neutrality) against potential *risks* (being caught).

FIG. 1. Guided visualization scenario and call to action.

displaying high internal consistency ($\alpha_{\text{risk}} = 0.765$ for three items, $\alpha_{\text{payoff}} = 0.856$ for eight items). As such, we formed single composites of each of these variables to yield an Overall Risk and Overall Payoff variable. Henceforth, any mention of Risk or Payoff will refer to these composite variables. The descriptive statistics of these risk and payoff variables, as well as their final composites, are displayed in Table 1. All variables were measured on 0- to 100-point scales, and they were captured in 10-point Likert-type intervals.

Peer prevalence. Participants whose friends have carried out hacking attacks may have witnessed first-hand stories of risks and payoffs endured to complete the attack. If so, these accounts might confound participants' own estimations. We, therefore, used a single item—namely, *Outside of class assignments, how common do you think it is among your peers to carry out attacks such as this one?*—to capture peer prevalence of hacking behaviors on a scale of 0 to 100 ($M = 27.16$, $SD = 23.05$, $N = 454$). We also investigated whether a participant's self-reported gender (reported before seeing the scenario) impacted perceived risks, payoffs, or attack likelihood. An intercorrelation matrix of all variables of interest is shown in Table 2.

Results

We predicted that participants would follow a general game theoretic framework of decision making in which subjective risks, payoffs, and any interaction of the two should inform their perceived likelihood of carrying out a hacktivist attack. The overall ordinary least-squares regression model (see Eq. 1) was highly significant, $F(3, 465) = 67.915$, $p < 0.001$, $\eta_p^2 = 0.180$, with 30.5 percent of the observed variance in our dependent variable attributed to participants' subjective risks, payoffs, and the interaction between the two; however, only subjective payoffs were a significant predictor of attack likelihood ($\beta = 0.551$, $t = 14.218$, $p < 0.001$). There was no significant two-way interaction ($\beta = 0.005$, $t = 0.141$, $p = 0.888$, ns) or main effect of subjective risk ($\beta = -0.040$, $t = -1.038$, $p = 0.300$, ns).

Table 1. Descriptive Statistics: Risk, Payoff, and Attack Likelihood Variables

Variable	M *(SD)*	N	*Variable*	M *(SD)*	N
Risk	58.73 (26.21)	472	Challenge	46.91 (30.12)	456
CaughtLikely	46.03 (27.42)	469	BoostStatus	41.64 (28.39)	446
PunishSevere	57.76 (27.15)	473	Fight NN	43.05 (29.68)	455
SuccessLikely	33.13 (23.92)	460	Raise Aware NN	51.04 (29.74)	463
SizePayoffs	33.76 (26.30)	457	Risk Overall	54.12 (22.16)	473
BenefitPride	58.81 (26.50)	464	Payoff Overall	44.06 (19.84)	473
BenefitOthers	45.49 (29.62)	461	AttackLikely	43.93 (22.48)	471

Table 2. Intercorrelation Matrix of Variables of Interest

	RiskOverall	*PayoffOverall*	*AttackLikely*
PayoffOverall	0.031	—	—
AttackLikely	–0.022	0.551***	—
PeerPrevalence	0.020	0.410***	0.368***

***$p < 0.001$.

We then broke subjective risks and payoffs down into (a) the probability of a success or failure, and (b) the magnitude or severity of the benefits or detriments posed. The resulting equation would, therefore, appear thus:

$$\begin{aligned} \mathrm{AL} = {} & \beta_0 + \beta_1(\mathrm{R}) + \beta_2(\mathrm{CL}) + \beta_3(\mathrm{SL}) \\ & + \beta_4(\mathrm{SP}) + \beta_5(\mathrm{R} \times \mathrm{CL}) + \beta_6(\mathrm{SL} \times \mathrm{SP}) \end{aligned} \tag{2}$$

where R indicates risks, CL denotes likelihood of getting caught, SL signifies likelihood of success, and SP stands for size of payoffs. This regression model was significant, $R^2 = 0.230$, $F(6, 441) = 22.006$, $p < 0.001$, $\eta_p{}^2 = 0.130$; however, mirroring our previous model, only likelihood of success ($\beta = 0.344$, $t = 5.946$, $p < 0.001$) and size of payoffs ($\beta = 0.128$, $t = 2.247$, $p = 0.025$) were significant predictors of attack likelihood. The impacts of risks, likelihood of getting caught, and both two-way interactions were nonsignificant.

Contrary to expectations, as per a game theoretic framework, the perceived likelihood of carrying out a hacktivist attack was driven by payoffs, but not risks. These findings suggest that college-aged students majoring in computer science would expect a third-person subject similar to themselves to behave less than rational decision makers—that is, to weigh benefits, rather than costs, when choosing to carry out a potentially self-threatening action, possibly overestimating others' recklessness or risk-taking tendencies when faced with a threatening scenario.

Peer prevalence

We next examined whether peer prevalence of hacking played a role in participants' estimations of risks, payoffs, or attack likelihood. The degree to which hacking behaviors were common among participants' peers was positively related to each of our eight payoff variables ($0.175 \leq r[481] \leq 0.411$, $p < 0.001$), as well as attack likelihood ($r[494] = 0.368$, $p <$ 0.001). However, it was not related to all but one of our risk measures ($0.03 \leq |r[491]| \leq 0.05, p > 0.05, ns$); namely, those whose peers had engaged in hacking behaviors in the past estimated a higher likelihood of getting caught ($r[489] = 0.103$, $p = 0.02$), but not a higher degree of overall risk or punishment severity (Table 2). Taken together, these findings suggest that among participants whose friends carry out hacking behaviors, the risks involved in a pending attack are irrelevant to the subsequent decision to hack; instead, payoffs seem consistently greater when peers had hacked in the past, or continue to do so in the present.

Sex differences

When we examined whether sex differences existed in our game theoretic model, we found that regardless of gender, the effects were the same: Again, only subjective payoffs were a significant predictor of estimated attack likelihood. We then examined Equation (2), in which this game theoretic approach is further broken down to comprise risks, likelihood of getting caught, size of payoffs, and likelihood of success. Our findings, in which only likelihood of success and size of payoffs were significant predictors, were mirrored across genders.

Discussion

The purpose of the present study was to examine the likelihood of carrying out a hacktivist attack as a product of costs and benefits. Contrary to what one might expect through a game theoretic framework of decision making, computer sciences students' estimations of attack likelihood stemmed solely from expected payoffs, rather than their interplay with subjective risks. Among this population, highly risky decisions may be made based solely on possible benefits, rather than the ratio of benefits to potential costs—a decision that would not, by game theory standards, be considered rational. Alternatively, this finding might suggest that when asked to predict a third party's action when presented with a risky (but potentially advantageous) situation, we are more likely to view others as risk-takers or, perhaps, as illogical and payoff driven.

This study also seeks to construct a descriptive profile of potential hacktivists. That we found no sex differences in the prediction of attack likelihood is noteworthy. Although men are generally higher in risk-taking behaviors,[22] it is possible that because hacktivism exists in a private space without concrete personas, such sex differences might lessen in intensity. These findings might reflect the changing demographics among hacktivists: The gender of this group of individuals should not, by default, be assumed male. Nonetheless, this finding should be replicated in future work.

We further found that the extent to which hacking is prevalent among one's peers has a positive impact on estimations of payoff and attack likelihood—that is, participants who may be more familiar with hacking through their peer groups may perceive higher payoffs and, therefore, a higher likelihood of carrying out an attack regardless of the risks involved.

We believe that this research poses a series of theoretically and societally important implications, and we hope that future research will further investigate this topic with methodologies that can, perhaps, more accurately examine true behaviors, rather than hypothetical estimations. One study design that could assess these true behaviors might take the form of a fake call to action, which could be placed on Web sites that are conducive to audiences that might expect such messages (e.g., 4chan.org, reddit.com).

An additional direction for future inquiry is to more closely examine the influence of peer prevalence on perceived payoffs and attack likelihood. It is possible that by manipulating the degree to which participants believe it common for their peers to engage in hacking behaviors, participants' perceptions of the payoffs underlying hacktivist behaviors will be directly impacted and, therefore, their perceived likelihood of attack. Future research might likewise alter perceived peer prevalence with the intention of reducing not only hacker intentions but perhaps even illegal downloading and trolling behaviors.

The present findings contribute to our existing knowledge of game theoretic approaches toward decision-making strategies, such that when facing a risky, anti-government but pro-justice situation, college-aged participants weigh payoffs—rather than the ratio of payoffs to risks—before making an attack decision. Should this effect be replicated among a wider range of demographic groups and attack targets, this finding may suggest that we are not, in fact, rational decision makers in domains where hacking and online activism are involved. Furthermore, social activism exists in a concrete and identifiable world, whereas hacktivism remains unique in its degree of mystery: Because it exists solely in a private space that is devoid of social labels or expectations, demographic differences may cease to exist, pointing instead to the importance of social networks and a desire for justice that is so all-consuming that it may eclipse the significant dangers posed.

Author Disclosure Statement

No competing financial interests exist.

References

1. Antón PS, Anderson RH, Mesic R, et al. (2013) *The vulnerability assessment & mitigation methodology: Finding and fixing vulnerabilities in information systems* (Report No. OMB 0704-0188). RAND National Defense Research Institute.
2. Jordan T, Taylor PA. (2004) *Hacktivism and cyberwars: Rebels with a cause*? New York: Routledge.
3. Colesky MR, Van Niekerk J. (2012, September) *Hacktivism: controlling the effects*. Paper presented at The 2012 Annual Conference on www Applications. Durban, South Africa.
4. Albanesius C. (2011, September) Cyber crime costs $114b per year, mobile attacks on the rise. *PC Magazine*. www.pcmag.com/article2/0,2817,2392570,00.asp (accessed Nov. 13, 2014).
5. Pilkington E. (2013, November) Jailed Anonymous hacker Jeremy Hammond: 'My days of hacking are done'. *The Guardian*. www.theguardian.com/ technology/2013/nov/15/jeremy-hammond-anonymous-hacker-sentenced (accessed Nov. 13, 2014).
6. Vincent J. (2013, December) $183,000 fine for man who joined Anonymous attack for 'one minute'. *The Independent: Technology*. www.independent.co.uk/ life-style/gadgets-and-tech/183000-fine-for-man-who-joined-anonymous-attack-for-one-minute-8995609.html (accessed Nov. 13, 2014).
7. Dixit AK, Nalebuff BJ. (2008) *The art of strategy: A Game theorist's guide to success in business and life*. New York, NY: W. W. Norton & Company, Inc.
8. Liu P. (2005) *A game theoretic approach to cyber-attack prediction* (Report No. DOE/ER/25527). U.S. Department of Energy: Office of Science.
9. Dahan M. (2013) Hacking for the homeland: Patriotic hackers versus hacktivists. *Proceedings of the 8th International Conference on Information Warfare and Security*. Denver, CO: ICIW.
10. Rogers M. (1999) *Modern-day Robin Hood or moral disengagement: Understanding the justification for criminal computer activity*. The Center for Education and Research in Information Assurance and Security at Purdue University. http://homes.cerias.purdue.edu/~mkr/moral.doc (accessed Nov. 21, 2014).
11. Stryker C. (2012) *Hacking the future: Privacy, identity, and anonymity on the Web*. New York, NY: Overlook Press.
12. Massa FG. (2011, January) Out of bounds: Anonymous' transition to collective action. In *Academy of Management Proceedings*, *1*, 1–6. San Antonio, TX: Academy of Management.
13. Walgrave S, Rucht D, Van Aelst P. (2007) Sociodemographics: Typical new social movement activists, old leftists or normalized protesters? In Walgrave S, Rucht D, eds. *Protest politics: Antiwar mobilization in advanced industrial democracies*. Minneapolis: The University of Minnesota Press, pp. 78–97.
14. Coleman G. (2014) *Hacker, hoaxer, whistleblower, spy: The Many faces of Anonymous*. Brooklyn, NY: Verso Books.
15. Yar M. (2013) *Cybercrime and society*. New York, NY: SAGE Publications Limited.
16. Cialdini RB. Crafting normative messages to protect the environment. Current Directions in Psychological Science 2003; 12:105–109.
17. Ayres J, Hopf TS. Visualization: A means of reducing speech anxiety. Communication Education 1985; 34:318–323.
18. Ayres J, Hopf TS. The long-term effect of visualization in the classroom: a brief research report. Communication Education 1990; 39:75–78.
19. Knappenberger B. (Producer and Director). (2012) *We are legion: The Rise of the hacktivists* [Motion picture]. United States: Luminant Media.
20. Shakarian P, Shakarian J, Ruef A. (2013) Cyber attacks by nonstate hacking groups: The case of anonymous and its affiliates. In Shakarian P, Shakarian J, Ruef A, eds. *Introduction to cyber-warfare: A multidisciplinary approach*. Sebastopol, CA: Syngress, pp 67–110.
21. AnonymousOfficial24 (2014, May 5). *Anonymous—The monopoly [comcast, net neutrality]* [Video file]. www.youtube.com/watch?v=9CMFabT7Tyk (accessed Nov. 21, 2014).
22. Zuckerman M, Kuhlman DM, Joireman J, et al. A comparison of three structural models for personality: the Big Three, the Big Five, and the Alternative Five. Journal of Personality and Social Psychology 1993; 65:757–768.

Address correspondence to:
Dr. Jessica E. Bodford
Department of Psychology
Arizona State University
P.O. Box 871104
Tempe, AZ 85287-1104

E-mail: jbodford@asu.edu