

The social psychology of cybersecurity

John McAlaney, Helen Thackray and Jacqui Taylor

Online interactions are often based on trust, the sharing of information and a degree of interdependence, as found with many offline relationships. The recent events at companies such as TalkTalk and Ashley Madison are the latest in a series of high profile cybersecurity incidents that demonstrate what happens when this trust relationship is breached. Despite the common media depiction, incidents of this type may not be the result of the archetypal hacker using technological means to hack into a system. Instead cybersecurity attacks are increasingly based primarily on social engineering techniques, which refers to the use of psychological manipulation to trick people into disclosing sensitive information or inappropriately granting access to a secure system (Tetri & Vuorinen, 2013). One of the more well-known examples of social engineering that many of us will have encountered would be phishing emails, which attempt to fool the recipient into opening a link or attachment that will install malicious software onto their computer. These phishing emails draw upon many of principles of social psychology, consumer psychology and behaviour change by, for example, using a fear appeal or invoking a sense of scarcity or urgency if the recipient does not act quickly. These phishing emails may only be successful a fraction of the time, but with the ability to send out tens of thousands of emails at once, at no or zero cost to the sender, this can still be a productive means of gaining access to individuals' computers. Yet despite the intrinsically psychological nature of many cybersecurity attacks, research into the role of psychology in cybersecurity is still limited. Indeed, even research into social engineering is often conducted from the discipline of computing rather than social psychology.

The motivations of online adversaries are varied, from financial gain to acts of political and ideological protest or even just personal amusement. Large scale cybersecurity incidents appear to have one factor in common however, which is that they are often instigated by groups, as opposed to individuals acting alone. As such these incidents can be regarded as the result of group actions and group processes, although applying traditional concepts of group structures to the online world is problematic. For example, Anonymous, one of the more publically known hacktivist groups, is not a close cohesive group that works in cooperation, but an umbrella term for those who identify with the aims of a section. Both Lulzsec and Lizard Squad were distinct groups that evolved from the loose collective of Anonymous, each with their own profile and characteristics. It has been alleged that in some instances individuals have been coerced or manipulated into participating in online activities by members of these groups without an accurate understanding of the consequence of their actions or the risk of criminalisation (Olson, 2012). Tools and software that can be used to aide an individual with limited technological knowledge to hack into systems are increasingly available on the internet; as such barriers to participation in cybercrime and hacktivism are falling. Such individuals may be referred to as script kiddies, which is a label that is both used by the script kiddies themselves and as a derogatory term by those who see themselves as real hackers. An example of this type of software would be the LOIC software shared amongst users of the internet image board website 4chan. Named after a weapon called a low orbit ion canon in the video game *Command and Conquer* this open source software can be used to participate in a distributed denial of service (DDoS) attack which overloads a website and takes it offline. This software has been used in a number of group hacktivism campaigns, including one known as Project Chanology, carried out by Anonymous, in which the Church of Scientology was targeted. Instructions on how to download and use the software were disseminated to individuals who wished to participate in the campaign, not all of whom had a good understanding of the software or the possible risks use of it entailed. Following the incident a number of individuals who participated in the campaign were arrested (Olson, 2012).

The majority of psychological research that has been done to date in this area has focussed on prevention and mitigation strategies for the targets of cybersecurity incidents. There is little understanding of what motivates the adversaries to select the targets in the first place, particularly in cases of hacktivism and online protest where adversaries place themselves at risk of prosecution for no obvious financial gain. This is in contrast to other areas of psychology and behaviour change. In the alcohol field for instance, research and support services have been developed to help those affected by heavy drinking, but efforts are also made to understand why people drink heavily and to educate them on the possible consequences of doing so. There is the potential for similar approaches to be used in relation to participation in cybersecurity incidents. The need for work in this area has recently been highlighted by the National Crime Agency, who provide advice on how to stop young people from becoming involved in cybercrime (<http://tinyurl.com/jqraiq6>). In this article we discuss the contributions that social psychological research and theory can make to cybersecurity, and how individuals may be empowered to make informed decisions about their participation in cybersecurity incidents.

For the lulz

One motivation that is evident in a number of cybersecurity incidents, particularly those targeted at individuals, is personal enjoyment, or for the lulz to use the language of some online groups (Coleman, 2014). This type of cybersecurity incident is under-researched, but could perhaps be informed by some of the studies that have been conducted on cyberbullying and trolling, which refer to intentional, aggressive acts carried out repeatedly over time using electronic communication (Slonje & Smith, 2008). In many cases it may in fact be difficult to differentiate a cyberbullying incident from a cybersecurity incident. As would appear to be the case with a number of cybersecurity incidents some perpetrators identify their primary motivation for engaging in cyber bullying to simply be because they find it funny (Slonje & Smith, 2008). The influence of anonymity, disinhibition and deindividuation may be of particular importance in such incidents. The perception of anonymity afforded by online communications allows individuals to take actions that would otherwise result in legal or social sanctions, although as noted earlier there may of course be a gap between an individual's understanding of how likely they are to be identified and the actual reality of this. Disinhibition on the other hand refers to the sense that actions conducted online do not feel as real as those conducted offline – which, it has been argued, can lead individuals to lose self-control (Mason, 2008). Finally deindividuation may occur, in which individuals lose their sense of self-awareness when within a group (Reicher, Spears, & Postmes, 1995; Taylor & MacDonald, 2002). Again this is an under-researched area, but it would seem that some individuals become engaged with online groups to an extent which would seem to be particularly intense.

Hacktivism

The motivation behind a cyberattack may also relate to political and ideological motivations, which in some ways could be considered to be the antithesis of cyberattacks committed for the lulz. Yet there is not necessarily a clear distinction between the different motivations which may drive an individual or group to participate in cyber adversarial acts. Whilst Anonymous have been implicated in cyberattacks against apparently random targets they have also taken part in actions such as providing internet access to protestors in Tunisia during the 2011 uprising, after the Tunisian government attempted to block all internet traffic within the country. Cybercrime, cyberterrorism and hacktivism may share many of the same outcomes, and as Rogers notes groups may actively choose to present themselves as hacktivists rather than cybercriminals or cyberterrorists due to the greater social acceptance this will give them (Rogers, 2010). Nevertheless, there would appear to be cyberattacks which have a wider political goal, particularly those attacks instigated by groups. Alberici et al (2012) argue that the motivations that drive people to collective action include identification with a group involved in a conflict; a feeling that the situation of one's own group is

unfair and a shared belief that the group can bring about change. Underpinning this there is also often a sense that core moral principles have been violated and must be defended and reinstated. These motivations would seem to fit with the tendency of hacktivism campaigns to be directed at organisations that are perceived to be suppressing the freedom of information. This conflict between politically motivated, social justice campaigns and committing cybersecurity attacks for the lulz may undermine some groups. In the case of Anonymous it has been suggested that some of the splintering and in-fighting within collective arose with from some members feeling that ideologically driven campaigns were not consistent with the type of anarchy and random actions by which the group had previously defined itself (Olson, 2012).

Impression management and social identity

A characteristic of hacktivist groups and cyberattackers is the mystery around the identity of their members and the actual reach and capabilities of the group. This is of course necessitated in some cases so that individuals can avoid being identified by law enforcement agencies, but many of these groups have become adept at brand management for the promotion of themselves and their cause. Anonymous have become particularly known for the Guy Fawkes mask, taken from a graphic novel on which the film *V for Vendetta* was based on. This symbol has become co-opted and replicated across many groups and used as a way of expressing discontent with the Establishment. This type of sophisticated marketing and self-promotion is at odds with what could be considered the stereotype as a computer hacker. Rogers argues that such stereotypes of hackers as young, socially awkward males have become a substitute for actual research (Rogers, 2010). Individuals may be skilled socially when taking part in online communications, even if they are less so in offline situations, and these skills are valued within the groups. Online communication does allow for a far greater degree of impression management than is the case offline, and it is easy to see why individuals who are less socially skilled offline may be drawn towards interacting with others online (Fullwood, 2015).

Groups can also engage in forms of collective impression management. It has been claimed for example that Anonymous engaged in impression management by overstating their capabilities to journalists (Olson, 2012) who investigated some of the early cybersecurity incidents instigated by the group. The use of impression management by such groups may be a factor as to why people are drawn towards them. The annual DEF CON hacking convention held every year in Las Vegas is attended not only by those who engage in hacking, but also by those who identify with the excitement and glamour that they perceive hacking to have, despite have little or no personal involvement in actual hacking activities. Their association with groups and individuals who do engage in hacktivism and cybercrime may be an example of what Cialdini refers to as basking in reflected glory (Cialdini et al., 1976). In order to help prevent future cybersecurity incidents the media could also, as argued by Rogers (Rogers, 2010) take a more responsible approach to the reporting of cybercriminals, to avoid glamourizing individuals and setting them up as role models. There are also claims via social media accounts, on platforms such as Twitter, that the hype surrounding multiple “ops” are a deliberate attempt to undermine the achievements and popularity of the hacktivist groups. It has been suggested that by creating many ops on social media, detractors of hacktivist groups dilute the power of numbers (which is a large part of Anonymous’ success) and then use the “failed ops” as examples of how the hacktivist groups are declining and far less influential than they once were.

Group processes

As previously discussed, many cybersecurity incidents would appear to be orchestrated by groups. As is established in social psychology being a member of a group can alter individual behaviour and thought processes in a number of fundamental ways. For example, cognitive processes such as decision making, planning, judging and problem solving may be undertaken at a group level rather

than an individual level (Hinsz, Tindale, & Vollrath, 1997). Similarly, emotions can spread throughout groups, even to members who were not involved in the original event that prompted the emotion (E. R. Smith, Seger, & Mackie, 2007). This may be particularly relevant to cybersecurity incidents that are linked to online social protest and hacktivism, where there is a sense of anger at a perceived injustice or suppression of freedom. It has also been found that individuals will make riskier decisions when in groups than when alone, even if the decision is made privately after the group discussion has taken place (Wallach, Kogan, & Bem, 1962). However, individuals are often unaware of the influence that the group is having on their behaviour (Darley, 1992).

As predicted by intergroup attribution research (Hewstone & Jaspars, 1982) the success of group actions by hacktivism and other cyber adversarial groups could be expected to strengthen individual members' beliefs that they are highly skilled, and that any successes of opposing groups such as law enforcement are more attributable to external circumstances and luck. This may have emboldened the group to take further actions against other organisations, particularly if that group identity is reinforced by media reporting. It has been commented that early news reports about Anonymous generally overstated both the level of cohesiveness between group members and organisational structure of the group (Olson, 2012), an observation which in fact has been a factor in the group becoming more cohesive and organised. The cohesiveness of groups however was also affected by the shock of the "Sabu-tage" when a high ranking member of Lulzsec was revealed to have been an informant for the FBI, leading to the arrests of other prominent group members. There are notable changes to the group behaviours since, with increased dislike of what is termed within the group as 'leader-fags', suspicion of new or unknown members, and those who are perceived to be looking for attention.

Empowering informed decision making

Hacktivism and cybercrime have the potential to create ethical conflicts for psychologists and other behaviour change experts. Hacktivism campaigns vary in how much impact they have, but often include at least some degree of illegality. What is seen as a form of social protest through hacktivism in one country may be viewed as a clear case of cybercrime in another (R. G. Smith, Cheung, & Yiu-Chang Lau, 2015). As social engineering based attacks become more common and hacktivism groups make greater use of social media it is inevitable that psychologists will be increasingly drawn into the sphere of cybersecurity. Conflicts will arise between the need to help individuals and organisations protect themselves, and the rights of a population to protest against their governments.

Attempts to dissuade people, and especially young adults, from becoming involved in hacktivism and cybercrime by instructing them they should not do so, would seem destined to fail. As experienced across a range of health and social behaviours, such direct and blatant attempts to change behaviour can easily result in a reactance response, in which an individual or group resent the perception that their choices are being removed from them. In some cases the individual may adopt an attitude directly contrary to that which they feel is being pushed upon them, in a process known as negative attitude change (Fuegen & Brehm, 2004). Given that many hacktivism groups define themselves on the basis of being anarchistic it would seem especially likely that attempts to bring about behaviour change in these groups would result in this type of negative response. It is important to also understand the role that being part of these groups has for individuals. If we discourage them from interacting with groups associated with hacktivism and cybercrime then we may also be asking them to abandon something that is an important basis of their social identity and self-esteem.

An alternative approach is to aim to empower young people to make more informed decisions about their participation in hacktivism and other online activities. By doing this no comment is being made to the young people about what they should be doing; instead it just prompts them to consider the factors that may be influencing their behaviour and what the consequences of participation may be.

This does of course have parallels with other areas of behaviour change, where there is a move away from using scare tactics and 'health terrorism'. As discussed above there are a number of group processes and biases which young people could be educated about in order to help them develop resilience against being misled into participating in actions which could result in their criminalisation. Making them aware that they may be misled could trigger a reactance response, which could be a protective factor by making them be more critical and sceptical of people they encounter online.

As the internet and digital technologies become increasingly pervasive in our lives it is important that psychologists gain a better understanding of hacktivism and cybercrime through engagement with people involved in such groups; and take caution in accepting the negative stereotypes as fact. Whilst there will always be trolls and those who are only in it for the lulz, their motivations may not be as straightforward as first thought. There also appears to be a number of young people involved in hacktivism groups who are intelligent, skilled and passionate about social justice. It is better to work with such individuals to explore ways of bringing about positive change than to leave them to become criminalised.

References

- Alberici, I. A., Milesi, P., Malfermo, P., Marzana, D., & Canfora, R. (2012). Comparing social movements and political parties' activism: The psychosocial predictors of collective action and the role of the internet. *Contention, 0*(0).
- Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in Reflected Glory - 3 (Football) Field Studies. *Journal of Personality and Social Psychology, 34*(3), 366-375. doi:Doi 10.1037//0022-3514.34.3.366
- Coleman, E. G. (2014). *Hacker, hoaxer, whistleblower, spy : the many faces of Anonymous*. London ; New York: Verso.
- Darley, J. M. (1992). Social organization for the production of evil. *Psychological Inquiry, 3*(2), 199-218. doi:10.1207/s15327965pli0302_28
- Fuegen, K., & Brehm, J. W. (2004). The intensity of affect and resistance to social influence. *Resistance and Persuasion, 39-63*. Retrieved from <Go to ISI>://WOS:000222053500003
- Fullwood, C. (2015). The role of personality in online self-presentation. In A. Attrill (Ed.), *Cyberpsychology* (pp. 9 - 28). Oxford: Oxford University Press.
- Hewstone, M., & Jaspars, J. M. F. (1982). Intergroup relations and attribution processes. In H. Tajfel (Ed.), *Social Identity And Intergroup Relations* (pp. 99 - 133). Cambridge: Cambridge University Press.
- Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin, 121*(1), 43-64. doi:Doi 10.1037/0033-2909.121.1.43
- Mason, K. L. (2008). Cyberbullying: A preliminary assessment for school personnel. *Psychology in the Schools, 45*(4), 323-348. doi:10.1002/pits.20301
- Olson, P. (2012). *We Are Anonymous*. New York: Back Bay Books.
- Reicher, S. D., Spears, R., & Postmes, T. (1995). A social identity model of deindividuation phenomena. *European Review of Social Psychology, 6*(1), 161-198. doi:10.1080/14792779443000049
- Rogers, M. K. (2010). The psyche of cybercriminals: A psycho-social perspective. In G. Ghosh & E. Turrini (Eds.), *Cybercrimes: A Multidisciplinary Analysis*.
- Slonje, R., & Smith, P. K. (2008). Cyberbullying: another main type of bullying? *Scandinavian Journal of Psychology, 49*(2), 147-154. doi:10.1111/j.1467-9450.2007.00611.x
- Smith, E. R., Seger, C. R., & Mackie, D. A. (2007). Can emotions be truly group level? evidence regarding four conceptual criteria. *Journal of Personality and Social Psychology, 93*(3), 431-446. doi:Doi 10.1037/0022-3514.93.3.431
- Smith, R. G., Cheung, R., & Yiu-Chang Lau, L. (2015). *Cybercrime Risks and Responses: Eastern and Western Perspectives*.
- Taylor, J., & MacDonald, J. (2002). The effects of asynchronous computer-mediated group interaction on group processes. *Social Science Computer Review, 20*(3), 260-274. doi:10.1177/08939302020003004
- Tetri, P., & Vuorinen, J. (2013). Dissecting social engineering. *Behaviour & Information Technology, 32*(10), 1014-1023. doi:10.1080/0144929x.2013.763860
- Wallach, M. A., Kogan, N., & Bem, D. J. (1962). Group influence on individual risk-taking. *Journal of Abnormal Psychology, 65*(2), 75-&. doi:Doi 10.1037/H0044376