The Social Psychology of Computer Viruses and Worms[*]

Jonathan J. Rusch
Georgetown University Law Center
United States of America
Jonathan.Rusch@worldnet.att.net

Paper Presented at INET 2002, Crystal Gateway Marriott, Crystal City, Virginia, June 21, 2002

Introduction

When the defenders of Troy first saw the Trojan Horse outside their walls, legend has it, the gods on Mount Olympus did not compel them to bring it inside the city. The Trojans' decision to do so, though wholly voluntary, was strongly influenced by the Greek army's clever manipulations of their perceptions. The Greeks not only hauled the horse by night to the gates of Troy, but spread a rumor that the horse had a benign purpose: appeasement of the war goddess Minerva to ensure a safe return home. They also sailed all of their warships away from Troy to a hidden anchorage. They chose, however, to leave behind one Greek, named Sinon. Situating himself where he would be found easily by Trojan forces, Sinon pretended to have escaped wrongful imprisonment after being designated for sacrifice by his own people.

While a few Trojans presciently warned against the horse, most accepted Sinon's solemn confirmation that its purpose was benign. Many even participated in breaching the city walls to ensure that the outsized horse could be hauled inside.[1] The bitter remarks of Aeneas, who survived the fall of Troy, suggest that the Trojans were, in effect, the earliest known victims of "social engineering":

> This fraud of Sinon, his accomplished lying,
> Won us over; a tall tale and fake tears
> Had captured us, whom neither Diomedes
> Nor Larisaean Achilles overpowered,
> Nor ten long years, nor all their thousand ships.[2]

The "Trojan horses" we encounter today, along with the computer viruses and worms that bear them,[3] seem far more complex and sophisticated than the Trojan horse of legend. Yet virus and worm makers often show that they are as capable as the ancient Greeks of influencing people to open their computers and networks to malicious code or even mistakenly to destroy their own data.

Since February 2001, a spate of reports indicate that “social engineering” techniques – some more sophisticated than others – are becoming increasingly popular in virus and worm writing. Here are a dozen of the more widely reported viruses, worms, Trojan horses, and virus-related hoaxes during that period, listed in approximate chronological order:

- *February 2001* The now-famous “Anna Kournikova” virus was one of the pioneering uses of social engineering in virus propagation. It accompanied e-mails with the subject line, “Here you have,;0)”. The attachment bore the filename AnnaKournikova.jpg.vbs, to make it appear that the attachment was a jpg picture of the tennis player Anna Kournikova. In fact, the attachment was a Visual Basic Script that infected Outlook and mailed itself to contacts in the target computer’s address book.[4]
- *May 2001* In the “sulfnbk.exe” hoax, an e-mail told recipients that the sender, without knowing it, had a file on his computer that proved to be a virus undetectable by anti-virus software. The sender then provided recipients with instructions on how to find and delete the file on their own computers. In fact, sulfnbk.exe was a standard executable Microsoft Windows file that serves as a utility to restore long file names. Many recipients of the e-mail who looked for the file and found it on their computers apparently concluded that the sender’s information was accurate and deleted the file in the mistaken belief that it was malicious code.[5]
- *July 2001* The “W32.LeaveB” worm appeared as an attachment to an e-mail that purported to be a Microsoft security bulletin. The spurious bulletin informed recipients that a “serious virus” was aimed at Windows computers and that they should protect their computers by downloading and installing an attached “security patch.” The bulletin also reported that the virus had the “complexity to destroy data like none seen before.” The worm in fact downloaded components from Websites and contained code to accept commands from Internet Relay Chat programs. The worm’s purpose was unclear, as it apparently was not damaging computers or facilitate the theft of data from target computers. One leading computer security expert speculated that it was intended to use target computers “to click ad banners and other sites as part of a money-making scheme.”[6]
- *December 2001* The “Reezak” worm appeared during the 2001 holiday season. The subject line of the e-mail transporting it was “Happy New Year” and the message text “Hi i can’t describe my feelings But all i can say is Happy New Year :) bye.” A recipient who clicked on the attachment saw a Flashmedia Christmas greeting card, with Santa and a reindeer against a background of snow. What the recipient did not see was Reezak’s efforts to delete the target

computer's Windows System directory and to disable antivirus software. Reezak also attempted to redirect the Internet Explorer home page to an infected Website with a message that declared both "Sharoon" [*sic*] and Bush a "war criminel" [*sic*]. The infected Webpage included an infected script called Outlook.vbs. This script sent a second message to contacts in the Outlook address book that urged people to visit the infected Webpage themselves.[7]

- *January 2002* The "Gigger" worm (actually a self-propagating virus) circulated as an attachment to e-mails. The e-mails bore the subject line "Outlook Express Update," and the attachment bore the name "mmsn_offline.htm," to encourage the impression that the e-mails were from Microsoft. If opened, the attachment could infect the target computer and delete all files on its hard drive.[8]
- *January 2002* The "MyParty" worm appeared as a purported link in an e-mail message. The e-mail, which bore the subject line "new photos from my party!," told recipients, "Hello! My party ... It was absolutely amazing! I have attached my web page with new photos! If you can please make color prints of my photos. Thanks!" The message also contained the link to what appeared to be a Yahoo! website, www.myparty.yahoo.com or myparty.photos.yahoo.com. In fact, clicking on the purported link caused a virus to copy itself to C:\Recycled\regctrl.exe and execute that file. The virus then retrieves the default SMTP server of the user's infected computer from the registry and launches itself to solicited addresses in the target computer's Outlook directory.[9] Some reports also indicated that the virus also left Trojan horses on infected machines before mailing itself to others.[10]
- *March 2002* According to the Computer Emergency Response Team (CERT) at Carnegie Mellon University, intruders are using Internet Relay Chat (IRC) and Instant Messaging (IM) to send messages that appear to come from the IRC or IM network. These messages are crafted to make the recipients believe that they already have a virus and must download an attached program to clean their computers or risk being banned from the network. For example:

  > You are infected with a virus that lets hackers get into your machine and read ur files, etc. I suggest you to [*sic*] download *[malicious url]* and clean ur infected machine. Otherwise you will be banned from *[IRC network]*.[11]

  CERT reported that tens of thousands of systems reportedly have been compromised in this way.[12]
- *March 2002* The "W32/Gibe@mm" virus circulated as an attachment to an e-mail purportedly from the "Microsoft Corporation Security Center." The e-mail warned recipients of vulnerabilities in Internet Explorer, Outlook, and Outlook Express programs, and noted that the attached security-update program would fix those vulnerabilities. In fact, the "Gibe" virus would attach itself to the target

computer's registry, e-mail copies of itself to addresses it could find, and open a port on the target computer for malicious code to enter.[13]

- *April 2002* In the latest variants of the MyLife virus, e-mails invited the viewer to click on an attachment that purports to be a screensaver. The screensaver allegedly poked fun as such prominent political figures as former President Clinton and Israeli Premier Ariel Sharon. Clicking on the attachment released a worm that spread from the victim's computer to addresses in the victim's Outlook address book or MSN Messenger contact list.[14]
- *April 2002* The "Jenna Jameson" virus, named for a well-known porn star, was attached to an e-mail with the subject line "Jenna Jameson pornostar free superfuck+photo addresses." The attachment bore the filename "JENNA-JAMESON-FREE-SUPERFUCK.TXT.vbs." To the hasty reader, this filename made it appear that the file was a text file; closer inspection showed that it was a malicious Visual Basic Script. Executing the attachment allowed the recipient to see a text document with a list of links to porn sites, but also launched the virus, which infected the target computer. The virus first sent itself to all names in the target's Windows address book and was set to display a message on May 12, 2002: "Your PC has been hacked by KaGra[ATZI virus ver 2.1]." On May 13, it was also set to delete the Windows folder on the target's C drive or, if the target is running Windows NT, the Winnt folder.[15]
- *April 2002* The most virulent e-mail virus, according to various sources, is the Klez.H virus.[16] The latest variants of Klez, particularly Klez.H, use a large number of subject lines and texts that present a wide range of messages. Klez.H is attached to e-mail whose subject line may contains any one of approximately 120 phrases, such as "Re: A Win XP patch," "Undeliverable mail–(random)," "Returned mail–(random)," "(random) (random) game," and "darling." Klez.H also spoofs an e-mail address found on the target computer, to make it appear that it has been sent from a familiar person or entity.[17] The e-mail text, which has numerous variations, consists of messages that may include "This is a special humour game"or "(virus name) is a dangerous virus that spread through email. (Antivirus vendor) give you the (virus name) removal tools. For more information, please visit http://www.(antivirus vendor).com." When the user clicks on the attachment, Klez.H uses its own Simple Mail Transfer Protocol (SMTP) server to send infected copies of itself to the target computer, bypassing e-mail software on that computer. It can copy itself to remote disk drives by creating random filenames and adding random suffixes such as .exe, .com, .bat, or .scr. It also contains an upgraded version of the Elkern virus, Elkern.c, which adds a hidden file to the Registry entry and can corrupt files without changing their size.[18]

- *May 2002* The "cute.exe" Trojan horse program is an attachment to an e-mail with the subject line "Thoughts" and the text "I just found this program, and, i don't know why . . . but it reminded me of you. check it out." The attachment uses a standard JPG icon, but is an executable file. If the recipient clicks on cute.exe, it will unpack itself and make various changes to system files "to ensure that the program ('kernel32.exe') will execute after a reboot."[19] In a variation of a standard "Floodnet" bot, cute.exe also contacts an Internet Relay Chat (IRC) server and joins a predefined IRC channel. The allows the attacker to obtain information about the target computer, to launch various denial of service attacks, and to instruct the program to update or remove itself.[20]

The increasing use of social influence techniques should be of great concern to computer security specialists. Greater use of these techniques can greatly complicate the tasks of preventing or reducing the spread of viruses, worms, and "blended threats" (i.e., code combining elements of worms and viruses) just when they are becoming increasingly ubiquitous on the Internet. According to a survey by ICSA Labs, in 2001 corporations were hit with a monthly average of 113 virus infections for every 1,000 computers they owned. The majority of the viruses identified in the survey were spread through e-mail, and mass mailers accounted for 80 percent of the viruses.[21] Another survey by Information Security Magazine found that 90 percent of the companies surveyed had been infected with worms or viruses.[22]

Antivirus specialists have tended to explain the success of social engineering viruses, in part, by casting aspersions on the intelligence of the victims, calling them "ignorant"[23] or suggesting that they needed to apply "common sense."[24] One researcher expressed surprise that variants of the MyLife virus were spreading "because the tricks used by the virus to fool people into double clicking on the attachment, and becoming infected, were crude."[25]

These comments reflect a significant gap in our understanding of viruses and worms. We know a fair amount about the technical operations of viruses and worms, thanks to the efforts of many computer security researchers around the world. We are taking tentative steps toward understanding the thinking of malicious code writers, through the work of antivirus researchers like Sarah Gordon.[26] In contrast, there has been no systematic analysis of the effectiveness of social engineering techniques in spreading viruses and worms.

One source of insight for such an analysis is the field of social psychology -- "the scientific study of how people think about, influence, and relate to one another."[27] Social psychology has developed a number of behavioral principles and concepts that help to explain many types of human interactions.[28] These principles and concepts can be applied to the topic of social engineering to show how social engineering can exploit what social psychologists term "the power of the situation" to influence online behavior.

This paper will identify some of the most pertinent principles and concepts in social psychology, and use them to develop likely explanations for the success of social engineering viruses and worms. As this paper will show, the psychological forces which malicious code writers are seeking to exploit have far greater strength and effectiveness in influencing behavior than antivirus researchers have realized. It will also suggest some possible measures for preventing or reducing the future success of "socially engineered" malicious code.

Principles and Concepts in Social Psychology

Routes to Persuasion and "Mindlessness"

A fundamental concept in social psychology is the idea that there are two routes to persuading someone to take a certain action: central and peripheral. Persuasion efforts that involve the central route use systematic arguments and sound reasoning, such as advertisements that present comparative prices and features of a product, to stimulate favorable responses. In contrast, persuasion efforts that involve the peripheral route provide cues, such as beautiful or pleasurable images, "that trigger acceptance without much thinking."[29]

The peripheral route to persuasion can be effective in many contexts because of our dependence on mental shortcuts in everyday life. It stands to reason that if we were to apply systematic, deliberate thinking to the hundreds and thousands of decisions we must make every day – from the path we walk to the bus or train station each morning to the television programs we watch each evening, our lives would quickly grind to a halt. Human beings need simple heuristics – mental shortcuts – to make many of those decisions quickly and get on with their lives. For that reason, our reliance on mental shortcuts is usually not a bad thing.[30]

In certain circumstances, however, reliance on mental shortcuts can lead us awry. Psychologists have noted that we should interpret and react to information with reference to the context in which they are receiving that information. When we begin to treat information as though it were context-free – i.e., true regardless of the circumstances – we enter a mode of behavior that Professor Ellen Langer of Harvard University has termed "mindlessness."[31] Mindlessness does not mean "stupidity" or "lack of common sense." Depending on the circumstances, well-informed, highly-educated, and intelligent people can fall into "mindlessness."

Two examples of "mindlessness" that can be triggered in certain social contexts come from Professor Langer's own research. One experiment involved mindlessness in response to oral cues. Researchers approached people using a copy machine at a university and made one of three requests: (1) "Excuse me, may I use the Xerox machine?"; (2) "Excuse me, may I use the Xerox machine because I want to make copies?"; and (3) "Excuse me, may I use the Xerox machine because I'm in a rush?". The first form of request, obviously, offers no reason that the person already using the machine should allow the requester to cut in; the third form of request offers a logical reason that the person using the machine can easily understand. The second form of request, in contrast, offers no real reason for the request. Although it uses the same sentence structure as the third form of request ("May I use the Xerox because I'm in a rush?"), its actual content is no different from the first form of request.[32]

What was the response of the Xerox users to each of the three forms of request? For those who asked simply, "May I use the Xerox machine?," the positive response was only 60 percent. For those who asked, "May I use the Xerox because I'm in a rush?", the positive response dramatically increased to 94 percent. But for those who asked, "Excuse me, may I use the Xerox machine because I want to make copies?", the positive response for this request-without-a-reason was virtually the same as the response for the request with a reason: 93 percent.[33] People, in other words, may comply without thinking – act "mindlessly," in other words – when the request for their compliance is structured as though it has a logical reason, even though the content of the communication does not.

In the second experiment, Professor Langer and her colleagues

> sent an interdepartmental memo around some university offices. The message either requested or demanded the return of the memo to a designated room – and that was all it said. ("Please return this immediately to Room 247," or "This memo is to be returned to Room 247.") Anyone who read such a memo mindfully would ask, "If whoever wanted such a memo wanted it, why did he or she send it?" and therefore would not return the memo. Half of the memos were designed to look exactly like those usually sent between departments. The other half were made to look in some way different. When the memo looked like those they were used to, 90 percent of the recipients actually returned it. When the memo looked different, 60 percent returned it.[34]

In this experiment, the mental shortcut on which the memo's recipients evidently relied was the overall appearance of the memo. So long as it looked like the kind of memo they were used to seeing, they carefully followed its instructions without considering the logic behind those instructions.

Exacerbating Factors

A variety of psychological factors may enhance the state or duration of mindlessness and exacerbate its effects. Six of these seem particularly relevant to persuasive messages sent via the Internet.

(1) *"Flow"*

Professor Mihaly Csikszentmihalyi adopted the term "flow" to describe a state of mind in which someone invests great energy and uses his skills to pursue a goal by immersing himself in some activity that he finds absorbing.[35] A person can experience "flow" during activities such as sports or games, [36] or even in navigating Websites.[37] At such times, the person typically loses consciousness about events in the external world, and the passage of time.[38] While this sensation of "flow" can be pleasurable for the participant, it may also have the effect of hampering the participant's ability to engage in reasoned decisionmaking or recognizing that changed circumstances may require him to consult other sources of information before acting.

(2) *Context Confusion*

Professor Langer has identified a phenomenon that she calls "context confusion." Context confusion can arise when people "confuse the context controlling the behavior of another person with the context determining their own behavior. Most people assume that other peoples' motives and intentions are the same as theirs, although the same behavior may have very different meanings."[39] In other words, people who are the recipients of a persuasive communication may have difficulty in believing that the sender of that communication has entirely hostile intentions when the form of the communications seems to offer the recipients something benign, pleasing, or even helpful.

(3) *Arousal and Repetition*

One phenomenon well-established in experimental psychology is that arousal – such as excitement or even the presence of others –

> facilitates whatever response tendency is dominant. Increased arousal enhances performance on easy tasks for which the most likely – "dominant" – response is the correct one. People solve easy anagrams, such as *akec*, fastest when they are anxious. On complex tasks, for which the correct answer is not dominant, increased arousal promotes *incorrect* responding. On harder anagrams people do worse when anxious.[40]

Psychologists also know that the more we repeat actions or statements in a particular context, the more likely we are to believe that we should continue to take the same type of action or make the same kind of statement in that context.[41] Taken together, these two concepts suggest that in a situation where we engage in a small number of highly repetitive actions, those actions tend to become our response tendency, and that this response tendency will be facilitated by whatever mechanism triggers arousal.

(4) *Distraction*

Arousal also affects a distinctive aspect of human behavior in reacting to persuasive communications: that verbal persuasion increases "by distracting people with something that attracts their attention just enough to inhibit counterarguing" – i.e., thinking meaningfully about the proposition before us and developing responses to the arguments presented.[42] While political advertisements contain their share of distractions,[43] so do communications that contain sexual content or warnings that may make us worried or apprehensive.

(5) *Claims of Authority*

Professor Robert Cialdini has identified a number of persuasive influences that he has termed "weapons of influence" because they can be so powerful in inducing compliance without the use of force.[44] One of these "weapons of influence" is the concept of authority. Psychologists in numerous experiments have documented that human beings can be extraordinarily deferential in the face of claimed authority.

In one study by psychologist Leonard Bickman, a confederate of the experimenter – sometimes dressed in street clothes, sometimes in a security guard's uniform – would stop passersby on the street

> and point to a man standing by a parking meter 50 feet away. The requester, whether dressed normally or as a security guard, always said the same thing to the pedestrian: "You see that guy over there by the meter? He's overparked but doesn't have any change. Give him a dime!" The requester then turned a corner

> and walked away so that by the time the pedestrian reached the meter, the requester was out of sight.[45]

Even though the requester was no longer in sight, 92 percent of the pedestrians complied with his request when he wore the guard's uniform, while only 42 percent complied when he wore street clothes.[46]

People can display extreme deference to a claim of authority even when the claim is solely in writing. For example, in one case documented by two pharmacology professors, a physician had prescribed ear drops for a patient who had pain and infection in his right ear. In writing out the prescription, however, the doctor abbreviated the word "right" so that the instructions for administering the drops read "place in R ear." The duty nurse who received the prescription proceeded to "put the required number of ear drops into the patient's anus."[47]

(6) *Confirmation Bias*

Finally, another universal trait of human behavior is confirmation bias: that is, the tendency, when we test our beliefs, to "seek information that will verify them before we seek disconfirming behavior."[48] For example, a person may form an erroneous belief about a situation and then search for evidence that will confirm his belief rather than attempting to disconfirm it.[49]

Analyzing the Social Psychology of Viruses and Worms

When we examine the 12 malicious codes described earlier through the conceptual lenses of social psychology, we can see several distinctive features that merit discussion.

Routes to Persuasion

Each of these exploits seeks to exploit the peripheral route to persuasion, by relying on emotionally salient cues, principally fear of harm or sexual desire. These cues prompt us to click on the malicious attachment (or, in the sulfnbk hoax, deleting the identified file) with little or no conscious thinking. At the same time, half of these viruses and worms – sulfnbk, LeaveB, Gigger, the IRC/IM messaging, Gibe, and Klez – presented messages that were crafted to appeal to the central route to persuasion. In each case, the message purports to inform the recipients of an imminent threat to their computers or data and suggests an ostensibly logical response, the opening of an antivirus file that can protect the recipients.

“Mindlessness” and “Flow”

If there is any social environment that seems likely to foster “mindlessness,” it is the Internet, particularly online messaging such as e-mail and IRC or IM. Conducting online activities such as surfing or reviewing e-mails may encourage a sense of “flow” because those activities can be highly involving and appeal to our fascination with the range and variety of information and other sensory attractions we find there. At the same time, we conduct these activities by engaging in the most minimal of physical movements: scanning text with our eyes, moving a mouse a few centimeters this way or that, and causing perhaps one or two fingers to left- or right-click the mouse. Few of our daily activities seem more likely to induce a state of virtual physical immobility than website and e-mail viewing.

Moreover, the minimal physical movement we need to perform these online activities is extremely repetitive. In contrast, opening a typical array of physical mail we receive at home or work usually requires us to vary our physical movements substantially, in accordance with the size, weight, and shape of each piece of mail, as well as the degree of difficulty in opening each piece and our various sensory perceptions of the relative robustness or fragility of its contents.

Under these circumstances, the online environment can be seen as a nearly perfect setting for virus and worm writers to try to influence us. Since the acts of pointing and clicking are highly repetitive actions, conducted over the course of many minutes or even hours, it stands to reason that those acts tend to become our response tendency for our online activity in general. Furthermore, when the writers’ messages are designed to trigger strong forms of arousal, such as fear or sexual desire, they are highly likely to facilitate that response tendency.

Context Confusion

Many people who participate in online activities such as surfing and e-mail viewing are highly likely to fall into the trap of context confusion. Data from the UCLA Center for Communication Policy indicates that the online population is becoming more trusting of what they see online. In 2001, 58 percent of the online population (compared to 54.8 percent in 2000) believed that most or all of online information is accurate, and only 5.7 percent (compared with 7.5 percent in 2000) believed that little or none of online information was accurate.[50]

If, as these data suggest, a significant percentage of the online population is highly trusting of online content, it stands to reason that many of those same people would assume that others whom they encounter online are as trusting and trustworthy as they. In these circumstances, malicious contacts such as the sulfnbk.exe hoax, the Reezak worm, the MyParty worm, the MyLife virus, and the cute.exe Trojan may succeed, in part, because they adopt a friendly manner and purport to offer something pleasing to the recipient, but in language that suggests the recipient knows or has reason to know the sender. More trusting recipients may infer, wrongly, that if they think they know the sender the sender's message is unlikely to include malicious code.

Arousal, Repetition, and Distraction

Several of the malicious code examples in this paper, such as the Anna Kournikova and Jenna Jameson viruses, appear to be successful in persuading recipients to click on them because the promised content -- photographs of attractive women, in varying amounts of clothing -- is so highly likely to produce a state of arousal involving sexual attraction. Other malicious code examples, such as the LeaveB worm, the IRC and IM messages, and the W32/Gibe@mm virus, can produce a different form of arousal, by trying to persuade the recipients that their computers are at immediate risk of harm.

Either type of social engineering exploit can be highly effective at facilitating a recipient's dominant response: that is, clicking without thinking. In addition, some of these same viruses, such as Jenna Jameson and LeaveB, may prompt people to click because the distracting nature of their message (i.e., the promise of sexually appealing jpg files or anxiety-producing messages) inhibits any prospect of counterarguing or consciously considering the invited action.

Claims of Authority

A number of the 12 examples, such as LeaveB, W32/Gibe@mm, and Klez.H, undoubtedly derive some of their persuasive power from the assertion that reliable sources of authority and expert knowledge (i.e., Microsoft or prominent antivirus vendors) are responsible for the messages. Precisely because these messages often purport to warn recipients about the existence of certain malicious code from which those recipients need protection, many recipients are likely to infer that the source is indeed a trustworthy source, or at least that a person would not send such a message unless that person were genuinely informed about the issue and had no interest in trying to get the recipients to cause harm to themselves or their systems.

Confirmation Bias

Confirmation bias is a notable influence in a number of the social engineering codes we have discussed. The sulfnbk hoax, for example, was largely dependent on confirmation bias for its success. Many people who received the message that the sulfnbk.exe file was malicious code looked for that filename and, having found it, assumed that what they were seeing confirmed the message and was therefore sufficient to warrant deleting the file. But other malicious codes, such as LeaveB and MyLife, also benefit from confirmation bias. LeaveB, which invited recipients to download a "security patch," can succeed without our actually seeing the code that constitutes the alleged "patch." Because we cannot actually see what code is entering our computers when we download real security patches, we take it as a matter of course that things are working properly so long as our computers appear to be downloading something in response to our clicking on the offered link. In other words, the initiation of the downloading process constitutes confirmation that what we were told we would receive is in fact what we are receiving.

~ ~ ~

It is important to note that each of the social engineering exploits discussed above reflects the application of multiple persuasive influences. Even the simpler varieties of social engineering viruses like Anna Kournikova involves, at a minimum, use of the peripheral route to persuasion, mindlessness, flow, context confusion, arousal and facilitation of dominant responses, and distraction. Other more sophisticated invitations to click on malicious code links, such as LeaveB and W32/Gibe@mm, make use of all of these influences as well as spurious claims of authority. One of the reasons for Klez.H's extraordinary success may be that it incorporates multiple types of messages, and multiple e-mail addresses of known e-mail users, in ways that draw on all of the persuasive influences discussed here.

When so many persuasive influences are simultaneously brought to bear on members of the public – many of whom do not directly receive official bulletins, or routinely seek out media reports, about the latest variations on "social engineering" exploits – we should not be surprised that so many people become victims of these exploits.

Implications for Prevention and Public Education

The preceding analysis does not purport to be a comprehensive or experimentally based approach to explaining the success of social engineering viruses and worms. It nonetheless provides some insights into the problem of "socially engineered" code that can help in identifying some possible responses to the problem.

Among other things, we need to create a typology which distinguishes malicious codes that exploit only the peripheral route from codes that exploit both central and peripheral routes. In other words, we need to move beyond treating all "social engineered" viruses and worms as though they exploit the same kinds of human vulnerabilities in the same way.

We also need to keep in mind the distinction between messages that exploit both central and peripheral routes and messages that exploit only the peripheral route. With respect to the central-peripheral messages, the preceding analysis should give us – though it may sound odd to say so – some reason for optimism. When so many people have responded to messages that conveyed plausible warnings about malicious code by clicking on attachments that purported to protect them from such code, it seems clear that some parts of the general public have gotten part of the message that they need to get: i.e., that there are such things as viruses and worms, that those things are bad, and that they need to do something to protect themselves and their computers. At the same time, much of the public seems not to understand or accept other messages we want them to take to heart – for example, that they should not trust or download any attachments to e-mails that purport to be "security patches." This suggests that we need to rethink what kinds of messages government and the private sector are sending the public about social engineering malicious code, how widely those messages are being disseminated, and how persuasive those messages are.

At present, many of the public warnings about social engineering exploits come from highly credible institutional sources, such as CERT and the National Infrastructure Protection Center. The content of these warnings, however, tends to fall into one of two categories: (1) highly general warnings about a broad category of exploits; or (2) fairly specific warnings that deliver detailed information to a technologically sophisticated audience, such as systems administrators and computer security experts. Neither of these types of messages would be particularly helpful to much of the general public, who regularly use computers at work or home without having any real understanding of their technology. Warnings about social engineering exploits, therefore, may need to be crafted to provide more focused messages for a less sophisticated audience, in part by giving them information that they can understand and take to heart without having to absorb technologically complex details.

Social psychologists have found that people are more willing to change attitudes if they think a message contains new information than if they think a message merely repeats previously encountered information. They also have found that repeating highly similar messages has a positive effect on immediate attitude change, but that mere repetition of the same message does not produce more immediate attitude change than a single presentation of the message.[51]

These observations suggest that, in informing the public about social engineering viruses and worms, we need to strike a balance between repeating old information (e.g., "Viruses and worms are bad") and providing new information that is neither too generic nor too detailed. As one example, here are some concepts for a public-service campaign incorporating print and electronic advertisements:

- "People aren't the only ones who can get viruses." This type of advertisement, which could include an image such as a cartoon or an animation of a computer with a thermometer in its mouth, would go on to explain that computer viruses can do real damage to files and programs, and that people can help to prevent the spread of viruses just by ignoring e-mails that purport to have antivirus programs or patches attached.
- "Before you click on a link, think." This type of advertisement would adopt a theme that could encourage people to refrain from mindlessly clicking on links from sources that have not been verified as trustworthy. It also could include examples of the kinds of malicious code that people unwittingly download when they click on links without confirming the source or the validity of the message.
- "Some people will do anything to give you a computer virus. Don't let them." This type of advertisement, for television, could begin with the sound of a doorbell and the view of a house front door being opened. Outside the door stands a teenage boy wearing a Bill Gates mask. The teenager says, in an unmistakably teenaged tone and manner, "Hi, I'm Bill Gates. I think you may have a new virus on your computer. Just let me in so I could, like, put some antivirus stuff on your computer?" The camera then turns toward the door, and shows the real Bill Gates, who says, "No thanks" and closes the door. The next segment could begin the same way, with the doorbell sound and the opening of the door. This time, outside the door stands a heavyset, sweaty-looking older man who says to the unseen person, in a smarmy tone of voice,"Hi, I got some great photos of Anna Kournikova. Could I come in so I could put 'em on your computer?" The camera then turns toward the door, and shows the real Anna Kournikova, who says, "Uh, no, sorry," and closes the door. The ad would then use both text and voiceover to say, "If you *wouldn't* let them in your house, *don't*

> let them in your computer. Don't click on any e-mail attachments unless you're sure that someone you know really did send it to you. If you're not sure, call them up and ask. Your friend won't mind hearing from you, and it just might save you from a real computer virus." The advertisement would conclude with the caption: "Some people will do anything to give you a computer virus. Don't let them."

A campaign of this type would have at least four advantages. First, it would feature the same basic messages about the risks of computer viruses and how to recognize them, but would present them in different ways, with different balances of humor and serious advice. Second, it would offer people some positive strategies for how to respond to suspicious e-mails, rather than simply telling them what not to do. Third, it would encourage people to adopt more "mindful" behavior online and to think about the possible consequences of initiating downloads from potentially hostile sources. Finally, by including some widely recognized public figures, it would likely enhance the credibility and appeal of the messages.

These suggestions are hardly the only way, or necessarily the most effective way, to present public-service advertisements directed at social engineering viruses and worms. The essential point is that government and the private sector are failing to address this issue with any of the persuasive approaches and techniques with which Madison Avenue has long been familiar. If we want to reduce the success rates of socially engineered malicious code, we need not only to encourage greater use of antivirus programs, but also to make better use of social psychology in developing persuasive antivirus messages. To do otherwise will only ensure that myriads of computer users will continue to repeat the Trojans' mistake and fall victim to the Sinons of cyberspace.

~ ~ ~

Copies of this Paper will be available through the Internet Society Website, www.isoc.org.

Endnotes

*. Copyright Jonathan J. Rusch 2002. All rights reserved. Mr. Rusch is an Adjunct Professor of Law at Georgetown University Law Center, and Special Counsel for Fraud Prevention, Fraud Section, Criminal Division, U.S. Department of Justice. The views expressed in this presentation are solely those of the author, and do not necessarily reflect the views of Georgetown University Law Center or the United States Department of Justice or any component or officer thereof.

1. *See* VIRGIL, THE AENEID, Book II, lines 1-260, 311-335, at 33-40 (Robert Fitzgerald trans.) (1992 ed.).

2. *Id.* Book II, lines 268-272 at 40. Diomedes was one of the greatest of the Greeks who fought at Troy. *See id.* 445 n.137. Larisaean means a person from Larisa, a town in Thessaly that is sometimes listed as the home of the great Greek warrior Achilles. *See id.* 432, 449 n.271.

3. *Cf. id.* Book II, lines 69-70 at 35 (statement of the priest Laocöon, “Timeo Danaos et dona ferentes,” usually translated as “I fear Greeks bearing gifts”).

4. *See* James Middleton, *Anna virus the work of ‘script kiddies’*, vnunet.com, Feb. 13, 2001, wysiwig://127/http://www.vnunet.com/Print/1117639.

5. *See* George A. Chidi Jr., *E-mail virus hoax makes users do the dirty work*, InfoWorld.com, May 30, 2001, http://www.idg.net/go.cgi?id=509782.

6. *See* Todd R. Weiss, *Internet worm disguised as security alert*, CNN.com, July 17, 2001, wysiwig://34/http://cnn.career.printthi...43571004330329&partnerID+2007&expire=- .

7. *See* Robert Vamosi, *Help & How To: Reezak*, ZDNet [UK}, Dec. 20, 2001, http://news.zdnet.co.uk/story/0,,t269-s2101307,00.html.

8. *See* Reuters, *Worm posing as Microsoft update moving slowly*, SiliconValley.com, Jan. 14, 2002, http://www.siliconvalley.com/docs/news/svfront/061193.htm; John Leyden, *Gigger worm can format Windows PCs*, The Register, Jan. 11, 2002, http://www.theregister.co.uk/content/56/23652.html.

9. *See* Brian Fonseca & Sam Costello, IDG News Service, *MyParty virus features dangerous snapshot*, ITworld.com, Jan. 28, 2002, http://www.itworld.com/Net/3271/IDG020128mypartyvirus/pfindex.htm; *E-mail virus crashes the party*, BBC News, Jan. 28, 2002,

http://news.bbc.co.uk/hi/english/sci/tech/newsid_1787000/1787265.stm.

10. *See* Ellen Messmer, *Hackers, vendors put camouflage to use*, NetworkWorldFusion, Feb. 4, 2002, http://www.nwfusion.com/cgi-bin/mailto/x.cg.

11. CERT Incident Note IN-2002-03, Social Engineering Attacks via IRC and Instant Messaging (March 19, 2002), http://www.cert.org/incident_notes/IN-2002-03.html.

12. *Id.*

13. *See* Rob Pegoraro, *Be Ready to Repel Viruses, Old and New*, Wash. Post, Apr. 21, 2002, at H7.

14. *See* John Leyden, *Clinton worm variant makes fun of Sharon*, The Register, Apr. 12, 2002, http://www.theregister.co.uk/content/56/24831.html.

15. *See* James Middleton, *Porn star debuts vicious virus*, vnunet.com, Apr. 23, 2002, wysisig://82/http://www.vnunet.com/Print/1131174.

16. *See* Andy McCue, *Klez worms its way into history*, vnunet.com, June 6, 2002, wysiwig://131/http://www.vnunet.com/Print/1132339; Michelle Delio, *Klez: Hi Mom, We're No. 1*, Wired News, May 24, 2002, wysiwig://64/http://www.wired.com/news/print/0,1294,52765,00.htm.

17. *See* National Infrastructure Protection Center, Propagation of the W32/Klez.h@mm Worm and Variants, Alert 02-002 (Apr. 26, 2002).

18. *See* National Infrastructure Protection Center, Propagation of the W32/Klez.h@mm Worm and Variants, Alert 02-002 (Apr. 26, 2002); Robert Vamosi, *Klez worm spreading rapidly*, ZDNet.com, Apr. 25, 2002, wysiwig://15/http://zdnet.com.com/2102-1105-891854.htm.

19. SANS Institute, *The Handler's Diary*, May 4, 2002, http://www.incidents.org/diary/diary.php?id=151.

20. *Id.*

21. *See* Sam Costello, IDG News Service, *Survey: Virus problem grows in 2001, future*, ITworld.com, March 7, 2002, http://ww.itworld.com/Net/3271/020307virusgrowth/pfindex.htm.

[22] *See* Sam Costello, IDG News Service, *Web Attacks Have Doubled, Survey Says*, PCWorld.com, Oct. 10, 2001, wysiwig://80/http://www.pcworld.com/resource/printable/article/0,aid,65526,00.a

[23] *See* James Middleton, *supra* note 15 (comments of Mihaela Stoian, virus researcher at BitDefender).

[24] *See* Damian Carrington, *Are computer viruses unstoppable?*, BBC News, May 5, 2000, http://news.bbc.co.uk/hi/english/sci/tech/newsid_737000/737396.stm.

[25] John Leyden, *supra* note 14 (reported comments of Andre Post, senior researcher at Symantec Antivirus Research Center).

[26] *See, e.g.,* Sarah Gordon, *Distributing Viruses*, New Architect Magazine, May 2002, http://www.newarchitectmag.com/print/documentID=24768; Rachel Konrad, *Deciphering the hacker myth*, CNET.com, Feb. 5, 2002, http://news.com.com/2102-1082-829812.htm; Kevin Anderson, *Why write computer viruses?*, BBC News, May 6, 2000, http://news.bbc.co.uk/hi/english/sci/tech/newsid_738000/738348.stm.

[27] DAVID G. MYERS, EXPLORING SOCIAL PSYCHOLOGY 3 (1994) (italics omitted).

[28] *Id.* 131.

[29] *Id.* 150-51. Professors John T. Cacioppo and Richard E. Petty are responsible for first propounding and proving the "routes-to-persuasion" concept in numerous publications. *See, e.g.,* John T. Cacioppo, Richard E. Petty, Chuan Feng Kao, and Regina Rodriguez, *Central and Peripheral Routes to Persuasion: An Individual Difference Perspective*, 51 J. Pers'y & Soc. Psych. 1032 (1986).

[30] *See* ROBERT B. CIALDINI, INFLUENCE 7 (rev. ed. 1994).

[31] *See*, e.g., ELLEN J. LANGER, MINDFULNESS (1989).

[32] *Id.* 14.

[33] *See* ROBERT B. CIALDINI, *supra* note 30, at 4-5.

[34] ELLEN J. LANGER, *supra* note 31, at 15 (footnote omitted).

[35] *See* MIHALYI CSIKSZENTMIHALYI, FLOW 6, 210-11, 214 (1990).

[36] *See id.* 6.

[37] *See* John Geirland and Eva Sonesh-Kedar, *What Is This Thing Called Flow? Think Nirvana on the Web*, Los Angeles Times, July 6, 1998.

[38] *See id.* (remarks of Professor Thomas Novak of Vanderbilt University).

[39] Ellen J. Langer, *supra* note 31, at 40.

[40] David G. Myers, *supra* note 27, at 173.

[41] *See id.* 21.

[42] *See* David G. Myers, *supra* note 27, at 158.

[43] *See id.*

[44] *See* Robert B. Cialdini, *supra* note 30.

[45] *Id.* 226-27.

[46] *Id.* 227.

[47] *Id.*, citing Michael Cohen and Neil Davis, Medication Errors: Causes and Prevention (1981).

[48] David G. Myers, *supra* note 27, at 49.

[49] *Id.*

[50] *See* UCLA Center for Communication Policy, Surveying the Digital Future (November 2001), http://www.ccp.ucla.edu.

[51] *See* Jonathan J. Rusch, *The "Social Engineering" of Internet Fraud*, Paper Presented at 1999 Internet Society Annual Conference, http://www.isoc.org/isoc/conferences/inet/99/proceedings/3g/3g_2.htm.