

Identity Theft

Markus Jakobsson

Palo Alto Research Center

Alex Tsow[†]

The MITRE Corporation

Identity theft is commonly defined as unwanted appropriation of access credentials that allows creation and access of accounts and that allows the aggressor to pose as the victim. Phishing is a type of identity theft that is perpetrated on the Internet and that typically relies on social engineering to obtain the access credentials of the victim. Similar deceit techniques are becoming increasingly common in the context of crimeware. Crimeware, in turn, is often defined as economically motivated malware. Whereas computer science has a long-term tradition of studying and understanding security threats, the human component of the problem is traditionally ignored. In this chapter, we describe the importance of understanding the human factor of security and detail the findings from a study on deceit.

Social engineering can be thought of as an establishment of trust between an attacker and a victim, where the attacker's goal is to make the victim perform some action he would not have wanted to perform had he understood the consequences. Attackers leverage preexisting trust between victims and the chosen false identities to spur dubious actions (illegally transferring money, remailing stolen goods, installing malware on computers, and recommending fraudulent services to friends).

To understand deceit in this context, it is worth recalling that people are more likely to install software on their computer if they believe it is manufacturer-distributed patch rather than a third-party enhancement. Similarly, Internet users are more likely to visit a website

when recommended by friends^{1,2} and may agree to signing up to services that appear to be recommended by their friends. Moreover, when site-content hinges on accepting third-party browser extensions, friend recommendations prove highly effective in inducing the required installation.³ When identity is used convincingly, these behaviors become social vectors for spreading crimeware and for causing users to opt in where they would not otherwise have.

Institutions and individuals project Internet identity through their Web sites and through email communication. How do attackers engineer contact with false identities? Clearly, email can be sent to anyone. Filters limit the quantity of unwanted messages, but spammers have successfully responded with increased volume and variation. Superficially, arranging contact with bogus Web sites appears to be a more difficult problem since legitimate content providers uncommonly link to spoofed Web hosts.

Roughly 50% of Web requests (by volume) are not the result of site-to-site linking, based on results from over 100,000 Internet clients hosted by the Indiana University campuses, according to the Indiana University Advanced Network Management Lab.

The other half of Web visits follow from bookmarks, direct address bar manipulation, or linking from external sources (email, word-processor documents, and instant-

[†]. The author's affiliation with The MITRE Corporation is provided for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions or viewpoints expressed by the author.

1. T. Jagatic, N. Johnson, M. Jakobsson, and F. Menczer, "Social phishing," *Communications of the ACM*, 2007, available at <http://doi.acm.org/10.1145/1290958.1290968>.

2. A. Genkina and L. J. Camp, "Phishing and countermeasures: understanding the increasing problem of electronic identity theft," chapter case study: *Net Trust*, John Wiley & Sons, 2007.

3. M. Gandhi, S. Stamm, M. Jakobsson, "verybigad.com: A study in socially transmitted malware," www.indiana.edu/~phishing/verybigad/.

messaging sessions). Social engineers influence these values through bogus links in email and by using domain names that are deceptive.

There are many studies of ways in which humans relate to deceit and treason, and there are many studies that focus on Internet security, but there is not an abundance of research on the combination of these two important fields. How do people relate to deceit on the Internet? This is an important question to ask—and to answer—for anybody who wants to improve Internet security.

1. EXPERIMENTAL DESIGN

This experiment tests the ability to identify phishing, an Internet scam that spoofs emails and Web pages to trick victims into revealing sensitive information. Although cast in terms of phishing, the results generalize to identity spoofing for purposes beyond information theft. This experiment tests the effects of several media features—sometimes in multiple contexts—on an individual's evaluation of its phishing likelihood. This experiment shows subjects six email screenshots followed by six Web page screenshots and asks them to rate their authenticity on a five-point scale: Certainly phishing, Probably phishing, No opinion, Probably not phishing, and Certainly not phishing (see Figure e49.1 for an example). The experiment was administered through SurveyMonkey.com,⁴ an online Web survey service. Subjects were required to rate each screenshot before advancing to the next stimulus. The survey provided the following instructions to subjects:

- Phishing is a form of Internet fraud that spoofs emails and Web pages to trick people into disclosing sensitive information. When an email or Web page fraudulently represents itself, we classify it as phishing.
- This survey displays a sequence of email and Web site screenshots. Assume that your name is John Doe and that your email address is johndoe1972@gmail.com. Please rate each screenshot's authenticity using the five-point scale: Certainly phishing, Probably phishing, No opinion, Probably not phishing, Certainly not phishing.

This style of testing, termed *security first*, measures fraud-recognition skills rather than habits. Subjects are not trying to accomplish other work but are merely instructed to rate a series of legitimate and illegitimate stimuli. For this reason, security-first measurements place a plausible upper bound on fraud detection habits in normal computer usage. Even though security-first

4. SurveyMonkey.com, "Surveymonkey.com-powerful tool for creating web surveys. online survey software made easy!" www.surveymonkey.com/, retrieved December 2006.

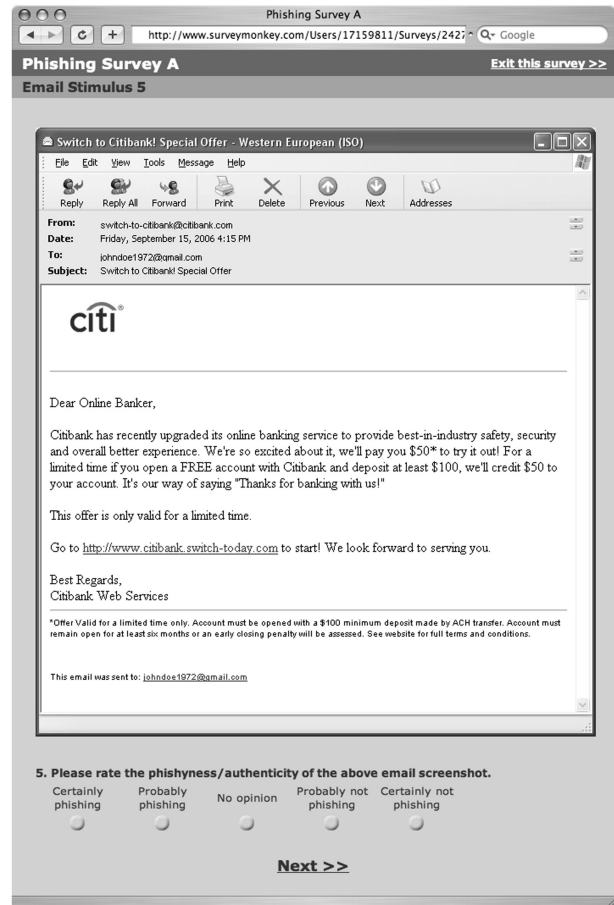


FIGURE e49.1 Subjects evaluate authenticity based on Screenshots using a five-point scale. The survey required a judgment before proceeding to the next stimulus.

evaluations have shown high susceptibility to phishing,^{5,6} role-playing experiments designed to measure fraud detection habits^{7,8} demonstrate even more serious vulnerability.

Subjects were recruited from an undergraduate introductory noncomputer-science-major class on computer usage and literacy. Of a class size exceeding 600 students, 435 began this study. All but 12 subjects were between

5. A. Genkina and L. J. Camp, "Phishing and countermeasures: understanding the increasing problem of electronic identity theft," chapter case study: *Net Trust*, John Wiley & Sons, 2007.

6. M. Jakobsson, A. Tsow, A. Shah, E. Blevis, and Y.-K. Lim, "What instills trust? A qualitative study of phishing," In submission, 2006.

7. J. S. Downs, M. B. Holbrook, and L. F. Cranor, "Decision strategies and susceptibility to phishing," In *SOUPS '06: Proceedings of the Second Symposium on Usable Privacy and Security*, pp. 79–90, New York, 2006, ACM Press.

8. M. Wu, R. C. Miller, and S. L. Garfinkel, "Do security toolbars actually prevent phishing attacks?" In *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 601–610, New York, 2006, ACM Press.

17 and 22 years old; the gender split was 40.0% male to 57.9% female (2.0% did not respond to this question). Although the test population is not demographically representative of general computer users, their enrollment in the introductory course suggests that their computer skill level is generally representative. The class's only prerequisite is high-school algebra. Almost all students in this class had used computers before but had no particular expertise.

The experiment divided the population into two sets through random selection. The two sets completed different versions of the survey. For 10 of the stimuli, the two versions differ only by a target collection of test features. Our primary analysis compares the impact of the feature changes, by using the χ^2 to represent the difference between response distributions. In two other question sets, subjects evaluate the authenticity of messages and Web pages under third-party administration (a potential vector for social engineering). We further designed the test to simulate a roughly equal number of authentic and phishing stimuli to avoid effective use of a trivial rating strategy: If there were significantly more phishing stimuli than authentic stimuli, subjects could employ an “always phishing” strategy that would correctly evaluate most of the stimuli without exercising due consideration.

Since the stimuli are only screenshots, their inauthentic features were designed to be evident on examination (rather than mouse-over or source analysis). For instance, incorrect domains are apparent in email hyperlinks; they are not disguised by an inconsistent *href* attribute. The domains we chose to simulate inauthentic URLs were not in use at the time of testing, but some are owned by their respective companies, others are owned by unrelated companies, and the rest appear to be unregistered, according to the Whois database. Our use of these domains as representations of inauthentic URLs is still valid because none of these URLs exist with the content we present. We outline the stimuli, their relevant features, and what we hope to learn in Figures e49.2a and e49.2b.

Authentic Payment Notification: Plain versus Fancy Layout

These two email messages use actual payment notification text from Chase Bank (see sidebar, “A Strong Phishing Message”). The text personalizes its greeting and references a recent payment transaction; there are no hyperlinks. One version uses the original layout (a one-color header containing the company logo followed by the message text); the other version uses an enhanced

layout (a header that includes a continuous tone shiny logo and a photograph of a satisfied customer, a smooth gradient footer graphic that spans the page with a gentle concave arch, hyperlinks to Privacy and Terms of Use, and a copyright notice; these graphics were adapted from the Web page at www.bankone.com).

A Strong Phishing Message

Dear John Doe,

JPMorgan Chase & Co. is proud to serve you as a former Bank One client.

Chase Online's patented ePIN technology is used both for eDebit transactions and for physical ATM access. While we will support the legacy 4 digit PIN for the remainder of the year, until December 31, 2006, we invite you to register for the ePIN program through our secure online server:

<https://www.chase.ePIN-simplicity.com>

If you have any questions about this or other Chase programs, do not hesitate to call our toll-free customer service number, 1-877-CHASEPC.

Sincerely,

Client Services

JPMorgan Chase & Co.

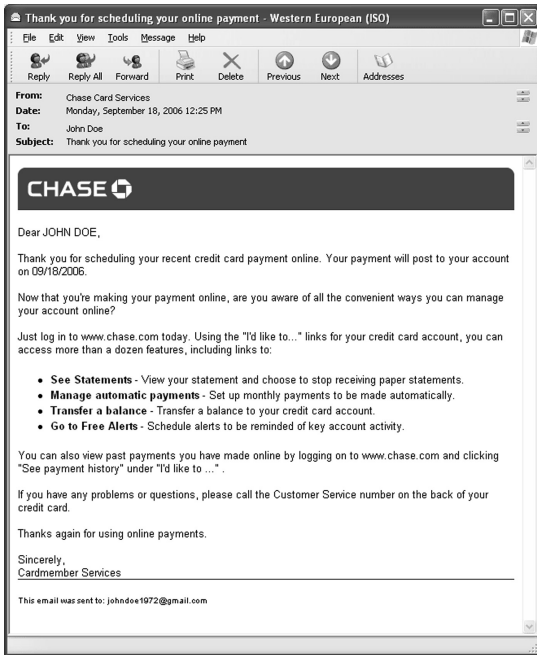
Strong Phishing Message: Plain Versus Fancy Layout

We constructed the phishing message text to sound as plausible as possible. Opening with a personalized greeting, the message explains that former Bank One customers will need to register for Chase's ePIN program—a replacement for ATM PINs that is also bundled with a new eDebit online service. It implicitly threatens service discontinuation by supporting “legacy 4 digit PINs for the rest of the calendar year.” The bogus registration hyperlink uses the made-up URL <https://www.chase.ePIN-simplicity.com>. The message closes with a bogus phone number to call for assistance. The two versions of this message use the same plain and fancy layout schema described earlier, with one exception: The fancy layout adds shiny letters proclaiming “Bank One is now Chase” (see Figure e49.3) between the header and message text.

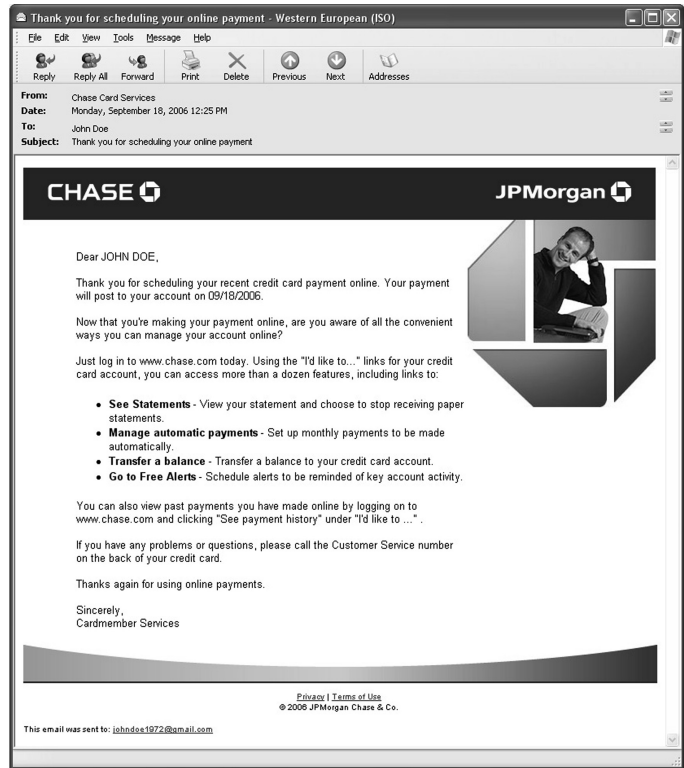
Authentic Promotion: Effect of Small Footers

Figure e49.4 shows an authentic message from the AT&T Universal card that promotes the company's paperless billing system. It personalizes the greeting and includes the last four digits of the account number. There are multiple company logos, a blue outline around text, an Email

9. R. Dhamija, J. D. Tygar, and M. Hearst, “Why phishing works,” In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 581–590, New York, 2006, ACM Press.



(a)



(b)

FIGURE e49.2 (a) Plain layout. (b) fancy layout.

Security Zone box, and a small-print footer filled with trademark, copyright, and contact notices as well as various informational and administrative hyperlinks. The principal login hyperlink conceals its destination. The test pair consists of the original message and a modified version that excludes the small print footer.

Weak Phishing Message

The sidebar “Phishing Message” shows a phishing message promising \$50 for opening an account with Citibank. There is a simple company logo in the header; a footer contains legal disclaimers about the offer. There is no personalization, and the lone hyperlink is a made-up domain (actually owned by an unrelated organization), <http://www.citibank.switch-today.com>. Figure e49.5 contains its screenshot. The two versions differ only by the presence of a center-aligned “VeriSign Secured” endorsement logo that follows the footer’s legal disclaimers.

Authentic Message

The test determines the impact of a “VeriSign Secured” logo added to the footer of an authentic message, as

Phishing Message

Dear Online Banker,

Citibank has recently upgraded its online banking service to provide best-in-industry safety, security and overall better experience. We’re so excited about it, we’ll pay you \$50* to try it out! For a limited time if you open a FREE account with Citibank and deposit at least \$100, we’ll credit \$50 to your account. It’s our way of saying “Thanks for banking with us!”

This offer is only valid for a limited time.

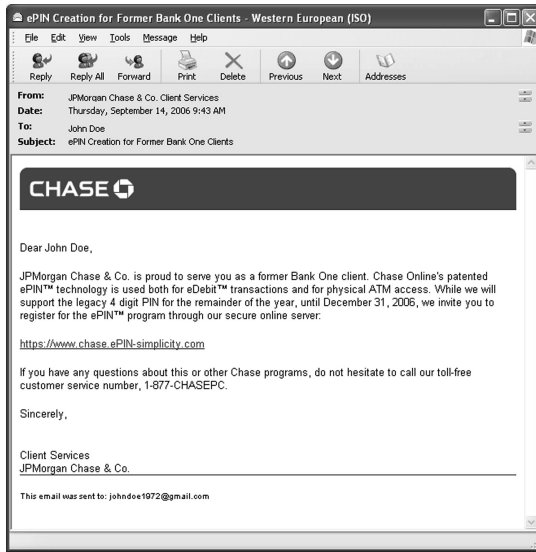
Go to <http://www.citibank.switch-today.com> to start!

We look forward to serving you.

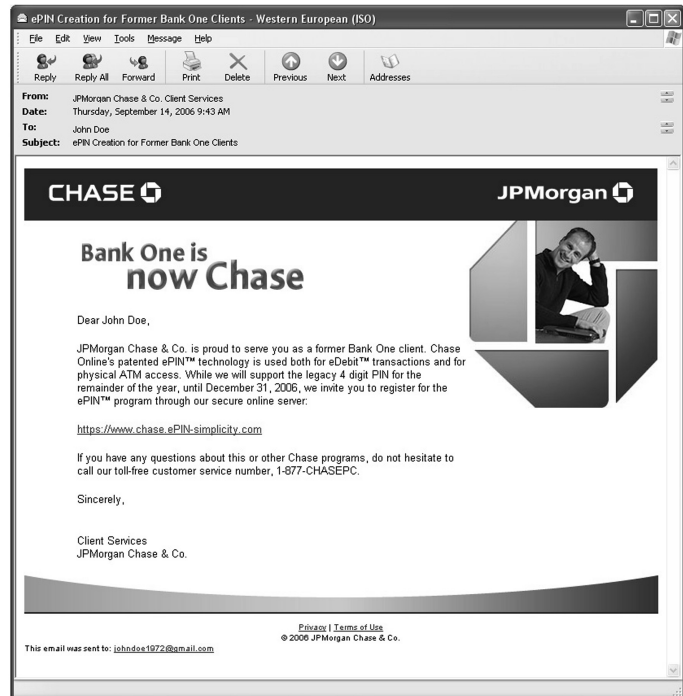
Best Regards,
Citibank Web Services

shown in Figure e49.6. The notice begins with a personalized greeting and informs the client about changes in PayPal’s logo insertion policy. The message body is considerably longer than all of the other messages except for

*. Offer valid for a limited time only. Account must be opened with \$100 minimum deposit made by ACH transfer. Account must remain open for at least six months or an early closing penalty will be assessed. See website for full terms and conditions.



(a)



(b)

FIGURE e49.3 (a) Using a plain layout schema (b) using a fancy layout schema; and (c) using a plain and fancy layout schema.

the Netflix stimulus. The primary message contains no hyperlinks, but their small-print footer furnishes a hyperlink to unsubscribe from their newsletter. One interesting feature of the message is a boldfaced statement: “If you do not wish to have PayPal automatically inserted in your listings, you must update your preferences by 9/25.” Though genuine, this message parallels the account shut-down threats brandished by many phishing messages. The header contains a monochrome company logo and a two-tone horizontal separation bar.

Login Page

We say that the URL is strongly aligned with the content of the page when these two “belong together.” Imagine that one would look at some ten Web pages (without associated URLs), and then some ten randomly ordered URLs, each one belonging to one of the ten Web pages. The easier it is for a potential reader to correctly pair up Web pages and URLs, the stronger the alignment. If any Web page and associated URL is not correctly matched up, then the alignment is very weak.

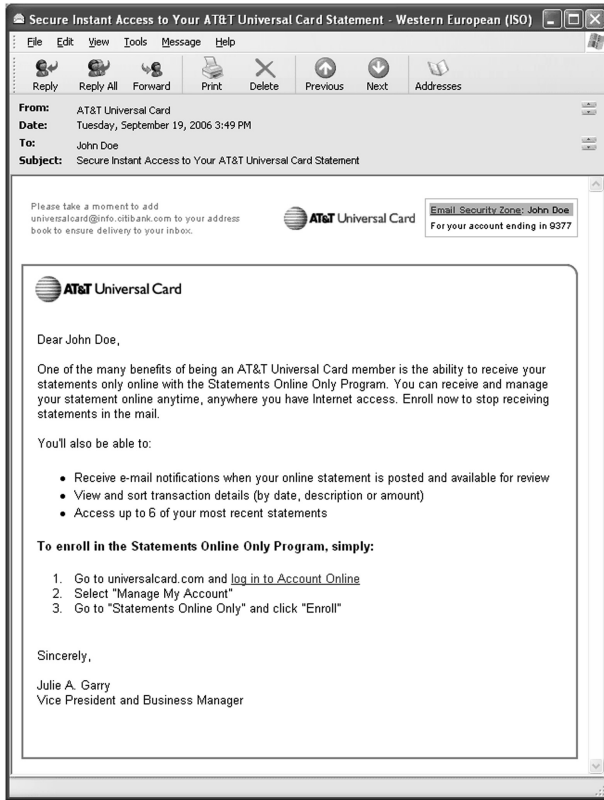
Let’s now turn to an example, as shown in Figure e49.7. This browser’s content window displays an exact copy of the AT&T Universal card login page. Like most Web login pages, it displays a high level of layout sophistication: photographs of happy clients, navigation

bars, product pictures, a sidebar, promotional windows, and small-print legal disclaimers. It also displays a “VeriSign Secured” site endorsement logo. The two versions of this stimulus differ by their address bar contents: version (a) uses <https://www.accountonline.com/View?docId=Index&siteId=AC&langId=EN> and consequently displays a browser frame padlock; version (b) uses the unencrypted URL <http://www.attuniversalcard.com/> (owned by AT&T but not in use).

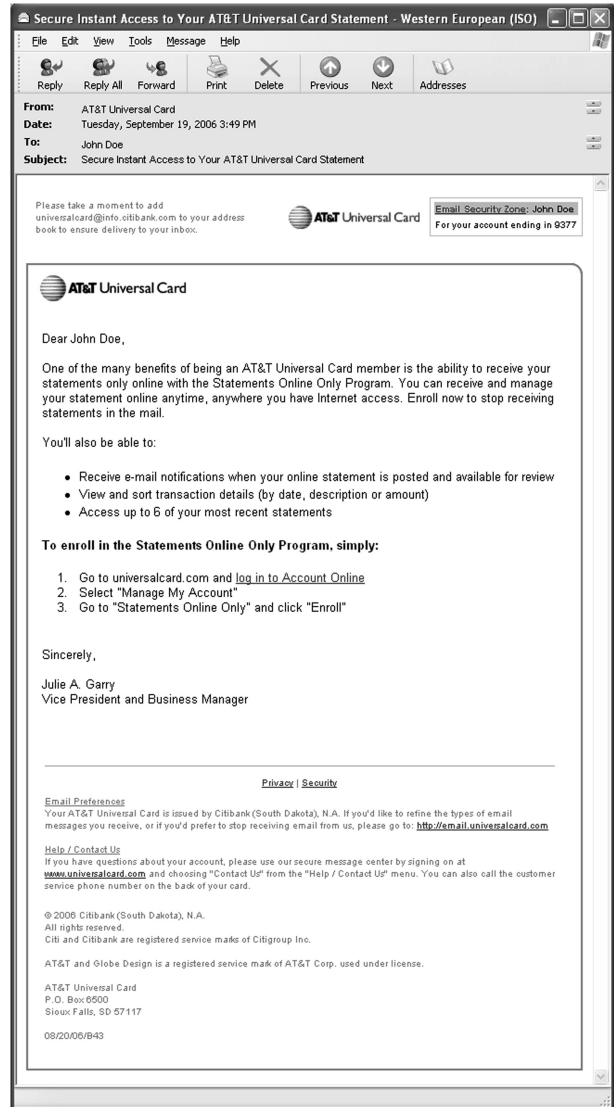
Login Page: Strong and Weak Content Alignment

This next set takes the alternative approach to aligning the address bar URL with content: change the content. Both pages (see Figure e49.8) use the unregistered URL www.citicardmembers.com/. Version (a) displays a precise copy of the authentic Citi Credit Cards login page in its content window, whereas the version (b) content window displays modified logos and links (see Figure e49.9) for better alignment with the URL.

Figures 49.8a and 49.8b use a sophisticated layout with nearly all the same identifiable features of the AT&T Universal Card login: photographs of happy clients, navigation bars, product pictures, sidebars, promotional windows, and the “VeriSign Secured” logo.



(a)



(b)

Please take a moment to add
 universalcard@info.citibank.com to your address
 book to ensure delivery to your inbox.

(c)

Email Security Zone: John Doe
 For your account ending in 9377

(d)

To enroll in the Statements Online Only Program, simply:

1. Go to universalcard.com and [log in to Account Online](#)

(e)

FIGURE e49.4 (a) An authentic message from the AT&T Universal card that promotes the company’s paperless billing system. (b) an authentic message from AT&T Universal card that personalizes the greeting and includes the last four digits of the account number; (c) header detail left side; (d) header detail right side; and (e) hyperlink detail.

Login Page: Authentic and Bogus (But Plausible) URLs

These two stimuli test the impact of changing a well-aligned authentic URL (see Figure e49.9), <http://www.paypal.com/ebay>, to a reasonably well-aligned bogus URL, <http://www.ebaygroup.com/paypal> (domain owned by eBay but not in use). The main content window is the same for both: The eBay decorated version of the PayPal login page, which contains an eBay logo to the lower right of the primary PayPal logo. The page layout contains all the main features of the previous login pages but includes a more thorough set of third-party endorsement

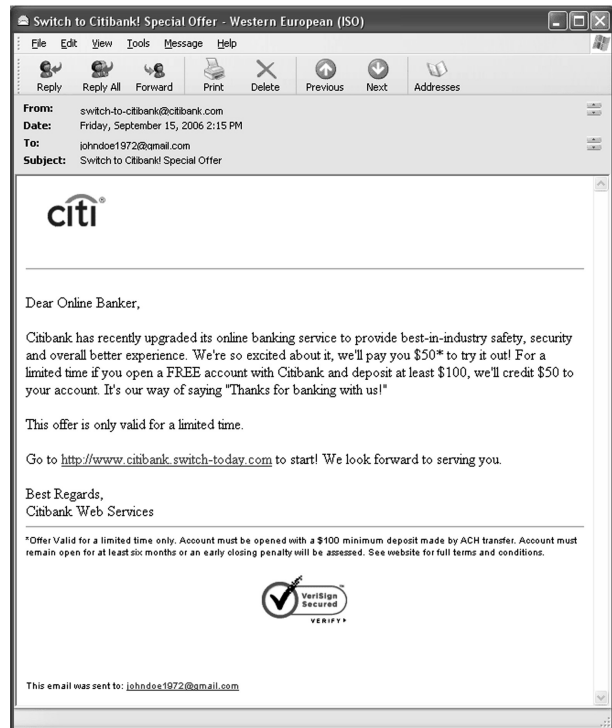
logos: “VeriSign Secured,” “Reviewed by TRUST-e,” and “Privacy: BBB OnLine.” SSL is not used in either stimulus.

Login Page: Hard and Soft Emphasis on Security

Figure e49.10 tests whether it is possible to undermine confidence in an authentic login page with excessive concern about security and online fraud. These two stimuli represent an extreme but real-world case. Clients of the



(a)



(b)

FIGURE e49.5 (a) Effect of endorsement logo. (b) center-aligned “VeriSign Secured” endorsement logo.

Indiana University Employees Federal Credit Union (IUCU) were targeted by a phishing attack in early August 2006. In response, the credit union altered its Web page to include a large banner, as shown in Figure e49.10f.

They further augmented the news section with a similar message: “Warning! Phishing Scam in progress (learn more).” Finally, a section named “Critical Fraud Alerts” contained the exact same warning as the one from the news section. The twin page eliminates all phishing warnings (including the banner) and changes the “Critical Fraud Alerts” section heading to read “Fraud Prevention Center.” Generally, the language was changed to sound “in control” rather than alarmist.

Bad URL, with and without SSL and Endorsement Logo

Can an endorsement logo and SSL padlock overcome a bad domain name? This next set, as shown in Figure e49.11, tests these two features on a Wells Fargo phishing site based on the bogus domain www-wellsfargo.com. The login page is similar in layout to the others but does not feature photographs of people. The only continuous tone graphic is an image of a speeding horse-drawn carriage that evokes a Wild West money

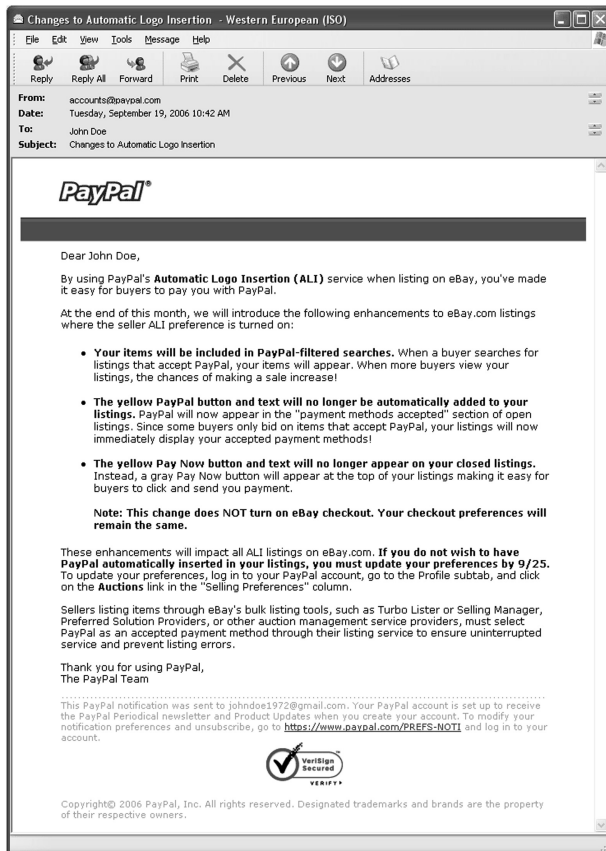
transfer service. One screenshot contains the original page content using an unencrypted connection; the other stimulus uses SSL and adds a center-aligned “VeriSign Secured” logo to the page’s footer.

High-Profile Recall Notice

At the time of testing (10/02/2006–10/12/2006), the press had been alive with recent reports¹⁰ of laptop battery recalls from both Dell and Apple. Sony, the source of the batteries, ultimately issued a direct recall for the same batteries, by adding several more brands on 10/23/2006.

As shown in Figure e49.12, this test set does not follow the controlled-pair format of the previous stimuli. One stimulus is a screenshot of the official Dell Battery Return Program Web page. The page layout is substantially simpler than all other Web stimuli. A four-color header logo adorns the top of the page. It presents the content as a letter to “Dell Customer,” explaining the danger and how to determine eligibility for exchange. Notably, there is a single column of content, no photos, and no promotional content of any kind. They use the third-party domain dellbatteryprogram.com. Use of this

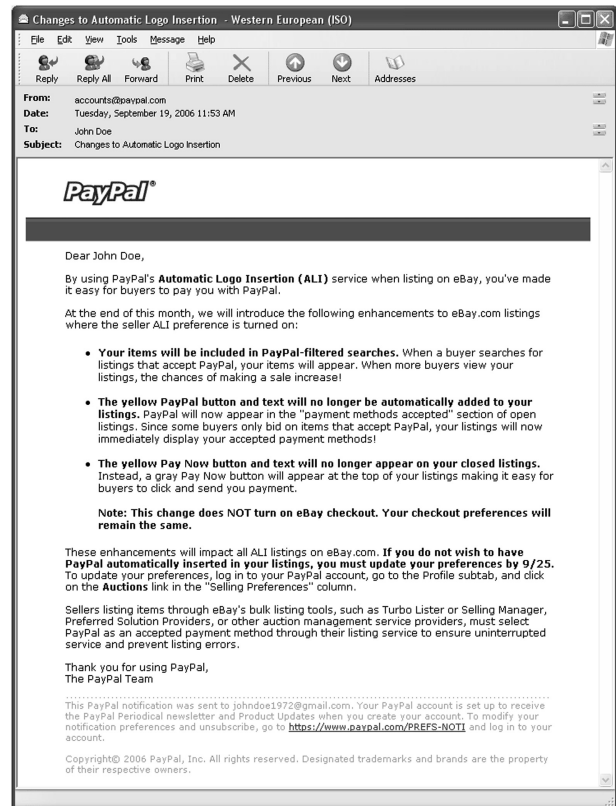
10. D. Darlin, “Dell will recall batteries in PCs,” *New York Times*, 15 August 2006, <http://select.nytimes.com/search/restricted/article?res=F10A1FF83C5A0C768DDDA10894DE404482>.



(a)

These enhancements will impact all ALI listings on eBay.com. **If you do not wish to have PayPal automatically inserted in your listings, you must update your preferences by 9/25.** To update your preferences, log in to your PayPal account, go to the Profile subtab, and click on the **Auctions** link in the "Selling Preferences" column.

(c)



(b)

This PayPal notification was sent to johndoe1972@gmail.com. Your PayPal account is set up to receive the PayPal Periodical newsletter and Product Updates when you create your account. To modify your notification preferences and unsubscribe, go to <https://www.paypal.com/PREFS-NOTI> and log in to your account.



Copyright© 2006 PayPal, Inc. All rights reserved. Designated trademarks and brands are the property of their respective owners.

(d)

FIGURE e49.6 (a) “VeriSign Secured” logo added to the footer of an authentic message. (b) authentic message—effect of endorsement logo; (c) shared body detail; (d) endorsement footer detail.

domain for the official page makes the replacement service ripe for phishing.

We constructed a phishing email message using the header, footer, and textual content from this Web page. The phishing message omits the middle section on how to identify eligible batteries and instead requests that the recipient go to the bogus Web page at <http://www.dellbatteryreplacements.com>.

Low-Profile Class-Action Lawsuit

As shown in Figure e49.13, this last set of stimuli also follows the third-party email and Web page form of the previous set. Both stimuli are authentic, but they use altered dates to appear relevant at the time of testing. The email message is a lengthy notice that describes a class-action lawsuit against Netflix, a settlement to the lawsuit,

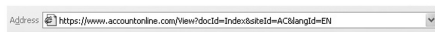
and options for claiming benefits. There is no greeting, signature, color, or graphics. The only hyperlinks direct the user to the authentic third-party URL, <http://www.netflixsettlement.com>. The Web page has a similarly bare appearance but with much less text. It behaves as a hyperlink gateway for more information under the URL <http://www.netflix.com/settlement/> (the result of redirection from <http://www.netflixsettlement.com>).

2. RESULTS AND ANALYSIS

Our experiment directly controls for the effect of several design features. There are some surprises in the direct results from these tests, including the stunning impact of a detailed small print footer on an otherwise well-conceived legitimate message; however, the experiment reveals an unexpected, but in retrospect obvious, lesson



(a)



(b)



(c)

FIGURE e49.7 (a) Strong and weak URL alignment; (b) weak alignment URL detail; (c) strong alignment URL detail.

about email messages: The “story” of the message is critical. Messages with strong and succinct narrative components rated highly and their ratings appear to be less susceptible to changes in graphic design. On the other hand, authenticity perception changed significantly for messages that say little (such as a service promotion) under document feature variances. Two of the five sets of “twins” did not change significantly, according to the metric, when augmented with the very same features that produced significant changes in other messages. Subjects judged these messages principally on their narrative content.

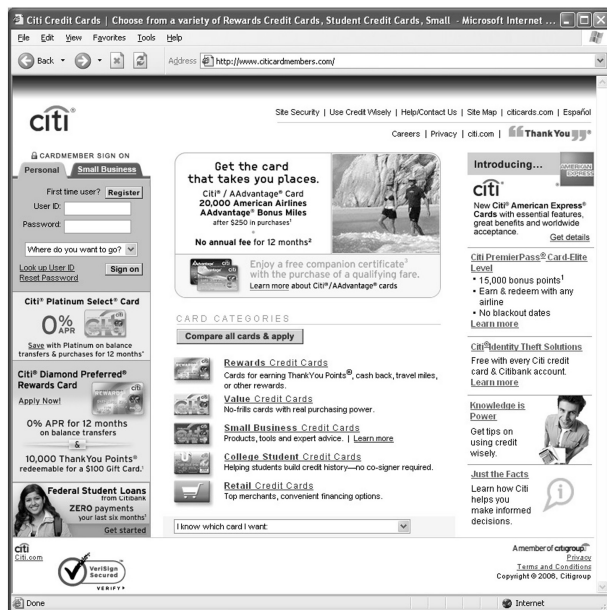
The Chase phishing message uses the company’s recent acquisition of Bank One as a pretext for imminent

service change to ATM card authentication; the story further bundles this change with the addition of a new service, eDebit, and implicitly threatens discontinuation of service by claiming to “support the legacy 4 digit PIN for the remainder of the year.” This consequence is much less direct than the standard “Your account will be suspended in 48 hours if ...” strategy used by many phishers. It works on a less urgent time scale and is pitched as a convenience to all clients rather than an anomaly specialized to a specific client. There was no significant difference in subject evaluation between the plainly formatted version of this message with the monochrome corporate header and the one with several customized continuous tone graphics.

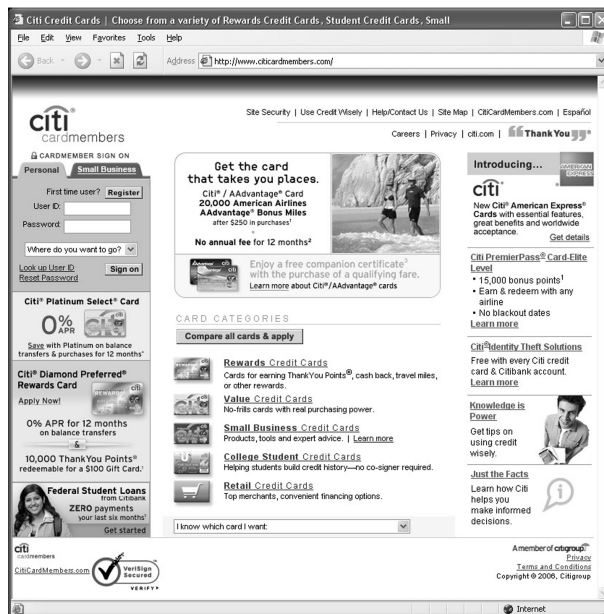
Yet these same graphics, minus the shiny “Bank One is now Chase” banner, produced a significant change ($p = 0.018$) in evaluations of Chase’s authentic payment notice. The payment notice thanks the receiver for a recent online payment, then goes on to inform the user about features of their online account management interface. In terms of relevance, there is less potential impact on the user. Ignoring this message won’t expose the client to any changes (good or bad). Assuming that the name and recent online interaction are correct, the message communicates little that would surprise the average client. So, in place of strong narrative components, the subject looks to formatting cues to further inform confidence. This message rated highly in its simple form and was pushed higher by the improved graphics. We attribute its initially high rating to its exclusion of hyperlinks, informative nature, and well-contextualized message.

Subject reactions to the presence of the “VeriSign Secured” logo differed dramatically between the two test messages. One message, a change of policy notice from PayPal, experienced no statistical difference in subject evaluations of its endorsed and unendorsed forms. The policy change notice shares several narrative features with the Chase phishing message: Both messages have a customized salutation, both inform users about an institution wide change (in this case, the particulars of their logo insertion policy), and both claim that inaction will result in a change of service. The PayPal message has no hyperlink in the message body but does contain a link in its small-font gray footer to manage user preferences; we think that this forecasts a potential phishing strategy.

The other message that tests the impact of the “VeriSign Secured” logo is a phishing message that exploits the Citibank brand. The experiment shows a statistically significant change in subject evaluations ($p = 0.047$) due to this single feature change. Of all the messages, this message makes the weakest connections to the receiver. It begins with a generic salutation. Worse, the first sentence promotes the goodness of their online service but fails to involve the receiver in any way. Not



(a)



(b)



(c)



(d)



(e)



(f)



(g)



(h)

FIGURE e49.8 (a) A precise copy of the authentic Citi Credit Cards login page in its content window. (b) content window displays modified logos and links; (c) location of detail window; (d) original and modified detail window; (e) header menu of detail window; (f) Cardmember sign on detail window; and (h) Cardmembers of detail window.

until the second sentence of the message's relevance become evident to the reader: They are offering "\$50* to try it out!" These two messages had the lowest average ranking of all the email stimuli. Ignoring this message has no impact on the user except for failing to miss out on an offer of dubious value. Ultimately, the stimulus fails to engage the reader, and so subjects base more of their evaluation on nonnarrative factors such as the endorsement logo and the bogus URL.

The Dell battery replacement program message presents a compelling story, but not directly. The incident received a high level of media coverage due to spectacular reports of exploding and burning laptop computers. The Dell message, which we manufactured, benefits from other sources spreading the story. Without this third-party validation, this message could have bordered on implausibility, but instead our subjects produced ratings that were statistically indistinguishable from the two most highly ranked email messages in the batch. This message contains a slightly nicer-than-average layout (multicolored header, footer graphic with links) but less personalization and a fraudulent (but semantically plausible) URL. The story was so powerful and present in the subjects' minds

that they were willing to discount the suspicious link and generic greeting.

The other third-party attack (Netflix) did not benefit from recent or high-profile media coverage (see Table e49.1). We may have further lowered its rating by altering the dates to appear relevant at the time of testing.

Subjects could have perceived the timeline as implausibly long (even for legal action) or may have been familiar with the case and known that the dates were incorrect. In addition to these changes, the original message is particularly poorly conceived. Though it has strong narrative elements that present lawsuit context, the elements of the settlement, and response options, the message is entirely too long. Message length and detail create an incentive for users to quickly evaluate according to nonnarrative features. The most visually obvious features are the inclusion of blue hyperlinks—the only non-black-and-white symbols—that point to <http://www.netflixsettlement.com>. Though this is the legitimate domain, it should raise suspicion because it is an apparent "cousin domain" to the parent company's Web site. The lack of strong design features seals its poor evaluation. There is no company header—a feature present on every other stimulus—and



(a)



(b)



(c)



(d)

FIGURE e49.9 (a) A reasonably well-aligned bogus URL. (b) logo detail; (c) bogus URL detail; and (d) authentic URL detail.

no opening salutation or signature. The contact address appears to be an afterthought that does not even specify a division of the company, let alone an appropriate administrator.

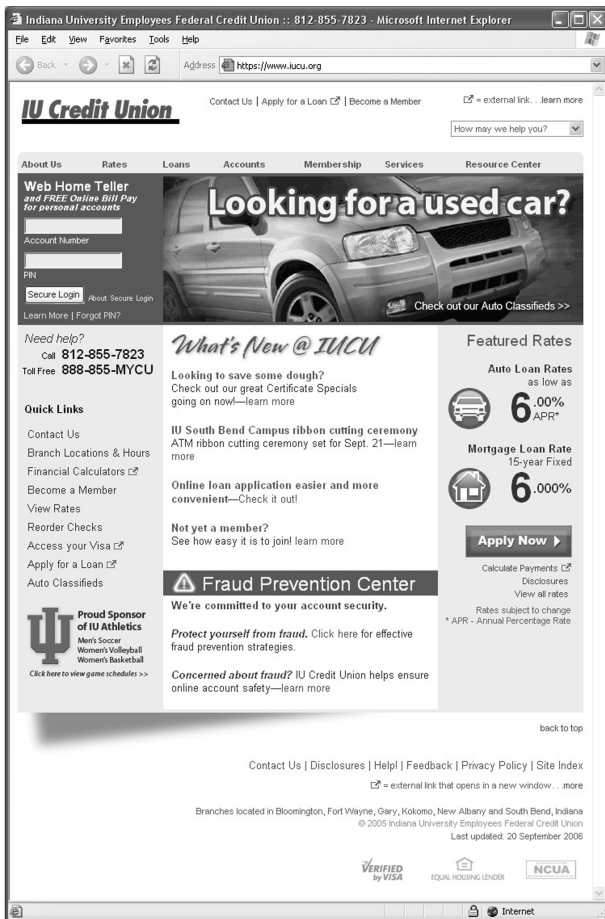
The biggest surprise of the test appears in a pair of authentic AT&T Universal card messages. In many ways, this is the polar opposite of the Netflix settlements message: it has strong design elements and a short, weak narrative. The message promotes AT&T's Statements Online Only program without bundling it with a recent action (as Chase does with its payment notification). Ignoring this message will not produce any change in the receiver's

service, nor does enrolling in the program provide any obvious benefit to the client; in fact, enrollment could result in unintended late payments due to imperfect spam interdiction of electronic billing notices. What the message lacks in narrative appeal it makes up for in design strength. It customizes the message to the receiver both in the opening salutation and in an "Email Security Zone" header box that displays client name and client account number suffix. The header also has a spam awareness message and company logo. A blue outline that complements the company logo encloses the rest of the content. The corporate logo appears a second time within the blue content box and the letter opens and closes with personalized salutations: "Dear John Doe," and "Sincerely, Julie A. Garry." The two versions of this message differ in the presence of a detailed small-print footer below the signature, which contains hyperlinks to privacy and security policies, as well as hyperlinks bearing the universalcard.com domain. The footer uses a small gray font and presents text for adjusting "Email Preferences" and a "Help/Contact Us" section containing the postal address and various trademark and copyright notices.

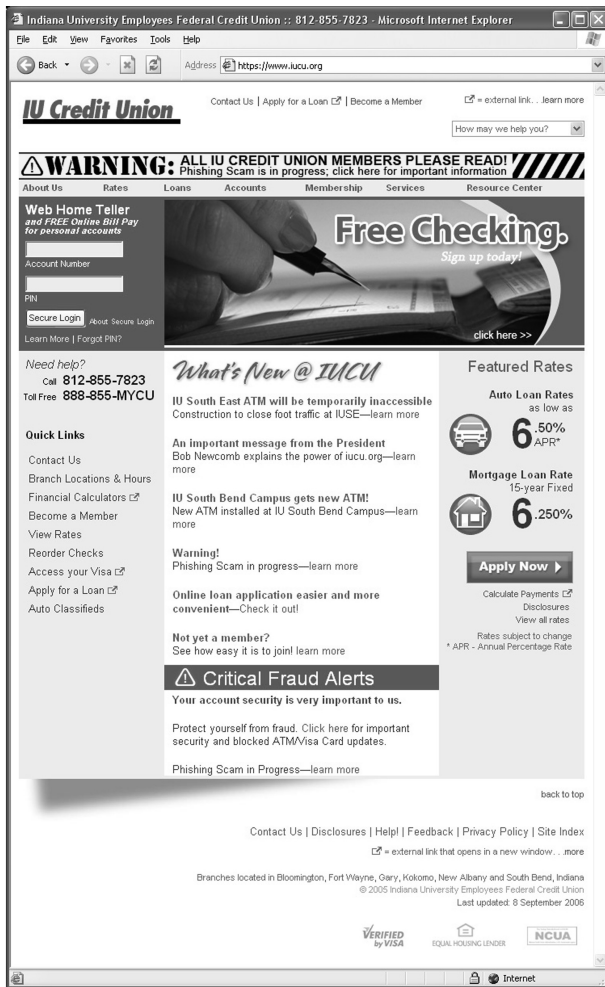
The one ambiguous feature of this email is a centrally located hyperlink, labeled "log in to Account Online." It does not indicate the URL in the text. Phishers frequently employ this sort of hyperlinking strategy to conceal the bogus server's URL. The footer may add confidence because its hyperlinks appear to reference URLs with legitimate and semantically aligned domain names; none of the hyperlinks outside the header indicate a target domain. Alternatively, the contact, copyright, and trademark notices themselves may improve confidence in the message. It is particularly interesting that even though the footer-less message displayed the last four digits of the credit-card number, customized the greeting, and employed generally strong design elements, except for the Netflix settlement, it was still ranked lower than any other legitimate message. This supports the experimental results in¹¹ that indicate indifference to customized greetings in certain contexts. Yet adding the footer boosts the evaluations to the point where it is statistically indistinguishable from the other two most highly ranked messages—the Chase payment notification with fancy graphics and the Dell battery recall notice.

Web sites, particularly the login and information collection pages associated with phishing scams, do not present a story the way email messages do. For this reason, their credibility depends much more on document features and graphic design. Subjects assigned significantly

11. M. Jakobsson and J. Ratkiewicz, "Designing ethical phishing experiments: a study of (rot13) ronl query features," In WWW '06: Proceedings of the 15th International Conference on World Wide Web, pp. 513–522, New York, 2006, ACM Press.



(a)



(b)



(c)



(d)

Warning!
Phishing Scam in progress—learn more

(e)



(f)

FIGURE e49.10 (a) Soft emphasis on security; (b) hard emphasis on security; (c) soft security detail; (d) hard security detail; (e) soft security detail 2; and (f) Web page warning.

different ratings to three of the five sets of twin stimuli. The results show that address bar alignment with page content, overwrought concerns about fraud, and third-party endorsements substantively change authenticity assessments.

The biggest rating difference among the Web page stimuli was measured between the two versions of the “AT&T Universal Card Sign-on” page. The official version, which uses the address URL <https://www.accountonline.com/View?docId=Index&siteId=A> C&langId=EN, has the lowest average rating of the ten Web stimuli in the paired testing. With an average rating

of 2.76, it rated lower than a simulated phishing Web site based on the suspiciously formed domain [www-attuniversalcard.com](http://www.attuniversalcard.com) (avg. rating: 3.17). Subjects who saw the unused domain <http://www.attuniversalcard.com> in the address bar of the AT&T card login page rated it significantly higher ($p = 0.004$) than the authentic page. The page content strongly aligns with the URL text <http://www.attuniversalcard.com>: the phrase “AT&T Universal” appears no less than seven times on the login page, whereas the phrase “AT&T Universal Card” appears four times in the content window. Interestingly, the official page uses HTTPS and displays an SSL padlock in the



(a)



(b)



(c)



(d)



(e)



(f)



(g)

FIGURE e49.11 (a) Without SSL and endorsement logo. (b) with SSL and endorsement logo; (c) no SSL/endorsement logo; (d) SSL/endorsement detail; (e) Internet no SSL/endorsement status bar; (f) Internet SSL/endorsement status bar; and (g) VeriSign logo.

lower right-hand browser frame, whereas the attuniversal-card.com domain does not use SSL and consequently does not display the padlock on the browser frame. Subjects found the semantic alignment of the URL to be a much stronger indicator of authenticity than SSL utilization. In fairness, subjects could not examine the certificates, nor would they have been subject to the “Unable to verify the identity of host” pop-up window in the case of a self-signed certificate. Nevertheless, other user studies have found that in practice, subjects rarely consider these factors.^{12,13}

Much to our surprise, the phishing simulation based on the URL www-wellsfargo.com rated significantly higher ($p = 0:00001$) than the official AT&T Universal card login page. Subjects valued semantic alignment between content and host domain more than domain well-formedness. Syntactically there is nothing wrong with the domain, but replacing the dot with a dash is clearly an

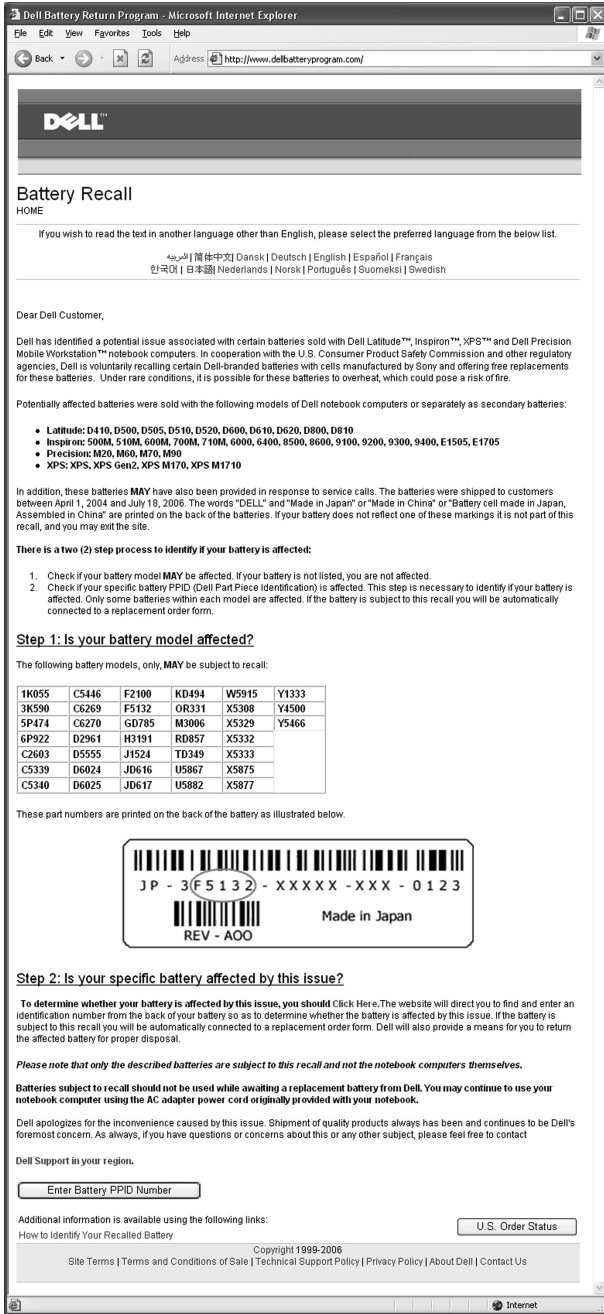
attempt at deception. Adding SSL and a “VeriSign Secured” logo to the Wells-Fargo phishing page produced a significant improvement in authenticity ratings ($p = 0:029$). It’s worth noting that the authentic login page (not in the test) does not display a VeriSign logo but does use SSL. In spite of subjects either failing to notice the dash-for-dot exchange or not thinking that it was suspicious, they did notice the presence of either SSL or the VeriSign logo. Note that the AT&T login page also used SSL and displayed a VeriSign endorsement, but neither of these features could overcome the mistrust of the accountonline.com domain.

The last statistically significant difference between twins in the Indiana University Credit Union homepage ($p = 0:022$) shows that too much concern about security can reduce customer confidence. Subjects responded positively to use of less fearful language and rated the softer, more constructive content significantly higher than the page that displayed stark warnings. Note that correct domain names and SSL were used on both stimuli. This is a case where a good-faith effort to educate clients about phishing undermines confidence in the Web site’s authenticity. Login pages are no place for fear-provoking messages.

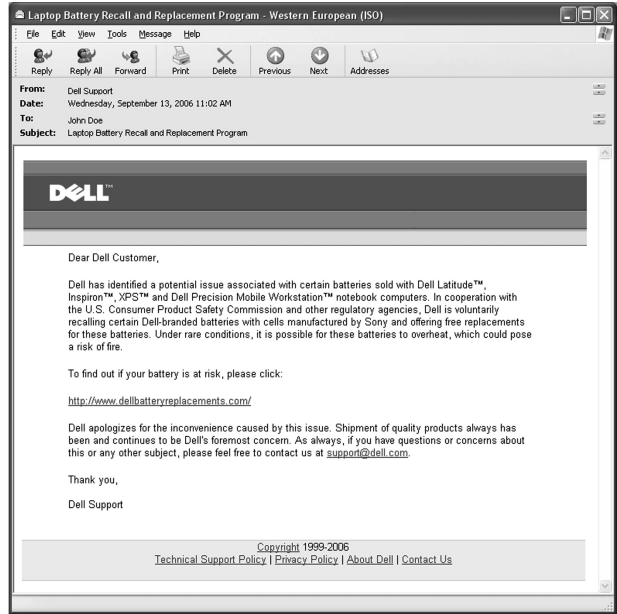
One way phishers align page content with URLs is by choosing an apt domain name; the other way is to change

12. R. Dhamija, J. D. Tygar, and M. Hearst, “Why phishing works,” In *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 581–590, New York, 2006, ACM Press.

13. M. Wu, R. C. Miller, and S. L. Garfinkel, “Do security toolbars actually prevent phishing attacks?” In *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 601–610, New York, 2006, ACM Press.



(a)



(b)



(c)

To find out if your battery is at risk, please click:

<http://www.dellbatteryreplacements.com/>

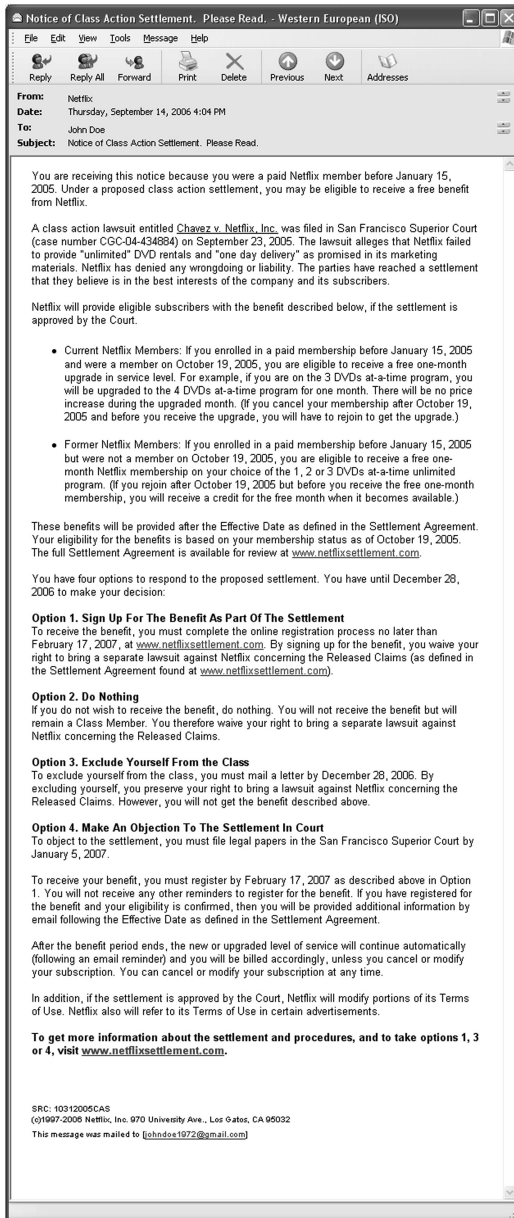
(d)

FIGURE e49.12 (a) The official Dell Battery Return Program Web page. (b) a letter to Dell customer; (c) genuine Web site URL detail; (d) bogus email body detail.

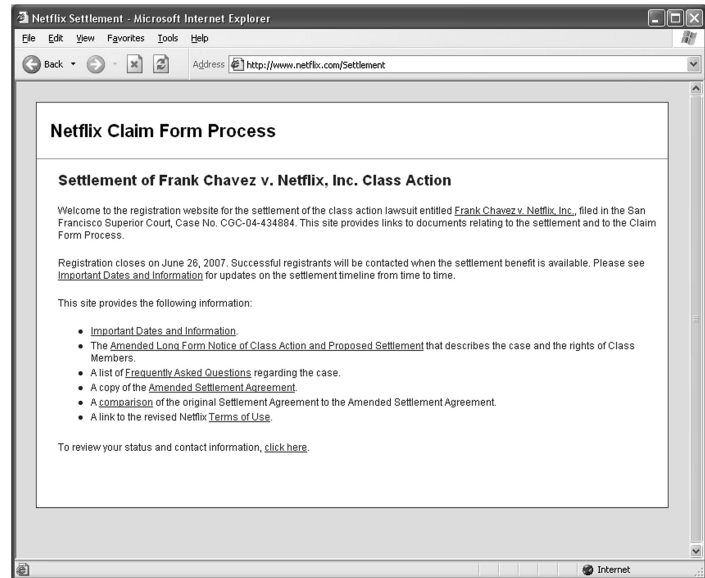
the page content. Though subjects gave a higher average rating to our modified Citibank cardmembers login page, the test showed that the two distributions were not significantly different ($p = 0.133$).

The last test pair was nearly significant ($p = 0.65$) but does not confirm that the two ratings come from different distributions. This pair compared the effects of a plausible

parent company domain and subsidiary subdirectory, <http://www.ebaygroup.com/paypal/>, with the authentic URL that reverses their positions, <http://www.paypal.com/ebay/>. Although the ebaygroup.com domain is unbound, eBay has registered it. Nevertheless, this test shows a certain flexibility in user acceptance of domain alignment. No PayPal client has seen the PayPal page



(a)



(b)

FIGURE e49.13 (a) Low-profile class-action suit. (b) a lengthy notice that describes a class-action lawsuit against Netflix, a settlement to the lawsuit, and options for claiming benefits.

displayed under the <http://www.ebaygroup.com/paypal/> address, yet their authenticity ratings are not significantly different. This result furthers our conviction that semantic alignment between content and URL is a principal factor in authenticity evaluations.

The last two Web stimuli sets are not twins; they are the Dell battery replacement page and the Netflix settlement page. Both pages are authentic, although the content of the Netflix page was altered to appear relevant at the time of testing. They received polar opposite ratings. The

Dell battery page was statistically indistinguishable from the highest-rated page (the authentic PayPal site), and the Netflix page rated dead last—significantly lower than the second lowest rating ($p = 1:20 \cdot 10^{-7}$). As mentioned before, the Dell battery program stimuli benefit from a high visibility news story. It's noteworthy that the URL in the Web page (authentic) is different from the URL in the email (phishing), which subjects saw first. Subjects did not penalize the Web page for this inconsistency. Subjects may have had difficulty constructing a phishing

TABLE e49.1 The Other Third Party Attack.

| Stimulus Description | Mean | Diff. | χ^2 | p |
|---|------|-------|----------|-------|
| Chase card payment statement (legit)—plain layout | 3.40 | 0.36 | 11.89 | 0.018 |
| Chase card payment statement (legit)—fancy layout | 3.76 | | | |
| Chase phish-fancy layout | 3.19 | 0.02 | 6.31 | 0.177 |
| Chase phish-plain layout | 3.18 | | | |
| AT & T Universal Card statement without legal notices | 3.05 | 0.62 | 30.18 | 0 |
| AT & T Universal Card statement with legal notices | 3.66 | | | |
| PayPal policy change + VeriSign | 3.19 | 0.11 | 5.75 | 0.219 |
| PayPal policy change – no VeriSign | 3.30 | | | |
| Citibank phish – no VeriSign | 2.40 | 0.29 | 9.62 | 0.047 |
| Citibank phish + VeriSign | 2.69 | | | |
| AT & T Universal card login | | | | |
| https://www.accountonline.com/View?docId = Index & siteId = AC & langId = EN | 2.76 | 0.49 | 15.46 | 0.004 |
| http://www.attuniversalcard.com | 3.25 | | | |
| Citibank (phish); URL = http://www.citicardmembers.com/ | 3.11 | 0.32 | 7.06 | 0.133 |
| Copy of original site; | 3.43 | | | |
| Logos modified to better match domain | | | | |
| PayPal Web site displaying eBay logo: | | | | |
| URL = http://www.ebaygroup.com/paypal/ | 3.35 | 0.35 | 8.83 | 0.065 |
| URL = http://www.paypal.com/ebay/ | 3.70 | | | |
| Indiana University Credit Union homepage: | | | | |
| Deemphasizes security language; no mention of “ attacks “ | 3.69 | 0.32 | 11.45 | 0.022 |
| Phishing attack banner + strong fraud warnings | 3.37 | | | |
| Wells Fargo phishing page: | | | | |
| Reproduces original content; | 3.17 | 0.31 | 10.83 | 0.029 |
| URL = http://www-wellsfargo.com/ | | | | |
| Adds VeriSign endorsement; uses SSL; | 3.48 | | | |
| URL = https://www-wellsfargo.com/ | | | | |
| Netflix class-action settlement email (authentic) | 2.72 | | | |
| Netflix class-action settlement homepage (authentic) | 2.55 | | | |
| Dell battery replacement email (phishing) | 3.61 | | | |
| Dell battery replacement Web page (authentic) | 3.54 | | | |

Note: The first section of the table reports on the differences between email messages; the next section reports on the Web pages; and the last section gives the average rating for the third-party attacks.

scenario based on the informational nature of the page; there is no request for personal information.

Similarly, the Netflix settlement page does not make any overtures for personal information. Even more surprising is that the URL <http://www.netflix.com/>

settlement/ aligns well with the content. Subjects may have dismissed the page based on mistrust of the email stimulus, which they viewed prior to (several screenshots before) the Web stimulus. The Netflix page is notable for its brevity and unsophisticated layout. It is the only page

An Agenda for Action for Social Engineering Tactics

Our results forecast the following social engineering tactics: (Check All Tasks Completed):

- ____ 1. Construct messages with weak narratives (bordering on innocuous) but use strong design elements (graphics, small-print footer, endorsement logos) and identifying information to improve authenticity impressions.
- ____ 2. Use softer bait. Messages that do not encapsulate an imminent request for information, such as the Dell battery bait, rated highly in the test.
- ____ 3. Use plausibly unfamiliar administration pages; for example, the Dell Battery Return Program Web site

provides a service that is not typically seen, such as a login page (so visitor expectations are less concrete).

- ____ 4. Leverage high-profile news to produce messages with credible and strong narratives. Personalization will be less important in these cases.
- ____ 5. Align domain names with page content. Although subjects were turned off by semantic mismatches between domain names and content, they were insensitive to malformed links (<http://www-wellsfargo.com>).

without graphics or logos of any kind. There are no apparent links back to the primary Netflix page. With the exception of the blue underlined hyperlinks and gray margins, the page is black and white. We take from this rating that utilizing minimalist design is a poor strategy for unsolicited communications, even for important and serious matters such as law.

3. IMPLICATIONS FOR CRIMEWARE

The experiment focused on design features and their effect on people's ability to distinguish authentic email and Web sites from malicious forgeries. Although presented in the context of phishing, we do not measure how often subjects disclose passwords or other sensitive data; rather, we identify design principles that convey authenticity. Just as phishing bait promises resolution upon revealing information, social crimeware bait may promise resolution contingent upon installing browser extensions or accepting self-signed Java applets. Presenting a convincing false identity to the victim is essential in both contexts (see checklist, "An Agenda For Action For Social Engineering Tactics").

With respect to the final point, the "rock-phish" gang has proven that effective domain alignment can be achieved through deceptive subdomains,¹⁴ the control of which is delegated to the domain owner rather than the registrar. The following URL, from a social engineering attack in the wild, illustrates this tactic:

`www.paypal.com.cgi.bin.account.webscr.cmd.login.run.php.draciimasi.info/webscr.php?cmd = Login`

The registered domain is draciimasi.info, but the owners have prepended it with a deep subdomain. Since

subjects accepted the substitution of a dash for a dot in `www-wellsfargo.com`, they could easily accept a dot for a slash, as above. Moreover, the preponderance of subdirectory names such as `cgi`, `bin`, `webscr`, and the like further clouds the issue for the technically uninformed. This tactic may be particularly effective for download pages because they tend to be buried several directories deep; login pages, on the other hand, are frequently in the root of a domain.

Example: Vulnerability of Web-Based Update Mechanisms

Legitimate Web sites often make their services contingent upon changing settings, installing extensions, or accepting certificates. One important example is Microsoft's Windows Update Web site. It scans the client for installation detail through ActiveX extensions. When accessing the Web site through a professionally managed client at Indiana University, an update is not possible because the administrators have disabled the service. However, the Web site (see Figure e49.14) suggests workarounds involving settings changes.

None of these suggestions will enable remote update for this professionally managed computer, but subtle

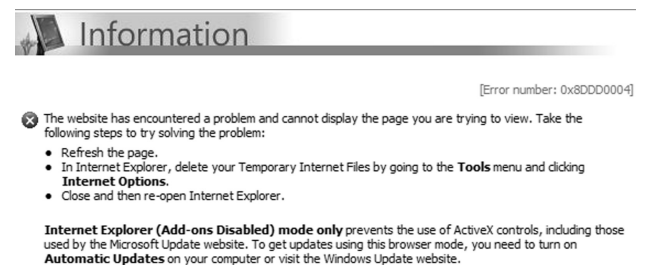


FIGURE e49.14 Suggested workarounds involving settings changes.

14. T. Moore and R. Clayton, "An empirical analysis of the current state of phishing attack and defense," The 2007 Workshop on the Economics of Information Security (WEIS 2007), 7–8 June 2007.

changes to the instructions could cause unsophisticated users to disengage important access controls. For example, the user could have been instructed to enter Addons Enabled mode. Subsequent installation of malicious addons will lead to a compromise. As long as the host's identity has been convincingly spoofed, users will be vulnerable to these kinds of attacks.

Example: The Unsubscribe Spam Attack

This attack leverages the first two tactics we discussed: weak narrative combined with strong design elements and softer bait. Some of the most highly rated email messages avoided hyperlinks in the main message text. The Chase account payment was completely devoid of hyperlinks and instead directed receivers to type `www.chase.com` into their address bar. Similarly, the PayPal message had no hyperlinks in its body, but it included a hyperlink in the footer to change preferences. The highly rated AT&T Universal card promotion also contains links in its small-print footer.

The attacker will send out promotional email that appears to come from the spoofed institution. The promotion would employ a weak narrative to shift user attention to a plethora of design features (graphical header, footer, small print, personalization, genuine but unlinked URLs in the body, and so on). The body will generate the perception of authenticity by referring the receivers to the phone number on back of her credit card or by requiring users to manually type in the promotional URL. Among the design features is a small-print footer with an unsubscribe hyperlink. This link will take users to a Web page that spreads crimeware simply by loading malicious JavaScript code, like a “drive-by pharming attack.”

The bait message gets users to click on the link indirectly: annoyance with the volume of unsolicited messages. No suspicion is aroused through directions to change settings; the malware spreads on load.

The Strong Narrative Attack

The strong narrative attack engages the receiver with a plausible story, often bundling actions to well-known news stories. The Chase phishing message that promotes ePIN to incoming Bank One customers is such an example; it leverages in the news of the Bank One acquisition. The Dell battery program stimuli gain most of their credibility from the story's media coverage. The message maintains this credibility by deferring the request for personal information; standard attacks request an “account login” or “settings update” in the message body. This battery exchange program could have been turned into a “patch-now” attack by claiming that a firmware or operating system fix would prevent overheating.

Though scams that exploit strong narratives and current events are not new (many fraud cases capitalized on the September 11, 2001, and Hurricane Katrina tragedies^{15,16}), our research suggests that they are less influenced by design features. This finding is supported by the persistence of the Nigerian code 419 advance fee scams.¹⁷ One widespread form of this attack entices victims with a story of the death of a foreign dignitary and the need to move large amounts of money (allegedly to protect it from corrupt enemies); they offer the victim a cut for moving the money. After drawing the victim into this illusion, the scammers request advance fees to enable the transfer of money. These messages break many design rules that promote trust: They use poor spelling and grammar, email messages are often plaintext, return addresses are essentially anonymous using free email accounts. Yet these scams still account for large amounts of Internet fraud, exceeding \$3 billion in losses according to some estimates.¹⁸

SUMMARY

This study tested the impact of several document features on user authenticity perceptions for both email messages and Web pages. The influence of these features was context dependent in email messages. We were surprised that this context was shaped more by a message's narrative strength, rather than its underlying authenticity. Third-party endorsements and glossy graphics proved to be effective authenticity stimulators when message content was short and unsurprising. The same document features failed to influence authenticity judgments in a significant way when applied to more involving messages. Most surprising was the huge increase in trust caused by a small-print footer in a message that already exhibited strong personalization with its greeting and presentation of a four-digit account number suffix.

The data suggest a link between narrative strength and susceptibility to trust-improving document features, but the experiment was not designed to test this hypothesis. Future work should characterize more precisely what kind

15. Fraud Section, Criminal Division, U.S. Department of Justice. Special report on possible fraud schemes, www.usdoj.gov/criminal/fraud/WTCPent-SpecRpt.htm, 27 September 2001, retrieved December 2006.

16. B. Krebs, “Katrina phishing scams begin,” *WashingtonPost.com*: Security Fix, 31 August 2005.

17. M. Zuckofi, “The perfect mark: How Massachusetts psychotherapist fell for a Nigerian e-mail scam,” *The New Yorker*, 15 May 2006, www.newyorker.com/fact/content/articles/060515fa_fact.

18. Ultrascan Advanced Global Investigations, “Advance fee fraud in 37 nations,” www.ultrascan.nl/html/aff_37_countries.html, 25 March 2006, retrieved December 2006.

of messages can benefit from these features and what kind of messages are resistant to their sway.

Since spoofed Web page content need not differ from the authentic pages, we focused three Web page tests on the effects of semantic alignment between address bar URLs and page content. The first showed a clear statistical preference for a simulated Web page whose domain name matched its content rather than the genuine page whose domain was only weakly aligned with the same content. The second test, which created better alignment with a bogus domain name by altering company logos, failed to register a statistically significant change in authenticity ratings. The third test compared an authentic page (and URL) with an authentic version of the same page content paired with a well-aligned but bogus URL; the results which favored the genuine URL were just shy of statistical significance. In conclusion, we find that URL can change authenticity ratings.

This experiment also verified that it is possible to overuse well-intended notices about security and fraud. We observed a statistically significant negative effect of genuine, but heavy-handed, fraud warnings. Another test showed a statistically significant improvement in authenticity perception when using SSL and a third-party endorsement logo on a fraudulent Web page showing a suspiciously formed, but semantically well-aligned, domain name.

The experiment simulated two sequences (one email and one Web page) that appeared to be third parties charged with handling embarrassing incidents for their corporate clients. Though separated by many variables, one turned out to be among the most trusted stimuli in the test, whereas the other rank among the lowest. The poorly ranked one, though authentic, broke all the rules: poor publicity, long and rambling message, use of third-party domain names, and no graphics. The highly ranked one (whose bogus email message was concocted by the authors) benefited from a widely publicized recall message. The story overrode the message's poor personalization, illegitimate URL, and relatively simple layout.

These factors offer a glimpse into what kinds of social engineering tactics may be deployed in the future. We describe an unsubscribed attack which contains an innocuous message, many authenticity stimulating document features, and an unsubscribe link that leads to a noninteractively infectious Web site. Our tests with third-party administration suggest that organizations in the process of correcting an embarrassing incident are highly vulnerable to social engineering attacks. Finally, our findings suggest some common pitfalls for legitimate Internet communications to avoid: overuse of fraud warnings, utilization of poorly aligned domain names, failure to use HTTPS for rendering login pages, and long or rambling email messages.

Finally, let's move on to the real interactive part of this Chapter: review questions/exercises, hands-on projects, case projects and optional team case project. The answers and/or solutions by chapter can be found in the Online Instructor's Solutions Manual.

CHAPTER REVIEW QUESTIONS/EXERCISES

True/False

1. True or False? Identity theft is commonly defined as unwanted appropriation of access credentials that allows creation and access of accounts and that allows the aggressor to pose as the victim.
2. True or False? Phishing is a form of Internet fraud that spoofs emails and Web pages to trick people into disclosing sensitive information.
3. True or False? Web sites, particularly the login and information collection pages associated with phishing scams, do present a story the way email messages do.
4. True or False? One way phishers align page content with URLs is by choosing an apt domain name; the other way is to change the page content.
5. True or False? Legitimate Web sites often make their services contingent upon changing settings, installing extensions, or accepting certificates.

Multiple Choice

1. An attacker will send out _____ that appears to come from the spoofed institution.
 - A. Privacy-enhancing technology
 - B. Location technology
 - C. Promotional email
 - D. Executable policies
 - E. Data controller
2. What message gets users to click on the link indirectly: annoyance with the volume of unsolicited messages?
 - A. Policy enforcement
 - B. Location technology
 - C. Valid
 - D. Environmental data
 - E. Bait
3. What attack engages the receiver with a plausible story, often bundling actions to well-known news stories?
 - A. Data minimization
 - B. XACML
 - C. Privacy risks
 - D. Strong narrative
 - E. Security

4. What is a type of identity theft that is perpetrated on the Internet and that typically relies on social engineering to obtain the access credentials of the victim?
 - A. Privacy metrics
 - B. Greedy strategy
 - C. Sensitive information
 - D. Phishing
 - E. Taps
5. What is often defined as economically motivated malware?
 - A. Irrelevant
 - B. Sensor nodes
 - C. Crimeware
 - D. Disclose-to
 - E. Server policy

EXERCISE

Problem

You should never respond or reply to an e-mail, phone call, or text message that?

Hands-on Projects

Project

How can one enhance the security of their transactions and experiences on the Internet and better control any related risks?

Case Projects

Problem

Occurrences of identity theft continue to increase. Please explain identity theft, and how you might avoid becoming a victim?

Optional Team Case Project

Problem

What should one do if he or she have become a victim of identity theft?