Broken Windows, Bad Passwords: Influencing Secure User Behavior via Website Design

Abstract

The broken windows theory of crime deterrence suggests that features in a community such as broken windows, graffiti, and petty crime send signals that criminal behavior is accepted as the norm. We extend the broken windows theory to information security by suggesting that poor web site design sends signals that insecure behavior is the norm. Extending the theory of planned behavior, we hypothesize that visual appeal and trust influence subjective norms such that users on a low quality web site will exhibit reduced intentions and behaviors of security. In a laboratory experiment we manipulate web site quality while participants create user accounts. Password entropy and self-reported password strength are compared for accounts created on the high and low quality sites. Our findings suggest that organizations may be able to promote secure behavior by creating systems that are high in visual appeal and trustworthiness.

Keywords

Security, Passwords, Theory of Planned Behavior

Introduction

Throughout the 1970’s and 1980’s, violent crime in New York City was a major and growing problem, peaking in 1990 with 1,181 incidents of violent crime per 100,000 inhabitants – over 61% higher than the national average (The Disaster Center 2013). To combat this trend, criminologist George Kelling was hired by the New York Transit Authority in 1985 to reduce violent crime in New York City’s subway system. Throughout the late 1980’s and early 1990’s, Kelling’s approach was to focus on minor crimes such as graffiti, fare-beating, and aggressive panhandling rather than directly addressing more serious or violent crimes (Gladwell 2006 pp. 140–151). This approach, known as the broken windows theory of crime deterrence, works on the premise that features such as broken windows, graffiti and petty crime send signals to individuals that criminal behavior is acceptable. By rigorously enforcing laws related to petty crimes, the signal is sent that criminal behavior is not the norm and criminals, whether opportunistic or professional, will be less likely to engage in undesirable behavior (Harcourt 1998; Wilson and Kelling 1982). We suggest that the broken windows theory applies not only to criminal behavior, but to any situation in which the environment may influence what behavior individuals deem acceptable.

In this paper, we explore the application of the broken windows theory of crime deterrence to information systems security to answer the following research question: Does perceived web site quality influence intentions and practice of secure behavior?
It has been suggested that elements of a web site’s design send signals to users regarding the quality, credibility, and care of the site. For example, broken links on a web site may be interpreted by users as a signal that “the site has been neglected or that the site was not carefully created in the first place” (Fogg 2003 p. 723). We hypothesize that these signals influence users’ subjective norms of acceptable behavior for interaction with the web site, thereby influencing their intentions of secure behavior. Web sites that send signals of low quality and credibility – analogous to broken windows in an abandoned warehouse – prime users to behave less securely. We propose that by carefully crafting the signals a web site is sending by using good design practices, users will behave more securely than when elements of good design are neglected. In this way, organizations may be able to encourage secure behavior with minimal cost and impact to users.

Users are often cited as the weakest link in information systems security (Boss et al. 2009). Persuading users to behave securely is an important issue for information systems and significant research has been conducted to identify effective ways to influence user security behavior. Security policies (Bulgurcu et al. 2010; D’Arcy et al. 2009), training (Charoen et al. 2008; Ives et al. 2004), and reminders (Jenkins and Durcikova 2013; Jenkins et al. 2013) are three methods that are commonly used for influencing secure behavior. While each of these methods positively influences secure user behavior, they also require end users to expend additional effort to carry out the secure behavior. For policies to work, users must be familiar with the policies, believe in the efficacy of the policies and consciously adhere to them (Bulgurcu et al. 2010). Security training may be time consuming and expensive for a company to implement, and users must take time from their schedule to participate in the training, often on multiple occasions. Likewise, reminders may be intrusive and may have a negative impact on the user experience or may simply be ignored by users. An ideal aide to influence secure behavior would subconsciously influence behavior and require no additional action from the user. While there are many aspects of secure user behavior, in this paper we focus on one specific aspect – the creation of strong passwords.

Building on the theory of planned behavior (TPB; Ajzen 1991), we suggest that visual and trust elements of a web site will influence the subjective norms of secure behavior for users of the site. TPB posits that changes in subjective norms, along with attitudes and perceived behavioral control, lead to a change in intentions and ultimately a change in behavior. The balance of this paper provides a review of literature relevant to creating websites high in visual appeal and trust, describes a laboratory experiment used to test our hypotheses, and discusses our results and the implications of our findings.

**Literature Review**

Significant research in the area of web site design provides a basis for this work. In this section, we discuss some of the factors that have been found to influence visual appeal and trust in web sites.

Prominence-interpretation theory (Fogg 2003) suggests that people assess credibility online by noticing something (prominence) and making a judgment about it (interpretation). If one or the other does not happen, no assessment of credibility is made. Therefore, many of the elements proposed in the literature on web site quality are naturally interrelated and dependent on one another. For example, Alsudani and Casey (2009) find that dominant images (prominence) leads to higher ratings of credibility, however, as Hynes (2009), Lowery et al. (2014) and Belanger (2002) point out, the prominent symbols must have a positive interpretation to increase credibility. Table 1 presents a brief overview of some of the features that impact trust or visual appeal of web sites.

Just as broken windows and graffiti send signals to people regarding the norms for behavior in that environment, we suggest that website visual appeal and trust send signals to users regarding website behavioral norms. We expect that sites with high visual appeal and trust will trigger norms related to security best practices.
Influencing Secure User Behavior via Website Design

<table>
<thead>
<tr>
<th>Citation</th>
<th>Features</th>
<th>Influence on Trust or Visual Appeal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbasi et al. 2010</td>
<td>Web page text, URL attributes and links</td>
<td>Fraudulent sites contain incorrect spelling/grammar, have more complex URLs and tend to have fewer working links</td>
</tr>
<tr>
<td>Al Sudani and Casey 2009</td>
<td>Balance, harmony, contrast dominance and image dominance</td>
<td>Contrast dominance and image dominance lead to higher ratings of credibility</td>
</tr>
<tr>
<td>Bauerly and Liu 2008</td>
<td>Symmetry and number of elements</td>
<td>High complexity and low symmetry of elements leads to lower aesthetic appeal</td>
</tr>
<tr>
<td>Belanger et al. 2002</td>
<td>Symbols of security</td>
<td>Security symbols increase trustworthiness</td>
</tr>
<tr>
<td>Bonnardel et al. 2011</td>
<td>Color</td>
<td>Blue and Orange sites were considered more “usual” and received higher user appreciation than grey sites</td>
</tr>
<tr>
<td>Cyr 2008</td>
<td>Navigation, visual and information design</td>
<td>Navigation, visual and information design all positively related to trust and satisfaction</td>
</tr>
<tr>
<td>Cyr et al. 2010</td>
<td>Color</td>
<td>Increased color appeal results in greater trust and greater satisfaction</td>
</tr>
<tr>
<td>Fogg 2003</td>
<td>Prominence and interpretation</td>
<td>Highly prominent elements with positive interpretation lead to higher credibility</td>
</tr>
<tr>
<td>Fogg et al. 2003</td>
<td>18 elements including: design look, structure, information focus, name recognition, tone of writing and functionality</td>
<td>Sites with good design, organization, functionality and information rated as more credible. Name recognition and formal tone lead to higher ratings of credibility</td>
</tr>
<tr>
<td>Hynes 2009</td>
<td>Color and Logos</td>
<td>Globe logo and blue/brown associated with protection, stability and reliability</td>
</tr>
<tr>
<td>Lowry et al. 2014</td>
<td>Logo and website design</td>
<td>Consistency, stability, reassuring colors and good design increase trustworthiness</td>
</tr>
<tr>
<td>Robins and Holmes 2008</td>
<td>Aesthetic qualities</td>
<td>High aesthetics lead to higher ratings of credibility</td>
</tr>
</tbody>
</table>

Table 1: Site Features and impact on user perception

Research Model

We use TPB to explore why individuals carry out security behaviors. TPB suggests that three constructs influence behavioral intention: behavioral attitude, subjective norms, and perceived behavioral control (Ajzen 1991). Each of the constructs in TPB has features that are stable over time and features than may be influenced by situational factors. For example, one’s overall feeling of perceived behavioral control for secure behavior may be high (high self-efficacy), however, they may feel a low level of behavioral control in certain situations (low response-efficacy). In the current study, we suggest that web site quality, operationalized here as visual appeal and trust, will not affect attitudes nor perceived behavioral control, but will have a positive relationship with subjective norms (Figure 1).
We therefore propose the following hypotheses:

- **H1**: Trust will be positively influenced by visual appeal of a website
- **H2**: Higher (a) visual appeal and (b) trust of a website will positively influence subjective norms of security
- **H3**: Subjective norms of security will be positively related to security intentions
- **H4**: Attitudes regarding security will be positively related to security intentions
- **H5**: Perceived behavioral control of security will be positively related to security intentions
- **H6**: Intentions of secure behavior will be positively related to secure behavior as measured by password entropy

**Methodology**

A laboratory experiment was conducted at a large public university in the Southwestern United States. 189 students (100 male) from a junior level MIS class participated in exchange for class credit. Sixteen participants were removed from the sample for providing invalid survey responses, leaving 173 participants for the analysis. Upon arriving at the lab, participants were instructed to create an account on an experiment management website similar to other systems used in the college for students to register for experiments. The stated purpose of creating an account was so participants could log in at a future date (in 5-7 days) to complete the second part of the experiment. The experiment following the account creation, including the second part, are unrelated to this study and are discussed in no further detail here. Since participants knew they would have to log in again in the future to complete the experiment, this incentivized them to use a password they would be able to remember rather than using a “throw away” password for the account. To discourage password reuse, a fear appeals message was used to instruct participants to not reuse a password they have used on other systems (Herath and Rao 2009; Jenkins et al. 2013; Johnston and Warkentin 2010) (Figure 2).
Due to the recent data breaches at several large retailers

**Please create a unique password for this website**

This will help to protect your online accounts as well as our system

**Figure 2: Unique password solicitation**

While creating their account, participants were exposed to either a high quality site (Figure 3) or a low quality site (Figure 4). Several attributes suggested by the research in Table 1 were used to manipulate website quality as described in Table 2.

<table>
<thead>
<tr>
<th>Feature</th>
<th>High Quality Manipulation</th>
<th>Low Quality Manipulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbols of security (Belanger et al. 2002)</td>
<td>RSA and Verisign logos present, the word “secure” in the name of the system</td>
<td>No symbols of security</td>
</tr>
<tr>
<td>Quality of images (Alsudani and Casey 2009)</td>
<td>Visually appealing, dominant, high quality images</td>
<td>Antiquated web site images with warped aspect ratio</td>
</tr>
<tr>
<td>Symmetry of page (Bauerly and Liu 2008)</td>
<td>Appealing symmetrical layout</td>
<td>Amateurish asymmetric layout</td>
</tr>
<tr>
<td>Links (Abbasi et al. 2010; Lee and Kozar 2006)</td>
<td>Working links for help, contact and about us</td>
<td>Broken links for help, contact and about us, broken image link</td>
</tr>
<tr>
<td>Grammar and spelling (Abbasi et al. 2010)</td>
<td>Proper use of grammar and no spelling mistakes</td>
<td>Incorrect grammar and misspelled words</td>
</tr>
<tr>
<td>Color (Bonnardel et al. 2011; Bottomley and Doyle 2006; Cyr et al. 2010; Hynes 2009)</td>
<td>Use of bold, contrasting colors (red, blue and white)</td>
<td>Use of unappealing, low contrast colors (yellow, grey and white)</td>
</tr>
<tr>
<td>Logos (Hynes 2009; Lowry et al. 2014)</td>
<td>University logo a exudes an air of quality and research</td>
<td>Logo suggests an incomplete web site</td>
</tr>
<tr>
<td>Website URL (Abbasi et al. 2010)</td>
<td>Short and simple URL hosted on a university server</td>
<td>Long, complex URL including an IP address and cryptic words</td>
</tr>
</tbody>
</table>

**Table 2: Manipulations of site quality**
Password entropy was used to operationalize secure behavior. Entropy is defined as the amount of randomness a secret contains based on the length and potential character set and is usually stated in bits (Burr et al. 2011). While bits of entropy is a statistical measure of randomness formally stated for random strings as $\log_2(b^n)$ where $b$ is the size of the character set and $n$ is the length of the string, calculating entropy for passwords created by humans is significantly more complex as people typically generate passwords that roughly conform to rules of word composition (Burr et al. 2011). Therefore, the effective entropy of human generated passwords is significantly lower than for random strings (Shannon 1948). For these reasons, calculations of entropy for user-generated passwords are approximate at best and may follow a wide range of criteria. To calculate the effective password entropy of the participant generated passwords, an open source JavaScript software called passchk was used to estimate the entropy of each password (Akins 2014).

During the account creation process, the participant’s password was encrypted with SHA-256 encryption and associated with the user account to enable them to log in again in the future. A clear text version of the password was temporarily stored in a separate database in order to calculate the password entropy,
however, the clear text password was not directly associated with the user. Only the password entropy value was associated with the participant. At no time were clear text passwords associated with any personally identifiable information of the participants and all passwords were deleted from the system as soon as practical.

**Results**

Survey items for subjective norms, attitudes, perceived behavioral self-efficacy, and intentions were adapted from Anderson and Agarwal (2010). Items for website visual appeal and website trust were taken from Loiacono et al. (2002). An exploratory factor analysis was used to verify the measurement model. Structural equation modelling (SEM) was used to test the research hypotheses. R version 3.0.2 and the lavaan package were used to conduct the analysis (Rosseel 2012).

We assessed the constructs for convergent validity, reliability, and discriminant validity. One perceived behavioral control item was dropped because it loaded significantly poorer than the remaining four items. All of the remaining items loaded above the recommended 0.60 level (Bagozzi and Yi 1988). See Table 3 for the detailed loadings and reliability. Scales are typically regarded as reliable if the Cronbach’s alpha is above 0.70 and the AVE is above 0.50 (Bagozzi and Yi 1988; Straub et al. 2004). All latent factors met the criteria for reliability. Discriminant validity is demonstrated in part when items load on a single construct and no other constructs. No item loads on an unrelated factor more than 0.20. In addition, the square root of AVE of the constructs should be higher than the inter-construct correlations. All latent constructs satisfy these criteria for discriminant validity. Having established the validity and reliability of our model, we proceeded with the analysis.

<table>
<thead>
<tr>
<th>Constructs/items</th>
<th>Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Subjective Norm</td>
<td></td>
</tr>
<tr>
<td>Alpha=0.82</td>
<td>SN1</td>
</tr>
<tr>
<td>AVE=0.65</td>
<td>SN2</td>
</tr>
<tr>
<td></td>
<td>SN3</td>
</tr>
<tr>
<td>Perceived Behavioral Control</td>
<td></td>
</tr>
<tr>
<td>Alpha=0.82</td>
<td>PBC1</td>
</tr>
<tr>
<td>AVE=0.69</td>
<td>PBC2</td>
</tr>
<tr>
<td></td>
<td>PBC3</td>
</tr>
<tr>
<td></td>
<td>PBC4</td>
</tr>
<tr>
<td>Attitude</td>
<td></td>
</tr>
<tr>
<td>Alpha=0.79</td>
<td>ATT1</td>
</tr>
<tr>
<td>AVE=0.67</td>
<td>ATT2</td>
</tr>
<tr>
<td></td>
<td>ATT3</td>
</tr>
<tr>
<td>Intention</td>
<td></td>
</tr>
<tr>
<td>Alpha=0.81</td>
<td>INT1</td>
</tr>
<tr>
<td>AVE=0.55</td>
<td>INT2</td>
</tr>
<tr>
<td></td>
<td>INT3</td>
</tr>
<tr>
<td>Website Trust</td>
<td></td>
</tr>
<tr>
<td>Alpha=0.84</td>
<td>WTR1</td>
</tr>
<tr>
<td>AVE=0.65</td>
<td>WTR2</td>
</tr>
<tr>
<td></td>
<td>WTR3</td>
</tr>
<tr>
<td>Website Visual Appeal</td>
<td></td>
</tr>
<tr>
<td>Alpha=0.96</td>
<td>WVA1</td>
</tr>
<tr>
<td>AVE=0.93</td>
<td>WVA2</td>
</tr>
<tr>
<td></td>
<td>WVA3</td>
</tr>
</tbody>
</table>

Table 3: Latent factor loadings, Cronbach’s Alpha and AVE of measurement instrument
Hypothesis Testing Results

A manipulation check was performed to validate that the experimental treatment had the intended effect. As expected, participants reported on average significantly greater website visual appeal in the high quality condition ($M=4.94$, $SE=0.11$) than in the low quality condition ($M=3.32$, $SE=0.15$), $t(163.07)=-8.80$, $p<0.001$. Likewise, participants reported higher trust in the high quality site ($M=5.36$, $SE=0.11$), than in the low quality site ($M=4.61$, $SE=0.13$), $t(163.42)=-4.38$, $p<0.001$.

Initial model testing showed that some improvements to the model were necessary. Modification indices suggested a path from website trust to security intentions, a path from website visual appeal to intention, and residual covariances between select items for perceived behavioral control, attitudes, website visual appearance, and intentions. Four indices were used to estimate goodness of fit: chi-square ($\chi^2$), relative chi-square ($\chi^2/df$), comparative fit index (CFI), normed fit index (NFI), and root mean square error of approximation (RMSEA) (Holbert and Stephenson 2002). To demonstrate good fit, the $\chi^2$ statistic should be nonsignificant at the 0.05 level, CFI and NFI should be greater than or equal to 0.95, and RMSEA should be .05 or less (Hooper et al. 2008). It has been noted that NFI tends to report poorer fit even for good fitting models for sample sizes less than 200, which is the case in this study (Bentler 1990). The final model achieved a good fit ($\chi^2=181.18$, $df=152$, $p>0.05$, $\chi^2/df=1.19$, $CFI=0.98$, $NFI=0.91$, $RMSEA=0.03$).

As illustrated in Figure 5, our model explains 27 percent of the variance in website trust, 15 percent of the variance in subjective norms, 49 percent of the variance in intention to create secure passwords and 0 percent of the variance in actual behavior as measured by password entropy.

As proposed in H1, a significant positive relationship was found between visual appeal and trust. Visual appeal was found to have a significant positive effect on subjective norms, thus supporting H2a. Contrary to our hypothesis, however, trust did not have a statistically significant relationship with subjective norms, but rather influenced intentions directly. Therefore, H2b was not supported. As expected from the TPB model, subjective norms, attitudes and perceived behavioral control were found to have statistically significant relationships with intentions, supporting H3, H4 and H5. Finally, the relationship between intentions and observed behavior (H6) was not found to be significant. These results are summarized in Table 5. Though not originally hypothesized, the path from website trust to intention was significant ($\beta=0.42$, $SE=0.113$, $p<0.001$). The added path from website visual appeal to intention was not significant ($\beta=-0.221$, $SE=0.122$, $p>0.05$).

A follow-up survey of participants was conducted to investigate a) whether the participants reused a password they have used on other web sites and b) their perception of the strength of their password. We suggest that perception of strength may be a more appropriate measure of behavior than password entropy, since participants may have thought they were creating a strong password but did not have the skills to do so (Furnell 2007). In total, 172 participants responded to the follow up survey. Of those that
responded only 39 (18 in the high quality condition) indicated that they created a unique password for the experimental web site. Of those that created a unique password for the site, participants in the high quality condition rated their passwords using a ten point scale as significantly stronger ($M=5.44$, $SD=2.19$) than those in the low quality condition ($M=3.83$, $SD=1.84$), ($t(33.37)=2.40$, $p=0.011$). This effect suggests that a) participants that reused a password may have simply reused a password out of habit with no regard to their security intentions and b) for many participants their perceived strength of the password they created was not reflected in the actual strength of their password.

![Figure 5: Model results](image)

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Supported?</th>
<th>Path Coefficient</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: Website Visual Appeal $\rightarrow$ Website Trust</td>
<td>Yes</td>
<td>0.614</td>
<td>0.101</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>H2a: Website Visual Appeal $\rightarrow$ Subjective Norms</td>
<td>Yes</td>
<td>0.433</td>
<td>0.106</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>H2b: Website Trust $\rightarrow$ Subjective Norms</td>
<td>No</td>
<td>-0.034</td>
<td>0.088</td>
<td>0.697</td>
</tr>
<tr>
<td>H3: Subjective Norms $\rightarrow$ Intention</td>
<td>Yes</td>
<td>0.285</td>
<td>0.128</td>
<td>0.026</td>
</tr>
<tr>
<td>H4: Attitude Toward Behavior $\rightarrow$ Intention</td>
<td>Yes</td>
<td>0.566</td>
<td>0.148</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>H5: Perceived Behavioral Control $\rightarrow$ Intention</td>
<td>Yes</td>
<td>0.207</td>
<td>0.097</td>
<td>0.032</td>
</tr>
<tr>
<td>H6: Intention $\rightarrow$ Behavior</td>
<td>No</td>
<td>-0.029</td>
<td>0.059</td>
<td>0.617</td>
</tr>
</tbody>
</table>

Table 5: Summary of hypothesis testing results
Discussion

The data support the overarching idea that signals of quality presented by a website influence how users intend to interact with the site, however, the precise operationalization of these changes deviated slightly from the hypothesized relationships. While we predicted both visual appeal and trust would influence subjective norms, we found that only visual appeal directly influence subjective norms, while trust directly influenced intentions.

Although intention did not have a significant effect on password entropy, we suggest there are two main reason for this. First, the majority of participants neglected the instructions to create a unique password for this website, thus diminishing the impact of the website quality manipulation. Password reuse is a major problem in information systems security, as it can lead to a “domino effect” – i.e. if a user’s password is compromised on a single web site with poor security, that password could be used to compromise their account across multiple systems (Ives et al. 2004). To combat password reuse, administrators of websites might consider implementing means to identify and deter password reuse (see Jenkins et al. 2013) in addition to manipulations such as the ones presented in this paper to encourage users to create strong passwords. Second, using password entropy as an operationalization of behavior may not have been representative of the behavior participants believed they were engaging in. Participants that intended to engage in secure behavior may have genuinely thought they were constructing a secure password (engaging in secure behavior), despite their password having low entropy. Training in how to create secure passwords might be used to improve this disparity, and future research should consider exploring the relationship between how secure users believe their passwords are and how secure they actually are.

Conclusion

Encouraging secure behavior in users is an important, but difficult, task. While many methods such as training, reminders and security policies are currently used with varying levels of success, traditional methods of influencing secure behavior places additional load on the user. Furthermore, since behaving securely is rarely, if ever, the primary goal of users when interacting with a computer system or web site, secure behavior is likely to be one of the first goals neglected by users when exposed to additional stressors. This study does not suggest that priming secure user behavior is a substitute for policies, training, or reminders, but rather it should be considered a supplement to these methods to increase the protection afforded by passwords. This study demonstrates that subtle cognitive priming via website quality may be used to generate more secure user behavior without placing additional load or requirements on the user.

While creating quality web sites is already a focus of many practitioners, this study helps to illustrate the importance of good design and may provide additional incentive for companies to design high quality web sites. By exerting a modest amount of extra effort into designing sites to exude symbols of trust and visual appeal, firms may be able to increase secure behavior with minimal cost and impact to users.

References


Influencing Secure User Behavior via Website Design


