Geo-Phisher: The Design and Evaluation of Information Visualizations about Internet Phishing Trends

Leah Zhang-Kennedy*, Elias Fares†, Sonia Chiasson* and Robert Biddle*
*School of Computer Science, †School of Information Technology
Carleton University, Ottawa, Canada
leah.zhang@carleton.ca, elias.fares@cmail.carleton.ca, chiasson@scs.carleton.ca, robert.biddle@carleton.ca

***Abstract*—We designed an information visualization about phishing trends and phishing prevention for the general public to examine the effects of interactivity on information finding, user perceptions and security behaviour intentions, and effectiveness of learning. In an user study (*N* = 30) with two experimental conditions (*HI* – high interactivity, and *LO* – low interactivity control condition), the results show that the *HI* interactivity condition supported more accurate information finding, resulted in greater perceived interactivity and perceived knowledge than the *LO* interactivity condition, but did not affect attitudes toward the visualization and security behaviour intentions for proactive awareness. Furthermore, the *HI* interactivity condition led to greater learning effects and a deeper understanding towards phishing prevention than the control condition.**

I. Introduction

Phishing is a significant type of Internet crime that tricks users into revealing their personal and financial information. To combat phishing, browser manufacturers, software vendors, and organizations have compiled repositories of phishing URLs (blacklists). For example, during the third quarter of 2014, the Anti-Phishing Working Group (APWG) received approximately 50,000 unique phishing e-mail reports from consumers monthly, targeting more than 500 unique brands [2]. Phishing blacklists enable the analysis of reported phishing attacks to be shared among anti-phishing communities to gain awareness of evolving phishing trends, and to assist anti-phishing communities and organizations to take down, block, or warn users when they attempt to visit these URLs. The largest blacklists are operated by major browser vendors and by organizations like Phishtank and the APWG. The databases contain manually verified phishing URLs reported by users.

Data about phishing trends is widely available on the Internet, but it is not easily comprehensible to the general public. We therefore propose an information visualization tool called Geo-Phisher (Figure 2) to make the information more accessible to end-users. The application features a scatterplot map interface that plots the temporal and geographical information of phishing URLs. Applied to blacklist data from the APWG [4], the prototype reveals several interesting patterns in phishing URLs hosting locations and distributions of the top phished brands across the globe.

978-1-5090-2922-8/16/$31.00 ©2016 IEEE

The goal of the visualization tool is to spark curiosity in the data, and to get the general public acquainted with the problems of phishing on a global scale to raise awareness about the issue. Additionally, we aim to educate the public from falling for phishing attacks by providing phishing prevention advice. The Geo-Phisher visualization provides context for phishing crimes, which may help to bring the public's attention to the issue and support users in making sense of how phishing works. Since many users lack motivation to learn about security related information [51], it is beneficial to use visual methods of communication to motivate users to pay attention [29], [56], [57].

The work is inspired by current trends in the information visualization (infovis) community to present and share publicly available data to raise awareness about particular issues [12]. Information visualization is the process of presenting data, information, and knowledge into a visual form that works with humans' natural visual capabilities [16]. Recently, various infovis tools were developed in a variety of domains to disseminate data and making it comprehensible for the general public. Some examples include visualizations developed for health trends detection (e.g., [30]), medical data (e.g., [8]), environmental sustainability (e.g., [48]), decision making (e.g., [25]), and awareness about local issues (e.g., [9]). Infovis is regularly praised for its ability to tell stories within the data. Segel and Heer [41] suggest that graphical elements help to direct the narrative flow while the discovery of the story is often achieved through interactive exploration.

The benefits of visualizing information include the discovering hidden insights, patterns and trends, improved efficiency, reduced cognitive loads, and increased interactivity [30]. Patterns can be perceived readily through the representation itself, or further facilitated by interactive mechanisms acting on the underlying representation to make different features of the data salient. Since interaction with the data allows multiple representations to be linked cognitively through interactivity, it is suggested that interactivity could be used to overcome the limitations of static images [11].

The paper presents the design and evaluation of Geo-Phisher, an interactive poster visualization to raise awareness about phishing. First, we give an overview of information visualization (infovis). We build on the foundation of persua-

Zhang-Kennedy, L., Fares, E., Chiasson, S., & Biddle, R. (2016). Geo-Phisher: The design and evaluation of information visualizations about internet phishing trends. 2016 APWG Symposium on Electronic Crime Research (ECrime), 1–12. https://doi.org/10.1109/ECRIME.2016.7487941

sion in visualization from communications literature, where interactivity is found to affect the persuasiveness of the message tested with websites (e.g. [32]) and advertisements (e.g., [46]). Second, we explain the design process of the interactive Geo-Phisher visualization. Two experimental conditions were used in the evaluation: the interactive interface of the Geo-Phisher visualization (*HI* – high interactivity condition), and a static representation of the interface as the control condition (*LO* – low interactivity condition). We outline the differences between the two interfaces both in the design (Section III-B) and methodology (Section IV-C). Third, we discuss the results for six aspects of the visualization that we evaluated: effects of interactivity on information finding, perceived interactivity, perceived knowledge, attitudes toward the visualization, security behaviour intentions for proactive awareness, and effectiveness of learning. The results show that the *HI* interactivity condition supported more accurate information finding, resulted in greater perceived interactivity and perceived knowledge than the *LO* interactivity condition, but it did not affect attitudes toward the visualization and security behaviour intentions for proactive awareness. Furthermore, the *HI* interactivity condition led to greater learning effects and a deeper understanding towards phishing prevention than the control condition.

II. Background and Related Work

The most widely used definition of information visualization is *the use of computer-supported, interactive visual representations of data to amplify cognition* [6]. Various interaction techniques are used in infovis (see Yi et al. [54] for taxonomies relevant to interaction techniques), but central to user interaction in infovis is the benefit of interactivity to augment the human cognitive processes for thinking and analysis.

Recently, the field of traditional information visualization has expanded to include sub-domains of artistic [49], narrative [41], causal [36], ambient [42], ecological [48], and social visualization [19]. Within these sub-domains, a common aim is to engage people around a wide range of social issues to raise awareness and outline visions for change [12]. Zambrano and Engelhardt [55] dubbed information visualization used to raise public awareness "Diagrams for the Masses" and divided them into three categories: *View*, *Interact & Explore*, and *Create & Share*. In the *View* category, the public passively views the data and the designer chooses the type of visualization. The *Interact & Explore* category, in contrast, enables the user to play an active role in raising his or her own awareness about the topics from a larger dataset provided by the designer through interaction and exploration. In the *Create & Share* category, the user freely chooses data from any source and views it through a selection of visualization types provided by the designer. In this paper, the focus is on the first two categories, *View* and *Interact & Explore*.

Information is moving into the public space. There is transformation of deploying large interactive displays in urban environments, malls, transportation stations, and stadiums in place of traditional posters [31]. These interactive posters enable new forms of multimedia presentation and user experience. Public displays have the potential to attract user attention without any intention for information and interaction [31]. This may be beneficial for disseminating security-based information because users are typically uninterested in learning about security [51], and are unlikely to seek out the information on their own.

A. Interactivity

Interactivity has various definitions in literature. Early research of interactivity focused on user-machine interaction with an emphasis on human-computer interaction [24]. Communication networks and the Internet enabled user-user interaction, where interactive communication in a computer-mediated environment resembles interpersonal communication [24]. The third perspective is user-message interaction, which is defined as the ability of the user to control and modify messages [43]. In 2002, Liu and Shrum [24] consolidated the three aforementioned aspects of interaction into a three-dimensional construct: 1) active-control – the voluntary and instrumental action that directly influences the controller's experience; 2) two-way communication – the ability for reciprocal communication, and 3) synchronicity – the degree to which users' input into a communication and the response they receive from the communication is simultaneous. Sundar et al. [45] found that the interface is considered more interactive when messages are contingent on previous messages, one after another. It is suggested that this contingency view of interactivity helps to support the three-dimensional construct of interactivity previously described by Liu and Shrum [24]. Based on these definitions, we posit the following hypothesis:

> **H1:** The interactive representation of the phishing visualization will be perceived as having more interactivity (active control, two-way communication, and synchronicity) than the control condition.

B. Amplify Cognition

Findings from information theory and psychological studies provide evidence that visualization provides perceptual support of the human cognitive system. Vision is the most efficient sense for transmitting information to the brain [50]. At the low-level, some information can be processed pre-attentively. In pre-attentive theory, it is believed that the body can automatically process some sensory information very rapidly and accurately before the conscious mind starts to pay attention [50]. Some examples of visual features that can be detected pre-attentively are hue, intensity, orientation, size, and motion [18].

In a study of website interactivity [32], highly interactive websites led to greater cognitive absorption than static websites. Reeves and Nass [39] suggested that interactivity may enhance cognitive absorption due to "perceptual bandwidth", the notion that users can interact with the interface via multiple sensory channels, such as the user's motor response of selecting an visual object of interest while perceptually coding the

visual changes on the visualization and cognitively processing them. These notions form our second hypothesis.

> **H2:** The interactive representation of the phishing visualization will lead to greater learning about phishing than the control condition.

C. Interactivity and Persuasion

Modern theories of persuasion have evolved to consider the persuasive effects of communication. Among them is the well-known Elaboration Likelihood Model [34] that describes two distinct routes for persuasion-based decision making. Users take the cognitively intensive "central route" characterized by careful, logical, and conscious thinking about the communication when they are motivated to process the message. In computer security, however, end-users are generally unmotivated to focus their attention on security-specific information [51], and security communication tends to be processed through the "peripheral route". The elaboration likelihood through this route is determined by users' perception of the persuasiveness of the message influenced by surface attractiveness, perceived credibility, and the production quality of the message. In security education, surface attractiveness of the message through the "peripheral route" could provide a valuable access point to direct unmotivated users into a temporary state where they are more susceptible to persuasion [58].

The effects of interactivity on persuasion are studied in communications literature. For example, Oh and Sundar [32] found that highly interactive websites led to more positive assessment of the interface and more positive attitudes among those with low involvement in the message topic. On political websites, a higher number of interactive features is correlated to positive attitudes toward the political candidate featured on the site [45]. Interactivity in advertisements is positively associated with ad and product attitudes [46]. Others showed that interactivity led to more positive attitudes toward the portal [21], increased credibility [14] and user involvement [5]. Based on these effects that interactivity have on persuasion, we form our third and fourth hypotheses:

> **H3:** The interactive representation of the phishing visualization will lead to more positive attitudes toward the phishing visualization than the control condition.
>
> **H4:** The interactive representation of the phishing visualization will lead to more favourable attitudes toward security behaviour intentions for proactive awareness than the control condition.

D. Phishing Visualization

The academic work on phishing is diverse, ranging from work on phishing detection, phishing indicators, and phishing education. We focus on work that uses visualizations to communicate about phishing threats.

In helping end-users to make security decisions, Stoll et al. [44] proposed a security user interface called Sesame that provided users with visualized system level information through a spatial desktop metaphor. The tool facilitated users' comprehension in making security related decisions like phishing. In one scenario, the tool enabled the user to visually detect a spoofed banking site by seeing which process is connected to their web browser window. This enabled the user to notice that a remote computer connected to the process is located in an unfamiliar geographic location and the owner's name appears to be unrelated to the user's bank.

In phishing indicators, visual cues are commonly used to warn users about phishing sites. Yee et al. [53] proposed a tool called Passpet that creates user-assigned pet names as site labels and an animal icon as a visual indicator within the browser to show that users are on a previously trusted website. Other works in phishing indicators rely on visual changes in the browser chrome to warn users about phishing websites. Ye et al. [52] proposed the Synchronized Random Dynamic Boundaries system that modified the browser chrome to blink at a random rate. A trusted path is established if the blink rate matched the trusted window. In a similar solution, Dhamija and Tygar [10] proposed a scheme called Dynamic Security Skins that used coloured patterns to customize the browser window as an indicator of a trusted path between the user and the server to prevent spoofing of the window.

Visualization is an effective method to educate non-technical end-users about phishing threats. In one work, PhishGuru [22] used illustrated comics to teach users after they have responded to a fake phishing message. Users received the training material through simulated phishing emails and an intervention message was provided if the user fell for the email. Results from the PhishGuru user studies suggest that embedded training can effectively teach people how to avoid phishing attacks. The APWG/CMU-Cylab's phishing education landing page program [3] used an infographic approach to deliver anti-phishing training messaged in place of a phishing website that has been taken down. During the first six months of the landing page program, approximately 70,000 Internet users were redirected from phishing URLs to the landing page, where they received the educational material. In an underwater-themed online computer game, Anti-Phishing Phil [22] taught users how to avoid falling for phishing attacks. Participants who played the game identified fraudulent web sites more successfully than those who learned from other training activities, suggesting that interactive games can be an effective way of educating people about phishing and other security attacks.

Government and various organizations have created anti-phishing campaigns that use diagrams and infographic posters to spread awareness. An infographic created by GetCyberSafe [17] depicts cyber pirates on a "phishing" trip to tell the story of email phishing scams. PhishMe [35] is a threat management company offering a range of infographics on phishing. The APWG [2] use a variety of phishing trend diagrams in their quarterly reports. Lavasoft [23] is an anti-malware company who provided a map of the geographic distribution of phishing URLs.

Discovered	Brand	%	URL	IP
Jan 29, 2015 12:37 am UTC	Grupo Bancolombia Bancolombia Personas	50	http://www.dogstrainingsecrets.com/olb/1/bloqueo/ol	198.1.72.238
Jan 29, 2015 12:37 am UTC	Grupo Bancolombia Bancolombia Personas	50	http://www.dogstrainingsecrets.com/olb/1/bloqueo//c	198.1.72.238
Jan 29, 2015 12:36 am UTC	Grupo Bancolombia Bancolombia Personas	50	http://www.dogstrainingsecrets.com/olb/191.90.231.2	198.1.72.238
Jan 29, 2015 12:30 am UTC	Lloyds Bank Retail	50	http://online.llydsbanks.co.uk-claim-account-security-a	192.185.217.242
Jan 29, 2015 12:29 am UTC	MUFG Bank Of Tokyo Mitsubishi	50	http://100.net059085188.t-com.ne.jp/ibg/dfw/APLIN/c	59.85.188.100
Jan 29, 2015 12:27 am UTC	Yahoo! Inc.	50	http://apsmiles.com/wp-content/plugins/stats3/fedex/	216.227.221.77
Jan 29, 2015 12:26 am UTC	Yahoo! Inc.	50	http://apsmiles.com/wp-content/plugins/stats3/fedex/	216.227.221.77
Jan 29, 2015 12:15 am UTC	Lloyds Bank Retail	50	http://online.llydsbanks.co.uk-claim-account-security-a	192.185.217.242
Jan 29, 2015 12:15 am UTC	Yahoo! Inc.	50	http://apsmiles.com/wp-content/plugins/stats3/fedex/	216.227.221.77
Jan 28, 2015 11:35 pm UTC	Wells Fargo Personal Banking	50	http://quaideazam.com/images/identity.php	192.254.234.161
Jan 28, 2015 11:33 pm UTC	EBAY	##	http://www.paypal-updatess.com/89e3a77e6c3de1c09	50.63.202.54,inactive
Jan 28, 2015 11:24 pm UTC	EBAY	##	http://update.billing-paypal.info/webapps/home/793a	160.153.44.197,inactive
Jan 28, 2015 11:21 pm UTC	EBAY	##	http://update.billing-paypal.info/webapps/home/dcfcd	160.153.44.197,inactive
Jan 28, 2015 11:12 pm UTC	EBAY	##	http://update.billing-paypal.info/webapps/home/d385l	160.153.44.197,inactive

Fig. 1. A small sample of the raw dataset from the APWG

III. Design Process of the Interactive Geo-Phisher

The main purpose of the Geo-Phisher information visualization is to raise awareness about the global scale of phishing by making phishing blacklist data more accessible to the general public in an easy to use interactive-visual format. In the current prototype, users can explore the relationship between IP addresses, geo-locations, and targeted brands from more than 40,000 records from the APWG database [4] collected during the month of January 2015. The APWG database contained multiple data dimensions, such as a timestamp, the URL, the targeted brand, and an IP address that enabled us to obtain geoIP information. Geo-Phisher supports touch-interaction using p5.js [37], a JavaScript interpretation of the Processing language [38]. Users could interact directly and immediately with the elements on the screen using direct manipulation.

We applied the four basic stages of information visualization proposed by Ware [50] as a guideline to understand the dataset from the APWG's phishing URL block list [4].

1) Stage I: The collection and storage of data: The data collection stage involves finding and extracting the raw data. Various organizations maintain large phishing URL block lists containing information about reported phishing websites. We initially sought out two sources of phishing data: Phishtank [33] and the APWG's URL block list (UBL) [4]. We decided to use the APWG dataset because it contained data about IP addresses and targeted brands. Figure 1 shows a small subset of the APWG raw dataset. Columns represent the date and time of when the phish is reported, the targeted brand, confidence level of whether it is a phish, the phishing URL, and the IP address.

2) Stage II: Preprocessing: The preprocessing stage aims to transform the data into something that is easier to manipulate. The APWG database contained more than 27,973,000 lines of data (accessed February 2015). To maintain a workable data size for the beta version of our system, we extracted one month of data, which contained approximately 40,000 lines. The data was downloaded as XML files and converted on the local machine to comma-separated values (CSV), one of the recommended file formats [15] for integrating data with Processing. The data was further cleaned using Microsoft Excel. To enrich the original APWG dataset, we included latitudes and longitudes by looking up the geoIP data from `ipinfo.io` and `freegroupip.net`. A PHP script was used to automatically download geoIP data for each IP address found in the original APWG dataset.

3) Stage III: Mapping the data to a visual representation: The two main aspects of the data to visualize spatially are geo-location and time. We acquired the latitude and longitude for each IP address, and scaled it to a pixel location on the computer screen using an equirectangular map projection[1]. The representation of location is inspired by Ahlberg and Shneiderman's "Starfield display" [1], an interactive two-dimensional scatterplot map. A scatterplot displays data as a collection of points on the x and y axis, illustrating the relationship between two variables [7]. A scatterplot is useful to represent the spread of points over a range of data and whether there are differences in variables located in separate regions of the scatterplot [20]. The main advantage of this representation is that it supports viewing of hundreds or thousands of items as points on the map. Plotting points may be more convenient than large areas because points are small yet highly visible, could be colour coded, could be made selectable objects, and could display large datasets rapidly [1]. Each point in the scatterplot supports the minimum tap size on touch screens to ensure that it is easy to press.

Time is visualized on a time series line graph that maps changes in the data over one month, where time (by day) is represented on the x-axis and the number of phishing URLs is represented on the y-axis. Other months could be added to the dataset[2]. Ridges on the graph represent the number of phishing URLs by day, charted as a line graph. Time-based data is central to many datasets such as temperature, population change, or stock data, and it is a vastly utilized visualization technique [47].

4) Stage IV: The human perceptual and cognitive system: Contrast, density, and colour are detected pre-attentively. Light intensity of points on the map display creates a high contrast against the dark background. The density of points represents the relative frequency of attacks coming from the same geographical areas. Clusters of colours allow the user to distinguish the locations of the targeted brands selected by the user. We used 10 out of 12 colours recommend by Ware [50] for use in colour coding, except the colours black and grey which are used for the background. These colours have widely accepted category names and are easily distinguishable due to their wide colour space [50].

We applied principles of Gestalt as the basis for designing visual representations in Geo-Phisher. Gestalt theory outlines certain "laws" that the human brain uses to understand an image. They include *proximity, similarity, continuity, symmetry, closure,* and *relative size* [50]. For example, data on the map relies on the principles of proximity, relative size, and similarity to formulate a visual representation of the relationships between GeoIP locations and the targeted brands.

[1]The equirectangular projection is a cylindrical equidistant projection that is most widely used for mapping the relationship between a pixel on the screen and its corresponding geographic location.

[2]The beta version of the visualization shows a daily aggregation of data over the month of January 2015.

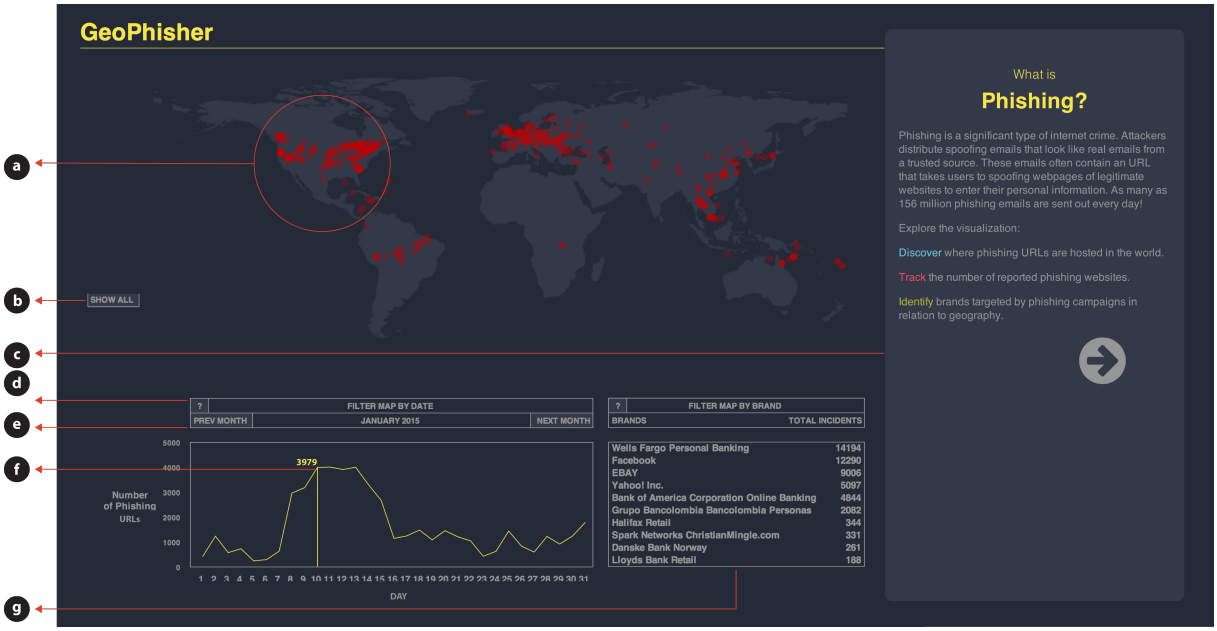

Fig. 2. Interactive interface (*HI* interactivity condition): (a) Map display, (b) Reset map, (c) Information pane, (d) Tool tip, (e) Next/previous month (disabled), (f) Filter-map-by-date feature, (g) Filter-map-by-brand feature.

A. *Overview of the interactive and static Interface*

We first built the interactive Geo-Phisher interface using the method described above. For evaluation purposes, we created a second static version of the same interface to use in the study as the control condition. The graphics used in the static interface were generated from the interactive interface by outputting the visualization as an image. Images from the various stages of interaction were then stitched together in Photoshop to compose a poster-like graphical interface.

In the results section, we refer to the interactive interface as the *HI* (high interactivity) condition and the static interface as the *LO* (low interactivity) condition. We chose to describe the static interface as having "low" interactivity instead of "no" interactivity because users were able perform some minor navigational interactions like scrolling the page and zooming-in. Both of the interactive and static interfaces support the following information-finding tasks:

1) *Tracking the number of reported phishing websites:* The visualizations display one month of data. The line graph shows that phishing attacks occurred between 250 times and 4000+ times a day during the month of January 2015. The highest number of attacks occurred near the middle of the month.
2) *Discovering where phishing URLs are located in the world:* The visualizations map phishing URLs by latitude and longitude. Phishing URLs are concentrated mostly in North America, western Europe, and southern Asia. Concentrations of points appear mostly over major cities, showing a possible correlation between where phishers are located and population density.
3) *Identifying brands targeted by phishing campaigns in relation to geography:* The visualizations show that the top 10 phished brands in January were: Wells Fargo, Facebook, Ebay, Yahoo, and Bank of America, Grupo Bancolombia, Halifax Retail, Spark Networks, Danske Bank Norway, and Lloyds Bank Retail. Financial institutions were targeted most frequently, followed by E-commerce and social media. The targeted brands had varying distributions across the map. For example, Facebook was targeted mainly from North America, but Ebay or Yahoo was targeted across the globe.

B. *User interaction between the interactive and the static interfaces*

We describe the differences in the visual representation and the level of user interaction between the interactive and static prototypes.

1) World map display and timeline graph: The interactive interface shows a scatterplot display over a world map (Figure 2). Points on the scatterplot map represent the hosting locations of phishing URLs. The points are applied an alpha value to show high versus low concentrations of light to represent the frequency of phishing URLs at the various locations. By default, the scatterplot map displays all points (phishing website hosting locations) over the course of one month. Users can move the vertical needle to a date of interest (Figure 2(f)) to filter the display of the hosting locations of phishing URLs by the selected date. Dragging the needle from left to right over the timeline results in consecutive filtering of hosting locations and produces an animated effect of the scatterplot map showing the changes of locations between days. Users can rollover the ridges on the timeline graph to display a number representing the number of phishing sites detected on that date.

Fig. 3. A portion of the quiz feature and phishing advice displayed on the information pane in the visualization. (Cropped to show details.)

Pressing the "show all" button (Figure 2(b)) resets the filters applied to the map.

Users viewed the scatterplot map in the static interface with one month of data with no filter-by date-options. The alpha value the frequency of phishing URLs appearing at the various locations is preserved (darker means more frequency). The number of phishing sites detected each day over the course of the month is displayed above the line graph.

2) World map display and the targeted brands List: (Figure 2(g)) in the interactive interface shows a list of the top ten brands targeted by phishers during the month. Users can select brand(s) of interest to filter the scatterplot map display. For example, users could select 'Ebay', 'Paypal', and "Facebook" to filter the scatterplot map to display only the locations where phishing URLs target these three brands. Each selected brand is assigned a unique colour that corresponds to the colour of the points on the scatterplot map. Users are shown different views of the scatterplot map based on what brands they select.

In the static interface, users viewed multiple displays of the scatterplot map simultaneously, each corresponding to a brand from the top ten brands list. We chose this representation because it enabled the user to see all ten brand filtering effects from the interactive interface. Users are able to zoom-in on the small multiple views and scroll the interface.

3) Quiz and phishing advice: (Figure 2(c)) shows an information pane with an introduction about phishing and the visualization. This page is followed by a quiz feature that prompts users to answer three multiple-choice questions: (1) Which continent do you think hosts the most phishing websites? (2) How many phishing websites are created by attackers worldwide per day from January 8-15? (3) What industry do you think is targeted the most by attackers? The purpose of the questions is to encourage users to ask questions and interact with the visualization to discover the answers. Then, related phishing advice shows users what to do to defend themselves. A portion of quiz feature and phishing advice is shown in Figure 3. The quiz feature responds to users' interactions and provides feedback based on the selected answer. For example, if the user selects an incorrect answer, a message stating why they have selected incorrectly appears followed by the correct answer. Users clicked on a forward button to go to the next question. After the quiz feature, four tips for phishing prevention were provided. For each tip, users had the option to click on a "why" button to learn about the advice rationale. When they are ready to view the next tip, they clicked on a forward button to continue.

In the static interface, users viewed the information in the exact wording. Users viewed phishing tips in block text along with an explanation of why users should follow the advice. Since the quiz is not interactive, the interface could not provide feedback on whether the users chose the answers correctly. To compensate for this, we provided the correct answers at the end of the page.

IV. Method

A. Participants

Thirty university students and staff volunteered to participate in our REB-approved study. The participants were recruited through email mailing lists and a university email newsletter. Each participant was compensated $15.

In the between-subject design, half of the 30 participants were randomly assigned to the *HI* (high-interactivity) condition (N = 15) , and the other half to the *LO* (low-interactivity) condition (N = 15). Gender, age, and education in our sample were fairly evenly split between the two conditions. In the *HI* interactivity condition, 47% of participants were male and 53% were female; 80% held bachelor's degrees or above; 67% were between 25 to 29 years old, 20% were between 18 to 24 years old, and 13% were 50+ years old. In the *LO*

interactivity condition, 53% of the participants were male and 47% were female; 87% held bachelor's degrees or above; 67% were between 25 to 34 years old, 20% were between 18 to 24 years old, and 13% were 50+ years old.

B. Study Procedure

The study was conducted in-person in a laboratory. All participants (N = 30) signed an informed consent form prior to starting the study and were debriefed at the end of the experiment. All questionnaires were completed online using Lime Survey software on a laptop computer. Each session took approximately 30 to 45 minutes to complete.

In the first part of the study, participants completed a 10-minute pre-test that consisted of a *demographics questionnaire*, a *pre-knowledge questionnaire* to assess knowledge and perception of phishing, and a *pre-proactive awareness questionnaire* to assess prior security behavioural intentions.

In the second part of the study, the participants were randomly assigned to either the *HI* or *LO* interactivity condition. Both prototypes were displayed on a wall mounted 55-inch Samsung touchscreen with PQ Frame that ran on a Dell computer with Windows 8. The participants were instructed to take as much time as they liked to explore the interface and follow the tasks provided on the interface. On average, participants spent 5.6 minutes viewing the *HI* condition and 5.3 minutes viewing the *LO* condition.

In the third part of the study, participants completed a 15 minute post-test that contained a *post-knowledge questionnaire* to assess post-knowledge and perception of phishing, a *post-proactive awareness questionnaire* to assess future security behavioural intentions, an *attitude towards the visualization questionnaire* to evaluate the user experience, and a *perceived interactivity questionnaire* to measure the degree of interactivity perceived by the participants.

C. Experimental Conditions

We used two independent variables, *HI* and *LO* interactivity conditions. The *HI* interactivity condition embodies defining characteristics of interactivity identified by various researchers [24], [28], such as giving greater control over the content and navigation, providing immediate feedback and a feeling of two-way communication, and offering a positive sense of system responsiveness and flow. The *LO* interactivity condition is treated like a static online information graphic that featured all the content on one scrollable page. It contained identical information to the interactive version. Both conditions included three multiple-choice questions that prompted users to perform information-finding tasks on the visualization.

D. Manipulation Validation

To ascertain the degree of interactivity in the *HI* and *LO* interactivity conditions is perceived differently by the participants, we included a 5-point semantic differential scale question in the post-test after the participants explored their prototype. The question asked, "On a scale of 1 to 5 (1 - not interactive, 5 - interactive), how interactive would you rate this visualization?" We used a Mann-Whitney's U test to evaluate the differences and found a highly significant effect between the degrees of interactivity perceived by the participants. The mean ranks of the *HI* and *LO* interactivity conditions were 20.6 and 10.4 respectively, with $U = 36, Z = -3.33, p = 0.001, r = 0.61$. Participants clearly recognized that the *HI* interactivity condition offered a more interactive experience.

E. Dependent Measures

1) Perceived interactivity: Additional measures for perceived interactivity included three 5-point Likert-scale questions derived from Liu and Shrum [24] on three dimensions of interactivity: active control, two-way communication, and synchronicity. Two-way communication refers to the communication via user input and feedback from the visualization; active control is the ability to control the user experience through navigation, the pace of the interaction, and the content being accessed; synchronous communication refers to the speed and ease of obtaining information, and the visualization's responsiveness. To ensure that the participants understood the meaning of active control, two-way communication, and synchronicity, we provided examples with each statement.

The specific questions are as follows. On a scale of 1 to 5 (1 - Strongly Disagree and 5 - Strongly Agree), please indicate to what extent you agree or disagree with each of the following statements about the visualization: 1) It enabled two-way communication between me and the visualization (e.g., how well the visualization gave feedback and enabled you to provide input); 2) It enabled me to actively control my experience (e.g., how well the visualization enabled you to control the pace of the interaction, the content being accessed, and the site navigation); 3) It enabled synchronous communication (e.g., how well the visualization performed in terms of the speed and ease of obtaining information, and system responsiveness).

2) Information finding: Participants performed three information-finding tasks on the visualization in both conditions to answer the questions provided in the multiple-choice quiz described in Section III-B3. We observed and took notes of which answer the participants selected. To confirm our observations, the participants' were also instructed to enter their answers in the post-test questionnaire.

3) Security knowledge and behaviour: We developed two 5-point semantic-differential scale questions (e.g., extremely knowledgeable/not at all knowledgeable) to assess pre- and post-knowledge about phishing threats and phishing prevention. To determine what participants know and if any misconceptions about phishing exists, we included 2 additional short answer questions: 1) what is a phishing attack?; 2) can you describe what you know about how to protect yourself from phishing? Technical details in participants' explanations are less important than a demonstration of overall understanding. For example, an answer such as "emails that trick people into going to a vulnerable website asking for personal information

by seeming like a well known website" is an acceptable response for showing awareness and understanding of phishing. We allocated each response a score of 2 for a correct answer, a score of 1 for a partially correct answer, and a score of 0 for an incorrect answer. These questions are included in both the pre- and post-test.

To inquire about the perception of the visualization on behaviour change, a question in the post-test asked, "What online habits would you change?" Participants were instructed to write "none" if the visualization had no effect.

4) Security behaviour intentions for proactive awareness: Security intention was assessed via a 5-point scale in the post-test: Never (1), Rarely (2), Sometimes (3), Often (4), and Always (5), obtained from a five item sub-scale on proactive awareness from the Security Behaviour Intentions Scale (SeBIS) developed by Egelman and Peer [13]. The scale evaluates users' intentions to behave securely by taking proactive measures, such as checking links before clicking them and verifying that the website is secure before sending information.

5) Attitude towards the visualization: The twelve five-point semantic differential items are adapted from the Attitude Towards the Ad scale [26] from communications. The three-part components and sub-items of the scale are: hedonism, (fun to see/not fun to see, pleasant/unpleasant, entertaining/not entertaining, enjoyable/not enjoyable), interestingness (important/not important, helpful/not helpful, informative/not informative, useful/not useful), and utilitarianism (curious/not curious, boring/not boring, interesting/not interesting, keeps my attention/does not keep my attention). Attitudes toward the ad is an important construct mediating the effects of advertising on brand attitude and purchase intention [26], and is often used as a dependent variable to study the effects of interactivity on attitudes toward a website (e.g., [32]). We included this scale because our visualization has some similarities to advertisements and websites, and it aims to achieve positive effects on attitudes and intentions towards security learning and behaviour.

V. Results

We used the non-parametric Mann-Whitney significance test (also known as Wilcoxon Rank sum test) to evaluate the ordinal data from the Likert-scale results to identify whether differences between the two conditions exist. For each result. we report the Mann Whitney value (U), standard deviation away from the mean (Z), the p-value (p), and the effect size (r). In all cases, $p < 0.05$ is considered significant.

A. Information Finding

Based on the selected answers from the multiple-choice quiz, we found that the *HI* interactivity condition supported more accurate information-finding than the *LO* interactivity condition for one of three tasks.

In the first challenge (see Figure 4), 93% of participants from the *HI* interactivity condition were successful at finding the correct answer (North America) compared to 60% from the *LO* interactivity condition. A Mann-Whitney's U test showed a significant effect ($U = 75, Z = -2.12, p = 0.03, r = 0.39$), with a mean rank of 18 for the *HI* and 13 for the *LO*.

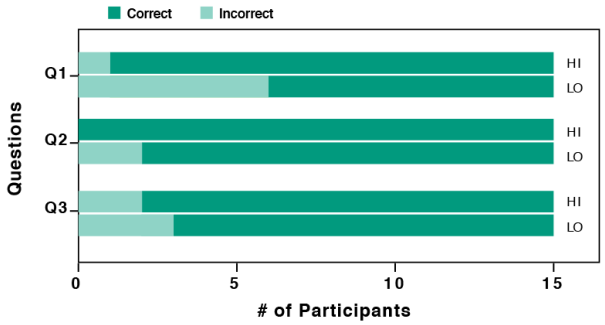

Fig. 4. The number of participants' who responded correctly or incorrectly to three information-finding tasks. Q1: Which continent do you think hosts the most phishing websites?; Q2: How many phishing websites are created by attackers worldwide per day from January 8-15?; Q3: What industry do you think is targeted the most by attackers?

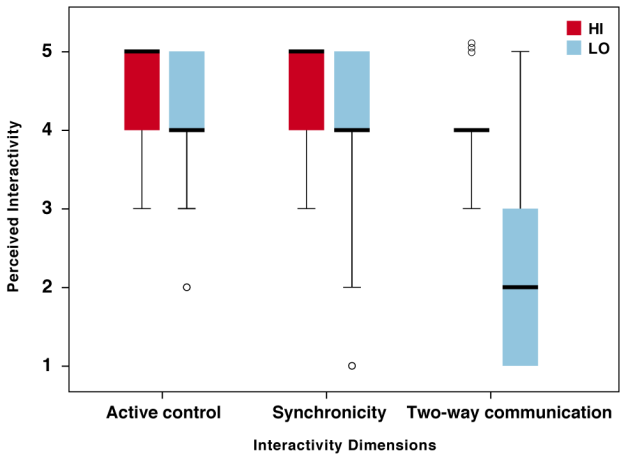

Fig. 5. Perceived interactivity based on the three dimensions using a 5-point Likert-scale (1 = strongly disagree, 5 = strongly agree).

Participants performed slightly better on the the second and third challenges in the *HI* interactivity condition, but the difference was not statistically significant. All of the participants from the *HI* interactivity condition were successful at finding the correct answer (around 4000) to the second question compared to 87% from the *LO* interactivity condition. For the third challenge, 87% of participants from the *HI* interactivity condition were successful at finding the correct answer (banking services) compared to 80% from the *LO* interactivity condition.

B. Perceived Interactivity

Participants evaluated the prototype on three dimensions of interactivity (see Figure 5): 1) active control, 2) synchronicity, and 3) two-way communication. We found no significant difference between the two conditions for active control and

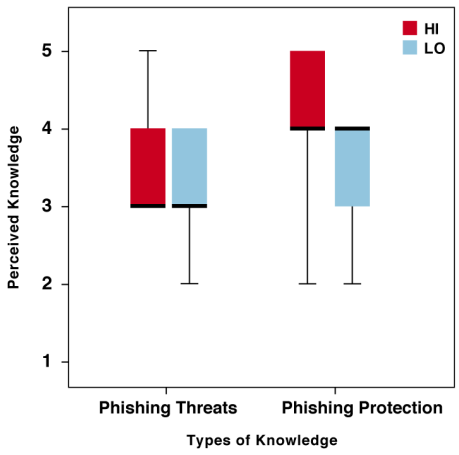

Fig. 6. Perceived knowledge about phishing threats and phishing prevention during the post-test on a 5-point Likert-scale (1 = strongly disagree, 5 = strongly agree).

synchronicity. The evaluations were fairly consistent for these two dimensions, both has a median of 5 for *HI* and 4 for *LO*. However, a significant difference was found on two-way communication ($U = 37, Z = -3.23, p = 0.001, r = 0.58$), with a median of 4 for *HI* and 2 for *LO* interactivity conditions.

H1 is only partially supported because the *HI* interactivity condition was perceived to be more interactive for two-way communication, but not for active control and synchronicity.

C. Perceived Knowledge

Participants' of the *HI* and *LO* conditions had near equal levels of self-evaluated pre-test knowledge of phishing. Mean values for participants' Likert-scale ratings (1 = not at all knowledgeable, 5 = extremely knowledgeable) to the question "How knowledgeable are you about phishing threats?" were 2.5 and 2.6 for the *HI* and *LO* interactivity conditions respectively. Mean values for their Likert-scale rating to the question "How knowledgeable are you about protecting yourself against phishing threats?" were 2.7 for both conditions. We therefore believe our samples between the two groups are comparable.

We repeated the two questions in the post-test questionnaire to assess their perceptions of knowledge gained after learning about phishing. Participants had similar perceived knowledge of phishing threats after learning, but we found a significant effect on perceived knowledge for prevention against phishing between the two conditions. A Mann-Whitney's U test showed a mean rank of 18.5 for the *HI* interactivity group and 12.5 for the *LO* interactivity group ($U = 67.5, Z = -1.99, p = 0.05, r = 0.36$). A comparison of participants' knowledge in the pre-test and post-test is summarized in Figure 6.

Therefore, H2 is supported for perceptions of learning gains about phishing prevention, with participants in the *HI* condition feeling more knowledgeable about protecting themselves against phishing.

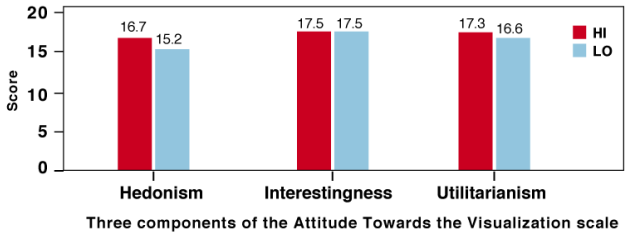

Fig. 7. Attitude towards the visualization results summary. The mean scores represent the sum of responses to 4 questions in each component.

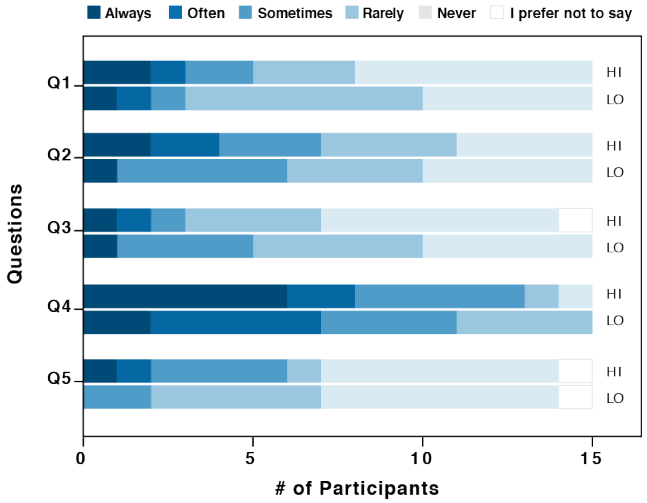

Fig. 8. Participants' security Intention responses for both conditions. The statements are Q1: When someone sends me a link, I would open it without first verifying where it goes; Q2: I would know what website I'm visiting based on its look and feel, rather than by looking at the URL bar; Q3: I would submit information to websites without first verifying that it will be sent securely (e.g., SSL, https://, a lock icon); Q4: When browsing websites, I would mouseover links to see where they go, before clicking them; Q5: If I discover a security problem, I would continue what I was doing because I assume someone else will fix it.

D. Attitude Towards the Visualization

We did not find a significant effect on participants' attitudes toward the visualization on the dimensions of hedonism, interestingness, and utilitarianism. Figure 7 shows that the participants responded positively towards both conditions. Thus, H3 is not supported.

E. Security Behaviour Intentions for Proactive Awareness

Participants' responses based on a 5-point scale (1 = always, 5 = never) for security intention are fairly consistent between the two conditions in the post-test. In all cases, 1 is considered the most negative and 5 the most positive. See Figure 8 for a summary.

Based on our statistical analysis, we did not find that interactivity significantly affected participants' security behaviour intention on proactive awareness. Therefore, H4 is not supported.

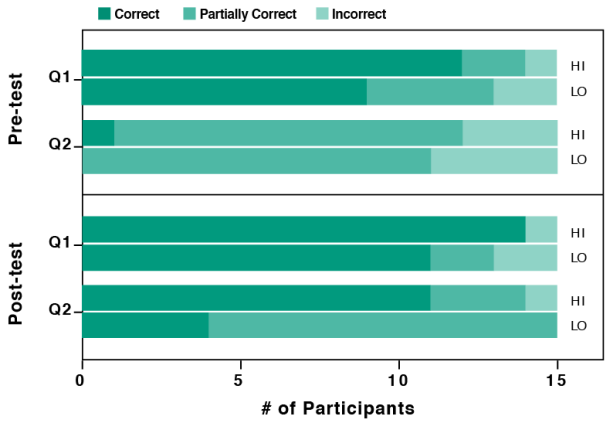

Fig. 9. The number of participants who correctly, partially correctly, or incorrectly responded to two questions about phishing threats and phishing prevention in the pre- and post-test. Q1: What is phishing?; Q2: Can you describe what you know about how to protect yourself from phishing?

F. Learning Effects

1) Phishing threats: When asked about "what is phishing?" in the pre-test, many participants from both groups demonstrated basic prior knowledge. A sample correct answer typically given is "a type of online activity in which someone sends an email or creates a website that is meant to look like a legitimate site, targeting individuals who are asked to provide personal or financial information for the gain of the phisher." 80% of participants from the *HI* interactivity condition and 60% from the *LO* interactivity condition provided a response similar to this example. An additional 13% from the *HI* interactivity condition and 27% from the *LO* interactivity condition provided a partially correct answer such as "an attempt to gain personal information for the purpose of defrauding the individual monetarily or of their personal identity/information," but did not describe how phishing works. 7% from *HI* and 13% from *LO* gave an incorrect answer.

Participants again explained what is phishing in the post-test after learning about it from the visualization. We found that participants from both conditions were more knowledgeable about phishing, where 88% of participants from the *HI* interactivity condition and 73% from the *LO* interactivity condition provided a correct answer, but no significant effect was found between the two conditions in the post-test.

2) Phishing prevention: When asked about "can you describe what you know about how to protect yourself from phishing?" in the pre-test, 20% from the *HI* interactivity group and 27% from the *LO* interactivity group showed a complete lack of understanding. Most participants (73% from both *HI* and *LO* conditions) showed a shallow understanding and gave a partially correct answer such as "only entering information on official websites", and "don't click on links or pictures that look *off*". Only 7% from the *HI* interactivity condition gave a correct answer and none from *LO*. An answer is attributed as correct when the participants described specific strategies on how to identify legitimate websites (e.g., https, a lock icon), fraudulent emails (e.g., urgent emails that request personal information), or other phishing prevention approaches (e.g., type URL in address bar instead of clicking on a link, not sending personal and financial information via email).

Participants from both conditions demonstrated increased post-test knowledge after seeing the visualization (see Figure 9). Interestingly, a significant effect was found between the two conditions using a Mann Whitney's U test ($U = 65.5, Z = -2.22, p = 0.03, r = 0.41$) with mean ranks of 18.63 for *HI* and 12.37 for *LO*. These results confirm the findings for participants' self-evaluated perceived knowledge of phishing prevention, where a significant effect was also found. Therefore, H2 is further supported, with greater cognitive absorption of phishing prevention information in the *HI* interactivity condition.

VI. Discussion

Both visualizations were rated positively by participants and helped them acquire information about phishing and preventative strategies. Although not as pronounced as we had initially anticipated, the interactive version of our visualization performed significantly better on some measures, and never worse than the static version.

Users performed better on one out of the three information seeking tasks. There was a large discrepancy between the success rates (93% for *HI* vs. 60% for *LO*) at answering which continent hosts the most phishing sites. The correct answer was North America, but several participant chose Europe. When the map displayed the entire month of scatterplot data in the *LO* condition, Europe appeared to have denser concentrations of phishing sites than North America. The *HI* condition also supported the month view, but the filter-map-by-date feature (Figure 2 (f)) in the *HI* condition enabled users to view a day-by-day aggregation of the scatterplot map. On most days, the data clearly showed that North America was the top hosting continent, but it is more difficult to differentiate when the scatterplot maps overlap to display the entire month. Both *HI* and *LO* conditions were able to support the other two information seeking tasks, which involved more direct readings of visual information. The observation suggests that filtering techniques in infovis help users to connect data segments and to make sense of the entire dataset as a whole, particularly for temporal data.

The *HI* interactivity condition had a significant effect on users' perceived and actual knowledge of phishing prevention. The result is consistent with prior work in website interactivity [32], where highly interactive websites led to greater cognitive absorption than static websites. Participants' described more ways to protect themselves against phishing in the post-test from the tips provided in the *HI* interactivity condition than the *LO* interactivity condition. This result is consistent with participants' self-evaluated perceived knowledge from the post-test. Since the two conditions contained identical information, we attribute the difference to participants' interaction with the information in the *HI* condition, where they received one advice one at a time, pressed a button to

reveal the rationale for the advice, then proceeded to the next advice. This information is displayed in stacked text blocks as a continuous unit in the *LO* condition. The result suggests that interactivity could help to facilitate deeper learning by providing learners with opportunities to pause and process the information before continuing to the next step. This effect of segmentation information is observed in education literature, where one study [27] found that broken a narrated animation into segments and pressing a button to continue increased students' learning performance.

In our study, participants from both groups gained knowledge about phishing threats and phishing prevention from our visualizations. No significant effect was found for knowledge of phishing threats because many participants had some prior knowledge, but a significant effect was found for phishing prevention, suggesting that users could benefit from educational efforts on what users should do to prevent phishing.

Interactivity had no significant effect on participants' security behaviour intentions for proactive awareness. Nevertheless, when inquired about what behaviours they would change during the post-test, the majority (93%) from both conditions stated at least one behaviour that they would change. The behaviours described include mousing-over links to confirm the source, ensuring that the website is secure via https, and typing web URLs directly into the address bar instead of clicking on links. All of the reported behaviours were recommended in the phishing advice section in the visualizations.

A. Limitations

Generalizability of our findings will need to be confirmed since our study was conducted in a lab environment with a small sample of university students and staff. Our study did not account for environmental factors that could influence user engagement.

Our current implementation displays a small segment of available phishing data. Adaptations of the entire dataset over months and years would yield more accurate and general patterns. Implementation that supports real time data would make phishing patterns more exact and relevant.

Our *LO* interactivity condition allowed users to scroll and zoom-in on the interface. The effect could be perceived as being very interactive for some users, even though the level of interactivity is much less than the *HI* interactivity condition. This may affected our results on the three dimensions of perceived interactivity, where our participants thought the *HI* condition to be more interactive than the *LO* condition for two-way communication via user input and feedback from the visualization, but not for active control and synchronous communication. We suspect that this is because in both conditions, users were in control of the pace and direction of how the information is consumed, whether through interaction with the visualization content in the *HI* condition, or through scrolling in the *LO* condition.

Participants reported intent to change security behaviour in the post-test. However, the behaviours will need to be confirmed in a follow-up study due to the likelihood of discrepancies between what users know and what they actually do in computer security [40].

Surprisingly, our hypothesis that the *HI* interactivity condition would result in more positive attitudes toward the visualization was not supported by our data; the participants responded positively towards both conditions. This could be due to our small sample size. In particular, the evaluations implied a slightly more positive response for the *HI* interactivity condition for the sub-items under hedonism. Although the difference was not significant in our sample, a study with a larger sample size may yield a more positive effect.

VII. Conclusions

The purpose of the study was to examine the effects of interactivity in a phishing information visualization tool on information finding, user perceptions and security behaviour intentions, and the effectiveness of learning. We compared two information-equivalent visualizations; one with multiple interactive features and a low-interactivity control version. We studied whether interactivity added significant value over the control condition, and conclude the novelty of interactivity had no significant effect over the control condition on participants' positive attitudes toward the visualization and their security behaviour intentions for proactive awareness. However, interactivity in visualizations do assist users with information finding, which may result in greater learning than from a static representation.

VIII. Authors and Acknowledgments

Robert Biddle is co-theme leader for NSERC SurfNet, a Digital Surface Software Application Network of researchers, government, and industry partners. Leah Zhang-Kennedy and Elias Fares are funded through SurfNet. Sonia Chiasson is Canada Research Chair in Human-Oriented Computer Security.

References

[1] C. Ahlberg and B. Shneiderman. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 313–317. ACM, 1994.

[2] APWG. Phishing activity trends report, March 2015. http://docs.apwg.org/reports/apwg_trends_report_q3_2014.pdf/.

[3] APWG. Phishing education landing page program, Accessed February 2015. http://phish-education.apwg.org.

[4] APWG. URL Block List (UBL), Accessed February 2015. https://ecrimex.net.

[5] E. P. Bucy. The interactivity paradox: Closer to the news but confused. *Media access: Social and psychological dimensions of new technology use*, pages 47–72, 2004.

[6] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in information visualization: Using vision to think*. Morgan Kaufmann, 1999.

[7] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey. *Graphical methods for data analysis*. Wadsworth/Brooks Cole, 1983.

[8] L. Chittaro. Information visualization and its application to medicine. *Artificial intelligence in medicine*, 22(2):81–88, 2001.

[9] S. Claes and A. Vande Moere. Street infographics: Raising awareness of local issues through a situated urban visualization. In *Proceedings of the ACM International Symposium on Pervasive Displays*, pages 133–138. ACM, 2013.

[10] R. Dhamija and J. D. Tygar. The battle against phishing: Dynamic security skins. In *Proceedings of the Symposium on Usable Privacy and Security*, pages 77–88. ACM, 2005.

[11] A. Dix and G. Ellis. Starting simple: Adding value to static visualization through simple interaction. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pages 124–134. ACM, 1998.

[12] M. Dörk, P. Feng, C. Collins, and S. Carpendale. Critical infovis: Exploring the politics of visualization. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 2189–2198. ACM, 2013.

[13] S. Egelman and E. Peer. Scaling the security wall: Developing a security behavior intentions scale. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 2873–2882. ACM, 2015.

[14] B. J. Fogg. *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann, San Francisco, 2003.

[15] B. Fry. *Visualizing data: exploring and explaining data with the Processing environment*. O'Reilly Media, Inc., 2007.

[16] N. Gershon, S. G. Eick, and S. Card. Information visualization. *Interactions*, 5(2):9–15, 1998.

[17] Get Cyber Safe. Phishing: How many take the bait?, accessed April 2015. http://www.getcybersafe.gc.ca/cnt/rsrcs/nfgrphcs/nfgrphcs-2012-10-11-en.aspx.

[18] C. G. Healey, K. S. Booth, and J. T. Enns. High-speed visual estimation using preattentive processing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 3(2):107–135, 1996.

[19] J. Heer and D. Boyd. Vizster: Visualizing online social networks. In *IEEE Symposium on Information Visualization*, pages 32–39. IEEE, 2005.

[20] D. R. Helsel and R. M. Hirsch. *Statistical methods in water resources*, volume 49. Elsevier, 1992.

[21] S. Kalyanaraman and S. S. Sundar. The psychological appeal of personalized content in web portals: Does customization affect attitudes and behavior? *Journal of Communication*, 56(1):110–132, 2006.

[22] P. Kumaraguru, S. Sheng, A. Acquisti, L. Cranor, and J. Hong. Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology*, 10(2):7, 2010.

[23] Lavasoft. Phishing URLs geographic distribution, April 2015. http://www.lavasoft.com/mylavasoft/securitycenter/whitepapers/detecting-malicious-urls-part-2-"where".

[24] Y. Liu and L. J. Shrum. What is interactivity and is it always such a good thing? implications of definition, person, and situation for the influence of interactivity on advertising effectiveness. *Journal of Advertising*, 31(4):53–64, 2002.

[25] N. H. Lurie and C. H. Mason. Visual representation: Implications for decision making. *Journal of Marketing*, 71(1):160–177, 2007.

[26] S. B. MacKenzie, R. J. Lutz, and G. E. Belch. The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing explanations. *Journal of Marketing Research*, pages 130–143, 1986.

[27] R. E. Mayer and P. Chandler. When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, 93(2):390, 2001.

[28] S. J. McMillan and J. Hwang. Measures of perceived interactivity: An exploration of the role of direction of communication, user control, and time in shaping perceptions of interactivity. *Journal of Advertising*, pages 29–42, 2002.

[29] C. Mekhail, L. Zhang-Kennedy, and S. Chiasson. Visualizations to teach about mobile online privacy. In *Persuasive Technology Conference, Adjunct Proceedings*. Springer, 2014.

[30] S. P. Moon, Y. Liu, S. Entezari, A. Pirzadeh, A. Pappas, and M. S. Pfaff. Top health trends: An information visualization tool for awareness of local health trends. In *Proceedings of the International ISCRAM Conference*, pages 177–187, 2013.

[31] J. Müller, F. Alt, D. Michelis, and A. Schmidt. Requirements and design space for interactive public displays. In *Proceedings of the International Conference on Multimedia*, pages 1285–1294. ACM, 2010.

[32] J. Oh and S. S. Sundar. How does interactivity persuade? an experimental test of interactivity on cognitive absorption, elaboration, and attitudes. *Journal of Communication*, 65(2):213–236, 2015.

[33] OpenDNS. PhishTank: Out of the net, into the tank, Accessed February 2015. http://www.phishtank.com.

[34] R. E. Petty and J. T. Cacioppo. The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19:123–205, 1986.

[35] PhishMe. Infographics, April 2015. http://phishme.com/resources/infographics/.

[36] Z. Pousman, J. T. Stasko, and M. Mateas. Casual information visualization: Depictions of data in everyday life. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6):1145–1152, 2007.

[37] Reas, C. and Fry, B. p5.js, Accessed February 2015. http://p5js.org.

[38] Reas, C. and Fry, B. Processing, Accessed February 2015. https://processing.org.

[39] B. Reeves and C. Nass. Perceptual user interfaces: Perceptual bandwidth. *Communications of the ACM*, 43(3):65–70, 2000.

[40] S. Riley. Password security: What users know and what they actually do. *Usability News*, 8(1), 2006.

[41] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010.

[42] T. Skog, S. Ljungblad, and L. E. Holmquist. Between aesthetics and utility: Designing ambient information visualizations. In *EEE Symposium on Information Visualization*, pages 233–240. IEEE, 2003.

[43] J. Steuer, F. Biocca, and M. R. Levy. Defining virtual reality: Dimensions determining telepresence. *Communication in the Age of Virtual Reality*, pages 33–56, 1995.

[44] J. Stoll, C. S. Tashman, W. K. Edwards, and K. Spafford. Sesame: Informing user security decisions with system visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1045–1054. ACM, 2008.

[45] S. S. Sundar, S. Kalyanaraman, and J. Brown. Explicating web site interactivity impression formation effects in political campaign sites. *Communication Research*, 30(1):30–59, 2003.

[46] S. S. Sundar and J. Kim. Interactivity and persuasion: Influencing attitudes with information and involvement. *Journal of Interactive Advertising*, 5(2):5–18, 2005.

[47] E. R. Tufte and P. R. Graves-Morris. *The visual display of quantitative information*, volume 2. Graphics Press Cheshire, CT, 1983.

[48] N. Valkanova, S. Jorda, M. Tomitsch, and A. Vande Moere. Reveal-it!: The impact of a social visualization projection on public awareness and discourse. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3461–3470. ACM, 2013.

[49] F. B. Viégas and M. Wattenberg. Artistic data visualization: Beyond visual analytics. In *Online Communities and Social Computing*, pages 182–191. Springer, 2007.

[50] C. Ware. *Information visualization: Perception for design*. Elsevier, 2012.

[51] A. Whitten and J. D. Tygar. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In *USENIX Security Symposium*, 1999.

[52] Z. E. Ye, S. Smith, and D. Anthony. Trusted paths for browsers. *ACM Transactions on Information and System Security (TISSEC)*, 8(2):153–186, 2005.

[53] K.-P. Yee and K. Sitaker. Passpet: Convenient password management and phishing protection. In *Proceedings of the Symposium on Usable Privacy and Security*, pages 32–43. ACM, 2006.

[54] J. S. Yi, Y. ah Kang, J. T. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6):1224–1231, 2007.

[55] R. N. Zambrano and Y. Engelhardt. Diagrams for the masses: Raising public awareness–from Neurath to Gapminder and Google Earth. In *Diagrammatic Representation and Inference*, pages 282–292. Springer, 2008.

[56] L. Zhang-Kennedy, S. Chiasson, and R. Biddle. Password advice shouldn't be boring: Visualizing password guessing attacks. In *APWG eCrime Summit*. IEEE, 2013.

[57] L. Zhang-Kennedy, S. Chiasson, and R. Biddle. Stop clicking on "update later": Persuading users they need up-to-date antivirus protection. In *Persuasive Technology Conference*. Springer, 2014.

[58] L. Zhang-Kennedy, S. Chiasson, and R. Biddle. The role of instructional design in persuasion: A comics approach for improving cyber security. (to appear). *International Journal of Human-Computer Interaction*, pages 1–19, 2015.