APPROVAL SHEET

Title of Thesis: Phishing in an Academic Community: A Study of User Susceptibility and Behavior

Name of Candidate: Alejandra Diaz
Master of Science - Computer Science, 2018

Thesis and Abstract Approved: Alan T. Sherman

Dr. Alan T. Sherman
Professor of Computer Science and Electrical Engineering
Department of Computer Science and Electrical Engineering

Dr. Anupam Joshi
Department Chair
Department of Computer Science and Electrical Engineering

Date Approved: 6-18-18

APPROVAL SHEET

Title of Thesis: Phishing in an Academic Community: A Study of User Susceptibility and Behavior

Name of Candidate: Alejandra Diaz
Master of Science - Computer Science, 2018

Thesis and Abstract Approved: _____

Dr. Alan T. Sherman
Professor of Computer Science and Electrical Engineering
Department of Computer Science and Electrical Engineering

_____

Dr. Anupam Joshi
Department Chair
Department of Computer Science and Electrical Engineering

Date Approved: _____

ABSTRACT

Title of Thesis: Phishing in an Academic Community: A Study of User Susceptibility and Behavior

Alejandra Diaz, Master of Science - Computer Science, 2018

Thesis directed by: Dr. Alan Sherman, Professor of Computer Science and Electrical Engineering
Dr. Anupam Joshi, Department Chair
Department of Computer Science and Electrical Engineering

We present an observational study on the relationship between demographic factors and phishing susceptibility. In spring 2018, we sent three phishing emails and a survey to examine user click rates and demographics within UMBC's undergraduate student population. This study, the first to investigate several demographic factors without prior user knowledge in a university setting, shows correlations between user susceptibility and college affiliation, age, cyber training levels, academic year progression, phishing awareness, cyber club or scholarship involvement, and amount of time spent on a computer. We observe no such relationship for gender.

We used the *Billing Problem*, *Contest Winner*, and *Expiration Date* phishing tactics. From March through May 2018, we performed three experiments that delivered phishing attacks to 450 randomly-selected students on three different days (1,350 students total). Unlike other studies, to simulate real phishing scenarios the participants were initially unaware of the study. Experiment 1 impersonated banking authorities; Experiment 2 enticed users with monetary rewards; and Experiment 3 threatened users with account cancellation. We then sent a survey that collected students college affiliation, age, cyber training levels, academic year progression, phishing awareness, cyber club or scholarship involvement, and amount of time spent on a computer.

We conclude that gender does not indicate student risk level ($\chi^2 = 0.43, p = 0.51, \alpha = 0.05$). Students within a technical field are less likely to click a link (39% students clicked), followed by Natural and Mathematical Sciences students (63% students clicked) second and Arts, Humanities and Social Sciences students most susceptible (78% students clicked) ($\chi^2 = 136.35, p < 0.0001, \alpha = 0.05$). Age ($\chi^2 = 16.25, p = 0.001, \alpha = 0.05$) and academic year progression ($\chi^2 = 15.67, p = 0.0013, \alpha = 0.05$) influenced susceptibility as well, with younger and less educated students having higher click rates to phishing schemes than did their older and more educated counterparts. There exists a correlation in level of cyber training and decreasing click rate ($\chi^2 = 19.47, p < 0.0001, \alpha = 0.05$), similar to the relationship of low click rates and cyber scholarship program involvement (28% students clicked), followed by cyber club membership (53% students clicked) and no involvement at all (73% students clicked) ($\chi^2 = 19.29, p < 0.0001, \alpha = 0.05$). Time spent on the computer is a significant factor in click rates as well ($Fisher's p < 0.0001, \alpha = 0.05$). Students that spend more time on the computer after 4 hours are documented to not click the phishing links as often (4-8 88% students clicked, 8-12 70% students, 12+ 52% students clicked). Contrary to our expectations, there exists a negative relationship between phishing awareness and students' resistance to clicking a phish link ($\chi^2 = 77.46, p < 0.0001, \alpha = 0.05$). Students who identified themselves as understanding the definition of phishing had a higher susceptibility rate (80% students clicked) than their peers who are merely aware of phishing attacks (43% students clicked) and those with no knowledge whatsoever (28% students clicked).

KEYWORDS: Phishing, spear-phishing, phishing scenarios, cyber demographics, user susceptibility, cybersecurity, *Billing Problem* tactic, *Contest Winner* tactic, *Expiration Date* tactic.

Phishing in an Academic Community: A Study of User Susceptibility and Behavior

by

Alejandra Diaz

Thesis submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfillment
of the requirements for the degree of
Master of Science in Computer Science
2018

ACKNOWLEDGMENTS

First, I want to thank my advisors Dr. Alan Sherman and Dr. Anupam Joshi, for I would have not been able to complete my thesis without their guidance and support.

Dr. Sherman's analytical mind and attention to detail is something that I have come to admire, and I would not have wished for a better advisor. I want to thank him for challenging me, because without his suggestions and questions I would not be the student I am today. He pushed me to be better — a better student, a better writer, a better researcher.

Dr. Joshi is another great advisor with whom I had the privilege to work. I have known him in various roles — as a faculty mentor, Cyber Scholar Director, professor, and now as my research advisor. He introduced me to research when I took a special topics class with him as an undergraduate student and opened my eyes to new possibilities.

A huge thank you to Dr. Sinha and Dr. Neerchal for their guidance on statistical tests and models. I would also like to thank Jack Seuss, Andy Johnston, Mark Cather, and all the DoIT staff that supported me throughout my research project. I am extremely thankful for their support and help as I was conducting my research project.

Other important supporters to mention are my friends — especially Merle Ferguson, who always believed in me and what I could accomplish. Thank you for pushing me and encouraging me on the days where the last thing I wanted to do was work on my thesis.

Last, I want to thank my parents, Carmen and Percy Diaz. They have sacrificed so much for my brother and me, and they are my inspiration and motivation. Thank you for always believing in me.

This would not be possible without each of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

We investigate if user demographics have a relationship to phishing susceptibility. By conducting simulated phishing tests, we observe age, college affiliation, academic year progression, time spent on a computer, cyber club or cyber scholarship program affiliation, cyber training, and phishing awareness as significant factors, with gender as an insignificant factor.

This observational study is the first to examine age, gender, college affiliation, academic year progression, time spent on a computer, cyber club or cyber scholarship program affiliation, cyber training, and phishing awareness demographics in one study. We incorporate a large and diverse sample group in a college setting. Furthermore, we simulate phishing scenarios by having the participants unaware of the experiments while conducting the simulated phishing attacks, unlike phishing studies that inform their participants beforehand that they are being tested on their ability to discern phishing attempts. This study offers significant advantages over conventional phishing tests by undergoing real-world phishing attacks, producing outcomes closer to a true representation of results for a phishing attack on a college campus.

We randomly select students into three Cohorts and create sub-groups by College: Arts, Humanities, and Social Sciences, Engineering and Information Technology, or Nat-

ural and Mathematical Sciences. The Cohorts are then sent phishing emails in three experiments: Experiment 1: PayPal, Experiment 2: Quadmania, and Experiment 3: DoIT on three separate days. These three experiments are created using the *Billing Problem*, the *Contest Winner*, and the *Expiration Date* phishing tactics [5]. Student click rates are collected using mail tracking software.

Once all experiments are concluded, we send a debriefing statement to all selected students and an additional optional survey to students who have opened the phishing emails. The debriefing statement informs the students of the study and ensures the students' confidentiality and anonymity. The survey, if the student chooses to fill out, collects age, gender, college affiliation, academic year progression, time spent on a computer, cyber club or cyber scholarship program affiliation, cyber training, and phishing awareness demographics.

Our motivation lies in whether demographic factors in a university setting may influence user susceptibility for phishing attacks. We wish to clarify our understanding of dependent variables in a student population such that a university's IT department may implement effective training tailored to the individual student.

We conclude that there are certain demographic factors that do indicate a student's susceptibility to a phishing scheme. We observe that gender does not indicate students' risk level ($\chi^2 = 0.43, p = 0.51, \alpha = 0.05$). Students within a technical field are less likely to click a link (39% students clicked), followed by Natural and Mathematical Sciences students (63% students clicked) second and Arts, Humanities and Social Sciences students most susceptible (78% students clicked) ($\chi^2 = 136.35, p < 0.0001, \alpha = 0.05$). Age ($\chi^2 = 16.25, p = 0.001, \alpha = 0.05$) and academic year progression ($\chi^2 = 15.67, p = 0.0013, \alpha = 0.05$) influenced susceptibility as well, with younger and less educated students having higher click rates to phishing schemes than did their older and more educated counterparts. There exists a correlation in level of cyber training and decreasing click rate ($\chi^2 = 19.47, p < 0.0001, \alpha = 0.05$), similar to the relationship of low click rates

and cyber scholarship program involvement (28% students clicked), followed by cyber club membership (53% students clicked) and no involvement at all (73% students clicked) ($\chi^2 = 19.29, p < 0.0001, \alpha = 0.05$). Time spent on the computer is a significant factor in click rates as well ($Fisher'sp < 0.0001, \alpha = 0.05$). Students that spend more time on the computer after 4 hours are documented to not click the phishing links as often (4-8 88% students clicked, 8-12 70% students, 12+ 52% students clicked). Contrary to our expectations, there exists a negative relationship between phishing awareness and students' resistance to clicking a phish link ($\chi^2 = 77.46, p < 0.0001, \alpha = 0.05$). Students who identified themselves as understanding the definition of phishing had a higher susceptibility rate (80% students clicked) than their peers who are merely aware of phishing attacks (43% students clicked) and those with no knowledge whatsoever (28% students clicked).

Chapter 2

BACKGROUND

When it comes to cybersecurity, various people perceive cyber threats to be malicious code or from nation-states. While these ideas may be true, the single most important and devastating vulnerability a company can have is its very own people [6]. The human factor, or human error, is what attributes to roughly 95% of security incidents [6]. Due to humans being the weakest link in an organization, various malicious actors aim to exploit users or employees into giving up valuable and confidential information. Dr. Jim Kent, Global Head of Security and Intelligence at Nuix, a global technology company, expresses that human behavior is prone to make mistakes. He states that even with security awareness training, employees will ”put their organizations at risk by opening malicious attachments and visiting suspect websites” [9].

With users identified as the greatest vulnerability in a system, companies are wary of social engineering tactics that target their employees. Social engineering is the process of exploiting human interaction and behavior to get a user to disclose sensitive information [8]. An extremely popular social engineering strategy is phishing. Phishing occurs when a malicious actor poses as an authority or person to get their victims credentials, personal information, or other confidential information [10]. Phishing emails, for example, take public information about a person, such as their name, employer, or friends group, and uses

it to entice the victim to click a link or fill out personal information. Common methods to allure a victim are to use urgency or scare tactics. The email body might impersonate a banking institution, contain references to a major event, provide winnings or a prize to the user, or to threaten the user in some capacity (such as account deletion for non-compliance) [8]. This study incorporates the *Billing Problem*, the *Contest Winner*, and the *Expiration Date* phishing tactics [5]. The *Billing Problem* tactic impersonates a well-known banking institution and presents the user with either a bill or banking statement that they have not ordered. The *Contest Winner* tactic uses monetary gain to entice the user to click. The email congratulates the victim for winning a prize as long as they fill out information. The *Expiration Date* tactic, however, uses fear and intimidation to fool the user. The phish usually demands an action be done under a certain amount of time to have a user's account not be deactivated or deleted [5].

We assume that students who are more technologically advanced are less susceptible to these tactics. Our expectations in this study are that students with higher phishing awareness and cyber training or affiliation are also less susceptible to phishing schemes. Likewise, students with no prior knowledge of phishing or no prior training or affiliation are expected to have higher click instances to the three Experiments.

Chapter 3

RELATED WORK

Studies are divided into two main categories: studies that incorporated unannounced phishing tests to their participants and studies that examined demographics as potentially significant factors to user behavior.

Dodge Jr., et al. conducted an unannounced phishing test on students of the United States Military Academy (USMA) [3]. The goal was to test how students, from freshmen to seniors, were able to identify phishing attempts directed at them [3]. This action was done to test their cyber training programs at USMA. Similarly, Aloul presented a project where a fake website portal recorded the students who fell for the phishing website [1]. The project concluded that security awareness could affect the user susceptibility for a phishing scheme [1]. In this study, multiple unannounced experiments are conducted in addition to a survey that captures demographics to determine relationships with user action. In contrast to the previous works, this observational study conducts phishing experiments directly to the users and analyzes more demographics than just overall student click rate or academic year.

There have been phishing studies based on demographics in the past. Sheng, et al. studied if age, sex, and education level determined phishing susceptibility [11]. Sun, et. al investigated links between gender and behavior [12]. However, these studies had partici-

pants aware of the phishing tests. The users knew that they were being tested on their ability to discern phishing attacks, with the studies themselves acknowledging possible limitations on how such methodologies may have affected their results. In this study, we not only include a more expansive list of demographics, but also simulate real-world phishing attacks by not informing the participants beforehand of the phishing Experiments.

Chapter 4

EXPERIMENTAL METHODOLOGY

4.1 Subject Population

The university used in this study is the University of Maryland, Baltimore County (UMBC), located in Baltimore, MD. UMBC has 11,234 undergraduate and 2,428 graduate students [13]. This study takes the undergraduate student population currently enrolled at UMBC as the target pool of possible phishing victims. UMBC's undergraduate student population is distributed along the following major College affiliations:

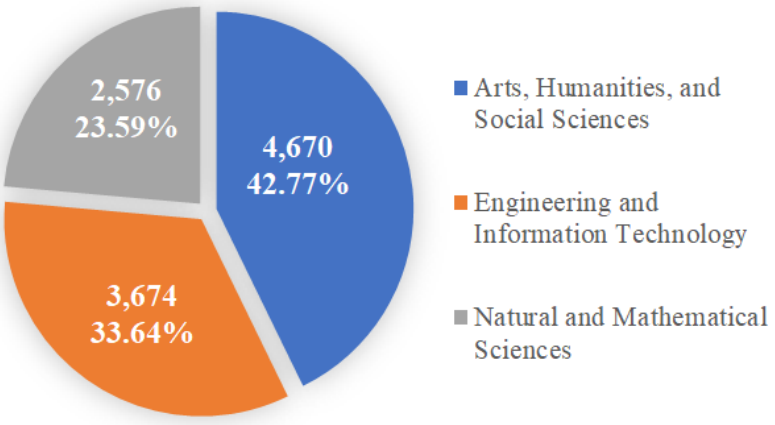


FIG. 4.1. UMBC Undergraduate Student Enrollment

UMBC holds three Colleges: the College of Arts, Humanities, and Social Sciences, the College of Engineering and Information Technology, and the College of Natural and Mathematical Sciences. UMBC boasts 48 majors, 38 minors, and 25 certificate programs within these colleges. In this study the college demographic focuses on the student's primary major, regardless of any subsequent major, minor, or certificate program [13].

Each phishing Experiment targets a pre-selected random set of students from the overall pool to create that Cohort. Each Cohort is divided up into three major subgroups, defined by the three main undergraduate colleges at UMBC. Each subgroup contains 150 students from each college, with 450 students per Cohort. In total, 1,350 students are targeted for the duration of the study.

The total number of students decreased from 11,234 to 10,920 due to disqualification purposes. Disqualifications include Interdisciplinary Studies track or marking 'Undecided' under major affiliation. These measures are set in place because Interdisciplinary Studies majors consist of multiple majors in potentially different Colleges, while the Undecided majors have yet to state which College they align themselves to.

Displayed in Table 4.1, the College of Arts, Humanities, and Social Sciences makes up roughly 43% of the total undergraduate enrollment, in part of their wide selection of majors. Due to the similarity in course design and requirements, the Management of Aging Services and Social Work majors have been included in this college. In the College of Engineering and Information Technology, over 40% of the entire college consists of Computer Science and Engineering majors, followed by the Information Systems majors group. The College of Natural and Mathematical Sciences has a majority under the Biological Sciences category and followed by the Chemistry category. The categorization of majors in all three colleges were followed under UMBC's HeadCount Enrollment documentation [14].

Major	Students	Within College	Overall
Africana Studies	7	0.15%	0.06%
American Studies	21	0.45%	0.19%
Ancient Studies	23	0.49%	0.21%
Asian Studies	46	0.99%	0.42%
Global Studies	107	2.29%	0.98%
Dance	38	0.81%	0.35%
Economics	486	10.41%	4.45%
Emergency Health Services	75	1.61%	0.69%
English	192	4.11%	1.76%
Gender and Women's Studies	8	0.17%	0.07%
Geography and Environmental Studies	225	4.82%	2.06%
History	186	3.98%	1.70%
Media and Communication Studies	236	5.05%	2.16%
Modern Languages	127	2.72%	1.16%
Music	127	2.72%	1.16%
Management of Aging Services	404	8.65%	3.70%
Philosophy	30	0.64%	0.27%
Political Science	237	5.07%	2.17%
Psychology	894	19.14%	8.19%
Social Work	382	8.18%	3.50%
Sociology	381	8.16%	3.49%
Theater	87	1.86%	0.80%
Visual Arts	351	7.52%	3.21%
Total	*4,670*	*100.00%*	*42.77%*
Computer Science and Engineering	1,500	40.83%	13.74%
Information Systems	1,132	30.81%	10.37%
Mechanical Engineering	585	15.92%	5.36%
Chemical Engineering	323	8.79%	2.96%
Pre-Engineering	134	3.65%	1.23%
Total	*3,674*	*100.00%*	*33.64%*
Biological Sciences	1,671	64.87%	15.30%
Chemistry	433	16.81%	3.97%
Math and Statistics	339	13.16%	3.10%
Physics	133	5.16%	1.22%
Total	*2,576*	*100.00%*	*23.59%*
Total	**10,920**	**100.00%**	**100.00%**

Table 4.1. UMBC College Major Breakdown

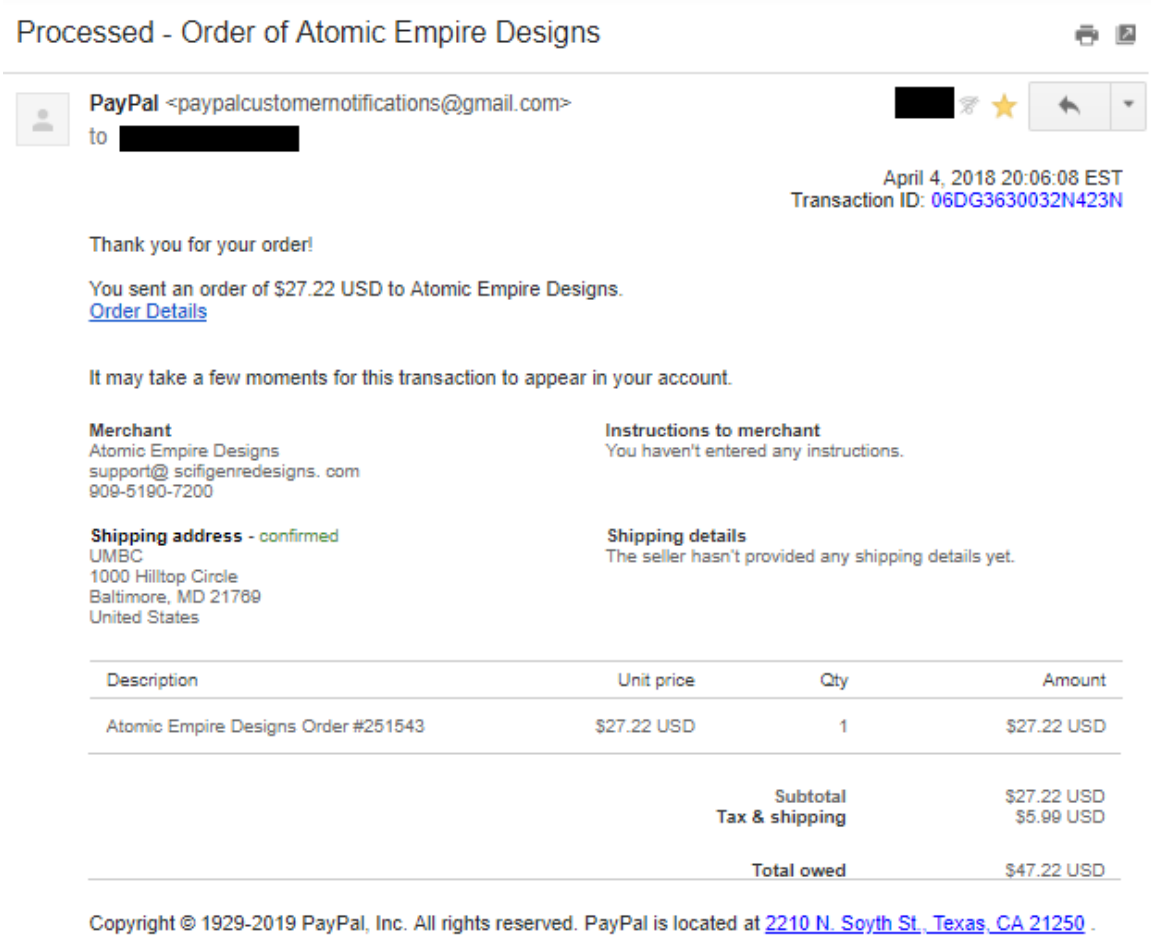
4.2 Experiment 1: PayPal

Experiment 1 is the first phishing email sent, using the popular *Billing Problem* tactic [5]. The *Billing Problem* is a very popular phishing technique, mostly due to the recipient trying to resolve any monetary issues as quickly as possible. This results in many users blindly clicking any links and giving away personal information before going over the email contents with care.

Shown in Figure 4.2, the fraudulent entity claims to be PayPal, a popular online payment company. The email tries to entice the user to click on the email link by claiming to have received an order from them and therefore billing their PayPal account. The fake order confirmation is problematic for the user because they have not ordered such an item, regardless if they have an existing PayPal account or not.

While this email is meant to look authentic, there are several red flags that indicate this email as illegitimate. There is no such company as Atomic Empire Designs. The fake company's customer service email isn't valid, nor is their phone number (which has an extra digit). Another detail is outlined in the "Shipping Address" section to UMBC. The address is vague enough that the package would not reach the student if they lived on-campus, or would be an incorrect address if the student lived off-campus. The zip code also does not pertain to Baltimore, MD but to another city altogether. The email time stamp is a time that hasn't passed yet, which marks the email as illegitimate. The total amount of money owed does not add up to the subtotal + Tax and shipping expenses. The last line of the email that states "Paypal is located at ..." is not a valid address, let alone Paypal's legitimate address.

Two less subtle markings are the link provided and the sender's email address. The email came from a "paypalcustomernotifications@gmail.com" address. However, any email from the PayPal business will have a @paypal.com address, not a gmail.com one. The link described as Order Details is also suspicious. If one hovers over the link, it does



Processed - Order of Atomic Empire Designs

PayPal <paypalcustomernotifications@gmail.com>
to [redacted]

April 4, 2018 20:06:08 EST
Transaction ID: 06DG3630032N423N

Thank you for your order!

You sent an order of $27.22 USD to Atomic Empire Designs.
Order Details

It may take a few moments for this transaction to appear in your account.

Merchant
Atomic Empire Designs
support@ scifigenredesigns. com
909-5190-7200

Instructions to merchant
You haven't entered any instructions.

Shipping address - confirmed
UMBC
1000 Hilltop Circle
Baltimore, MD 21769
United States

Shipping details
The seller hasn't provided any shipping details yet.

Description	Unit price	Qty	Amount
Atomic Empire Designs Order #251543	$27.22 USD	1	$27.22 USD

Subtotal	$27.22 USD
Tax & shipping	$5.99 USD
Total owed	$47.22 USD

Copyright © 1929-2019 PayPal, Inc. All rights reserved. PayPal is located at 2210 N. Soyth St , Texas, CA 21250 .

FIG. 4.2. Email used in the first Experiment. The email uses the Billing Problem tactic to inform the user of an unsolicited order from ”PayPal”.

not indicate any association with PayPal at all. Instead, it goes through a tracking url that contains a ”thisisnotmalware” string. Several of these indicators are subtle, but the combination of these red flags are enough where a user should be wary of this generic phishing email.

4.3 Experiment 2: Quadmania

In this Experiment the user is lured by monetary gain [5]. This gain, aptly named the *Contest Winner* tactic, congratulates the user for winning a contest that they did not sign up for. This email makes use of UMBC's Quadmania event, the university's big spring weekend festival. The email uses key information about Quadmania, including the Quadmania banner and the different festival activities found online at UMBC's website.

The Quadmania email congratulates the student on their $100 Amazon prize. It then urges the student to click the provided link so they could fill out private information to cash in on their prize. This email adds legitimacy by using the 2018 Quadmania banner. Also, the signature of the email proclaims it was sent by the UMBC Events Board. This name is similar to (seb), the Student Events Board that organizes Quadmania at UMBC.

Students at UMBC can recognize details that undermine the phish. For one, there was no prior UMCP survey at all, so the student would recognize that they have not participated nor selected they be included in the prize drawing. Futhermore, UMCP refers to another college — University of Maryland, College Park — instead of their own college. There are grammar and spelling inconsistencies, e.g., the keynote singer for the Quadmania concert, 21 Savage, is misspelled.

Another discrepancy comes from the email stating that the concert will be on Friday, even though the banner clearly states it takes place on Sunday. The inconsistency is a glaring red flag due to the promotion, marketing, and general student awareness of what is

Your our Quadmania Winner! Inbox x

UMBCEvents <nbrooks@umbcalerts.com>

to [redacted]

Dear UMBC Student,

Congratulations!!

If you are receiving this message, this means you were randomly selected to win one of our five **$100 Amazon** gift card prizes as part of our Quadmania 2018 prize giveaway contest!

Due to your earlier participation in the UMCP survey on March 11th, you indicated you wished to be included in this prize drawing.

Please click here to fill out your information to cash in your prize!

Congratulations and don't forget about our Quadmania festival and 21 Savege concert this Friday!!

- UMBC Events Board

FIG. 4.3. Email used in the second Experiment. The email uses the Contest tactic and congratulates the user on an Amazon gift card prize as part of ”Quadmania”.

arguably the most important social series of events that UMBC offers its students. The user can see the link redirects them to cnn.com after going through a tracking software. The email is sent from a @umbcalerts.com address, which causes suspicion because UMBC has every email address is under @umbc.edu, not any .com address.

4.4 Experiment 3: DoIT

The user receives a notice for account verification at the threat of account suspension [5]. The third email is a variation of the well known *Expiration Date* tactic. The authoritative entity, which mimics UMBC's Division of Information Technology, claims the user must verify their credentials to keep their data and UMBC account. A reference to the Quadmania phish adds validity to this email. The email states that this action is required under a small time window, creating an added sense of urgency.

As with all phishing Experiments, there are warnings in the content that alerts a vigilant student. There are several spelling and grammar errors, which is uncommon for an entity like UMBC to commit. The authority mimicking the Division of Information Technology names itself "Department of Institutional Technology", and later signs off with "UNCP DoIT". Apart from there being no Department of Institutional Technology or UNCP entity at UMBC, the inconsistencies are present.

Another inconsistency, shown in Figure 4.4, is the warning of having both 24 and 48 hours to do this action. An odd addition to this email is the quote near the end: *"New technology is not good or evil in and of itself. It's all about how people choose to use it"*. This quote is very out of character and unconventional for a university's IT department. The email address and link of this email are suspicious, just like in the previous two phishing Experiments. The link goes to the Google homepage after going through tracking software and the email address has a @umbcdoit.com email address instead of a @umbc.edu one.

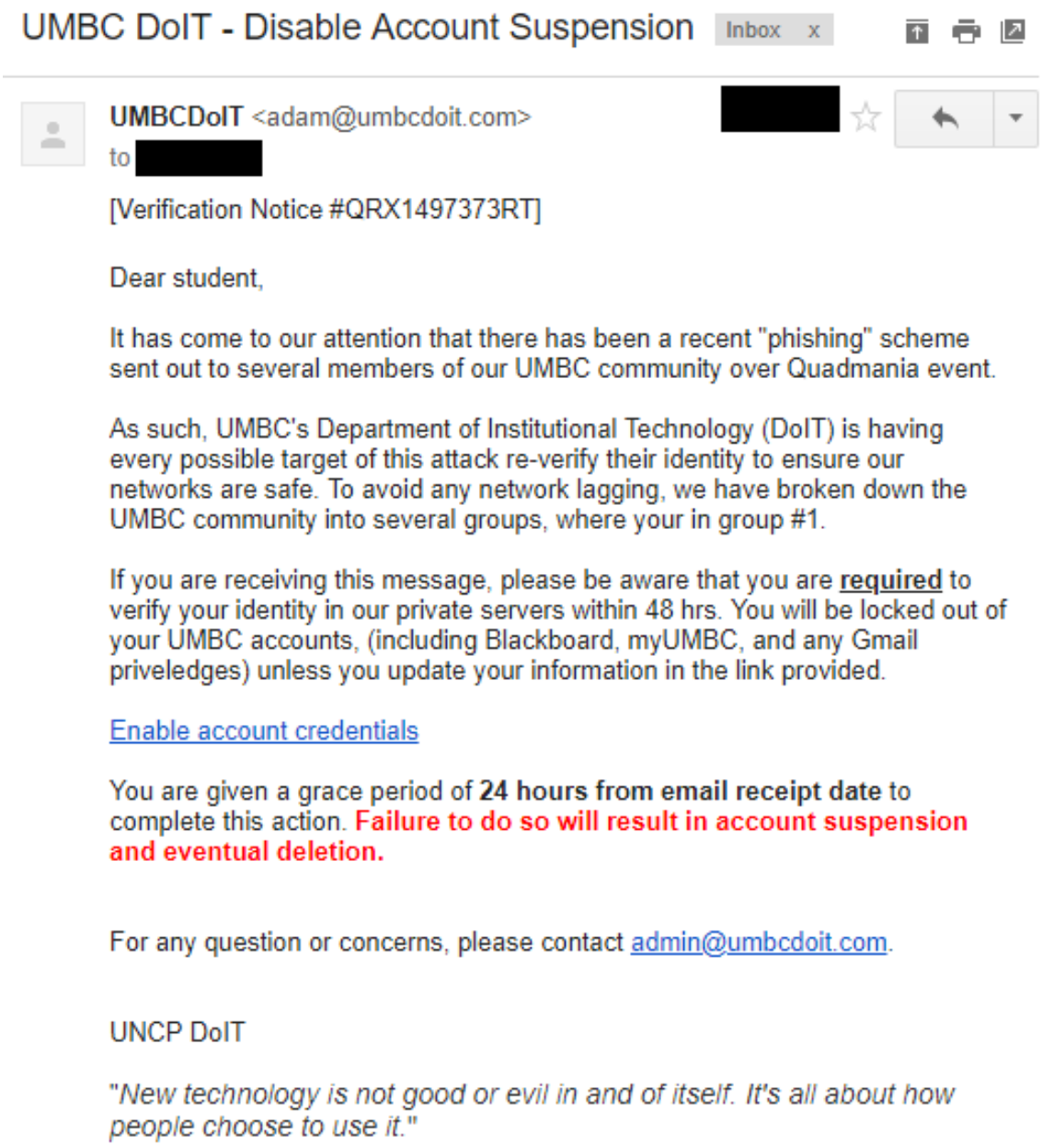


FIG. 4.4. Email used in the third Experiment. The email uses the Expiration Data tactic to threaten the user of account deletion while impersonation UMBC's DoIT team.

4.5 Demographic Survey

When all phishing Experiments are concluded and data gathered, an email detailing the study is sent to all students in Cohorts 1-3. Part of the Institutional Review Board (IRB) protocol is for this email to contain a debriefing statement that informs the all 1,350 selected students of the study and assuring that all data collected is kept anonymous and unable to identify them as an individual. The students are notified of emails that were sent throughout the semester that seemed suspicious or off. Students are made aware that they have been tested on their ability to identify phishing attacks directed at their UMBC emails as part of a graduate study.

This debriefing statement informs them specifically that they may have been deceived into clicking a link through a false banking or UMBC email directed at them. They are also informed that the purpose of this study is to see the effects of user susceptibility on user behavior through simulated phishing attacks and that the results are evaluated to see possible trends among the UMBC community.

Apart from the debriefing statement, students who were part of the 1,350 target group but had also opened a phishing email from Experiments 1-3 have the ability to participate in a survey. The purpose of the survey is to collect additional demographic data from the students in the Cohorts who interacted with the phishing emails. After asking for consent to provide data to this survey and ensuring that those answering the survey are at least 18 years of age, questions are asked on their academic year, major affiliation, gender, age, past cyber security training, participation in cyber clubs or cyber scholarship programs, phishing awareness, and time spent per day on the computer.

Once completing the survey, the student is given a brief definition of what a phishing attack is and quick tips on how to identify a phishing email. The user, if they so choose, is directed to the official UMBC phishing and spam FAQ page for more information.

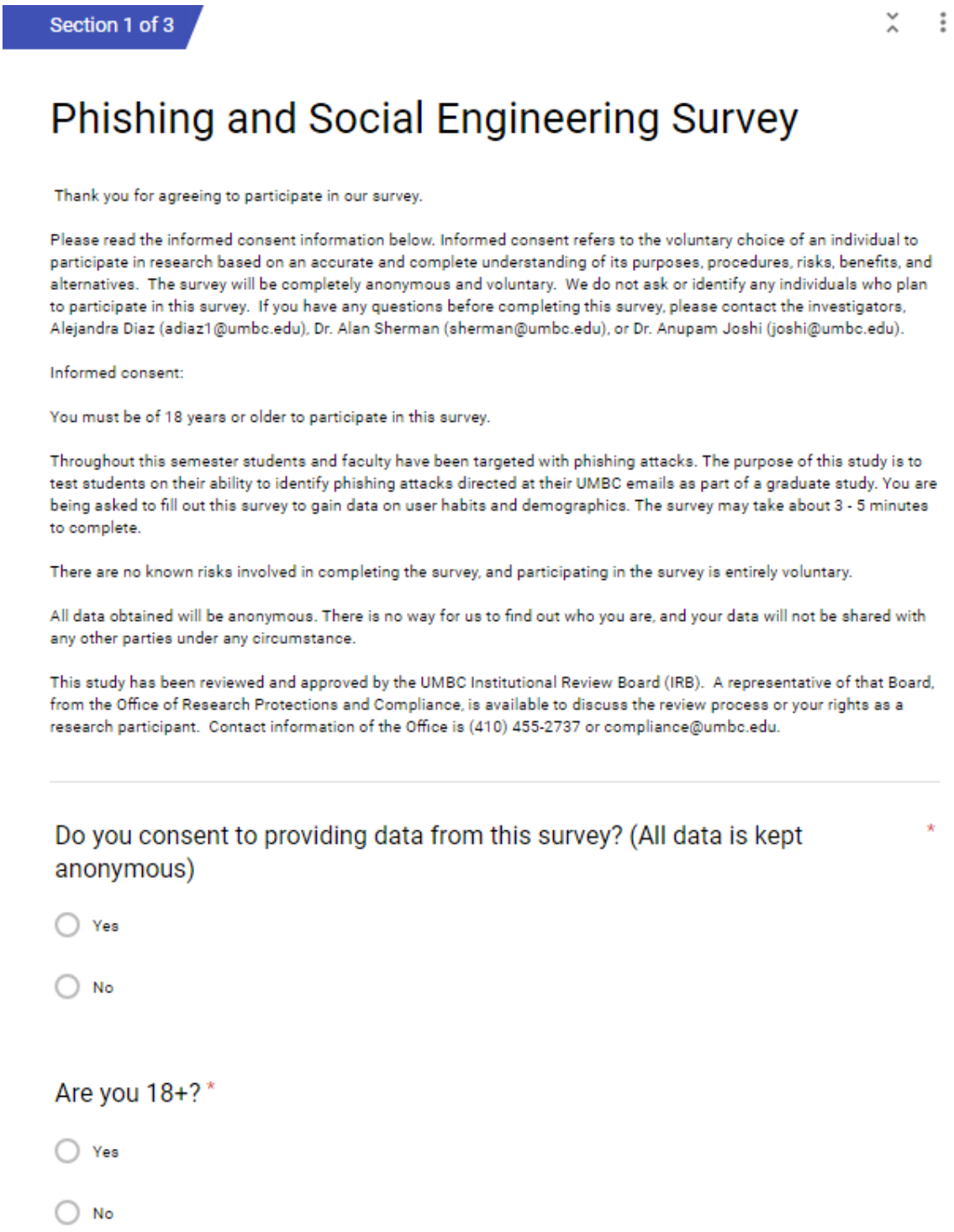


FIG. 4.5. First part of survey that informs the user of the study and required the user to consent to the collection of demographic data and be over 18 years of age

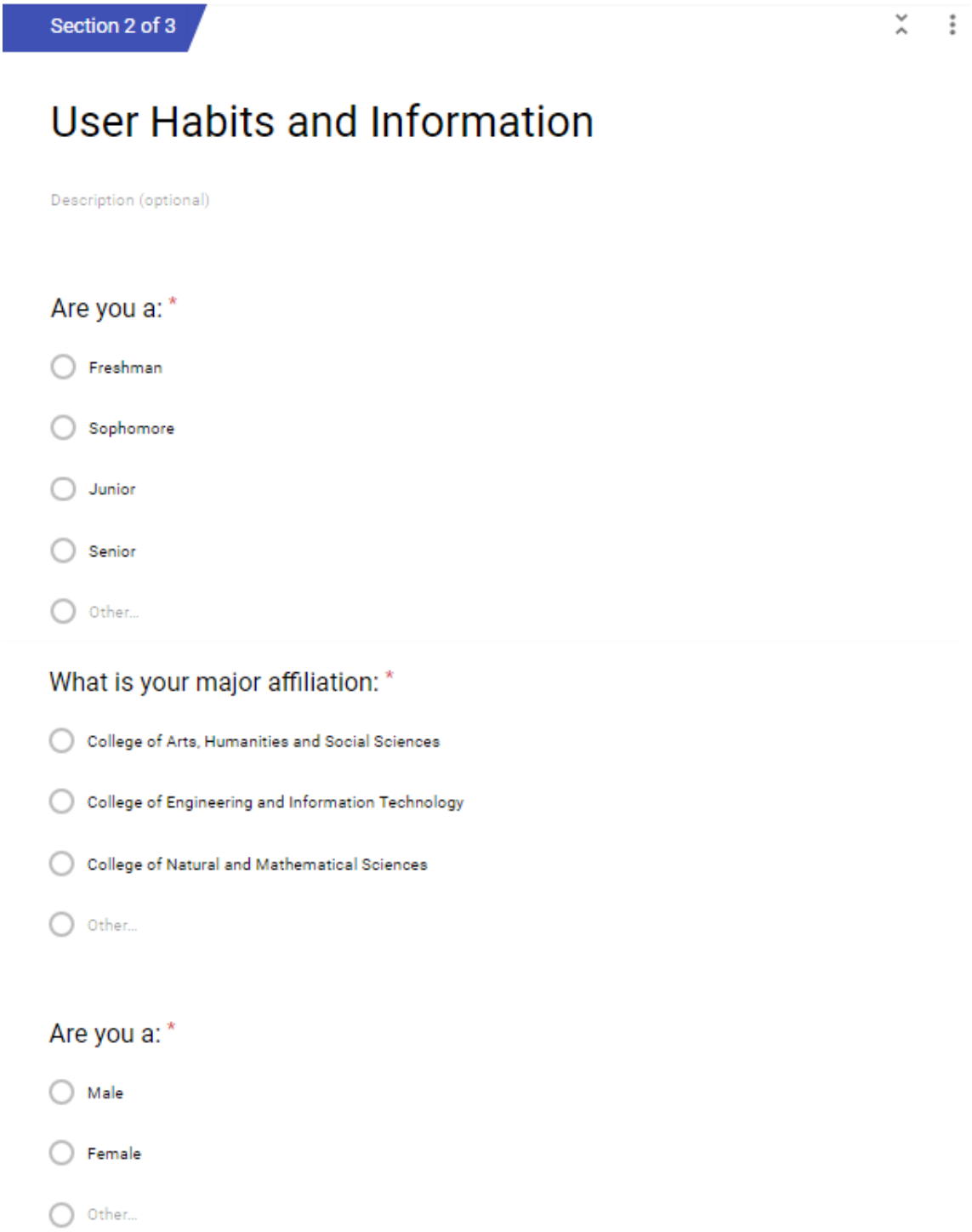


FIG. 4.6. Collects data on user habits and demographics

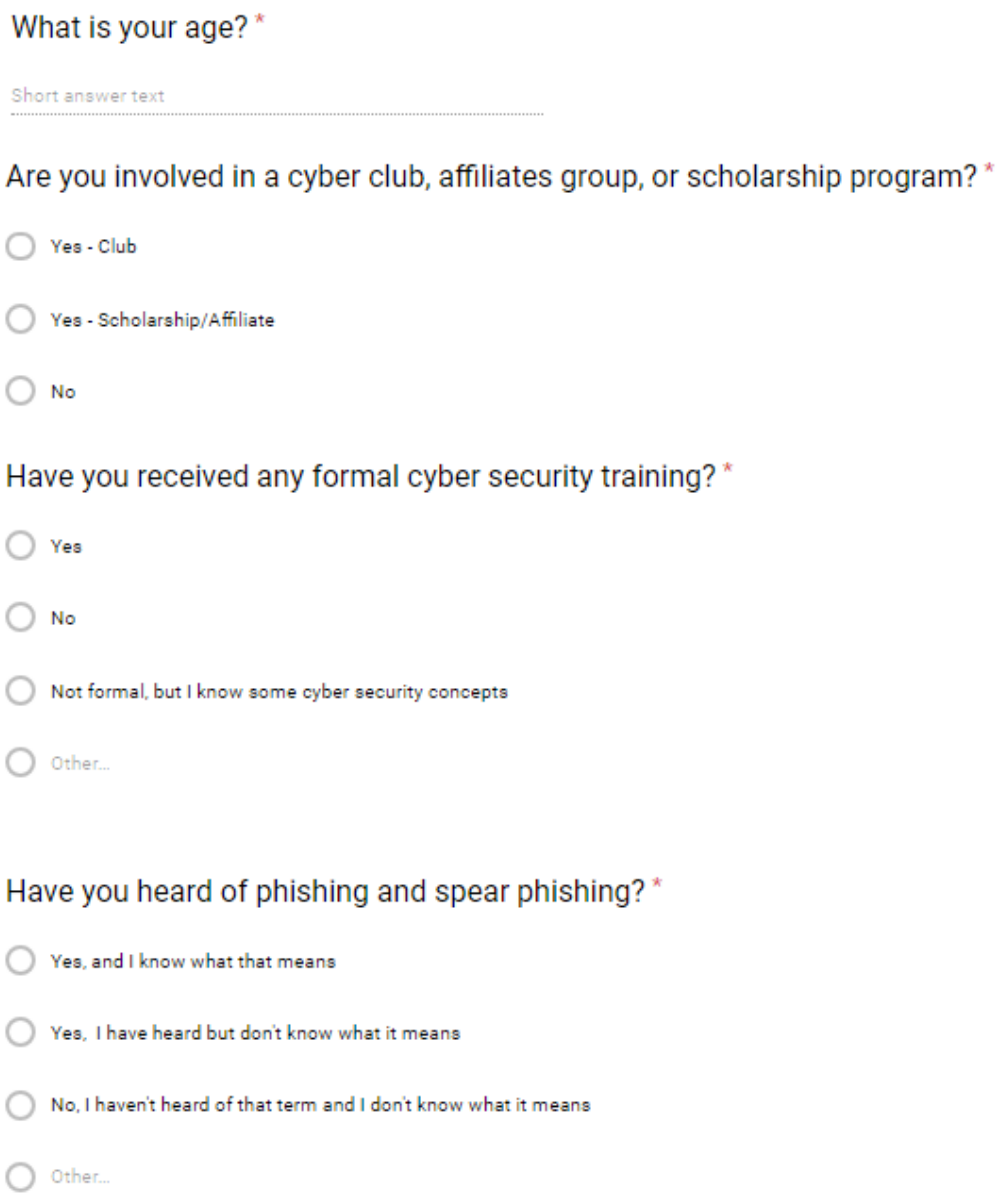


FIG. 4.7. Continued user habits and demographics questions

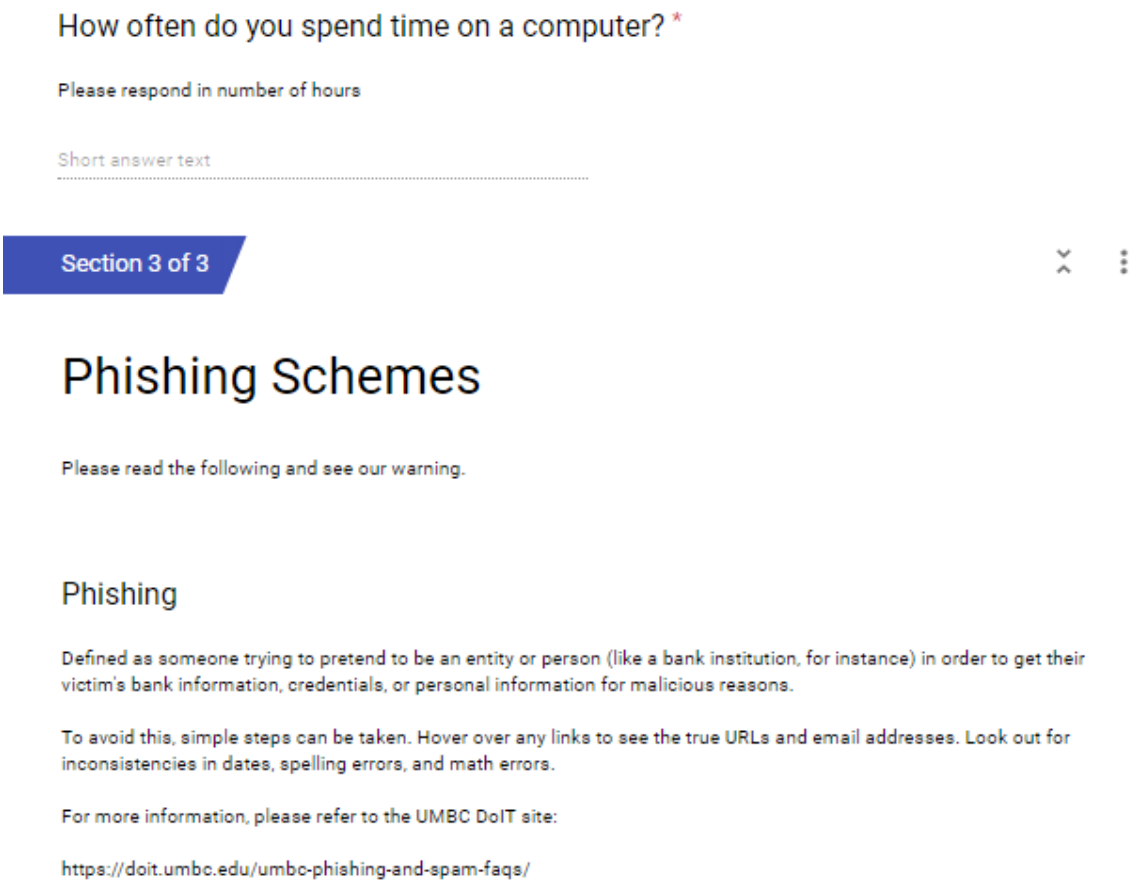


FIG. 4.8. The last section of the survey provides students with quick tips against phishing attacks and encourages students to learn more at UMBC's designated Phishing FAQs page.

4.6 Statistical Methods

To analyze the results, we use the free application MailTracker by Hunter and the EmailTracker by cloudHQ [2][7]. MailTracker by Hunter, when used as an add-on to Google's Gmail, is able to track down not only each instance that an email recipient has opened an email, but also the time, location, and device that they used to open the email. Another beneficial feature that MailTracker has is the ability to also track the instances and devices used to click any links within the emails sent. For EmailTrack by cloudHQ, this Chrome extension lets the email sender know when the email recipient last read the email, how many times the email has been read, whether they clicked any links or attachments within the email, and how many times said links or attachments were clicked. While these two add-ons are extremely similar in function, both are used to verify and confirm each other's recorded data.

We utilize Fisher's Exact test and Pearson's Chi-Square for significance testing and Cramer's V to test strength of that significance, with $\alpha = 0.05$. Fisher's Exact test is used in lieu of the Chi-Square test when an expected value is less than 5. We define the null hypothesis as there is no dependency between the demographic factors and student click rate. We use IBM's SPSS to create contingency tables and calculate these measurements.

Chapter 5

NUMERICAL RESULTS

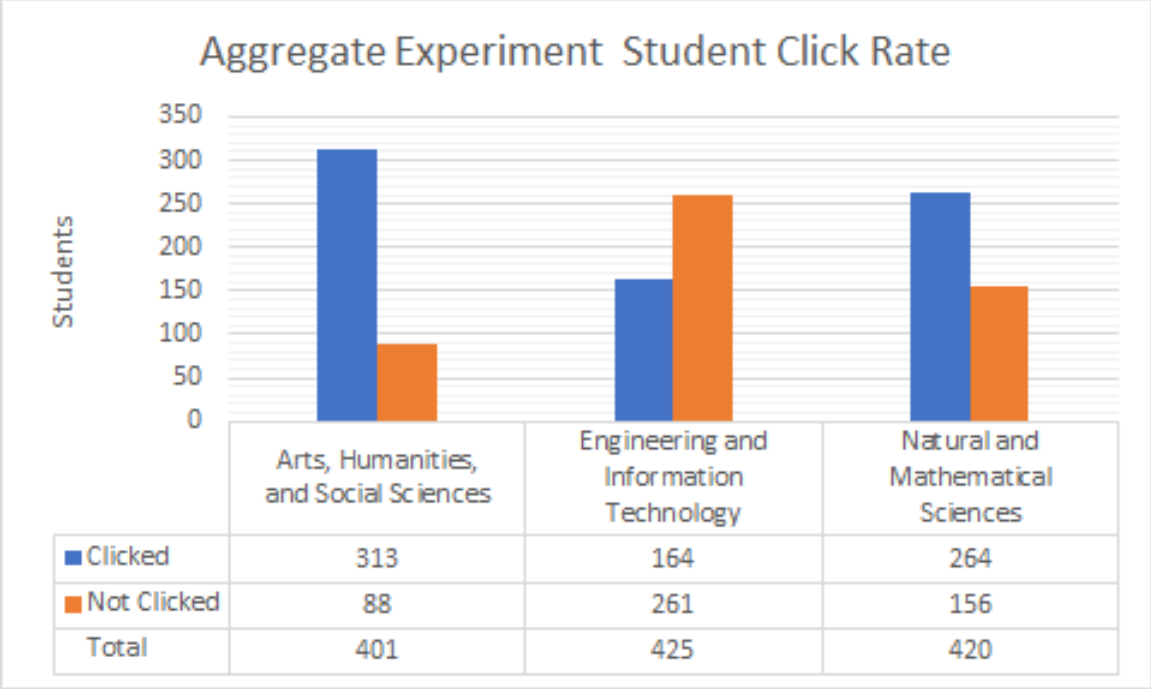


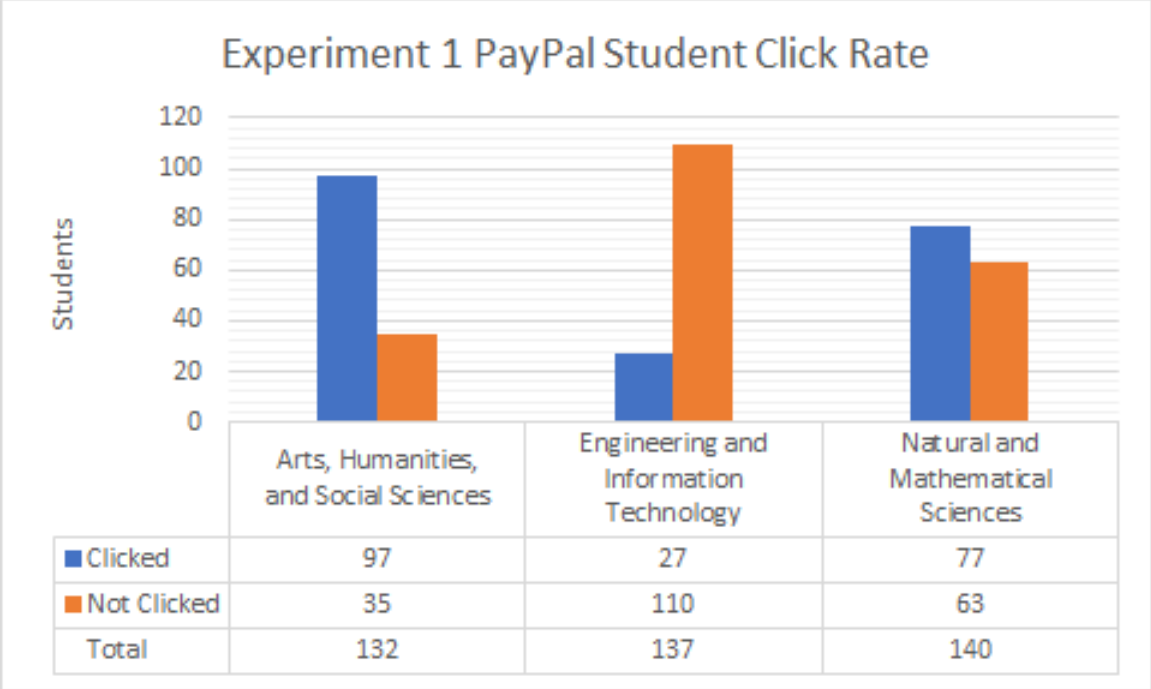
Table 5.1. Overall amount of student click actions for the three Experiments

Of the 1,350 students randomly selected for this study, 1,246 (92%) opened a phishing email throughout the three Experiments. Shown in Table 5.1, students in the Arts, Humanities, and Social Sciences College had higher click rates than their peers. Natural

and Mathematical Science majors followed, with Engineering and Information Technology students clicking the phishing link the least.

5.1 Experiment 1 Initial Data

Of the 450 students that are sent the PayPal phishing email, 409 (91%) opened the email. Of those 409 students, a majority of the Arts, Humanities, and Social Sciences majors seemed to click the link despite the warning signs. The Natural and Mathematical Sciences students were split roughly half and half for those that clicked vs not clicked, while the Engineering and Information Technology students not falling for the attempted phish as often.



	Arts, Humanities, and Social Sciences	Engineering and Information Technology	Natural and Mathematical Sciences
Clicked	97	27	77
Not Clicked	35	110	63
Total	132	137	140

Table 5.2. Amount of students from first Experiment (PayPal) divided into the three Colleges and user action

The PayPal email was sent to 150 students within each college. The total amount of students, however, was less than 150 because only students who have opened the email are eligible to be analyzed for associations. Arts, Humanities, and Social Sciences majors had 88% of targeted students reading the email, with 91.3% in Engineering and Information Technology and 93.3% in Natural and Mathematical Sciences.

Each college is broken down by major, shown in Figure 5.1. The major distribution leans towards the majors that have the highest headcount, which is to be expected. Within the College of Arts, Humanities, and Social Sciences, 132 students out of 150 (88%) opened the email. The Visual Arts, Psychology, and Political Science majors contributes the highest amount of click instances. However, Theater, Media and Communications, History, Geography and Environmental Studies, Dance, and Global Studies have a higher click per student rate in proportion to each majors' group size.

In a similar fashion, the College of Engineering and Information Technology has a major breakdown leaning towards the majors with a higher headcount. Of the 137 students (91%) who opened the email, around 110 (80%) did not click the link. Shown in orange in Table 5.1, only 27 out of the 137 (20%) were deceived by the PayPal email.

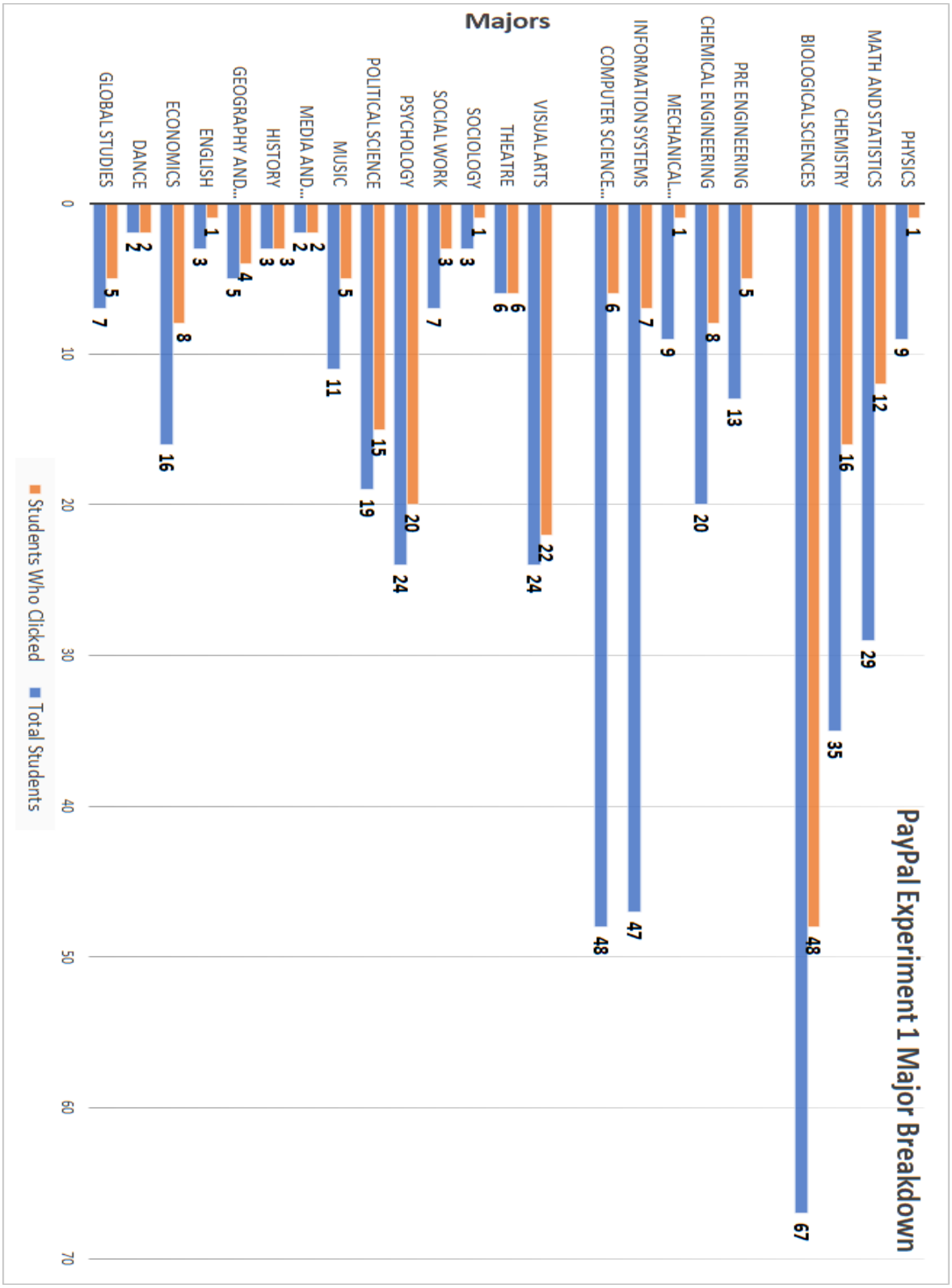
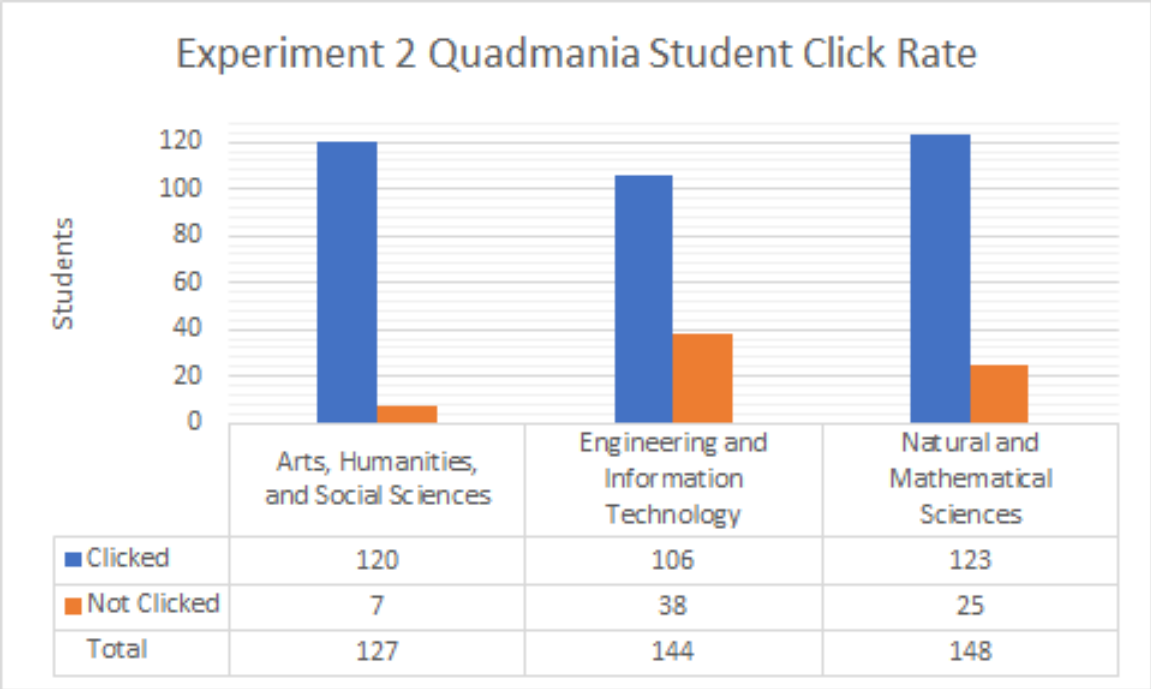


FIG. 5.1. Amount of students divided by major within the three colleges from first experiment.

5.2 Experiment 2 Initial Data

Included in Table 5.3 are the results for the Quadmania Phishing scheme for the second Cohort. Of the 450 students that were sent the Quadmania phishing email, 419 (93%) opened the email. only 70 of the entire 419 students (16.7%) did not click the Quadmania email. Of those 419 students, almost all of the Arts, Humanities, and Social Sciences majors (30%) seemed to click the link (95%), often within minutes of sending the email.



	Arts, Humanities, and Social Sciences	Engineering and Information Technology	Natural and Mathematical Sciences
Clicked	120	106	123
Not Clicked	7	38	25
Total	127	144	148

Table 5.3. Amount of students from second Experiment (Quadmania) divided into the three Colleges and user action

The two other Colleges also saw an increase in students who fell for the scheme than in Experiment 1. The College of Engineering and Information Technology had 74% of students clicking the link, while the College of Natural and Mathematical Sciences observed 83% click rates.

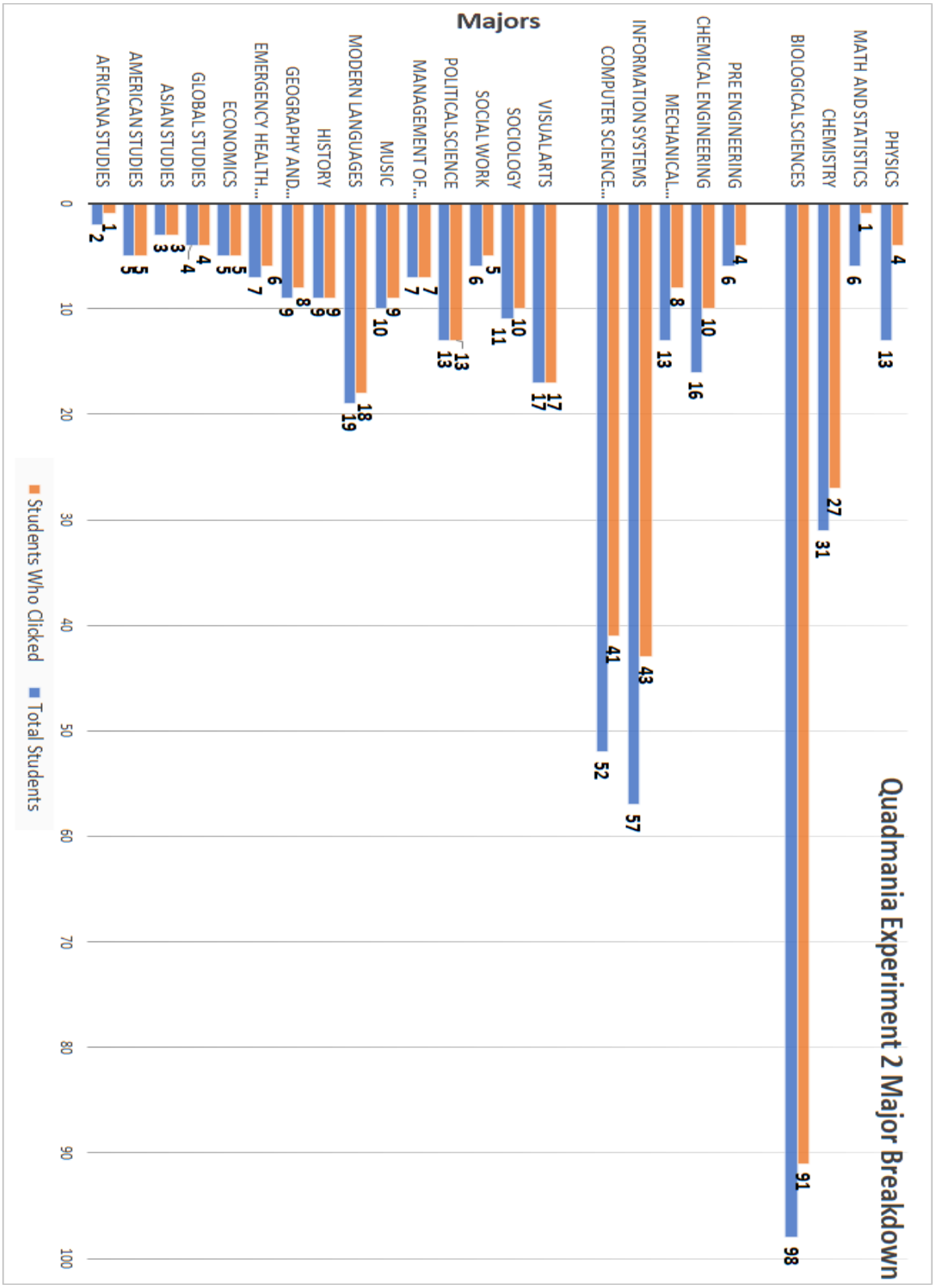
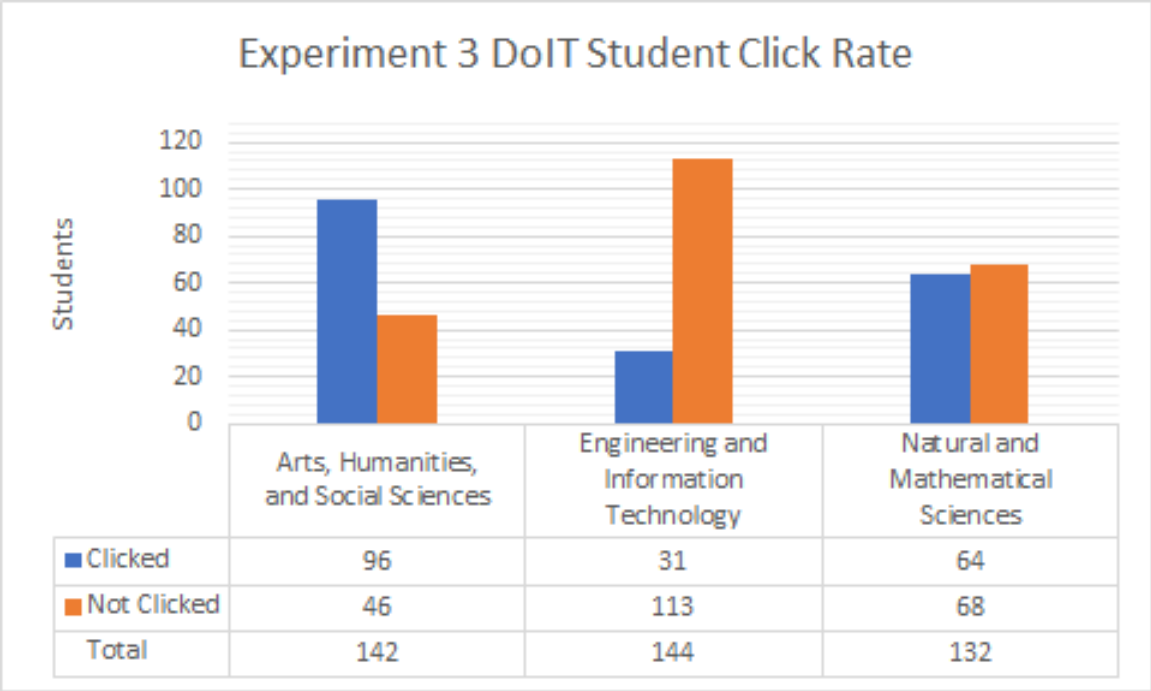


FIG. 5.2. Amount of students divided by major within the three colleges from second experiment.

5.3 Experiment 3 Initial Data

93% of students opened the third email. The Arts, Humanities, and Social Sciences and Natural and Mathematical Sciences majors seemed to have been fooled a moderate amount (48%). In contrast, the Engineering and Information Technology majors had a very low click rate within their college, where only 31 people were deceived into clicking the link (22%).



	Arts, Humanities, and Social Sciences	Engineering and Information Technology	Natural and Mathematical Sciences
Clicked	96	31	64
Not Clicked	46	113	68
Total	142	144	132

Table 5.4. Amount of students from the third Experiment (DoIT) divided into the three Colleges and user action

Similar to Experiment 1, the College of Arts, Humanities, and Social Sciences had its majors at an above average click rate (68%). Exceptions were Sociology and Ancient Studies majors. Some majors, like Visual Arts, Social Work, Philosophy, and Economics, were observed with a high click rate proportional to their sample size.

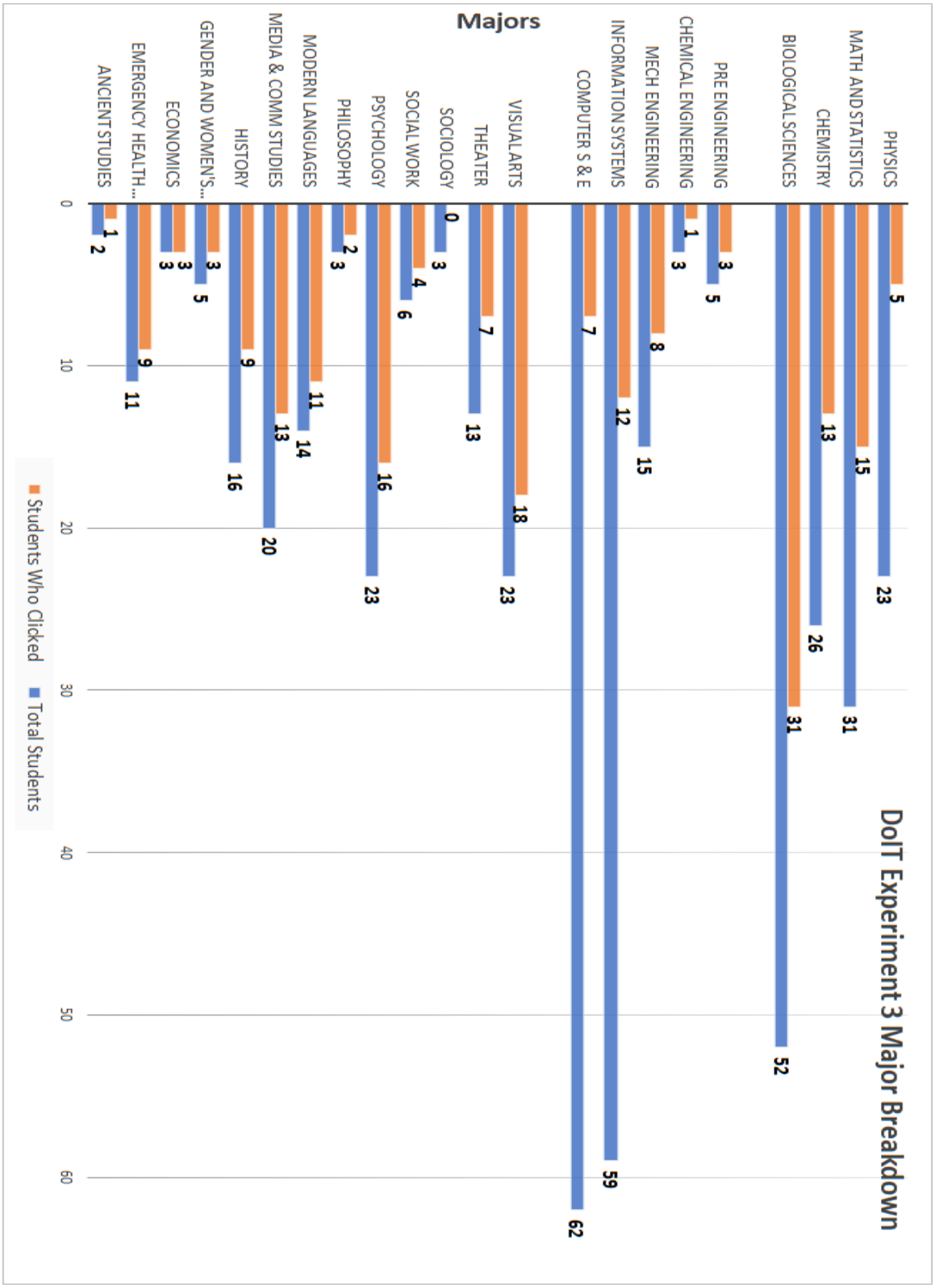


FIG. 5.3. Amount of students divided by major within the three colleges from third experiment.

5.4 Survey Results

Of the 1,246 students who had the option to complete the survey, 482 students (39%) responded within a 7 day period. Based on the survey alone, the following demographic breakdown was gathered:

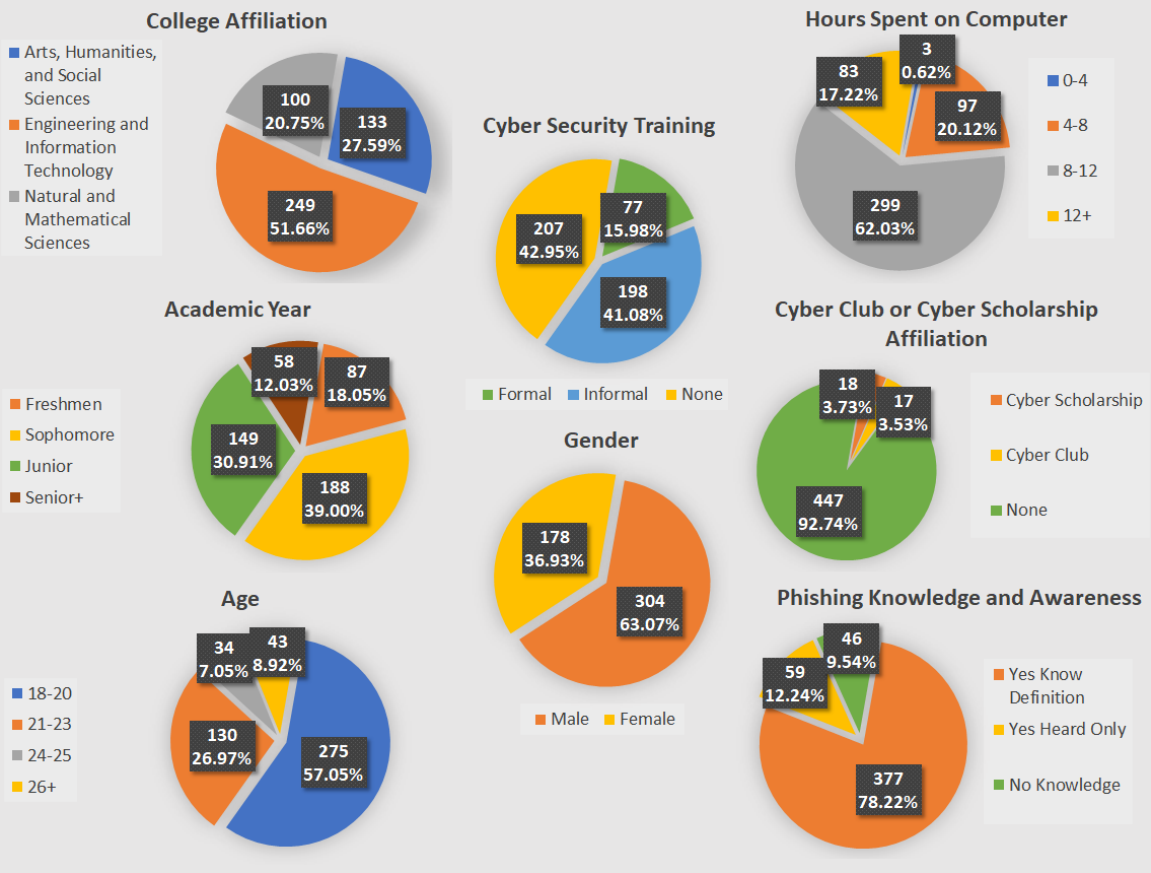


FIG. 5.4. Data Breakdown of Surveyed Demographics

Displayed in Table 5.5, each Cohort had at least 100 respondents who completed the survey. In Experiment 1, there was an even split between those who fell for the phishing

scheme and those who did not. For the last two phishing schemes, however, the respondents were heavily skewed towards those that were deceived by the emails.

Phase	PayPal	Quadmania	DoIT
Clicked	47	176	116
Did Not Click	55	49	39
Total	102	225	155

Table 5.5. Survey Experiment Breakdown of all three experiments.

5.5 Experiment 1 Results With Survey Data

College Affiliation

Between the three colleges, there were more Engineering and IT students that completed the survey, making up roughly 56% of the total number respondents in Experiment 1. The 102 students had varying click rates within their major, much like their Cohort counterparts. Overall, the College of Engineering and Information Technology had the lesser click rate compared to the other two colleges.

Academic Year

There were more Juniors (34%) than any other year for the first Cohort respondents. There were varying click rates, with Seniors (29% clicked) and above having the least amount of students falling for the PayPal email in comparison to its sample size.

Gender

There was an almost even distribution between the male and female survey respondents. Female survey respondents had a 50/50 click rate, while the male group had a similar click rate of 55/45.

FIG. 5.5. Experiment 1 surveyed demographic results. Each demographic is portrayed as percentages based on their individual sample size.

Age

The age category was split into four main groups. For the age groups between the ages of 18 - 23 and 26+, the click rate percentages were extremely similar. The discrepancy was in the 24 - 25 age group where more students clicking the PayPal link than not (65% clicked).

Hours Per Day Spent on Computer

Another question in the survey pertained to how many estimated hours a day a student spent on their computer. The answers ranged from 0 - 18 hours a day. The students within the first Cohort had no one within the 0 - 4 hour range. A majority of students spent from 8 hours up to 12 hours a day on the computer (74%), with a lesser click rate occurring as hours increased.

Cyber Club or Scholarship Affiliation

This demographic is their participation in a cyber security scholarship program or cyber club. Within UMBC, there are the Cyber Scholars and the Scholarship For Service (SFS) Scholars. These scholarship programs provide students extra support from faculty, staff, and peers with opportunities in research and internships within the cyber security field. Many students in these scholar programs also take cyber security electives during their time at UMBC. The click rates were very similar (club = 50%, None = 46% clicked), with the exception of the one Cyber/SFS Scholar that did not click the link.

Phishing Knowledge and Awareness

Students were asked if they have heard about phishing attacks and whether they understood what a phishing attack was. Shown in Figure 5.5, around 75% both heard about and knew what a phishing attack was, with 6% of students not knowing nor having heard anything about a phishing scheme.

Cyber Training

Much like age and gender, there was roughly an even division between those that clicked the link for Experiment 1 and those who did not. Not surprisingly, students with no training whatsoever had more students click the link (83%). Those who had formal training (51%) and informal training had similar click rates to one another (16%) had considerably lower click rates.

5.6 Experiment 2 Results With Survey Data

College Affiliation

There was a great increase in click rates across all three colleges for this Experiment. The 225 students had an overwhelming majority of students clicking the link. In this demographic, 78% of the students clicked the phishing link.

Academic Year

There was a similar click rate per student for the academic year demographic. All four years had a percentage rate between 15% to 25%.

Gender

In this demographic both male and female students clicked more often than not. Experiment 2 had similar results to Experiment 1 in difference between the male and female groups. Both male and female groups had a high click rate, although the rates were close together. Female survey respondents had a 20% click rate, while the male group had a close click rate of 33%.

Age

The age category had a slightly increasing rate of students who did not click the link as their age progressed. However, the rate of students who did not click increased substantially compared to its sample size for the 26+ age group.

Hours Per Day Spent on Computer

Much like in Experiment 1, the click rate of students declined as the number of hours per day spent on the computer increased.

Cyber Club or Scholarship Affiliation

There was a very small difference in percentages between students who joined a cyber club (75% clicked) versus students with no affiliation to a cyber club or scholar program (79% clicked). The Cyber/SFS scholars, however, had a low click rate for the Quadmania email (33% clicked).

Phishing Knowledge and Awareness

Surprisingly, the less students knew about phishing, the lower the click rates were. Students that had both phishing awareness and knowledge had a very high click rate (89%) compared to students who have only heard of phishing attacks (56%). Students who had no idea what a phishing attack was fared better against the Quadmania phish (27%).

Cyber Training

We collected that students with more cyber training had a lesser click rate. Students fared better with formal training (61%) versus informal training (70%), with both being superior in low click rates than students with no training whatsoever (88%).

FIG. 5.6. Experiment 2 surveyed demographic results. Each demographic is portrayed as percentages based on their individual sample size.

5.7 Experiment 3 Results With Survey Data

College Affiliation

Students in the College of Arts, Humanities, and Social Sciences (93%) and the College of Natural and Mathematical Sciences (95%) contributed to a high click rate. The College of Engineering and Information Technology (47%) had the lesser click rate, with a difference of over 45%.

Academic Year

There was a difference in click rates for students in separate academic years. The click rates decreased as students' year increased, with a jump for students in the junior year.

Gender

In this demographic female students clicked less than their male counterparts. Male students fell for the DoIT phishing scheme over 20% more than female survey respondents.

Age

Similar to the past Experiments, the age category had a slightly increasing rate of students who did not click the link as their age progressed.

Hours Per Day Spent on Computer

The click rate of students declined as the number of hours per day spent on the computer increased. The discrepancy to this trend are students belonging to the 0 - 4 hours per day group. This group had a high percentage due to the sample size being 3 students.

Cyber Club or Scholarship Affiliation

This Experiment saw a rise in club member students abstaining from clicking the DoIT link, surpassing the Cyber and SFS Scholar group. Students with no affiliation, however, stayed with a large click rate for their sample size.

FIG. 5.7. Experiment 3 surveyed demographic results. Each demographic is portrayed as percentages based on their individual sample size.

Phishing Knowledge and Awareness

The inversely proportional trend continued on to Experiment 3. Every student who responded that they had no prior knowledge to phishing attacks did not click the link. The click rates increased in this demographic as students reported that they understood and are aware of phishing schemes.

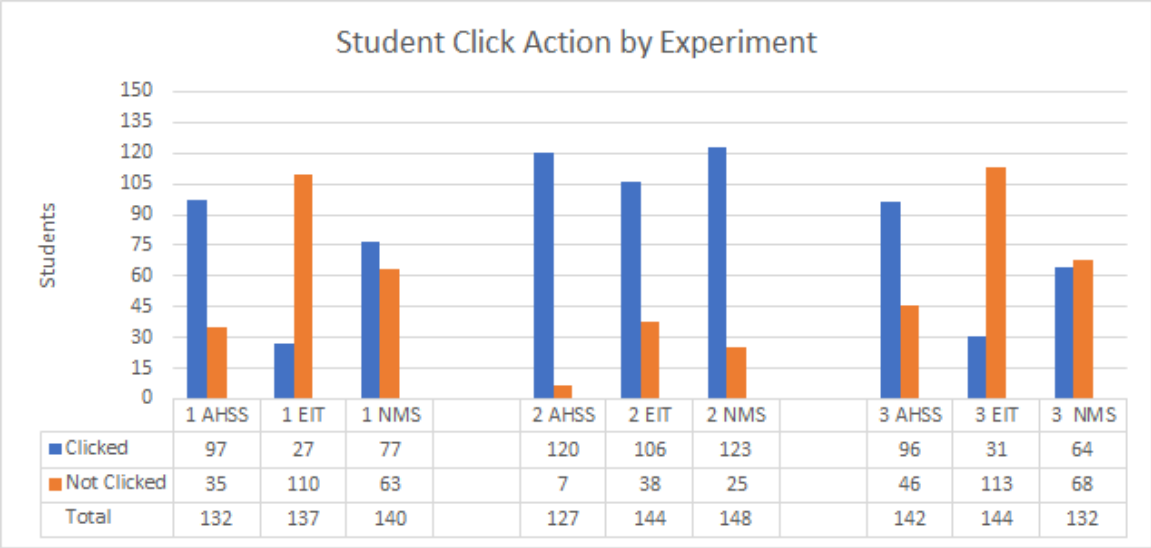
Cyber Training

Experiment 3 continued with the tendency of more cyber training and lesser click rates. Students fared better with formal training versus informal training, with both being superior in low click rates than students with no training whatsoever.

5.8 Aggregate Results

In previous sections each portion of the three experiments has been analyzed for association. However, the demographics from each Experiment have not been aggregated together to determine the impact the entire demographic could have on a student's probability on clicking a phishing email. Experiments 1 and 3 show very similar data for the three colleges. The College of Engineering and Information Technology has a high amount of students (79%) not clicking the links within the two emails. Conversely, the College of Arts, Humanities, and Social Sciences has high click rates for all three phishing schemes (30% not clicked).

The results gathered in Experiment 2 displays an astounding amount of students clicking the fake link. Despite the increase in student clicks, the colleges follow a trend. Non-STEM majors are shown to have higher click rates than their STEM counterparts, with Engineering and Information Technology students having the lowest click rates of the three colleges. This tendency is portrayed in the percentages comparative to the colleges' sample sizes. Each college has a similar ranking for each portion of the study. Despite Experiment



	1 AHSS	1 EIT	1 NMS	2 AHSS	2 EIT	2 NMS	3 AHSS	3 EIT	3 NMS
Clicked	97	27	77	120	106	123	96	31	64
Not Clicked	35	110	63	7	38	25	46	113	68
Total	132	137	140	127	144	148	142	144	132

Table 5.6. Student click numbers are presented by Experiment number and by College, with AHSS referring to the College of Arts, Humanities, and Social Sciences, EIT referring to the College of Engineering and Information Technology, and NMS referring to the College of Natural and Mathematical Sciences.

2's rise in click rate, Engineering and IT majors had less of a percentage in click rates than Mathematical and Statistics majors, who they themselves had a lesser percentage click rate than their Humanities peers. This conclusion can be more easily shown in Figure 5.9. This figure combines the numerical data of each college from each Experiment. The percentage breakdown displays clearly the aforementioned trends of the colleges.

Gender had very similar results, with female and male percentages between 26-32%. In contrast, the demographics shown in Figure 5.11 demonstrate click rate trends. Each demographic experienced a decrease in click rate as the student progressed in age, year, time spent on the computer, cyber training, and scholarship affiliation. Phishing awareness, however, had an inverse effect. Students who knew less tended to respond in a lesser click rate compared to students who are aware of phishing attacks and had prior knowledge.

FIG. 5.8. Academic year, age, hours spent on computer, cyber club or scholarship affiliation, phishing awareness, and cyber training aggregated demographics are shown from the survey.

Chapter 6

ANALYSIS

6.1 Experiments 1 – 3

Experiment 1 had 409 out of 450 (91%) students opening this email. Arts, Humanities, and Social Sciences majors had 132 (88%) of targeted students reading the email, with 137 (91%) in Engineering and Information Technology and 140 (93%) in Natural and Mathematical Sciences.

	Test	Value	Significance
	Critical Value	*5.991*	*$\alpha = 0.05$*
PayPal	Pearson Chi-Square	80.71	< 0.0001
	Fisher's Exact Test	-	< 0.0001
	Cramer's V	0.44	-
Quadmania	Pearson Chi-Square	21.14	< 0.0001
	Fisher's Exact Test	-	< 0.0001
	Cramer's V	0.23	-
DoIT	Pearson Chi-Square	61.78	< 0.0001
	Fisher's Exact Test	-	< 0.0001
	Cramer's V	0.38	-
Aggregate	Pearson Chi-Square	136.35	< 0.0001
	Fisher's Exact Test	-	< 0.0001
	Cramer's V	0.33	-

Table 6.1. Significance and strength of Significance results for Experiments 1 - 3

Table 6.1 shows a correlation between College affiliation and user susceptibility to the *Billing Problem* phishing tactic ($\chi^2 = 80.71, p < 0.0001, df = 2, \alpha = 0.05$). There exists a moderate strength in association for the tested variables ($\phi = 0.44$).

A majority of students (419 out of 450, 93%) opened the Quadmania email. The College of Arts, Humanities, and Social Sciences had 127 of the 150 (85%) students interacting with the email, along with 144 (96%) Engineering and Information Technology and 148 (99%) Natural and Mathematical Sciences students.

Similar to Experiment 1, we observe a correlation between College affiliation and user susceptibility using the *Contest Winner* tactic ($\chi^2 = 21.14, p < 0.0001, df = 2, \alpha = 0.05$). Experiment 2 has a weak strength of significance ($\phi = 0.23$), however, compared to the moderate strength found in Experiment 1 ($\phi = 0.44$).

Experiment 3 had almost the same amount of student interaction rate of 418 out of 450 (93%) students opening the email. Arts, Humanities, and Social Sciences majors had 142 (95%) of targeted students reading the email, with 144 (96%) in Engineering and Information Technology and 132 (88%) in Natural and Mathematical Sciences.

There is a correlation between College affiliation and user susceptibility for Experiment 3. Experiment 3 used the fear tactic *Expiration Date* approach. Table 6.3 shows not only associations, but also the level of strength of the association ($\chi^2 = 61.78, p < 0.0001, df = 2, \alpha = 0.05$). There exists a low to moderate strength ($\phi = 0.38$), closer to Experiment 1's ($\phi = 0.44$) value than Experiment 2's ($\phi = 0.23$) result.

Overall, each phishing experiment has shown that the college affiliation demographic as significant to a student's susceptibility to clicking a phishing link. When combining the click rates from all three Experiments, the null hypothesis is void as well, with the Chi-Square value of 136.35 exceeding the 5.991 critical value. There exists a low to moderate strength in this relationship as Cramer's V value is 0.33.

6.2 Experiment 1 With Survey Demographics

A total of 482 (39%) students out of the 1,246 completed the survey. There are 102 students (21%) from the first group that responded to the survey out of the 482 total respondents. Shown below in Table 6.4 are the calculated values for significance and strength of significance tests. Fisher's Exact test is calculated instead of Chi-Square whenever the expected values are < 5.

	Demographics	Strength of Significance	Significance			Critical Value	df
		Cramer's V	Fisher's p value	Chi-Square (χ^2)	χ^2 p value	$\alpha = 0.05$	
Significant	College Affiliation	0.37	0.0008	13.78	0.001	5.991	2
	Phishing Awareness	-	0.0027	-	-	5.991	2
	Academic Year	0.28	0.035	8.57	0.036	7.815	3
Insignificant	Age	-	0.41	2.85	0.42	7.815	3
	Cyber Training	-	-	0.62	0.73	5.991	2
	Gender	-	0.81	0.17	0.68	3.841	1
	Hours Spent on Computer	-	0.41	-	-	7.815	3
	Cyber Club/ Cyber Scholarship Affiliation	-	0.99	-	-	5.991	2

Table 6.2. Experiment 1 Significance and strength of Significance results for Survey Demographics

Demographic Relationships

We determine academic year progression of a student as a significant variable towards their susceptibility to a phishing attack ($\chi^2 = 8.57, p = 0.036, df = 3, \alpha = 0.05$). More educated students tend to click the link less often than students in lower grade levels. This relationship has low to moderate strength of association ($\phi = 0.37$). Another demographic relationship lies in a student's College affiliation($\chi^2 = 13.78, p = 0.001, df = 2, \alpha = 0.05$). College of Engineering and IT students have a substantially low click rate (19.7%)

compared to the other two colleges. The Natural and Mathematical Sciences students have a moderate click rate (55%) and the non-STEM college of Arts, Humanities, and Social Sciences has a high click rate (73.5%). Phishing awareness also is deemed significant, although in a negative manner (Fisher's $p = 0.0027, df = 2, \alpha = 0.05$). Surprisingly, the less a student knew or understood about phishing, the better they fared in not clicking the phishing link.

Insignificant Demographics

We discover several demographics as independent to phishing susceptibility in Experiment 1. Cyber training, gender, age, time spent on the computer, and involvement in cyber clubs or cyber scholarship programs give no bearing on whether a student is going to click the Paypal link.

6.3 Experiment 2 With Survey Demographics

225 students (46.7%) from the second group have responded to the survey. We utilize the same statistical testing in this experiment as in Experiment 1. When all preconditions for the Chi-Square test are met, the Chi-Square test is conducted. If the test does not have sufficient Expected Count values, Fisher's Exact test is calculated instead.

Demographic Relationships

We found relationships within phishing awareness, cyber training, college affiliation, amount of time spent on the computer, age, and cyber club/cyber scholarship affiliation. Similar to the first Experiment, STEM majors had a lower click rate than their non-STEM peers. Young students clicked the phishing link more often than older students, and students who spent less time on the computer tended to have a higher click rate. Those that indicated as having no cyber training whatsoever portrayed a lower click rate than students with either formal or informal training. Students who were less educated in phishing knowledge

performed better (less click rates) than their more knowledgeable counterparts, contrary to our expectations.

	Demographics	Strength of Significance	Significance			Critical Value	df
		Cramer's V	Fisher's p value	Chi-Square (χ^2)	χ^2 p value	$\alpha = 0.05$	
Significant	Phishing Awareness	0.53	<0.0001	63.16	<0.0001	5.991	2
	Cyber Training	0.25	0.0007	14.44	0.0005	5.991	2
	College Affiliation	0.16	0.035	6.08	0.048	5.991	2
	Hours Spent on Computer	-	0.001	-	-	7.815	3
	Age	-	0.001	-	-	7.815	3
	Cyber Club/ Cyber Scholarship Affiliation	-	0.045	-	-	5.991	2
Insignificant	Gender	-	0.18	2.23	0.14	3.841	1
	Academic Year	-	0.59	-	-	7.815	3

Table 6.3. Experiment 2 Significance and strength of Significance results for Survey Demographics

Insignificant Demographics

We determine academic year and gender as independent variables due to the insufficient amount compared to the corresponding critical value and high p value.

6.4 Experiment 3 With Survey Demographics

There were 155 students (32%) from the third Experiment who completed the survey. Shown in Table 6.6 are the results for the different demographics, where all but age showed to be significant factors.

Demographic Relationships

We show that academic year, college affiliation, gender, cyber training, phishing

	Demographics	Strength of Significance	Significance			Critical Value	df
		Cramer's V	Fisher's p value	Chi-Square (χ^2)	χ^2 p value	$\alpha = 0.05$	
Significant	College Affiliation	0.53	<0.0001	43.27	<0.0001	5.991	2
	Gender	0.38	<0.0001	22.73	<0.0001	3.841	1
	Cyber Training	0.32	0.006	15.38	0.0005	5.991	2
	Phishing Awareness	-	<0.0001	-	-	5.991	2
	Hours Spent on Computer	-	<0.0001	-	-	7.815	3
	Cyber Club/ Cyber Scholarship Affiliation	-	<0.0001	-	-	5.991	2
	Academic Year	-	0.019	-	-	7.815	3
Insignificant	Age	-	0.33	-	-	7.815	3

Table 6.4. Experiment 3 Significance and strength of Significance results for Survey Demographics

awareness, time spent on the computer, and cyber club/scholarship affiliation are significant variables to student click rates. Similar to the previous experiments, college affiliation demographics indicated that STEM majors — particularly those in the Engineering and IT fields — fell for the phishing link the least ($\chi^2 = 43.27, p < 0.001, df = 2, \alpha = 0.05$). Older students in regards to academic year (Fisher's $p = 0.019, df = 3, \alpha = 0.05$) tended to click less. Females had a 58% click rate as opposed to 91% of male students, showing a correlation for this demographic ($\chi^2 = 22.74, p < 0.001, df = 1, \alpha = 0.05$) Increased time on the computer (Fisher's $p < 0.001, df = 3, \alpha = 0.05$) and cyber training ($\chi^2 = 15.38, p = 0.006, df = 2, \alpha = 0.05$) also positively impacted students with lower click rates. Students within a cyber club or scholarship program also saw a drop in click rates compared to students with no such affiliation (Fisher's $p < 0.001, df = 2, \alpha$

$= 0.05$. Interestingly, there is a negative significance with phishing awareness (Fisher's $p < 0.001, df = 2, \alpha = 0.05$). Students who were unaware of phishing attacks performed better with lower click rates than students who identified that they were aware and understood what phishing attacks were.

Insignificant Demographics

We discover age is not a significant demographic factor since Fisher's p value exceeded the $\alpha = 0.05$.

6.5 Comparative Analysis

Demographic Relationships

We show that phishing awareness, hours spent on the computer, cyber training, cyber club or cyber scholarship affiliation, age, academic year, and college affiliation are significant variables to student susceptibility.

The aggregated college affiliation demographic indicated that STEM majors — with Engineering and IT majors in particular — had smaller click rates (EIT 65%, NMS 70%) compared to non-STEM majors (AHSS 80%) ($\chi^2 = 9.85, p = 0.0073, df = 2, \alpha = 0.05$). Increasing academic year progression influenced a rise in students who did not click the links ($\chi^2 = 15.67, p = 0.0013, df = 3, \alpha = 0.05$). Increased time on the computer (Fisher's $p < 0.0001, df = 3, \alpha = 0.05$) and cyber training ($\chi^2 = 19.47, p < 0.0001, df = 2, \alpha = 0.05$) also positively impacted students to lower click rates. Students within a cyber club or cyber scholarship program observed a drop in click rates compared to students with no such affiliation ($\chi^2 = 19.29, p < 0.0001, df = 2, \alpha = 0.05$). Within the cyber club and cyber scholarship group, students who were affiliated to a cyber scholarship program also had lower click rates compared to the cyber club students. Contrary to our expectations, there is a negative significance with phishing awareness ($\chi^2 = 77.46, p < 0.001, df =$

	Demographic	Strength of Significance	Significance			Critical Value	df
		Cramer's V	Fisher's p value	Chi-Square (χ^2)	χ^2 p value	$\alpha = 0.05$	
Significant	Phishing Awareness	0.40	<0.0001	77.46	<0.0001	5.991	2
	Hours Spent on Computer	-	<0.0001	-	-	7.815	3
	Cyber Training	0.20	0.0001	19.47	<0.0001	5.991	2
	Cyber Club or Cyber Scholarship	0.20	0.0001	19.29	<0.0001	5.991	2
	Age	0.18	0.0017	16.25	0.001	7.815	3
	Academic Year	0.18	0.0017	15.67	0.0013	7.815	3
	College Affiliation	0.14	0.0068	9.85	0.0073	5.991	2
Insignificant	Gender	-	0.536	0.43	0.512	3.841	1

Table 6.5. Aggregate Significance and strength of Significance results for Survey Demographics

$2, \alpha = 0.05$), as in Experiments 1 – 3. Students who were unaware of phishing attacks performed better (28% clicked) with lower click rates than students who identified that they were aware (42% clicked) and understood what phishing attacks were (80% clicked).

Insignificant Demographics

We discover gender is not a significant demographic factor. Despite gender being significant in Experiment 3, the aggregate data resulted in a Chi-Square calculation less than the critical value ($\chi^2 = 0.43$, critical value $= 3.841, \alpha = 0.05$).

Chapter 7

DISCUSSION

7.1 Campus Response

This study incorporated unannounced phishing tests. Because of this, students were unaware of the phishing study until all three Experiments concluded. As such, not only were students assessed on how they interacted with the phish emails, but UMBC was also adjacently assessed on how they reported such activity. The PayPal email from Experiment 1 received very little discourse throughout campus. This can be attributed to the generic layout of this phish, where the email itself can be regarded as spam and not targeting the user as part of the UMBC community exactly. The Quadmania phish received a lot of attention, however. This phish created many conversations and warnings by several entities at UMBC, including the Student Events Board (seb), the campus police, and the Division of Information Technology (DoIT). The warnings of a phishing scheme were sent out to the student body the very same day, a couple hours after the first email was sent. (seb), who the email impersonated, made it very clear to the student body that the Quadmania email was not from them. They employed not only the *myUMBC* dashboard to send out updates, but they also used social media to spread the word.

Their quick and efficient process reached several students within the Experiment 2 Cohort. Despite this, a vast majority of students had already clicked and ”fell” for the

phishing scheme. The students who were deceived by the phish, however, then reported this to DoIT or (seb) themselves, which enabled the quick turnover of warnings and cautionary updates.

7.2 Phishing Outcomes and Speculation

The first Phishing Experiment used the *Billing Problem* tactic to impersonate PayPal. This phishing scheme alerted the user of a recent order that would appear in their account. Regardless if they had a PayPal account, the unrequested order presented itself as a problem the user faced. Overall, the Arts, Humanities, and Social Science majors fell for this phish by 73%. The Natural and Mathematical Science majors also had a majority of students clicking the link, but with a lesser percentage of 55%. Lastly, the Engineering and Information Technology majors had a low click rate of 19.7%. The second phishing Experiment, known as the Quadmania email, used the classic *Contest Winner* tactic. This tactic was overwhelmingly successful. Each college had a click rate of over 70% for this Experiment. The College standings still persisted and passed significance testing, although in a lesser extent. The last Experiment, using the *Expiration Date* tactic to scare users of account closure, had less success amongst students. There was a noticeable decrease within the Engineering and Information Technology college in particular. The Arts, Humanities, and Social Sciences students still had a majority click rate, with the Natural and Mathematical Sciences students having an even distribution of students that clicked and not clicked.

Significance and strength of significance tests were conducted on both the individual Experiment demographics as well as the overall student demographics collected from the survey. A student was more likely to click a link if they belonged in the Arts, Humanities, and Social Sciences. Students within the College of Engineering and Information Tech-

nology had the least amount of probability to click the link, followed by the Natural and Mathematical Science students.

If a student had some form of cyber training, they would be more resistant to falling for a phishing scheme. Going even further, if the student had formal training, they would be less at-risk than their informal training peers. Age and academic year in college were other confirmed demographics. The older the student was or the higher academic year they were in, the less likely they would be to click the link. This was expected using our knowledge of past studies' results. On a similar note, generally the click rate decreased the longer a student spent time per day on the computer. The only exception was for students using a computer for less than 4 hours a day. Cyber club or scholarship affiliation demonstrated a correlation. Cyber and SFS scholars had low click rates, followed by students with no affiliation. Students with more exposure to material (be it from age, academic year, or simply content exposure) can be thought as be more educated.

The Phishing Knowledge and Awareness demographic also had a relationship with user susceptibility. Surprisingly, this demographic had an inverse, or negative, relationship to user susceptibility. If a student heard about phishing attacks and knew the definition, the click rate would increase. The less the student knew about a phishing scheme, the less likely the click rate would increase. This goes against our expectation that more phishing awareness would decrease click rates. Our thoughts, based on pure speculation, for why this demographic had such a negative correlation are based on two reasons. The first reason is the nature of the user-reported survey. The survey responses are taken as honest responses by the user. If a user did not reply honestly, then the results could be skewed. The second speculation could be that students who have clicked the false link in the phishing emails were now aware of what phishing was by the time that the survey was sent out. If they did not know what phishing was until they clicked the link and reported it to DoIT or researched for themselves after the fact, then the results could also have been impacted.

Of all the tested demographics, gender was calculated as independent. Past studies accredited gender as significant based on their small population size or scope of study. In Experiments 1 and 2, gender failed the significance test. Gender voided the null hypothesis only for the last Experiment, but ultimately upheld the null hypothesis when aggregate data was used.

7.3 Limitations

While this project was able to discern possible associations of what makes a user more at-risk to phishing attacks, there are some limitations that need to be addressed. One limitation was the time the experiments were sent out over the semester. While each of the phishing emails were sent in the same time frame to all the students, there was the prospect that one major or college had an exam or project due that same day or week compared to the other randomly selected participants. In that case, there is the uncontrolled factor of how busy or stressed the individual was when they were selected for one of the phishing Experiments.

Another limitation that could have impacted the results was the overall awareness of the student before they were selected for a phish experiment. For example, due to the campus reaction to the Quadmania phish, a student who was selected for Experiment 3 might have been more alert or prone to report the email due to their knowledge that their peers were targeted by a phishing scheme. Along these lines, if they were a student worker within (seb), they would be aware of the Quadmania phish. While these outside factors are uncontrolled, they must be considered when analyzing the collected statistics of this study.

Mentioned previously, another limitation that must also be considered is the truthfulness of the survey results. The demographic analysis was done under the pretense that the data collected was truthful and honest answers by the students. If students filled out their

surveys incorrectly or falsely, then the outcomes are effected.

The scope of this study must be taken into consideration as well. Many demographics have been deemed either dependent or independent. Gender, for example, was deemed as independent and not influencing user susceptibility using statistical testing, but only within a college campus setting. Gender itself cannot be said to be completely independent for all populations. Thus, these demographics and their conclusions are only valid within the college environment.

7.4 Future Work and Open Problems

Based on this demographic study's results, future research on why there is a negative correlation between phishing awareness and phishing click rate can be conducted. The demographics indicate that cyber knowledge and technical aptitude aid in lowering the risk factor, yet phishing knowledge by itself increases risk. A study on why this happens would be beneficial.

If this project could continue over several semesters, it would be valuable to analyze the differences and similarities between different groups of students over several semesters. In this way, data could be more normalized and assessed for congruency. The phishing emails could also be distributed over several semesters instead of a single semester. This would decrease the likelihood of a student being more alert of a phishing attack due to mere exposure. Another consideration would be to include faculty and staff to this study. By including faculty members, another dimension of demographic data can be assessed, as well as comparing those selected faculty against their learning pupils.

Chapter 8

CONCLUSION

Target students were randomly selected from a sample pool of over 10,000 students within the University of Maryland, Baltimore County (UMBC) undergraduate population. These students were selected into subgroups of 150 for the three College that UMBC offers, for a total of 450 students in each Cohort. The selected students were sent three phishing emails using modern-day phishing tactics, with data collected on student click rate. Near the end of the study, a survey was sent to all Cohorts for more demographic data.

The findings in this study indicate an association between several demographic factors and a student's susceptibility towards a phishing attack. Students within the College of Engineering and Information Technology had a lesser click rate than the students in College of Natural and Mathematical Sciences, and a much lesser rate than the College of Arts, Humanities, and Social Sciences students. Likewise, the older a student was in age and academic year, the lesser probability they had in clicking a phishing link. In contrast, gender provided no such association to phishing click rate.

If a student was part of a cyber club or scholars program, the student would have a lesser click rate then their peers. Within this group, the scholar group had fewer click instances than the cyber club students. The amount of time spent on the computer, if greater than 4 hours/day, positively impacted the click rate of students. This trend was also seen in

the amount of cyber training a student had before the study was conduced. Basic phishing knowledge and awareness, however, had a negative impact on a student's susceptibility. A student who understood the definition of phishing and was aware of phishing in general was more likely to click the phishing link than a student who didn't understand what a phishing attack was but had heard of the concept. Going further, students with no awareness or knowledge of phishing whatsoever had the least amount of clicks compared to their more knowledgeable counterparts.

We believe that these results could be beneficial to not only universities, but also to businesses in general. These associations provide insight to more successful and effective cyber approaches depending on students' affiliations. In doing so, a student population with similar cyber knowledge can be achieved by targeted at-risk students. Likewise, companies with different business sectors can improve their cyber training practices towards their at-risk employees.

REFERENCES

[1] Aloul, Fadi A. “The Need for Effective Information Security Awareness.” Journal of Advances in Information Technology, vol. 3, 2012, pp. 176–183.

[2] CloudHQ. “EmailTracker”. 2018. https://www.cloudhq.net/.

[3] Dodge Jr., Ronald C, et al. “Phishing For User Security Awareness.” Computers and Security, vol. 26, 2007, pp. 73–80.

[4] Downs, Julie S, et al. “Decision Strategies and Susceptibility to Phishing.” Carnegie Mellon University, 2006.

[5] Ellis, David. “Top 10 Types of Phishing Emails.” SecurityMetrics Blog, SecurityMetrics, 2014, blog.securitymetrics.com/2014/05/types-of-phishing-emails.html.

[6] Howarth, Fran. “The Role of Human Error in Successful Security Attacks”. *Security Intelligence*, 2014. https://securityintelligence.com/the-role-of-human-error-in-successful-security-attacks/.

[7] Hunter. “MailTracker”. 2018. https://hunter.io/.

[8] Norton. “What is Social Engineering?” *Symantec,* 2014. https://us.norton.com/internetsecurity-emerging-threats-what-is-social-engineering.html

[9] RIQ News Desk. “Human Behavior - A Major Threat to Organizational Cyber-Security”. *ReadItQik,* 2016. https://searchsecurity.techtarget.com/definition/phishing.

[10] Rouse, Margaret. “Phishing”. *SearchSecurity*, 2017. https://searchsecurity.techtarget.com/definition/phishing.

[11] Shen, Steve, et al. “Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions.” Carnegie Mellon University, 2010.

[12] Sun, Jerry Chih-Yean, et al. “The Mediating Effect of Anti-Phishing Self-Efficacy Between College Students' Internet Self-Efficacy and Anti-Phishing Behavior and Gender Difference.” Computers in Human Behavior, vol. 59, 2016, pp. 249–257.

[13] UMBC Admissions. “Student Enrollment and Persistence”. *UMBC*, 2018. https://umbc.app.box.com/s/1torn5ywqscyktvo48xnqdlin1rdhf2k.

[14] UMBC. “Undergraduate Admissions”. *UMBC*, 2018. https://undergraduate.umbc.edu/quicklinks/fast-facts.php.