Security Awareness Training
Through Experiencing the Adversarial Mindset

Jens Christian Dalgaard
jc@drdalle.com

Niek A. Janssen
itu@niek.dk

Oksana Kulyuk
and Carsten Schürmann
IT University of Copenhagen, Denmark
{okku, carsten}@itu.dk

***Abstract*—Cybersecurity concerns are increasingly growing across different sectors globally, yet security education remains a challenge. As such, many of the current proposals suffer from drawbacks, such as failing to engage users or to provide them with actionable guidelines on how to protect their security assets in practice. In this work, we propose an approach for designing security trainings from an *adversarial* perspective, where the audience learns about the specific methodology of the specific methods, which attackers can use to break into IT systems. We design a platform based on our proposed approach and evaluate it in an empirical study ($N = 34$), showing promising results in terms of motivating users to follow security policies.**

I. Introduction

IT security is becoming a more pressing issue with the increasing level of digitalisation. With recent legislation such as the European GDPR[1], or the Californian CCPA[2], proper IT security has formalized from being a good practice into a concrete legal requirement. At the same time, security awareness campaigns are often known to be lacking in effectiveness, due to issues such as participants not perceiving the contents as relevant to them or failing to provide actionable guidelines [3]. A number of approaches has been designed to implement the "learning-by-doing" principle, e.g. via gamification [30]. While such approaches are often considered to be promising in terms of engaging the user, they have so far been among the least applied in existing studies, having had limited development and evaluation [7], [30]. The evaluated approaches also suffer from drawbacks such as being targeted at security professionals rather than broader target groups of employees.

In our work, we rely on the principle of learning-by-doing to let the user experience the *adversarial mindset* in order to better understand the role of security protection measures. In this, we develop a training platform aimed at non-expert users with the goal of motivating the users to adhere to security policies of their companies by following established best security practices such as updating their system often and using unique strong passwords (facilitated e.g. by the use of a password manager) for each service. The core principle of our approach is letting users adopt a hacker's perspective and experience the exploitation of common vulnerabilities, followed by suggestions on what they can do to mitigate these vulnerabilities. We aim to answer the following research question:

Does experiencing IT security from an attacker's perspective motivate users towards better security behaviour?

For answering our research question, we have developed a prototype of a training platform[3], which we evaluated in a user study ($N = 34$). The evaluation has shown promising results in terms of increasing participants' motivation to follow the security policies of their organisation. We therefore conclude that while awareness and motivation is only a part of general measures for facilitating secure behaviour [3], our proposal can be a useful tool for encouraging users to think about their security.

II. Background

As the background for our work, we describe general issues of IT security awareness and discuss existing approaches for IT security awareness trainings. We summarise the section by putting our contribution in the context of the described related works and discussing how they contributed to the development of our prototype.

a) IT Security Awareness

IT security awareness has been considered in previous research from different aspects, see the literature review in [10]. As such, a number of works focused on studying the users' *knowledge* about IT security threats [13], [4], [5], [9], [18] and available countermeasures [27], [6], [25]. In particular some of the latest studies show that while the level of knowledge among the users can vary, the users in general tend to have at least a basic knowledge of security threats and how they should act in response to them [1], [21]. A number of works hence focused on further barriers towards secure behaviour, such as lack of motivation due to low risk appraisal, perceived (self-)efficacy or costs of secure behaviour (in terms of monetary and non-monetary resources such as time and cognitive effort) [26], [16], [17], [21].

[1] General Data Protection Regulation (https://data.europa.eu/eli/reg/2016/679/2016-05-04)

[2] California Consumer Privacy Act (https://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml)

[3] Source Code available on https://github.com/niekjanssen/coursecurity

Symposium on Usable Security and Privacy (USEC) 2023
27 February 2023, San Diego, CA, USA
ISBN 1-891562-91-6
https://dx.doi.org/10.14722/usec.2023.237300
www.ndss-symposium.org, https://www.usablesecurity.net/USEC/

Principle	**Description**
Segmenting	Segmenting lessons into learner-paced chunks rather than as a continuous unit helps people learn better.
Contiguity	Aligning words to corresponding graphics increases learning performance.
Reflection	Providing users with opportunities to reflect on what they learned increases learning.
Immediate Feedback	Giving immediate feedback helps users to assess how they are doing and provides efficient guidance in learning
Narrative	Presenting training material within the context of a story establishes an emotional connection with learners.
Signalling	Directing user attention to key messages in the lesson helps with information discovery and understanding.
Personalization	Using conversational style language and pedagogical agents create engagement and increase learning.
Multimedia	Combining words and images is more conductive to learning than words alone.
Conceptual Procedural	Showing causal relationships between conceptual knowledge and procedural knowledge increases understanding of a concept.
Learning-by doing	Practice (by doing) strengthens acquired knowledge and skills.

Table I: Instructional Design Principles [30]

b) Existing Approaches in IT Security Training/Education:

A number of different multimedia tools have been developed to teach IT security, see a systematic literature review in [30]. The review concludes that digital games are among the most widely used tools. However, only a few of these tools are aimed at adults and are focused on IT security in general instead of specific topics such as phishing. The paper furthermore provides recommendations in form of basic design principles that are according to the authors the most applicable when conducting IT security training (see Table I) and propose five criteria to evaluate the application of the principles, showing that less than 5% of the studied tools apply all these principles. In particular, the principles of *learning-by-doing* (i.e. enabling practical application of learned knowledge), *signalling* (i.e. directing user's attention to key messages in the lesson) and *contiguity* (i.e. aligning text to corresponding visual aids) have been the least commonly applied. [30] furthermore, reports that less than a third of the studied approaches (30%) have been evaluated empirically, despite the importance of evaluation in determining whether a proposed measure is actually effective.

c) Our contribution:

In light of the related work, we propose a tool that not only teaches users about threats and countermeasures, but aims to actively influence their behaviour by motivating users via demonstrating the ease of common attacks in absence of secure behaviour. Our proposal, targeted at non-expert users as called for by related research [2], [13], [20], follows the design principles suggested in [30], aiming to implement all of them and to evaluate the resulting prototype. Thus, following the recommendations by [30], our work contributes to further study and evaluation of simulation-based training approaches and their effectiveness in the security context.

III. SECURITY TRAINING CONCEPT

In the following section, we describe the design and development of the prototype of the platform, *Coursecurity*. Namely, we describe the concept of the platform, the story the users would be going through and discuss decisions made with regard to the learning design.

A. Concept

The general idea for the prototype is to use the adversarial mindset perspective to demonstrate common attacks, the vulnerabilities in the system that enable these attacks and their possible impact on organisation's security. Thus, the prototype should make it possible for individuals without a background in IT to gain hands-on experience with the world of hacking. In this, the users learn about the tools that hackers can use to get access to sensitive assets and the methodology the attackers follow in applying these tools. Thus, by using theses tools themselves the users would ideally understand the importance of countermeasures commonly applied to prevent these and similar attacks. We furthermore aim to integrate the recommended learning design principles [30] described in Table I into our prototype, see the summary in Table II.

In particular, the prototype attempts to motivate users to act more responsibly, building upon the theoretical foundation of the Protection Motivation Theory (PMT) model [24], which is a theory stemming from health psychology and widely adapted for use in various fields. PMT explains the coping with potential threats based on the *threat appraisal* (i.e. the appraisal of the likelihood as well as severity of the threat) and *coping appraisal* (i.e. efficacy of the recommended response to prevent or mitigate the threat). Despite existing criticism [23] and research showing that some constructs in PMT are of limited significance in studies in the IT security field [16], [11], [15], it is nonetheless one of the most used models in the field [15].

The prototype uses a story-telling approach to guide the user in hacking a fictive company, PIA, trying to steal information and break the system. In particular, we aimed to include attacks that are sufficiently common in the real world for them to be of relevance and to be relatively easy to explain to the user without technical knowledge as well as relatively easy to prevent if good security hygiene is followed. Thus, we decided to focus on demonstrating attacks based on *insecure credentials* and *outdated software*, as these attack vectors regularly appear in the list of common causes of security breaches (see e.g. [22], [12]), while the mitigation measures – that is, regular updates and using strong and unique passwords – are likely to be issues that users without technical expertise are commonly familiar with and understand, either as a part of organisational security policies or in their personal security-related behaviour.

In our proposal, the users conduct attacks exploiting these vulnerabilities by using real-life tools that are freely available, such as *Nmap*[4] and *Metasploit*[5]. Through the prototype, the users are provided with a guide aiding them in using these tools. The guide is structured in-line with the Cyber Kill Chain model [28] which has been proposed to represent a structured end-to-end approach followed by adversaries and divides the adversaries' process into stages such as "reconnaissance, weaponization, exploitation, command & control, etc.". Despite not being explicitly introduced to the participants, the model is used in our prototype for designing clear phases for the "attack", thus providing the participants with realistic and procedural insight into the adversary tactics and offers a holistic view of the adversaries' process. Additionally, the prototype also provides them with information on why they are executing different commands and what their effect is.

The users will go through a total of two challenges. Each challenge focuses on a specific topic and aims to inform the participant about the dangers surrounding the said topic and what to do. When starting a new challenge, the user will be met with introductory information for the upcoming challenge. Furthermore, the user will also be debriefed after completing a challenge, thereby providing the user with a combination of *conceptual and procedural knowledge*.

An accompanying introduction video was produced, that provides explanation regarding the basic functionality of the prototype. Aiding the participants by highlighting the most important features, thereby making it easier to use. The transcript of the video is included in the supplementary materials.

B. Story Development

The platform integrates a continuing story that is part of the challenges, providing a consistent red thread that guides the user through the platform. The story is further intended to engage the user in the platform, incorporating the *narrative* learning principle.

Throughout the development of the story, the aim is to incorporate topics that are understandable and relatable to the non-expert end-user and relate to relevant foundational security principles.

Despite not being an aspect the users are directly exposed to, the story is intended to follow common principles in IT Security research and education. The different *elements* of the story are based on the confidentiality, integrity, and availability security goals [29]. The overall story is designed to mimic the Cyber Kill Chain [28] approach.

The story contains two challenges. The division into two sections is made with regard to the learning principle of *segmentation*. Throughout the story, the user adopts the persona of someone trying to extract inside knowledge out of a corporation.

1) Challenge 1: Updates

The first part of the story takes the participant through a sequence similar to the Cyber Kill Chain [28].

Through initial reconnaissance and scanning of the target, the user is led on a search for vulnerabilities. At first, the user does not find any vulnerabilities on the target company's cloud platform. Next, the challenge continues through the discovery of a legacy application. The application appears to the user as not in active use, and thus will seem not to contain significant data. This application does contain vulnerabilities.

The information the user finds will become helpful further along in the story. This element is designed to challenge the users' evaluation regarding "sensitive" data and provide the insight that individual judgement might not be sufficient [17].

Since the entrance to this server is through a well-known vulnerability that could have been resolved through updating the software, the challenge showcases the value of regular software updates on any machine. Updating systems regularly should be a security practice that should be relatable to the user [21], [17].

2) Challenge 2: Passwords

The second part of the story highlights and contextualizes password security. This is a subject that users are familiar with, and according to [21] most users have a certain level of knowledge regarding this topic.

Rather than a theory- or risk-based discussion of password security users might be familiar with, they are guided through abusing insecure passwords themselves. With help from the information gathered in the previous challenge, the user will perform a variation on password stuffing. Since the information found in Challenge 1 is outdated, the passwords are no longer valid, though the user is encouraged (and helped) to abuse the clear password pattern of one of the fictional system users. The pattern then allows them to use the information gathered from an insignificant legacy system to access a well-secured application.

This *password stuffing* can show the user the risks associated with re-using passwords across multiple platforms, especially platforms with different security levels. Further, the *creative password stuffing* shows the user how easy a predictable password can be abused when one version is leaked. The story elements regarding passwords follow the recent changes in advice (e.g., [8]) that encourages longer passwords rather than frequent changes or intricate combinations of special characters.

[4] https://nmap.org/,lastaccessed16/11/2022

[5] https://www.metasploit.com/,lastaccessed16/11/2022

After gaining access to the application, the user can freely roam the file-sharing application and view various dummy documents. The application contains shared files between many departments, highlighting the impact the breach of a single account can have on an organisation. Understanding this could, through an increase in social motivation, lead to an increase in secure decision-making [17].

C. Learning Flow

As the main learning of the platform is provided through challenges, the users are guided through these challenges via multiple *tasks* they have to complete, each task consisting of one or more *steps*. The steps and tasks add another level to the principle of *segmenting*, while also supplying the user *immediate feedback* upon how they are moving along through the challenges. The guidance for the overall challenge, as well as for the individual tasks, is provided in the form of *introduction*, *instruction* and *debriefing* explained below.

1) Introduction

When entering the prototype or starting a new challenge, the users will be met with an introduction section of the current challenge. This introduction page is automatically shown to the user to ensure that the attention of the participant is drawn to key messages of the upcoming challenge, consequently, applying the *signalling* learning principle. The introduction furthermore aims to teach the users the *conceptual knowledge* relevant for the challenge.

2) Instruction

The purpose of the instruction is to help the user manage to complete the challenges/tasks, providing them with the *procedural knowledge*. The instruction for each challenge consists of the list of multiple tasks containing one or more steps that have to be completed (following the principle of *segmenting*). The instructions for the task contains a title and a description, thereby *signalling* the most important information regarding each task.

3) Debriefing

The debriefing is used to further elaborate upon what the users have managed to do, giving them the possibility to *reflect* upon what they just did and how they managed to do it. The debriefing furthermore brings *conceptual knowledge* to the users, explaining to them the reasoning and the consequences of their actions, as well as providing *immediate feedback* on the user's progress.

The guidance for individual steps is built with a similar structure. A step can contain three different kinds of text (description, instruction, and explanation) to indicate *why* the user is doing something, *what* the user is doing, and *how* they should do it. Consequently, the user is *signalled* what they have to do, as well as presented with *conceptual and procedural* knowledge. The usage of the short description sections is also to deepen the *narrative* further.

D. User Interface

The main page of the prototype platform can be seen in Figure 1a[6]. The main screen is divided into two parts: the *virtual computer* on the left and the *guide* on the right.

The virtual computer represents an attacker's machine which is used to break into the demonstrated system. To avoid confusing the users, any irrelevant applications has been removed from the interface. While the virtual computer is designed to run on a Linux operating system, the interface is styled to resemble a Windows operating system with the aim of making it more familiar to the non-expert users.

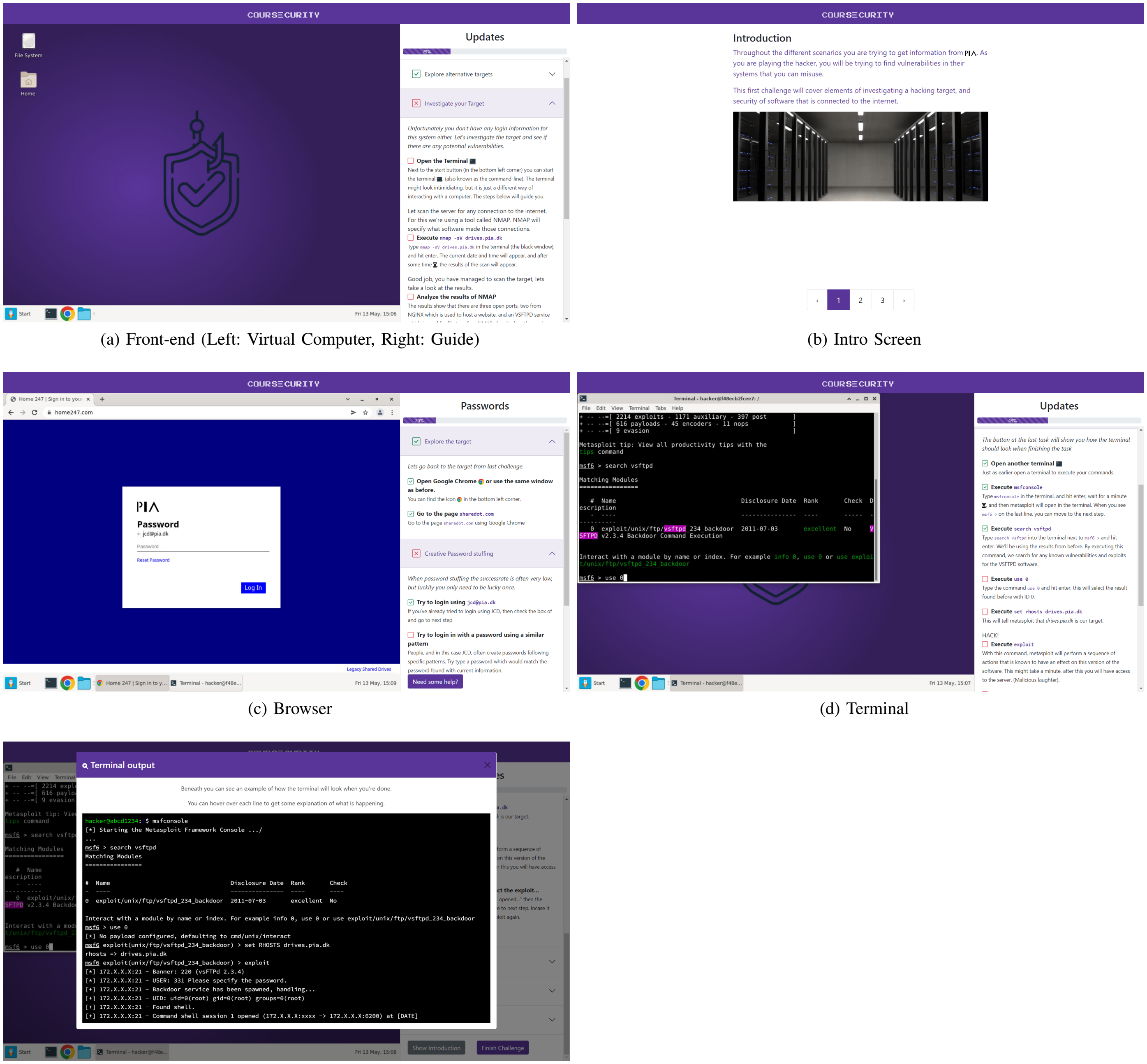
The guide consists of several components. At the top of the guide, there is the challenge's title and a progress bar, keeping track of the user's progress in the current challenge. The progress bar is giving the users *immediate feedback* of their progress. Beneath the progress bar, the actual guide with tasks is located. The steps needed to complete each task are furthermore included in the guide, and initially hidden from the user. Upon starting the training, only the current task is unfolded while all other tasks are closed. In this way, the users are being clearly *signalled* which task is the current one they need to solve. Both tasks and steps have a checkbox located next to them. These checkboxes allow the user to track what tasks are completed and which ones are not yet. When all steps in a task are checked off as completed, the task folds, and the next task unfolds, as shown in the two figures: Figure 1a and Figure 1c (p. 5). Each step within a task consists of the following components (see also Section III-C):

- **Description** (Optional) - A short text explaining the users where they are and what is going to happen next, maintaining the story *narrative*. For some steps, the description might also contain visual cues, e.g. a small icon for a program that the user would have to run or a button that the user would have to click. The usage of such cues, in particular, is aimed to serve *contiguity* by helping the user to identify the right actions to be performed on the virtual computer.
- **Instruction** - A short text clearly stating what the user has to do now (the action to perform).
- **Explanation** - A longer text explaining the user, *why* they are performing each action, as well as what the effect is, providing them with the *conceptual knowledge*.
- **Output explanation** (Optional) - An explanation of the output the participant might get for some of the steps, e.g. the steps that result in a lengthy output in the terminal (see Figure 1e for an example).

The texts used in describing the steps or providing instructions/explanations are also written, whenever appropriate, using more conversational style while maintaining narrative, thus ensuring *personalisation*.

As mentioned in Section III-C, the prototype platform also consists of introduction- and debriefing sections. These sections are developed as a full-screen modal, which appears

[6] The screenshots in higher resolution will be included in the supplementary material published online



Updates

29%

☑ Explore alternative targets

☒ Investigate your Target

Unfortunately you don't have any login information for this system either. Let's investigate the target and see if there are any potential vulnerabilities.

□ **Open the Terminal**
Next to the start button (in the bottom left corner) you can start the terminal, (also known as the command-line). The terminal might look intimidiating, but it is just a different way of interacting with a computer. The steps below will guide you.

Let scan the server for any connection to the internet. For this we're using a tool called NMAP. NMAP will specify what software made those connections.

□ **Execute `nmap -sV drives.pia.dk`**
Type `nmap -sV drives.pia.dk` in the terminal (the black window), and hit enter. The current date and time will appear, and after some time ⌛, the results of the scan will appear.

Good job, you have managed to scan the target, lets take a look at the results.

□ **Analyze the results of NMAP**
The results show that there are three open ports, two from NGINX which is used to host a website, and an VSFTPD service

(a) Front-end (Left: Virtual Computer, Right: Guide)

Introduction

Throughout the different scenarios you are trying to get information from PIA. As you are playing the hacker, you will be trying to find vulnerabilities in their systems that you can misuse.

This first challenge will cover elements of investigating a hacking target, and security of software that is connected to the internet.

(b) Intro Screen

Passwords

20%

☑ Explore the target

Lets go back to the target from last challenge.

☑ **Open Google Chrome or use the same window as before.**
You can find the icon in the bottom left corner.

☑ **Go to the page `sharedot.com`**
Go to the page `sharedot.com` using Google Chrome

☒ Creative Password stuffing

When password stuffing the successrate is often very low, but luckily you only need to be lucky once.

☑ **Try to login using `jcd@pia.dk`**
If you've already tried to login using JCD, then check the box of and go to next step

□ **Try to login in with a password using a similar pattern**
People, and in this case JCD, often create passwords following specific patterns. Try type a password which would match the password found with current information.

(c) Browser

COURSECURITY
Terminal - hacker@f48ecb2fcee7: /
File Edit View Terminal Tabs Help
Start
Terminal - hacker@f48e...
Fri 13 May, 15:07

```
+ -- --=[ 2214 exploits - 1171 auxiliary - 397 post       ]
+ -- --=[ 616 payloads - 45 encoders - 11 nops            ]
+ -- --=[ 9 evasion                                       ]

Metasploit tip: View all productivity tips with the
tips command

msf6 > search vsftpd

Matching Modules
================

   #  Name                                  Disclosure Date  Rank       Check  D
escription
   -  ----                                  ---------------  ----       -----  -
---------
   0  exploit/unix/ftp/vsftpd_234_backdoor  2011-07-03       excellent  No     V
SFTPD v2.3.4 Backdoor Command Execution


Interact with a module by name or index. For example info 0, use 0 or use exploi
t/unix/ftp/vsftpd_234_backdoor

msf6 > use 0
```

Updates

43%

The button at the last task will show you how the terminal should look when finishing the task

☑ **Open another terminal**
Just as earlier open a terminal to execute your commands.

☑ **Execute `msfconsole`**
Type `msfconsole` in the terminal, and hit enter, wait for a minute ⌛, and then metasploit will open in the terminal. When you see `msf6 >` on the last line, you can move to the next step.

☑ **Execute `search vsftpd`**
Type `search vsftpd` into the terminal next to `msf6 >` and hit enter. We'll be using the results from before. By executing this command, we search for any known vulnerabilities and exploits for the VSFTPD software.

□ **Execute `use 0`**
Type the command `use 0` and hit enter, this will select the result found before with ID 0.

□ **Execute `set rhosts drives.pia.dk`**
This will tell metasploit that *drives.pia.dk* is our target.

HACK!

□ **Execute `exploit`**
With this command, metasploit will perform a sequence of actions that is known to have an effect on this version of the software. This might take a minute, after this you will have access to the server. (Malicious laughter).

(d) Terminal

Terminal output

Beneath you can see an example of how the terminal will look when you're done.

You can hover over each line to get some explanation of what is happening.

```
hacker@abcd1234:/$ msfconsole
[*] Starting the Metasploit Framework Console .../
...
msf6 > search vsftpd
Matching Modules
================

#  Name                                  Disclosure Date  Rank       Check
-  ----                                  ---------------  ----       ----
0  exploit/unix/ftp/vsftpd_234_backdoor  2011-07-03       excellent  No

Interact with a module by name or index. For example info 0, use 0 or use exploit/unix/ftp/vsftpd_234_backdoor
msf6 > use 0
[*] No payload configured, defaulting to cmd/unix/interact
msf6 exploit(unix/ftp/vsftpd_234_backdoor) > set RHOSTS drives.pia.dk
rhosts => drives.pia.dk
msf6 exploit(unix/ftp/vsftpd_234_backdoor) > exploit
[*] 172.X.X.X:21 - Banner: 220 (vsFTPd 2.3.4)
[*] 172.X.X.X:21 - USER: 331 Please specify the password.
[+] 172.X.X.X:21 - Backdoor service has been spawned, handling...
[+] 172.X.X.X:21 - UID: uid=0(root) gid=0(root) groups=0(root)
[+] 172.X.X.X:21 - Found shell.
[+] 172.X.X.X:21 - Command shell session 1 opened (172.X.X.X:xxxx -> 172.X.X.X:6200) at [DATE]
```

(e) Help Modal

Figure 1: Front-end Screenshots

automatically upon starting and finishing a challenge. These sections are set up as pages, which the participant can scroll through. Each section contains conceptual information about the current challenge combined with related pictures. An example of the introduction modal is shown in Figure 1b.

Table II contains an overview of each learning principle and an example of how it's implemented in the prototype.[7]

IV. Evaluation

The following section describes the different components forming the basis for the research protocol. The evaluation will be based on theories described in Section II and on the recommendations of [30].

A. Recruitment

As our prototype targets non-expert users, for our evaluation we decided to focus on participants who were office workers who did not have extensive background in IT (either

[7] Some learning principles have been applied in multiple ways. Thus, the examples in table II may not be the only application of the principle.

Principle	**Example of application**
Segmenting	The story in the prototype is segmented into two challenges, with multiple sub-tasks and steps to complete.
Contiguity	Icons for different tools used are located in the guide when mentioned.
Reflection	Each challenge is followed by a debriefing section providing the possibility of reflection, through a break, and the content.
Immediate Feedback	Whenever a user finishes a task or step, they can check off the checkbox and follow their progress in the challenge.
Narrative	The story provides context around, for example, about why the user is attempting to hack PIA.
Signalling	Different styling of text in the guide signals to the user whether something is explanatory or instructive.
Personalization	The guide is set up in specific approach, in an attempt of applying conversational style language.
Multimedia	*Despite not exactly following the definition,* the prototype does include an interactive computer, which provides the users with live images while interacting with the introduced information.
Conceptual Procedural	The users are provided with textual information in the guide and introduction sections, while getting practical experience in completing the tasks.
Learning-by doing	The prototype is letting end users hack simulated targets. The prototype's core activity is learning-by doing.

Table II: Examples of application of the learning design principles.

educational or professional) yet had basic understanding of IT (e.g. being able to independently use an office package and/or administrative applications). Thus, the main grounds for a potential participant to disqualify from the experiment would be not classifying as an office worker (i.e., not regularly, independently, working on a computer with an office suite) or having an educational or professional background in IT (i.e., being an expert user).

For practical reasons, participants have been sourced through the researchers' professional networks following the snowball principle. Contacts were asked if they would be open to participating in the study, and with a positive response, contacted through e-mail with information and instructions. Participants got an e-mail address to contact in case of technical issues.

B. Ethics

While there is no mandatory Ethical Review Board in our institutions, we have considered the ethical issues of our study. As such, the main concern regarding the content of this study is the potential for harm through the exposure of participants to hacking tools, as the participants are introduced to and guided through the exploitation of vulnerabilities in web servers. As the tools they are shown are real and can be used to cause harm to production systems with vulnerabilities, there are ethical concerns when introducing the participants to such tools. Nonetheless, the researchers considered this an acceptable practice since the tools, including comprehensive tutorials on their use, are available to the general public online, and that the potential benefits to the participants (i.e. them learning about mitigation of the demonstrated vulnerabilities) outweigh the concerns. Furthermore, the specific vulnerabilities used – as part of the challenges – are especially simple to exploit and several years old. Additionally, more comprehensive tutorials are available freely. As such, participants do not get introduced to any information they would not be able to obtain on their own.

We furthermore took measures to ensure that the participants can provide their informed consent, being made aware that they can decide not to participate or stop the participation at any point also during the tests. Participants have furthermore been informed made aware that their answers are confidential and will be analysed anonymously and published only in aggregated form. While the participants were asked to provide a username to link their responses before and after completing the training, pseudonyms were accepted. The participants were furthermore invited to enter contact information if they consented to researchers contacting them after the study, being also told that providing such information is not mandatory and that it will not be shared with third parties. The participants were not reimbursed for their participation.

C. Hypotheses

We decided to conduct a pre/post study to evaluate the effect of using the platform on participants' attitudes towards security, as recommended in [14]. We furthermore relied on the constructs from the Protection Motivation Theory model [24], [11], namely, *security breach concern level*, *perceived probability of security breach*, *perceived severity of security breach*, *security policy compliance intention*, *response cost*, *security policy attitude*, *response efficacy*, *self-efficacy*.

Out of these constructs, we particularly focused on evaluating the effect of our proposal on the following constructs where we expected significant improvements: (1) *security policy compliance intention*, i.e. whether going through our training is effective in motivating the users to adopt secure behaviour; (2) *security breach concern level*, i.e. users' experiencing security breaches being performed with easily available tools and causing significant harm to the system; (3) *response efficacy*, i.e. users understanding how best practices in secure

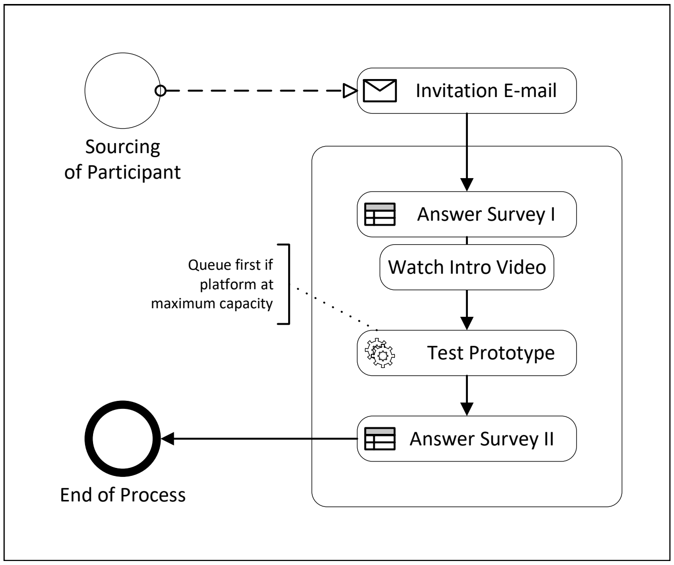


Figure 2: Study Structure

behaviour (e.g. regularly updating one's system) can prevent the demonstrated breaches. We therefore define the following hypotheses:

H$_1$ Participants' reported *security policy compliance intention* increases after the experiment (i.e. participants become more likely to intend to comply to the policies).

H$_2$ Participants' reported *security breach concern level* increases after the experiment (i.e. participants become more likely to report being concerned about potential breaches).

H$_3$ Participants' reported *response efficacy* increases after the experiment (i.e. participants being more likely to believe that them or other individual users/employees can positively affect the security of the systems they use).

We furthermore collect data on the rest of PMT constructs to better understand the impact of our intervention; for these, instead of conducting statistical analyses, we report on descriptive statistics only.

D. Study Procedure

A schematic representation of the study's structure can be seen in Figure 2. As part of the invitation e-mail, participants received a link to start the study process on their own devices, which would show them the first questionnaire, consisting of questions on selected PMT constructs (see Section IV-C) and questions about participants' demographics. Participants were furthermore asked to pick a username and use it consistently throughout the study in order to be able to link their responses. At the end of this questionnaire, the participants were presented with an introduction video that explained the platform's basic features and then redirected to the prototype.

After finishing the two challenges part of the prototype (see Section III-B), the participants were automatically brought to the final survey, consisting of the items evaluating the PMT constructs and an open question asking the participants about their experience with the platform.

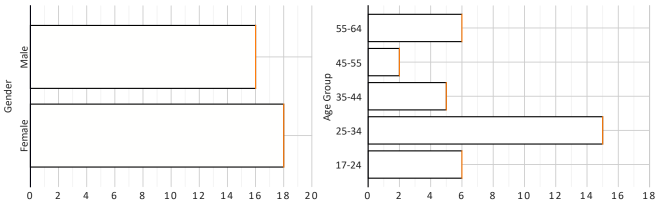


Figure 3: Age and Gender Distributions

The testing was done in person for the first participants in case of major technical issues. Afterwards, all tests were conducted without physical presence of the researchers, with participants still being able to reach out to the researchers via email in case of questions or issues. Despite the difference in remote e-mail and in-person support, the questions asked (and answered) were the same in both cases or situations. In particular, the researchers answered to the questions related to technical issues with the platform. In case of questions regarding the content or progress of the participants, the participants were advised to move on if they get stuck at any point, without providing explicit instructions. As some of the participants commented on the responsiveness issues during the first sessions, the maximum capacity regarding concurrent users was reduced in further sessions to ensure a smoother experience.

V. Results

A total of 40 participants completed the preliminary survey, of them, 34 participants have also completed the final survey of the study. Of these 34, 16 were men and 19 women, and the most common age group (15 out of 35) was between the ages of 25 and 34. The complete breakdown of the demographics is available at Figure 3.

The average completion time for the study did not provide a result, which rightfully gave an estimate of the completion time, as there were participants who have taken a break, and returned later. The median completion time was 54 minutes. The participants spent on average 5 minutes on the surveys, and the remaining time on interacting with the prototype.

The open-ended demographics questions, which focused on the participants' IT skills and employment, have been evaluated to ensure that the respondents fit within the target group. No samples had to be removed, as none of the answers suggested that a participant was an IT professional. Seven respondents stated they have good IT skills, often accompanied by the statement that they have no (professional) experience with IT (security).

We further describe the results of the analysis of the participants' responses. For analysis the hypotheses outlined in Section IV-C, the Wilcoxon test is used. The open-ended answers are analysed using open coding approach.

A. Quantitative Evaluation

The participants in our sample scored high on all the constructs before going through the intervention (see Table III), ranging from a mean score of 4.05 for the *cost* construct

	pre	post	changes
Perceived probability of security breach	5.07 (1.36)	6.23 (0.69)	1.16 (1.25)
Perceived severity of security breach	4.67 (1.31)	5.98 (0.93)	1.31 (1.27)
Security breach concern level	5.39 (0.83)	5.87 (0.77)	0.48 (0.83)
Response efficacy	5.67 (0.75)	6.01 (0.71)	0.34 (0.53)
Cost	4.06 (1.58)	4.5 (1.93)	0.44 (1.85)
Self-efficacy	5 (1.24)	5.59 (1.2)	0.59 (0.86)
Security policy compliance intention	6.13 (0.69)	6.44 (0.54)	0.31 (0.63)
Security policy attitude	6.16 (0.59)	6.54 (0.49)	0.38 (0.61)

Table III: Mean values (scale from 1 to 7) for each construct before the participants went through training (pre), after (post) and the difference between them. The values in brackets denote the standard deviation.

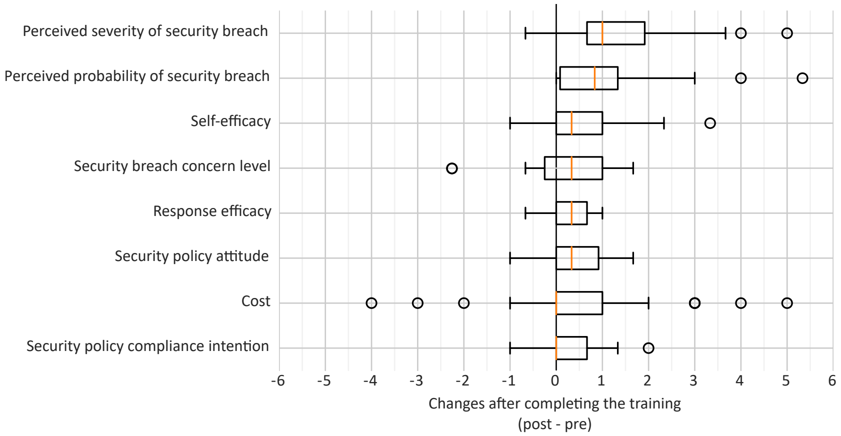


Figure 4: Distribution of changes in each construct (scale from -6 to 6)

(on a scale 1 to 7, with higher values indicating higher costs of security measures) to 6.16 for the *security policy attitude* construct (on a scale 1 to 7, with higher values indicating more positive attitudes towards security policies).

Nonetheless, going through the training has in general increased the participants' scores for all the constructs, see Figure 4. This increase were the most pronounced for *perceived severity of security breach*, with 89% of participants (24 out of 34) scoring higher after going through the training (i.e. being more likely to perceive potential breaches as severe) and 3% scoring lower; the least pronounced effect was on the perceived cost of the response, with only 38% (13 out of 34) scoring higher after the intervention (i.e. being more likely to estimate the costs of security measures as high) and 24% scored lower.

Prior to testing for hypotheses defined in Section IV-C, tests for normality have been conducted on the data, showing that the distribution of the difference between responses departed significantly from normality for the majority of the constructs. Based on testing for normality, the Wilcoxon Signed-Rank test was chosen as the primary test. The tests show significant improvements of all the constructs analysed according to the outlined hypotheses in Section IV-C, confirming H_1 ($p = 0.00453$, $V = 175$, effect size $r = .47$, moderate), H_2 ($p = 0.000276$, $V = 424$, effect size $r = .593$, large), H_3 ($p = 0.000164$, $V = 276$, effect size $r = .612$, large)[8].

B. Qualitative Analysis

The analysis of the qualitative part of the survey did not reveal any unexpected conclusions. For this analysis, the responses of the 34 participants were coded twice in parallel. After merging similar codes, there was 97% agreement in the coding between the two. The coding included three over-arching themes: **feedback** (recommendations), **positive** and **negative** comments. The final codes can be found in Appendix D.

While the majority of participants provided positive feedback (N=30), most of it was on a rather general level (N=11), such as *"A great tool to showcase how simply IT security could be compromised if an organization does not stay up-to-date"*. . More specific positive comments mentioned that the information was easy to follow (N=7) and that participants learned a lot (N=8)[9].

There were fewer participants that submitted negative comments (a total of N=3), though these comments were more detailed compared to the positive ones. Technical difficulties were commonly mentioned, especially the platform's responsiveness during the test (N=4) and issues with the keyboard (N=3) that

[8] The tests are conducted using R package "rstatix", the p-values are reported without adjustments for multiple comparisons. As we expected improvements in the constructs, the conducted tests are one-sided.

[9] Note, some of the participants' answers were assigned to multiple codes.

the first participants struggled with. Other negative comments included some instructions being hard to follow/understand (N=3).

A large part of the participants (N=17) did not give any recommendations or feedback regarding changes or additions to the platform. Multiple commented on grammar and spelling mistakes in the platform's content (N=4). A few comments were made regarding minor improvements in the user interface (N=2). Regarding content, more background information and technical explanations were proposed (N=2). Such recommendations included more on the possible consequences and the availability of this information to the general public (N=1). Examples of user-interface comments include a reset button in case the participant makes a mistake in the flow (N=1), more explicit visual changes throughout the challenges to indicate status and the alignment of some elements (N=1). The feedback also included recommendations outside the prototype's scope, such as follow-up activities by employers.

Finally, the participants were asked whether their perception of IT Security had changed. While a third of the participants responded that it either had not changed at all or did not change much (N=10), more participants reported that their perception has changed by realising how easy it is to hack/get hacked in the right circumstances (N=12), with people commenting: *”Pretty shocked how easy it was...”* and *”... It makes you see the danger”*. Participants furthermore reported becoming more convinced of the importance of proper passwords (N=10) and timely updates (N=3), as well as overall increased sense of urgency regarding IT security (N=6).

VI.Discussion and Conclusion

We have developed a prototype of a security awareness platform, aiming to utilise the adversarial perspective in demonstrating the feasibility of cyberattacks in case of common vulnerabilities such as outdated systems or weak passwords. In developing the prototype, we furthermore aimed to follow known best practices for effective security awareness, incorporating the learning principles outlined in [30]. Our evaluation of the prototype has demonstrated significant improvement in security policy compliance intention, security breach concern level and response efficacy. While the majority of qualitative feedback was positive, the majority of the negative references focus on performance issues of the implementation such as responsiveness, which can be improved via code optimisations or additional server capacity. Future work on the platform would include these and further technical improvements, as well as extensive security testing of the platform itself.

While our first evaluation of the prototype has shown promising results of the effectiveness of our proposed approach, we acknowledge that it has its limitations and there is a need for further extensive studies for a more conclusive evaluation of our proposal. In particular, evaluations with a larger sample size (also including older participants, as our sample tended to skew towards younger people) and a control group to compare the effectiveness of our approach to other proposals for security trainings would allow us to further understand the further validate our findings. Further evaluations that include measuring behaviour change as opposed to self-reporting, as well as further constructs that suggested in the Augmented PMT [19], would also be an important directions of future work. In this regard, the effectiveness of the feedback provided by the training in terms of suggested actions for protection against the attacks (e.g. using a password manager and/or multi-factor authentication to avoid password-related attacks) will also be evaluated.

A further direction of research would be working on the scalability and adaptability of our approach in terms of including further possible attack vectors. Resources such as OWASP Top 10 [22] as well as empirical studies involving a variety of stakeholders (e.g. security professionals) can be used as a guidance to ensure that the included attacks are sufficiently common to be relevant for the platform users. At the same time, ensuring that the selected attacks can be reasonably addressed by guidelines and security policies aimed at lay users is crucial to the response efficacy of the users.

Acknowledgment

The authors would like to thank Ingeniørforeningen (IDA) for participating in the experiment.

References

[1] L. A. Abuwardih, “Towards Evaluating Web Spam Threats and Countermeasures,” *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, 2018. [Online]. Available: http://thesai.org/Publications/ViewPaper?Volume=9&Issue=10&Code=ijacsa&SerialNo=65

[2] A. Alzubaidi, “Measuring the level of cyber-security awareness for cybercrime in Saudi Arabia,” *Heliyon*, vol. 7, no. 1, p. e06016, Jan. 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2405844021001213

[3] M. Bada, A. M. Sasse, and J. R. Nurse, “Cyber security awareness campaigns: Why do they fail to change behaviour?” *arXiv preprint arXiv:1901.02672*, 2019.

[4] Bulgurcu, Cavusoglu, and Benbasat, “Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness,” *MIS Quarterly*, vol. 34, no. 3, p. 523, 2010. [Online]. Available: https://www.jstor.org/stable/10.2307/25750690

[5] ENISA, *Unterstützung und Mittelbereitstellung durch die Unternehmensleitung während der Planung einer Sensibilisierungsinitiative*. GR: European Network and Information Security Agency, 2008. [Online]. Available: https://data.europa.eu/doi/10.2824/13275

[6] ——, *Information security awareness in financial organisations guidelines and case studies.* Heraklion: European Network and Information Security Agency, 2009.

[7] E. G. B. Gjertsen, E. A. Gjære, M. Bartnes, and W. R. Flores, “Gamification of Information Security Awareness and Training:,” in *Proceedings of the 3rd International Conference on Information Systems Security and Privacy*. SCITEPRESS - Science and Technology Publications, 2017, pp. 59–70, event-place: Porto, Portugal. [Online]. Available: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006128500590070

[8] P. A. Grassi, J. L. Fenton, E. M. Newton, R. A. Perlner, A. R. Regenscheid, W. E. Burr, J. P. Richer, N. B. Lefkovitz, J. M. Danker, Y.-Y. Choong, K. K. Greene, and M. F. Theofanos, “Digital identity guidelines: authentication and lifecycle management,” National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep. NIST SP 800-63b, Jun. 2017. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-63b.pdf

[9] B. Guttman and E. Roback, “An introduction to computer security :: the NIST handbook,” National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep. NIST SP 800-12, 1995. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-12.pdf

[10] N. Hansch and Z. Benenson, “Specifying IT Security Awareness,” in *2014 25th International Workshop on Database and Expert Systems Applications*. IEEE, Sep. 2014, pp. 326–330, event-place: Munich, Germany. [Online]. Available: https://ieeexplore.ieee.org/document/6974870

[11] T. Herath and H. R. Rao, “Protection motivation and deterrence: a framework for security policy compliance in organisations,” *European Journal of Information Systems*, vol. 18, no. 2, pp. 106–125, Apr. 2009. [Online]. Available: https://www.tandfonline.com/doi/full/10.1057/ejis.2009.6

[12] IBM, “Cost of a data breach report,” https://www.ibm.com/security/digital-assets/cost-data-breach-report/1Cost%20of%20a%20Data%20Breach%20Report%202020.pdf/, 2022, last accessed on 16.11.2022.

[13] Johnston and Warkentin, “Fear Appeals and Information Security Behaviors: An Empirical Study,” *MIS Quarterly*, vol. 34, no. 3, p. 549, 2010. [Online]. Available: https://www.jstor.org/stable/10.2307/25750691

[14] D. L. Kirkpatrick and J. D. Kirkpatrick, *Evaluating training programs: the four levels*, 3rd ed. San Francisco, CA: Berrett-Koehler, 2006.

[15] B. Lebek, J. Uffen, M. H. Breitner, M. Neumann, and B. Hohler, “Employees’ Information Security Awareness and Behavior: A Literature Review,” in *2013 46th Hawaii International Conference on System Sciences*. IEEE, Jan. 2013, pp. 2978–2987, event-place: Wailea, HI, USA. [Online]. Available: http://ieeexplore.ieee.org/document/6480203/

[16] P. Mayer, A. Kunz, and M. Volkamer, “Reliable Behavioural Factors in the Information Security Context,” in *Proceedings of the 12th International Conference on Availability, Reliability and Security*. ACM, Aug. 2017, pp. 1–10, event-place: Reggio Calabria Italy. [Online]. Available: https://dl.acm.org/doi/10.1145/3098954.3098986

[17] Michael Fagan and M. M. H. Khan, “Why Do They Do What They Do?” p. 18, Jun. 2016.

[18] J. D. Nosworthy, “Implementing Information Security In The 21st Century — Do You Have the Balancing Factors?” *Computers & Security*, vol. 19, no. 4, pp. 337–347, Apr. 2000. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167404800040219

[19] M. Oakley, S. Mohun Himmelweit, P. Leinster, and M. Casado, “Protection Motivation Theory: A Proposed Theoretical Extension and Moving beyond Rationality—The Case of Flooding,” *Water*, vol. 12, no. 7, p. 1848, Jun. 2020. [Online]. Available: https://www.mdpi.com/2073-4441/12/7/1848

[20] V. Pastor, G. Diaz, and M. Castro, “State-of-the-art simulation systems for information security education, training and awareness,” in *IEEE EDUCON 2010 Conference*. IEEE, Apr. 2010, pp. 1907–1916, event-place: Madrid. [Online]. Available: http://ieeexplore.ieee.org/document/5492435/

[21] A. Pollini, T. C. Callari, A. Tedeschi, D. Ruscio, L. Save, F. Chiarugi, and D. Guerri, “Leveraging human factors in cybersecurity: an integrated methodological approach,” *Cognition, Technology & Work*, vol. 24, no. 2, pp. 371–390, May 2022. [Online]. Available: https://link.springer.com/10.1007/s10111-021-00683-y

[22] O. W. A. S. Project, “Owasp top 10,” https://owasp.org/Top10/, 2021, last accessed on 14.09.2022.

[23] T. G. Reio, T. S. Rocco, D. H. Smith, and E. Chang, “A Critique of Kirkpatrick’s Evaluation Model,” *New Horizons in Adult Education and Human Resource Development*, vol. 29, no. 2, pp. 35–53, Apr. 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/nha3.20178

[24] R. W. Rogers, “A Protection Motivation Theory of Fear Appeals and Attitude Change1,” *The Journal of Psychology*, vol. 91, no. 1, pp. 93–114, Sep. 1975. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/00223980.1975.9915803

[25] D. W. Straub and R. J. Welke, “Coping with Systems Risk: Security Planning Models for Management Decision Making,” *MIS Quarterly*, vol. 22, no. 4, p. 441, Dec. 1998. [Online]. Available: https://www.jstor.org/stable/249551?origin=crossref

[26] M. Volkamer, K. Renaud, O. Kulyk, and S. Emeröz, “A Socio-Technical Investigation into Smartphone Security,” in *Security and Trust Management*. Cham: Springer International Publishing, 2015, vol. 9331, pp. 265–273. [Online]. Available: http://link.springer.com/10.1007/978-3-319-24858-5_17

[27] M. Wilson, D. E. de Zafra, S. I. Pitcher, J. D. Tressler, and J. B. Ippolito, “Information technology security training requirements :: a role- and performance-based model,” National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep. NIST SP 800-16, 1998. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-16.pdf

[28] T. Yadav and A. M. Rao, “Technical Aspects of Cyber Kill Chain,” in *Security in Computing and Communications*, J. H. Abawajy, S. Mukherjea, S. M. Thampi, and A. Ruiz-Martínez, Eds. Cham: Springer International Publishing, 2015, vol. 536, pp. 438–452. [Online]. Available: http://link.springer.com/10.1007/978-3-319-22915-7_40

[29] E. O. Yeboah-Boateng, *Cyber-security challenges with SMEs in developing economies: issues of confidentiality, integrity & availability (CIA)*. Aalborg: Department of Electronic Systems, Aalborg University, 2013.

[30] L. Zhang-Kennedy and S. Chiasson, “A Systematic Review of Multimedia Tools for Cybersecurity Awareness and Education,” *ACM Computing Surveys*, vol. 54, no. 1, pp. 1–39, Apr. 2021. [Online]. Available: https://dl.acm.org/doi/10.1145/3427920

Appendix

Find further supplemental material in the GitHub repository: https://github.com/niekjanssen/coursecurity

A. *Invitation Message*

Dear [Name],

Thank you for your interest in testing the Coursecurity platform. As part of the platform, you will be able to hack systems and applications. By using network and hacking tools, and social engineering. All tools that are part of the platform are used in the real-world, and are freely available to the general public online. On the platform, you will get to experience these tools first hand in a safe and closed environment with simulated systems.

For the evaluation of the platform, you are asked to fill out two similar questionnaires regarding your attitude towards IT Security. One when you start, and one once you have finished the challenges on the platform. Your answers will be treated confidentially, and only be discussed with others in anonymized and/or aggregated form. In total, the test takes approximately 30 minutes. Testing Process:

1) Answer questionnaire 1
2) Watch introductory video
3) Test the platform itself
4) Answer questionnaire 2

As part of the questionnaires and the platform, you are asked to enter your e-mail address. This allows us to contact you with potential follow-up questions (if you give consent to that), and allows us to combine your answers.

If you want to guarantee anonymity, you are welcome to put a random string instead of your e-mail address, but please ***make sure to put exactly the same*** *for the different stages of the test.*

Your participation is voluntary, by taking part you consent to the processing of your answers in anonymized and aggregated form as part of our Master Thesis. At any point during participation, you are able to withdraw.

In case you have technical issues while testing the platform, you can reach-out via [email]. You can start the test here: [link]

Thank you for participating!
Kind regards,
[Signature]

B. *Survey Instruments*

As part of the preliminary survey, the participants were requested to fill out some demographic information, due to the potentially sensitive nature these questions were not mandatory to complete the survey.

Demographics

- E-mail address (or pseudonym) to match the different elements of the study (Mandatory, open text field)
- Permission to contact them after in case of follow-up questions (Yes/No)
- Age Group (Not Mandatory, 10 year groups)
- Gender (Female/Male/Non-binary/Prefer not to say)
- Department/Organisation (Not mandatory, open text field)
- How would you describe your current skills within IT and IT Security? (Not mandatory, open text field)

The following questions makes use of a 7 point Likert-scale. The questions are adapted from [11], all questions related to constructs that align with Protection Motivation Theory were selected. The same questions were used to construct the preliminary- and post-survey.

Perceived probability of security breach

- How likely is it that a security violation will cause a significant outage that will result in loss of productivity?
- How likely is it that an security violation will cause a significant outage to the internet that results in financial losses to your organization
- How likely is it that your organization will lose sensitive data due to a security violation?

Perceived severity of security breach

- I believe that information stored on my organization's computers is vulnerable to security incidents.
- I believe the productivity of my organization and its employees is threatened by security incidents.
- I believe the profitability of my organization is threatened by security incidents.

Security breach concern level

- The information systems security issue affects my organization directly.
- The information systems security issue is exaggerated.
- I think information systems security is serious and needs attention.

Response efficacy

- Every employee can make a difference when it comes to helping to secure the organisation's information systems.
- There is not much that any one individual can do to help secure the information systems of my organization.
- If I follow the organization information systems security policies, I can make a difference in helping to secure my organization's information systems.

Cost

- Adopting security technologies and practices poses hindrance.

Self-efficacy

- I would feel comfortable following most of the information systems security policies on my own.

- If I wanted to, I could easily follow information systems security policies on my own.
- I would be able to follow most of the information systems security policies even if there was no one around to help me.

Security policy compliance intention

- I am likely to follow organisational security policies.
- It is possible that I will comply with organisational information systems security policies to protect the organisation's information systems.
- I am certain that I will follow organisational security policies.

Security policy attitude

- Adopting security technologies and practices is important.
- Adopting security technologies and practices is beneficial.
- Adopting security technologies and practices is helpful.

As part of the post survey, the participants were asked some open-ended feedback questions regarding their experience.

Feedback Questions

- What did you think of the learning platform as a whole?
- Has your perception of IT security changed? - If so, how specifically, and is there any particular element that caused it?
- Is there something you would improve or add?

C. Guide Text

Challenge 1 - Updates

Introduction - Text

Introduction section: Throughout the different scenarios you are trying to get information from PIA. As you are playing the hacker, you will be trying to find vulnerabilities in their systems that you can misuse.This first challenge will cover elements of investigating a hacking target, and security of software that is connected to the internet.

Background information: The internet is nothing more than many computers connected to each other. Computers like the one you work on, or larger ones dedicated to running websites or applications which are also called servers. Even though all computers are connected, it doesn't mean all computers on the internet are accessible. A well configured computer (or server) only allows the necessary connections.As a hacker the trick is to find out what connections are possible, and to misuse these entry points.

Challenge information: You are interested in knowing who works at PIA and what they earn. In order to get access to this information you will investigate any security holes and try to exploit them.You already know that:

1) They use Maxisoft Home247 for cloud services
2) Their usernames look like [the users initials] @ pia.dk
3) The information you want is in sharedot.com

Guide - Text

Tasks 1 - Explore the target: (Description) Lets go back to the target from last challenge.

Step 1:

1) (Instruction) Open Google Chrome
2) (Explanation) You can find the icon in the bottom left corner.

Step 2:

1) (Instruction) Go to the page sharedot.com
2) (Explanation) Go to the page sharedot.com using Google Chrome

Tasks 2 - Password stuffing: (Description) During the intro you learned about different ways of attacks towards passwords. Let's use password stuffing to try to breach the file. Remember Password stuffing involves using else where found credentials and try to stuff them into the system.

Step 1:

1) (Description) Remember, we stored the file on the desktop of the virtual computer.
2) (Instruction) Try to see if you can log in using any of the found users. Try SMF or VLA.
3) (Explanation) If you're not successful go to next step.

Tasks 3 - Creative Password stuffing: (Description) When password stuffing the success rate is often very low, but luckily you only need to be lucky once.

Step 1:

1) (Instruction) Try to login using jcd@pia.dk
2) (Explanation) If you've already tried to login using JCD, then check the box of and go to next step

Step 2:

1) (Instruction) Try to login in with a password using a similar pattern
2) (Explanation) People, and in this case JCD, often create passwords following specific patterns. Try type a password which would match the password found with current information.
3) (Dialog) explaining terminal output should be available

Tasks 4 - Explore sharedot.com: (Description) Great you managed to get in! Let's see what we can find of files on the internal drives.

Step 1:

1) (Instruction) Navigate around on the server and see what you can find.
2) (Explanation) You can see the folders on the left and the files, if any on the right. Simply click on them to navigate around. When you feel ready move to next task.

Tasks 5 - Let's find the personel files: (Description) Unless, you haven't already found them. Let's find the payslip files.

Step 1:

1) (Instruction) Go to the root folder
2) (Explanation) Click on the house icon

Step 2:

1) (Description) The HR folder seems like the one with highest possibility of containing the files we are looking for.
2) (Instruction) Go to the HR folder
3) (Explanation) Click on the HR folder in the left panel

Step 3:

1) (Description) The folder contains a temp folder. This indicates its a temporary folder, which probably is going to be deleted at some point.
2) (Instruction) Go to the Temp folder
3) (Explanation) Click on the Temp folder in the left panel

Step 4:

1) (Instruction) Open the file in the folder
2) (Explanation) To the right you can see the content of the folder. It contains one file (2019Q1.pdf). Let's open it!

Tasks 6 - Great you found the payslips!: (Description) You managed to get the information you were looking for and can now use this information to anything you want! E.g. fraud, phising email, etc.

Step 1: (Instruction) Good job!

Tasks 7 - Leave the crime scene: (Description) Great job, you've managed to hack the system and gain helpful information. Now let's leave the server.

Step 1:

1) (Description) Now move back to the terminal where you used Metasploit (where you wrote msfconsole)
2) (Instruction) As a last command before you leave execute wipe
3) (Explanation) When executing the command you will delete all data on the server. You can see the effect in Google Chrome if you have the page still open.

Step 2:

1) (Instruction) When you're ready to leave, simply close the two terminals.
2) (Explanation) This will also close the connection to the server automatically.

Debriefing - Text

Success section: You've managed to break-in to PIA! So even though one of PIA's systems was maintained by a large software company and up-to-date, an old — perhaps even forgotten — system was vulnerable because it had not been updated in a while. The information you've found so far may not be shocking or highly classified but it was not supposed to be public and it might help you get access to other information.

Danger section: Flaws are discovered in software all the time and security holes are normally fixed quickly after they've been found. This applies to both large servers, but also apps on your phone, or software running on your own computer. However, until a computer is actually updated the security hole will persist. With the update the problem often becomes public, so until the update is applied your device is vulnerable to a problem that the whole world knows about. Since many applications and devices are connected through a company network, even if a seemingly small or unimportant system gets compromised it can be the gate towards bigger issues.

Prevention section: Company-wide systems like this example are, of course, the responsibility of IT Specialists. However as mentioned before, the same principles apply on a smaller scale as well.As a hacker you only need one small mistake to get in, and often you can move further up the network from that moment. A single laptop that isn't updated can be enough to get access and do damage beyond that one device. Of course there are company wide security measures, but their effectiveness is greatly reduced if individuals don't take care of their IT Security. As an individual it's thus important to update your applications and devices as soon as possible, and to be careful with applications.

Challenge 2 - Passwords

Introduction - Text

Introduction section: During the last challenge you managed to gather information which might a potential aid in breaching sharedot.com using attacks on passwords. Throughout this challenge we'll show you how information obtained else where can be of use to gain access to other systems. Additionally we'll try to teach you different important factors to consider, when creating passwords for a system.

Background information: When attacking passwords there exist several different ways of doing so. Below we've explained only a few of them. Brute Force attack: Brute force attacks is when an attacker try to force there way through, by trying every single combination of passwords. A computer is incredible fast a can try a lot of combination very quick. It would for example take seconds to try every combination for a 4 digit pin. A Brute force attack is good against weaker passwords shorter passwords, but when the passwords is over 8 characters it starts to take to long. Dictionary attack: A Dictionary attack is when an attacker try to force there way through, by using a predefined list with combination as potential passwords. The list could be a dictionary with words, a list made out of combination, etc. Using a dictionary attack an attacker can breach passwords based on longer words which would take to long, when doing a brute force. A dictionary attack is good against users that are using patterns for creating there passwords. However, struggles against random combinations. Password stuffing: Stuffing is when an attacker uses information found else where, e.g. usernames and passwords, and try to stuff it into a different system. Password stuffing is good when users uses the same password across different platform, since a weakness is one platform results in a weakness in other.

Challenge information: You are still trying to figure out who works at PIA and see what they earn. During last challenge you gained valuable information, which might be able to aid you in your quest to access sharedot.com You will do this by using Password attacks as mentioned on last slide.

Guide - Text

Tasks 1 - Explore the target: (Description) Lets go back to the target from last challenge.

Step 1:

1) (Instruction) Open Google Chrome
2) (Explanation) You can find the icon in the bottom left corner.

Step 2:

1) (Instruction) Go to the page sharedot.com
2) (Explanation) Go to the page sharedot.com using Google Chrome

Tasks 2 - Password stuffing: (Description) During the intro you learned about different ways of attacks towards passwords. Let's use password stuffing to try to breach the file. Remember Password stuffing involves using else where found credentials and try to stuff them into the system.

Step 1:

1) (Description) Remember, we stored the file on the desktop of the virtual computer.
2) (Instruction) Try to see if you can log in using any of the found users. Try SMF or VLA.
3) (Explanation) If you're not successful go to next step.

Tasks 3 - Creative Password stuffing: (Description) When password stuffing the success rate is often very low, but luckily you only need to be lucky once.

Step 1:

1) (Instruction) Try to login using jcd@pia.dk
2) (Explanation) If you've already tried to login using JCD, then check the box of and go to next step

Step 2:

1) (Instruction) Try to login in with a password using a similar pattern
2) (Explanation) People, and in this case JCD, often create passwords following specific patterns. Try type a password which would match the password found with current information.
3) (Dialog) explaining terminal output should be available

Tasks 4 - Explore sharedot.com: (Description) Great you managed to get in! Let's see what we can find of files on the internal drives.

Step 1:

1) (Instruction) Navigate around on the server and see what you can find.
2) (Explanation) You can see the folders on the left and the files, if any on the right. Simply click on them to navigate around. When you feel ready move to next task.

Tasks 5 - Let's find the personnel files: (Description) Unless, you haven't already found them. Let's find the payslip files.

Step 1:

1) (Instruction) Go to the root folder
2) (Explanation) Click on the house icon

Step 2:

1) (Description) The HR folder seems like the one with highest possibility of containing the files we are looking for.
2) (Instruction) Go to the HR folder
3) (Explanation) Click on the HR folder in the left panel

Step 3:

1) (Description) The folder contains a temp folder. This indicates its a temporary folder, which probably is going to be deleted at some point.
2) (Instruction) Go to the Temp folder
3) (Explanation) Click on the Temp folder in the left panel

Step 4:

1) (Instruction) Open the file in the folder
2) (Explanation) To the right you can see the content of the folder. It contains one file (2019Q1.pdf). Let's open it!

Tasks 6 - Great you found the payslips!: (Description) You managed to get the information you were looking for and can now use this information to anything you want! E.g. fraud, phishing email, etc.

Step 1: (Instruction) Good job!

Debriefing - Text

Success section: Great job! You've managed to break into all systems and gather all the information needed! And now a well deserved break from hacking seems reasonable! You will not need to use the information gathered any further. However, a real hacker could use the information gathered for much more... A hacker could you use the information to: Create personalised phishing emails to specific targets (Spear phishing) Use the information gathered as a whole (about departments, employees etc.) to target important personnel by winning their trust with ”inside information” Black mailing ...And much more! However, all that is for another day.

Danger section: Authentication, such as a password, is essential for making the internet function probably. Today everyone uses many different systems, with different purposes. These systems greatly vary in the level of security protecting them. According to the sensitivity of the data in a system. Even though users are well aware that some systems are more important that others, re-using passwords across different platforms and systems is very common. Thereby making themselves vulnerable towards Password stuffing as seen in the prior challenge. A small webshop getting hacked may not be the end of the world. Though, if you use the same password for everything work-related, one data-leak at a small supplier can expose your username (email address) and password to way more important, and internal systems.

Prevention section: Luckily, more and more measures are created to enhance the security around authentication. Such as multi-factor authentication, where the users are asked to use an additional token or metric to login. E.g. a code they receive over SMS, or using a fingerprint. However technical solutions can only be effective when combined with careful users. When creating passwords you can follow the guidelines made by the American Government for it's secret services "https://pages.nist.gov/800-63-3/sp800-63b.html" The National Institute of Standards and Technology The most important points to remember when creating a password: Always make passwords as long as possible. Absolute minimum is 8 characters long. It's more effective to add a few more letters than to use special characters.Never use the same password across different platforms/websites. Never!When switching passwords don't follow a specific pattern. (Like seen in the assignment)

Transcript of Video

Thank you for filling out the first part of the survey, and welcome to this introduction video. After this short video you can do the practical part of our research. On the platform you will look at IT security from the perspective of a hacker. You will be using real life hacking tools to exploit servers and applications in this closed environment. Firstly, we will introduce you to the platform and give you a quick overview of how to use it. Initially, when entering the platform you'll be greeted by this screen. This screen will provide you with background information about the upcoming challenge. At your own pace, go through the information sections using the arrow buttons in the bottom of the screen. The last time you click the arrow button you'll close the information sections, and go to the actual platform. Here you see the challenge screen. On the left there is a virtual computer, which is located in a closed environment. Here we can safely try out everything. This is where you'll perform the tasks. You'll never have to do anything on your own computer. All tasks should be performed in the left side of this screen. It works just like your own computer, it just might be a little slower than you are used to. Also, important is that you cannot copy-paste from your computer to the virtual computer. You will have to type the text yourself when necessary. On the right side of the screen you will find the guide, which will guide you through all the steps you have to take. In the guide, there are different sections with subtasks. This is one section, called "Explore the target". As an example, the first task is open Google Chrome. Whenever you have to open something, you can normally find it in the bottom left corner of the virtual computer. Click it, and in a few moments Google Chrome will appear. When it has opened, you can check off this subtask, by clicking the checkbox. This will help you keep track of what tasks you've done and not. After you've checked off all subtasks, the section will be complete, and the next section will unfold. This way you progress down, until you've finished all tasks of the challenge. When you're done with the challenge. You can click "finish challenge" When doing so, you will be shown similar information screens, as explained during the start of this video. Afterwards you can move to the second challenge, which will start with an introduction again. When you're done with both challenges you'll be asked to fill out the second half of the survey. A last note, is that you also have to use the terminal, (or command line) some times. This might be intimidating, but simply follow all the steps in the guide. You type the command from the guide in the black screen. And click enter on your keyboard to execute a command. After a few moments the "result" of your command will appear on the lines below it. After that you can write a new command In case you get stuck at a certain task, it's better to simply skip a step than to give up completely. Thank you, and the best of luck.

D. Codebook

Name	*Description*	*ref.*
Feedback		**45 ref.**
- Addition	Feedback / Recommendation that proposes adding an element to the platform or training as a whole.	286 words (11 ref.)
- External	Feedback / Recommendation that proposes an external improvement / addition, e.g. recommendations that could only be executed by the participant's organisation.	24 words (2 ref.)
- None	Statement that participant has No Feedback or Recommendations.	63 words (16 ref.)
- Spelling	Spelling or Grammatical mistakes in the platform	134 words (7 ref.)
- User Interface	Feedback / Recommendation that proposes changes to the user interface, styling of the platform or guide.	184 words (7 ref.)
- Video	Feedback / Recommendations regarding the introduction video content and/or design.	18 words (2 ref.)
Negative		**27 ref.**
- Difficult	Negative comment regarding the difficulty (of understanding) the information / instructions on the platform.	107 words (10 ref.)
- *Technical*		
- Slow	Negative comment regarding the reaction time/speed/overall responsiveness of the platform.	90 words (9 ref.)
- Typing	Negative comment regarding the keyboard issues some participants faced.	83 words (8 ref.)
Perception Change		**102 ref.**
- Easier	Comment that the participant's perception has changed because hacking was easier than the participant expected.	399 words (31 ref.)
- General	Comment that the participant's perception has changed without an explanation why or how.	67 words (7 ref.)
- None	Comment that the participant's perception has not changed.	233 words (18 ref.)
- Passwords	Comment that the participant's perception has changed especially regarding the theme of password security.	300 words (23 ref.)
- Vulnerability	Comment that the participant's perception has changed especially regarding the vulnerability of them/their organisation.	115 words (9 ref.)
- Updates	Comment that the participant's perception has changed especially regarding the theme of updating software.	77 words (7 ref.)
- Urgency	Comment that the participant's perception has changed because a sense of urgency has occurred with the participant after the training.	93 words (7 ref.)
Positive		**134 ref.**
- Curious	Positive comment citing that it the training has sparked curiosity regarding IT Security with the participant.	40 words (2 ref.)
- Experience	Positive comment describing the training as a positive experience overall.	77 words (8 ref.)
- Fun	Positive comment describing the training as fun.	38 words (10 ref.)
- General	Positive comment about the platform without further explanation.	344 words (58 ref.)
- *Information*		
- Easy to Follow	Positive comment regarding how easy to follow along the training was.	172 words (20 ref.)
- Insightful	Positive comment regarding how insightful the training was.	366 words (34 ref.)
- Video	Positive comment regarding the introductory video.	6 words (2 ref.)