Digital-PASS: A Simulation-based Approach to Privacy Education

Kambiz Ghazinour
Center for Criminal Justice, Intelligence and Cybersecurity
State University of New York, Canton, NY, USA
ghazinourk@canton.edu

Ken Messner
Department of Computer Science
Kent State University
Kent, OH, USA
kmessne3@kent.edu

Sean Scarnecchia
Department of Computer Science
Kent State University
Kent, OH, USA
sscame1@kent.edu

David Selinger
Department of Computer Science
Kent State University
Kent, OH, USA
dselinge@kent.edu

ABSTRACT

With the increased proliferation of social media in the modern age, education on the potential dangers facing consumers in social media has not kept commensurate pace. Conventional education methods and standards have not proved to be effective in privacy education and increasing user awareness, and newer methods to bring safety knowledge to the public need to be introduced. In this paper, we propose that education on usable security and privacy in a social media context utilizing a simulation-based framework would bring promising results, as it has in other fields. We then describe the challenges in building such a system for educating people about privacy on social media and propose our own system named Digital-PASS, a simulation-based educational model for raising awareness. Our simulation model utilizes gamification as a principle means of user motivation. We then examine and discuss the effectiveness of our model through a detailed analysis of four case studies.

CCS CONCEPTS

• **Security and privacy** → Human and societal aspects of security and privacy; Social network security and privacy, Privacy protections;

KEYWORDS

Data privacy; social media; education; simulation; education; privacy-enhancing technologies

ACM Reference format:

Kambiz Ghazinour, Ken Messner, Sean Scarnecchia, David Selinger. 2019. Digital-PASS: A Simulation-based Approach to Privacy Education. In *Proceedings of ACM WPES conference (WPES'19). ACM, London, United Kingdom, 13 pages.* https://doi.org/10.1145/3338498.3358647

WPES'19, November 11, 2019, London, United Kingdom © 2019 Association for Computing Machinery. ACM ISBN 978-1-4503-6830-8/19/11…$15.00 https://doi.org/10.1145/3338498.3358647

1. Introduction

Social media has become an increasingly central component in the lives of a significant number of individuals. The lifestyle of the average individual has become more public, and large amounts of personal data are now accessible online by both corporations and the public, whether the data owner is aware of this access or not. For many, this represents a great benefit and opportunity to broaden their social circles and maintain deep social connections across long distances, but it comes with equally great risks. Social media has increasingly represented the primary means of internet-based communication for a growing share of users. Today, digital crime, identity theft, and misuse of personal information make online security and privacy even more crucial; however, education about security and privacy has not developed accordingly, particularly within the context of social networks. While many have been educated on potential dangers coming from a medium such as email, social networks as a threat vector are often not touched upon. We have mainly relied on individual social networks informing their users about the privacy policies of their own sites, with the intention that these users would read them, be aware of how their information is handled, and manage their own security.

For the most part, leaving social media consumers to their own recognizance has not been an effective strategy; as a result, a considerable portion of social media users are unaware of how the actions they take and the information they share online can endanger their personal privacy. Between this and the increased incentives social networks provide for interacting with their network at a deeper level, such as making more information public for the purposes of their directed marketing, adversaries have a wealth of information to exploit. Simple information commonly shared publicly, such as birth dates and zip codes, can be used by adversaries to obtain sensitive information through various identity theft techniques, so it is crucial that this sensitive data is protected. The principle areas of danger for consumers consist of:

1) Information compiled by adversaries that can be used to impersonate the target's friends or family in attempts to secure more data or engage in theft (commonly known as "phishing").
2) Publicly available data on targets that can be used by adversaries to query databases and find additional data.
3) Sensitive data that can end up being posted publicly due to user indiscretion and ignorance of available privacy settings.

Few social media users are aware of the dangers that commonplace data can potentially present to consumers. There is a clear and present need for increased education on social media behavior and awareness of the dangers to privacy on social media, so that consumers can prevent and avoid these potential dangers.

Traditional lecture-based education may be insufficient to the task due to lack of engagement, as well as lacking the immediacy of first-hand experience. However, given the harsh impact of a first-hand identity theft experience, consumers should not have to experience such an event without training.

We therefore believe that the ideal education experience for consumers would be one that incorporates first-hand experience of identity theft, without the dangerous consequences that go along with it. Based upon this principle, we are designing a new means of privacy-awareness instruction through simulation; that is, the creation of privacy-compromising experiences under safe, controlled conditions. An important aspect of developing any simulation is determining the conditions necessary to achieve the desired effect, and an important part of developing any educational curriculum is providing motivation to students to inspire them to learn. We endeavor to achieve both of these objectives simultaneously, through the use of a simulation with gamification principles as a centerpiece.

Gamification is defined as the use of game mechanics, dynamics, and frameworks to promote desired behaviors [8, 12]. Gamification has a twofold means of increasing student engagement in the educational process: 1) It creates an incentive for the participants to learn the lesson, as they want to win the game. 2) It provides an *organically* created experience by which people can learn lessons for themselves, without being overtly taught by educators. In the context of our simulation, gamification elements are utilized to form a consistent motivational element that will drive users through the simulation process and ensure repeated usage, which will further drive home the desired lessons and privacy practices.

2. Related Work

2.1 Gamification Methods

There has been a great deal of research indicating the efficacy of gamification as a learning tool when employed in conjunction with traditional methods.

Lee & Hammer speak highly of gamification when used properly, suggesting that it can motivate students to engage in the classroom, provide teachers with tools to track student progress, and blur the boundaries between formal and informal learning. [23]

Borges [2] indicates in his mapping of gamification endeavors that there were few initiatives focusing on Computer Supported Collaborative Learning (CSCL). This is the area we intend to focus on.

Kiryakova's study of gamification methods across education concluded that gamification represents an effective approach in education to make a positive change in student's behavior and attitude towards learning, which improves motivation and engagement. [22]

Hamari [15] indicates that engaging in a gamified education system socially, i.e. in a class or network environment, would increase positive attitude towards the system, and therefore towards learning.

There are also several studies that have presented literature reviews on gamification. Hamari's literature review on the effectiveness of gamification concluded that gamification was overall effective, and that the majority of studies covered by the review yielded positive effects on student education. [16]

Caponetto's review agreed, indicating that one of the principal reasons for employing gamification as an educational practice is to increase student engagement. [3]

Surendeleg's review builds upon the previous reviews, offering the suggestion that gamification in education should be expanded beyond the computer into practical motor skills, or onto phones. Given that a large amount of social network activity occurs in a mobile environment, this is a suggestion worth keeping in mind. [28].

2.2 Simulation-based General Education

Digital education has been utilized in many areas and fields in educating students, particularly when employed in conjunction with game-based incentive systems with great success in many fields across a wide age range.

Muntean's study on engagement in e-learning [26] shows that engagement is the most important metric for the overall success of an e-learning program. Specific means of generating engagement are also outlined, such as regular feedback, creating a social environment, and providing personal status for learning. We have endeavored to implement all of these attributes in our model.

In [19] the authors indicated methods of digital game-based learning to teach UNITY programming. Using webGL, they established a flipped classroom curriculum. Their aim was to provide students with additional motivation to learn computer programming.

[18] further indicates the effectiveness of game-based incentivization structures in the teaching of poetry to young children in particular. In Taiwan, young children are encouraged to memorize Chinese Poetry before they are able to understand the meaning. This project utilized game-based learning techniques to assist and motivate nursery school age children in memorization. Comparisons between the game-playing group and the control group showed that gamification provided a significant benefit to retention.

In [31] and [4], the authors have used gamification principles to develop a digital game-based learning system with a graduated prompting strategy for teaching math courses. In another research [17], the authors have created a 3D game-based learning system in a virtual world to educate lower achieving students in mathematics, and their results show a huge improvement in the outcome.

Ern in [10] shows the benefit of using gamification and serious games within interventions for children with autism spectrum disorder. Gamification is often used to help individuals learn how to use an application. For instance, Li et al. [24] introduce a model called GamiCAD, which is a gamified tutorial system for first time AutoCAD users.

There have also been studies that compared social networking and gamification in e-learning [6; 9], and found that both approaches presented better performance than a traditional e-learning approach in terms of academic achievement for practical assignments, but that when it came to assessing knowledge, the traditional e-learning approach was better compared to social networking and gamification for the particular course that they compared. Other studies such as [33] have studied the high impact of mobile and web apps games on enhancing the education experience of the students.

2.3 Using Gamification in Security Education

Aside from the areas discussed above, gamification has also been used in the realm of teaching digital security.

Capture the Flag (CTF) exercises are a common means of instruction about system security and protection from infiltration, both in attack-defense form and in pure defense form. Gondree and Peterson in particular used the CTF format to develop a long-term educational curriculum to be used over several weeks to increase engagement in Computer Science education by having students work together to accomplish challenges. This increased both the appeal of the work and student engagement. [14]

In 2006, the US Navy developed a scenario-based game for teaching organizational network security entitled CyberCIEGE [5]. It is used as an educational training tool for computer and network security technicians. It has been requested by over four hundred educational institutions worldwide, and can be played by any individual with access to a Windows operating system [29]. The game itself runs as a stand-alone application with a single player. The scenarios that the game provides are organized into campaigns, with each campaign addressing different network security topics. It also includes an online help facility called the encyclopedia that describes the security concepts.

The objective of CyberCIEGE is to create opportunities for students to learn about the decisions that go into maintaining network security. The game includes over twenty scenarios where students play the role of a decision maker for enterprises such as businesses or government divisions [29]. Their duties require them to purchase and configure workstations, servers, operating systems, applications, and network devices, as well as make important choices about security arrangements, balancing security and efficiency while under attack from hackers and potentially well-motivated professionals, with the consequences of their actions in full view. The game motivates players to make good decisions through an in-game rewards system, creating a feedback loop to help students take the lessons to heart.

3. Digital-PASS: Creating Privacy Awareness on Social Media Through Simulated Experience

When we decided to design our educational program based on gamification principles, it was important that we remained cognizant of what makes gamification effective in the first place. From the above examples of educational curricula utilizing gamification as an important component, we distilled three principal criteria that had to be met for the creation of our own educational game:

1) Users must be motivated to play the game and to continue playing the game through the use of a positive feedback system.

2) The desired educational concepts should be introduced in a manner that is not forced, and makes sense within the context of the game.

3) Rather than endlessly reiterating and spelling out the intended lesson to users, conditions should be created that allow users to come to the correct conclusions on their own.

Using these principles as a guide, we created the model for an educational simulation entitled Digital-PASS, used to create organic educational experiences in a simulated social network.

3.1 Summary of Game Mechanics

In our simulation model, users of a social network will be broadly defined into two groups: one that uses the social network for its intended purpose, who regularly create and view content, hereafter called *posters*. The second group uses social networks for *malicious* purposes. Identity theft is only one such purpose, but for simplicity, we will refer to this group collectively as *thieves*.

The general form of interaction users within our simulation would experience is relatively simple, as shown in Fig 1: Posters, acting according to their impulses and the various reasons they are on the system, generate *content*. This content sometimes includes information that can be compromising. Thieves, however, actively search around for prospective targets, reading such information and making use of it, often to look the target poster up in a more detailed data repository using the information that they already have. They can also actively attempt to compromise posters' identities by engaging in active con tactics such as phishing.

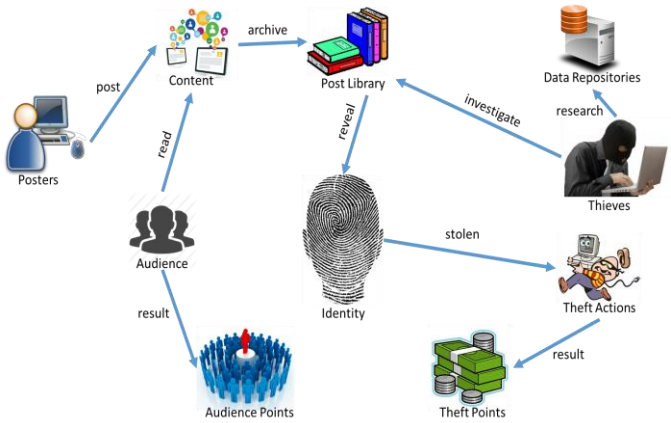

Fig. 1: Flowchart Illustration of Digital-PASS

The basic system of the game is as follows:

Players are divided into two groups, posters and thieves, with posters being the larger group. Posters create content, which is rated and scored by popularity with a computer-controlled audience, creating *audience points*. Their objective is to beat the other posters and reach the position of number one on the leaderboard. This competitiveness creates the incentive for posters to remain in the system, despite the risks. If they refuse to post, they do not gain any points and lose the competition. Through the cycle of posts, posters will build up a *post library*, and accumulate both followers through events and danger score through consequences (see Definition 5). Other players are able to read from this library and attempt to gain insight on the most lucrative category to post in. In addition, thieves will be able to read this post library to glean information on the identities of posters, slowly building their personal identity profile through investigating their posts and researching posters in data repositories.

Once a thief has built up enough information, they can execute certain simulated theft actions, such as hacking a poster's account and harassing followers, opening a credit card in their name, or stealing their identity to run cons on their friends and relations.

Theft actions will cause the posters to be *hacked*, therefore losing audience points (which consist of likes and followers). They also lose the post they made on the turn they were hacked. Subsequently, the thieves will gain the same amount of points in *theft points*.

The simulation concludes after a pre-set number of turns (the default value is set at 20 turns), after which the points of each poster are examined. The poster with the most points when the simulation ends is declared the winner.

To manage the storyline of possible scenarios, the game environment, and keeping a record of the player's activities, our simulation, Digital-PASS, has three main modules as shown in Fig. 2.

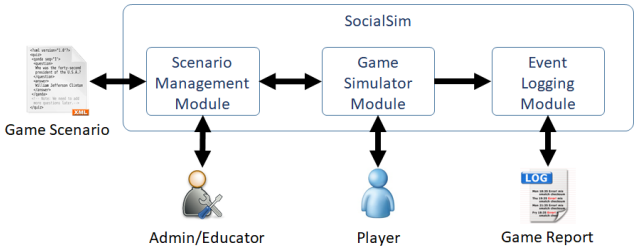

Fig 2. Simulation Control Modules

Game scenarios are written in a machine-readable format such as JSON, and using the *Scenario Management Module (SMM)*, an Admin or Educator can create, download or modify scenarios that are more appropriate for the players. For example, if the simulation is to be used in a high school, player's responses and posts would be very different than if it is used in a healthcare provider clinic to educate staff on the potential privacy risks in dealing with patient's health information on social media.

Once a scenario is selected, and modified if needed, the *Game Simulator Module (GSM)* loads the scenario and the player starts interacting with the game. For simplicity, we assume this is a singleplayer game and we will expand the simulator to be a multiplayer game in our future research direction.

The *Event Logging Module (ELM)* logs the simulation and generates a report which can be viewed as a supplemental teaching tool following the simulation to learn from possible mistakes and choices made.

The virtues of this particular model are threefold. It is organic, meaning that the learning experience unfolds as the simulation progresses rather than being a scripted experience. This gives users the impression that they are learning from their own actions, rather than being lectured to. Secondly, it is dynamic. No two games will play out exactly the same way, providing replay value that will help the lessons sink in. And thirdly, it is reviewable, meaning that each simulation can be analyzed and scrutinized, both as a source of data on personal social media habits, and as a larger source of research material on the habits of a population as a whole. In effect, Digital-PASS allows us to study security and privacy threats based on user behaviors without risking anyone's real identity or working with actual identity thieves, simply by leveraging gamification.

4. Digital-PASS's Architecture

This is an educational simulation designed to convey a learning experience, but naturally, the simulation must also provide ample motivation to the players. Therefore, when it came to devising its rules, we followed the standard principles of game design, including keeping ruleset simple, having a clear score system, and establishing an obvious "win" condition. The following section defines some terminology in the simulation.

For the purpose of clarification, as we cover this terminology, we will use the example of two potential users of the simulation, Alice and Bob.

4.1 Terminology and Formal Definitions

Definition 1 (Player) - The player is placed into one of two roles: content creator ("poster" for short), or identity thief ("thief" for short). Player p_i in the game belongs to a set of *P*, all players, including the computer generated ones, where:

$$\{ p_i \in P \mid i = p \text{ for poster} \vee t \text{ for thief} \}$$

Thieves: All thieves have 2 specific attributes that make up their in-game presence: a *theft score ts*, quantifying the security damage they have inflicted on posters, expressed in theft points, and their *Retrieved Identity RI*, a set of all pieces of the posters' identity they have uncovered. Hence, a thief is represented by the following tuple:

$$p_t = <ts, RI>$$

Imagine Alice is assigned the role of thief. At the beginning of the simulation, her theft score is zero, since she has not accomplished anything yet, and she has not uncovered any personal information, therefore her *RI* set is null. Thus, at the start of the game, her tuple can be represented as:

$$p_a = <0, \{\}>$$

Posters: All posters have these specific attributes that make up their in-game presence: their *post library L*, the set of every post (see Definition 3) they make over the course of the game; their *AP*, a set of *audience point, ap*, which is a quantified representation of their popularity with the audience (represented by AI-driven processes in this simulation) that reflects in their audience ranking; *DS*, a set of *danger score, ds*, representing their overall risk of having their identity compromised; and, they have their *identity profile IP*, a set of all pieces of information that they have disclosed during the game, either willingly or unwillingly. Hence a poster is represented by the following tuple:

$$p_p = <AP, DS, L, IP>$$

For example, Bob is assigned the role of poster. At the beginning of the simulation, he has zero audience points and zero danger score, since he has not posted anything. By the same token, his post library is a null set. His identity profile, though, is not null. His root identity values (see Definition 2) are already input into the system. Thus, his tuple can be expressed as:

$$p_b = <0, 0, \{\}, IP>$$

Definition 2 (Identity Profile *IP*) - At the beginning of the simulation, posters are assigned fictional identities by the system (rather than their actual identity). In the real world, of course, identity consists of a massive variety of statistics and numbers, but for the sake of simplification, the identity values will be boiled down to 8 attributes: *name, gender, birthday, address, zip code, email address, password, and pet name*. In the case of Bob, his root identity values could look something like this:

IP = {*Bob Smith, Male, January 1st, 1985, 123 Garden Drive, Anytown, NY, 12345, bob@email.com, swordfish, Buddy*}

These are the 8 *root identity values* that are added at the start of the game and if discovered by thieves, could be used directly in theft actions to cause serious damage. *Emergent Categorical Identity* is a secondary form of identity created over the course of the game by the choices that posters make. For example, if Bob decides to regularly post about politics, the information "very political" would likely show up in his identity profile, which could prompt thieves to search in repositories with political themes, such as party databases or board of elections websites. In this fashion, emergent categorical identity leads thieves in a linear path towards pieces of root identity. Posters, therefore, must remain cognizant of what reactions their chosen posts might have on their personal security. In this work, we will focus on the 8 root identity values in *IP* and will leave the Emergent Categorical Identity for future development.

Definition 3 (Post) – A Post is an individual unit of content that the poster player leaves or shares on social media. A post includes

a) A unique identifier, *ID*
b) Poster id, *pID*, that identifies who posted it
c) The medium *m* shows the type of the post such as text, image, link, video, etc.
d) Subject *s* where the post belongs to (e.g. sports, politics)
e) Visibility level *v* which shows to whom this post is available, rated from 1-4 (1: friends, 2: friends of friends, 3: network, 4: public)
f) A set of attributes *IP'* that is a subset of *Identity Profile IP*, that will be revealed as the result of this post
g) Potential audience points *ap* that will be gained by this post
h) A set of Consequences *C* that will be the result of this post (see Definition 5).

Hence a post $l \in L$ *post library*, is represented with the following tuple:

$$l = < ID, pID, m, s, v, IP', ap, C >$$

Let us consider an example of a post created by Bob. Bob is politically minded, and wants to post a video of himself talking about politics to share with his like-minded followers. So he creates a post with this video. The tuple representing this post could be displayed as follows:

$$l_b = <12345, 24, video, politics, friends, <Political, Supports candidate John Doe>, +573, <91, 1, 1>>$$

The post's *ID* is 12345, Bob's own pID is 24, the medium of this post is video, and the subject is politics. From the post, emergent categorical identity can be derived in the Political category that Bob supports candidate John Doe. The post increased Bob's audience points by 573, and it triggered a consequence of ID 91, with 100% likelihood that it will occur the next day.

Definition 4 (Event) – An event is an action/reaction to the player or system behavior. When the player takes an action, posts content, deletes a comment, etc., it potentially affects the state of the game and a set of events might be triggered to capture this change of state. Events might also be triggered when the system reaches a milestone (e.g. end of the game period) or having inactive users for a period of time.

An event *e* consists of the following predicates:

a) A unique identifier, *ID*
b) Description, *d*, which is in plain English and explains the event and its consequences to the admin/educator of the game
c) Scope, a set of *P'* players who will be affected by this event.
d) Points ±*ap* that will be given (+) or taken (-) from the *P'* players listed in the scope of the event (if applicable)
e) Type of the event, *y*. Events can be one of the types: Notifying (i.e. informing the player about the change of status); Questioning (i.e. prompts the player but requires an answer to continue and, depending on the answer, could trigger additional events); and Hidden (i.e. the player won't be notified of the event occurrence)
f) A set of consequences *C* that will result from this event (see Definition 5).

Hence, an event *e*, is represented with the following tuple:

$$e = < ID, d, P', \pm ap, y, C >$$

As an example, let us suppose that Bob receives a phishing email from a "Nigerian Prince" offering him money in exchange for his bank account number. Given that he has a choice, yes or no, this event would be of the type Questioning, with two possible results, yes or no. If he chose no, nothing would happen, but if he chose yes, he has a chance of being hacked, leading to a consequence. The tuple would look like this:

$$e_n = <74, \text{"Nigerian Prince"}, <Bob>, -1000, q, <5, 1, 0, .10>>$$

This indicates that the ID of the event in the library is 74, the description indicates the nature of the event, i.e. the Nigerian Prince, the scope is limited exclusively to Bob, the amount of points he stands to lose is 1000, the type of the event is question, meaning he'll be asked to answer yes or no, and if he answers no, the consequence with ID 5 will happen with 100% chance immediately, and his danger score will increase by 10%.

Definition 5 (Consequence) – Consequence *c* is the result of a post or event that has three main elements:

a) An event that could possibly be triggered, *eventID*
b) The probability *r* of the event occurring where $(0<r\leq 1)$
c) Timeline *t* that shows when the consequence is going to occur (i.e. *t*=0 means immediately, *t*=2 means two days from now)
d) danger score *ds*, which shows how this consequence has risked the identity of the poster who has posted or triggered an event.

Hence consequence c is represented with the following tuple:

$$c=< eventid, r, t, ds >$$

For example, consider the consequence listed in the previous definition if Bob fails the Nigerian Prince event:

$$c_n = <5, 1, 0, .10>$$

Because *r* = 1, there is a 100% certainty that the consequence will happen if it is triggered. The event referred to, with ID 5, has its own tuple of description, scope, points and consequences, but for simplicity's sake, it will be referred to here only by its ID.

As danger score DS for the poster (see Definition 1, Posters) increases, the poster has a greater likelihood of being the victim of a Theft Action. Danger score is modified in the following ways:

a) For each additional post made by a user, that user's danger score increases. The amount by which it increases depends on the score of the post and which consequence it raises.
b) Deleting a post will reduce the danger score by some of what the post originally created, but not all, as a way of representing information retention on the internet.
c) Changing the visibility of a previous post will cause the appropriate adjustment in danger score, with a proportional increase or decrease to the post's change in audience points.
d) If a user changes their password, it causes a significant reduction in danger score.
e) Making a choice during an event that endangers your security will cause a significant increase in danger score.

Definition 6 (Theft Action) - Theft Actions are how thieves cause serious inconvenience to posters and earn theft points, representing thieves' ill-gotten gains. Any theft action will require a certain subset of identity, and you will not necessarily know what identity you need before you try the action. For instance, in order to "hack" somebody by logging on to their account, you may have to answer a security question about a piece of their identity. If you don't know it, the poster will have the incursion reported and may change their password. Each thief p_t can use its identify profile *IP* about a particular player p_p to perform the theft action TA_i where *i* represents a theft action such as phishing or opening a credit card.

In the current version of Digital-PASS, for simplicity, the theft action is captured by hijacking the user's account, causing a loss of audience points and followers. Future theft actions may allow for more subtle means of meddling with players, which will be added to a future version of the simulation.

Definition 7 (Theft Point) - When a thief performs a successful theft action, they gain Theft Points equal to the amount of *audience points* they deducted from their target. Unlike posters, though, Theft Points are similar to experience points, rather than just a scoring value. By spending Theft Points, a thief will be able to gain access to new data repositories, and new forms of accessing those repositories, such as hacking them and directly extracting the contents rather than simply searching. Each theft action results in obtaining theft points TP, and the total number of TP from a poster is shown as $TP(p_p)$. It is important to note that the total number of theft points gained equals the total amount of audience points lost.

Definition 8 (The Audience) - The Audience is a representation of the masses on any social network that primarily consume content without producing it. In this case, the audience is represented by AI-driven pseudo-randomized processes and is responsible for scoring the popularity of the posts created by posters through the allocation of audience points. At the beginning of the game, the pool of audience members is distributed between the various post subjects; it is up to the posters to determine which subject is the most profitable in that game (ex: sports).

Definition 9 (Audience Points) - Audience Points are a generic representation of views and demonstrations of approval (likes, shares) from the audience of a social network. As part of the objective of the simulation, each poster p_p is attempting to collect the most audience points, *ap*, with respect to their posts. $ap(p_p)$ represents the total audience points of the poster player, and $ap(x)$ represents the audience points of a particular post *x*. Audience points are only valid for p_p, and not p_t.

Definition 10 (Message) - A message is a particular event that can occur depending on the most popular category of a user's posts. All messages are Events of type Q (Questioning), since they require a response. Each message has 4 components: a sending chance *c*, between 0 and 1, a subject *s*, equivalent to the subject in a Post, a text block *t*, representing the actual contents of the message, and a veracity *v*, that can either be true or false.
A message *m* is therefore represented by the following tuple:

$$m = <c, s, t, v>$$

The chance a poster has of receiving a message is determined by the number of posts they have in their most used category: as the number increases, so does the likelihood of receiving a message in that category. When a poster receives a message, they will be given a prompt, typically an offer to grant additional audience points in exchange for some kind of account access. Depending on the veracity of the message, the contents will be altered slightly to allow the poster to determine whether the message is real or fake, such as by examining context clues or the email address of the sender. The poster will then be prompted to accept or deny the offer. If the poster accepts a genuine offer, they will receive the benefit, and if they deny it, they will receive nothing. If the poster accepts a non-genuine offer, it causes a consequence that will dramatically increase their danger score.

As an example, consider an instance where Bob receives a message. He has been posting a lot about sports, so he receives a message from the World Sports Network, or so he thinks. The message is in fact a fake message sent by a thief to gain access to Bob's account. Bob will have to use context clues from the message itself in order to ascertain whether or not the message is legitimate: if it's full of spelling errors, or the email address seems sketchy, odds are it is not genuine. We can therefore represent this fraudulent message with this tuple:

m_b = *<0.75, Sports, "From: WSN01@Yahoo.com; To Bob: We've been following your profile and seeing you are a rising star. We see that you are very interested in sportes and we have an offer for you: We want to advertise our network with our larger audience. If you agre to do so, we will advertise your account on our website, which you believe will give you public access.All you have to do is periodically post our messages. You don't even have to write them, we do for you! Each move is automatically posted to our ads. What are you saying?", false>*

As shown above, the sender's email address, poor phrasing, and spelling mistakes are giveaways that the message is not genuine, and hopefully Bob notices them. Otherwise, he will experience a significant increase in his danger score, because he has given this fraudulent service access to his account.

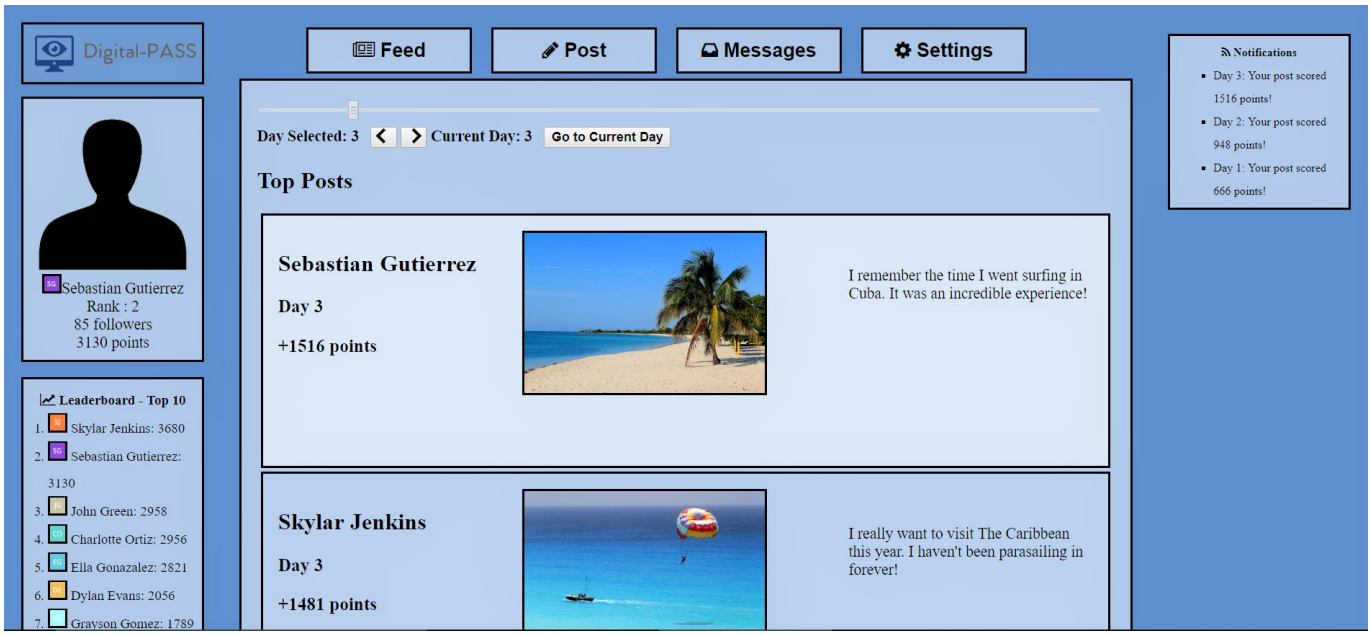

Fig 3. Digital-PASS Main Screen

5. Implementation

Digital-PASS is implemented by a team of 5 developers using multiple languages. The languages used to produce the simulation are JavaScript, CSS, HTML, and C++. The simulation is a web-based application that can run in most browsers. It uses a client-server connection to keep the core functionality hidden to the players. The log files created by these play sessions are pushed to the local machine running the simulation. The choice to not use a database was made so that if this application was deployed to outside organizations, they would have the confidentiality to run the simulation and gather their own data privately. Instead, a flexible editing interface allows simulation creators to add their own categories to the system, and tailor the simulation specifically to their needs.

6. Case Study 1

6.1. Participants

Our first case study was in the summer of 2018. The sample of the study group was 21 high school students who came to our university to familiarize themselves with a university environment. They volunteered to be a part of our experiment, which consisted of a brief explanation of the functions of the simulation before an unguided play session with very little intervention.

We observed how they interacted with our model, whether or not they were working with the simulation as intended, and whether or not our motivation scheme was proving effective. Once their time with the simulation was complete, we asked them to fill out a survey indicating their own perceptions about the simulation and social media privacy in general. Following the conclusion of the experiment, we analyzed both the surveys and the log files generated by their simulation sessions to more directly analyze their activity.

6.2. Interaction Observation

As the students were participating and engaging with the simulation, we were observing their behavior. On the whole, there was a great deal of focus on competitiveness and achieving the highest ranking possible, which was consistent with our predictions. Conversation between individuals about what rank they got in a certain playthrough or whether they were hacked during gameplay was a regular occurrence.

6.3. Survey Results

After the simulation session was complete, we presented a survey to the students and asked them to answer the following questions, rated on a scale from 1-5, 5 being the highest:

1) How fun did you find the experience?
2) How realistic did you think the simulation was?
3) How easy to use was the simulation?
4) How concerned were you about your online privacy before the simulation?
5) How concerned are you about your online privacy after using the simulation?

In our survey the students had the option to leave comments. We did not receive any negative comments towards the game and the overall tone was positive. The participants wanted to see more of the simulation. Some of the comments are as follows:

- “The simulation was really cool to play with though and thank you for letting us test it out.”
- “Good simulation. Needs a feature to compete against others.”
- “Needs a feature that allows one person to be the hacker.”

Fig 4 shows the results of the survey:

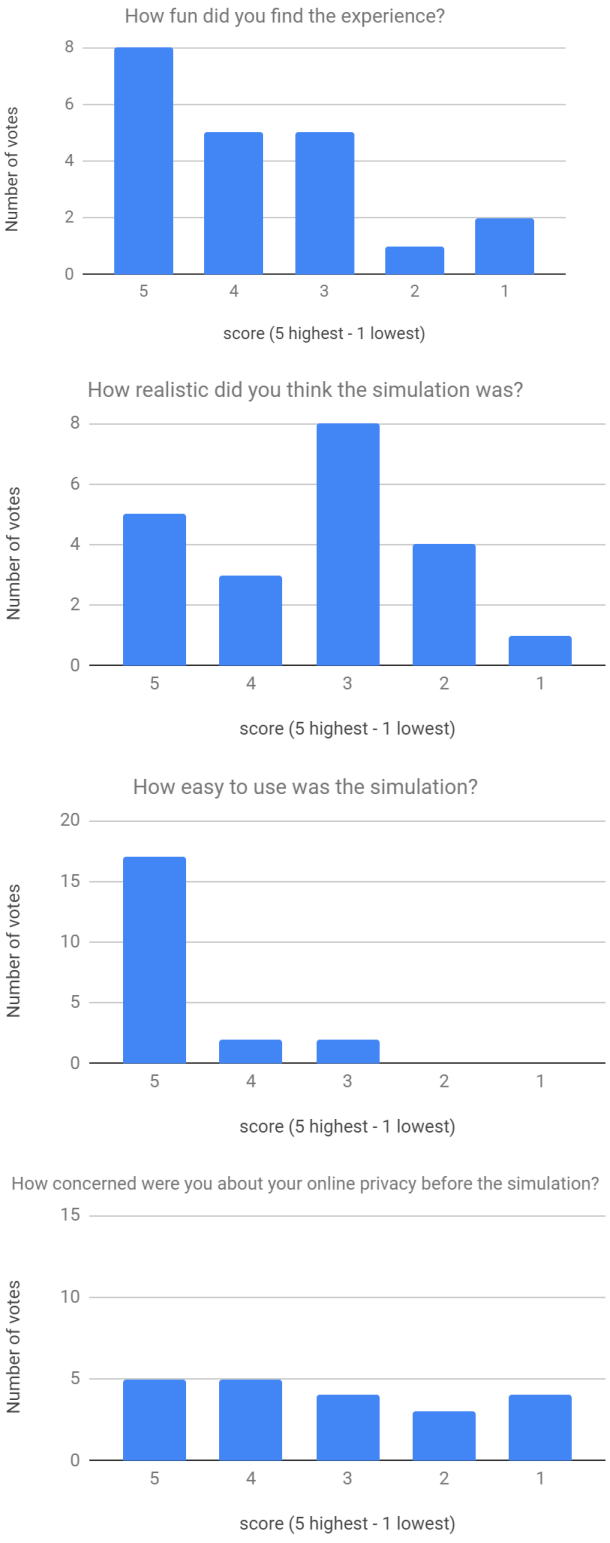

Fig 4. The results of the survey from Case Study 1

These are early indicators that suggest that the simulation will prove effective with students of this age range. Games are a context they are familiar with, and providing a simulation that works with similar structural elements will make it easier for them to become engaged. As shown in Fig 4, the participants had a good experience playing with the simulation, found it easy to work with, and somewhat realistic. In particular, the simulation strongly affected their perception of online privacy concerns, with the majority rating their concern after the simulation as 5, suggesting that the simulation achieved its mission of raising privacy awareness.

6.4. Log File Analysis

From the 21 players, we obtained 87 log files of separate instances of the simulation, of which 49 had over 20 turns completed. Of those 87 games, 63 included at least one instance of the player being hacked. From these log files, we have found the following data:

- The average time of a game was 2 minutes and 11 seconds, with an average turn duration of about 4 seconds.
- The average visibility of a poster's posts, rated from 1-4 (1: friends, 2: friends of friends, 3: network, 4: public) was 3.033, which is typical of this age-range as most of the content they post online is visible to the public or the entire social network.
- In games where player hacks occurred, the average visibility before the first hack was 3.39, but decreased to 3.21 after the first hack. This suggests that, in keeping with our hypothesis, the experience of being hacked makes people more aware of their vulnerability and leads them to further value privacy.

7. Case Study 2

7.1. Participants

With the lessons we learned from Case Study 1, we improved some features of the game, added more realistic posts and decided to test our simulation for the second time. The second case study happened in early 2019. We were invited to a local high school to let 36 students work with Digital-PASS and participate in a survey after they completed playing with the simulation. The survey was identical to the survey utilized in the first case study, in order for the results to be comparable.

Compared to case study 1, these students were a little older and were working with the simulation at their school (as opposed to group 1, who played with the simulation in our lab on campus). These participants were also given some points/credits by their school as an incentive to spend time and play with this simulation,

whereas in the first case study there were no compensation assigned. However, in both case studies, the participants had the option to refuse taking part in the experiment, and it was all done on a volunteer basis. The proper IRBs were filed, and permission was given by the parents/guardians prior to asking students to engage in the studies.

7.2. Interaction Observation

We noticed a significant difference between the overall engagement and investment compared to our first study. This might be caused by the points mentioned above as an incentive, as well as the familiarity of the environment to the participants. This provides support for the benefits of our web-based design, as students can play with our simulation from the comfort of their school environment, and do not need to come to our lab.

Similarly to the first case study, we observed that students were having engaging conversations with each other while playing Digital-PASS, and similar to the first group, students were giving each other advice on how not to get hacked, which emphasized the importance of peer-teaching in our simulation.

7.3. Survey Results

We asked our participants to take the survey after they finished playing with the simulation, and all of them participated. The questions were the same as our case study 1. The students also shared some comments with us. They all showed a positive experience and an eagerness to see what comes next. We present some as follows:

- "If I changed my password, watched what I was posting, and read over messages before accepting I was safe and gained followers. Overall I enjoyed the game and think it did a good job at teaching me the risk of posting on social media."
- "The game was fun to play but also educational."
- "This made me realize how easy it was to get hacked."
- "This was very cool and I would personally really enjoy using this again but with the ability to invite people."

Fig 5 shows the results of the survey from the second case study:

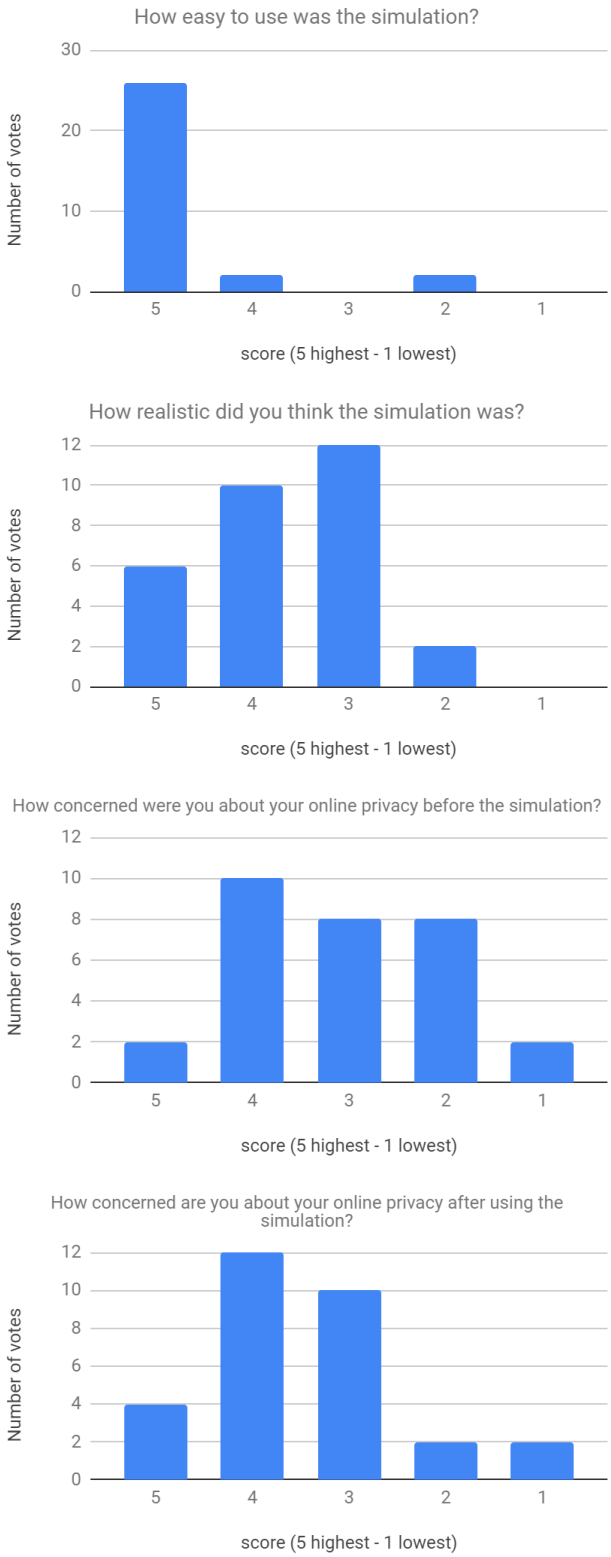

Fig 5. The results of the survey from Case Study 2

The overall scores are consistent with the first case study, but with a few notable differences. Privacy concern before the simulation in the second group was higher on average; however, the simulation

was still successful in raising awareness and improving their understanding of privacy risks on social media.

7.4. Log File Analysis

From the 36 students, 56 log files were generated. Of those log files, 46 were of simulations played to completion (20 turns). From those files, we derived the following statistics:

- Players were hacked on average 2.54 times per game, far more than the overall population of the game (including computer players), where the average hack rate was only 0.6 times per game. Therefore, although players scored on average far higher than their computer equivalents, they lost more points to hacks.
- Players changed their password on average 1.36 times per game, which was considerably lower than the computer players, who changed their passwords 2.97 times per game.
- The average visibility of human players' posts was 2.49, more public than the overall population (including AIs), which had an average visibility of 1.67, but lower than the visibility of the first case study group.
- In addition, players changed the visibility of their posts less often and deleted fewer posts than the overall population. This indicates they considered password changes to be their primary line of defense.

It is interesting to note some of the differences between the first and second case study groups. As the survey indicates, overall privacy concerns were higher among the second case study group, and the logs indicate that they were in fact more secure in their behaviors than the first case study group. They changed their passwords more often and kept their visibility at a lower level, which resulted in them being hacked, on average, less often.

7.5. Individual Simulation Examinations

Our log reports for the second play test contain a detailed list of every action the players took during each instance of the simulations. As a result, we have the ability to scrutinize their behavior and use their conduct to determine whether or not they learned from their experience. In a teaching environment, this detailed feedback would be invaluable to an instructor, who could determine to a certainty how well their students were responding to their lessons, and find any specific areas needing further instruction. Because the logs are anonymous, we can't determine for sure how one player learned from one simulation to the next, but we can compare a player's behavior within a simulation. For perspective, we have a few such logs available for review. Looking through these individual scenarios should demonstrate the effectiveness of the simulation in educating on unsafe privacy behavior.

Simulation #10

Player 10 was, on the whole, a sensible and secure player. They recognized that posts in the personal category might prove more damaging than posts in other categories, so when they chose to post personally, they did so at a lower visibility, usually friends only (1) or friends of friends (2). Otherwise, when posting in the vacation or celebrity category, they would post at the network level (3) or publically (4) for more points. As the game proceeded, posts in both categories steadily moved towards a higher visibility as Player 10 sought more points. All went well for a time, but on turn 15, Player 10 was hacked for the first time, and they lost quite a substantial amount of points.

The lesson was taken immediately, and they were on a defensive footing for the rest of the game. Turn 16 was spent on a password change, and during the last three turns of the game, the only posts that were made by Player 10 were of visibility 1, friends only. This shows that they recognized their error in making posts of high visibility, and sought to correct the action by making lower visibility posts in the future. The sharp shock of the hack taught them a valuable lesson. Anyone might experience a similar lesson if they had been using social media cavalierly for years, only to suddenly find their account ripped away from them. Player 10 had that experience without ever having their real account touched, and will hopefully profit from it.

Simulation #12

Player 12 was an interesting case. During the first 10 turns, they posted regularly, alternating between visibility 2 and 3. However, they lagged behind in the point total, so they decided to risk it and raise their visibility to public for increased point values, despite the danger. They were not totally devoid of caution, though: as a preemptive measure, they changed their password to rid themselves of the danger score they would have built up previously.

They then proceeded to make three public posts, but were unfortunately hacked. It seems as though they did not notice the notification that they were hacked, or they simply did not care, because they made another public post, and as luck would have it, they were immediately hacked a second time. At this point, they happened to be far enough ahead that even after two hacks, they were still in the first-place ranking, but only just. They changed their password on the next turn, and on the last turn they made yet another public post to keep their points up.

This instance of the simulation is a great example of the motivation loop luring a player who is otherwise sensible and secure into making an unwise decision. It is an accurate simulation of the pressure to be popular that many students this player's age face, a pressure that is compounded in the era of social media. Now that Player 12 knows what might happen if they succumb to that pressure in a simulation, they may be more likely to avoid succumbing in real life. And this is just from one simulation; the effect of an extended course utilizing the simulation could be considerably more lasting.

8. Case Study 3-4

8.1. Participants

The third and fourth case studies took place in June of 2019, with two separate but similar groups of 13 students. Both groups consisted of 13 high school students participating in tech-based summer camps being conducted locally. Our group was given time with each group of students, and after a brief lecture on the game mechanics, the students were allowed to freely interact with the simulation.

8.2. Interaction Observation

Perhaps due to the different environment, a camp setting rather than a school setting, the subjects were overall more easygoing and less committed to the work. However, for almost all the students, once they got started using the simulation, they became very engaged.

As projected, the subjects were much more conducive to working with a game rather than taking instruction from a teacher.

8.3. Survey Results

In this round of surveys, we asked for additional anonymous survey data than in previous case studies. In order to gain a deeper background and obtain context for observed behaviors, we inquired about the students' previous experiences with social networks, including what social networks they used, and if they had previously had any experience with identity theft.

The results were as follows:

- The three most used social networks in the group were Instagram (76.9%), Snapchat (69.2%), and Facebook (61.5%).
- Of 26 polled subjects, 3 subjects (12%) had experience with identity theft. In all 3 cases, it involved the subject being impersonated on social media.

Apart from these specific questions, Fig 6 shows the results of the pre- and post-game surveys:

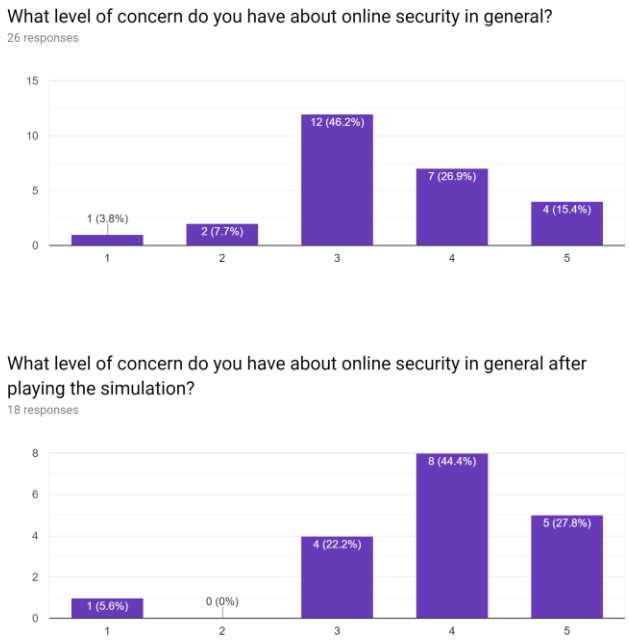

Fig 6. The results of the survey from Case Studies 3 & 4

Overall, there was a general increase in level of concern, which is particularly impressive in spite of some people missing out on the post-survey due to technical difficulties.

8.4. Log File Analysis

Between the two tests, a total of 62 games were played to completion (20 turns). From those games, we derived the following statistics and conclusions:

- All but 7 of the games involved at least one hack.
- The average duration of each game was approximately 9 minutes, which means players were spending on average 27 seconds per turn. Most players were seemingly careful with their overall decision-making, especially compared to our first case study with an average duration of only 4 seconds per turn. This is likely due to the increased features of the game creating more options for the users per turn and, therefore, increasing engagement in decision-making.
- The average visibility of a user player's post was 3.154, which on the 1-4 scale represents a visibility of Network. This is consistent with the privacy settings of most teenagers' actual social media posts, as well as our previous case studies. In contrast, the AI players were more cautious, with a population average post visibility of around 2.574 (friends/friends of friends).
- Players were hacked on average about 5.7 times per game, but that average is skewed by a few outliers who were hacked over 10 times per game.. However, the total population (including AI players) was hacked only 1.75 times per game; user players often sacrificed privacy for gaining points in an attempt to win the game.
- Players changed their password on average 2.68 times per game, but deleted or changed the visibility of old posts far less often. Therefore, many players seemed to be more focused on their future settings rather than fixing mistakes of the past.
- In contrast, our pseudo-random AI posters executed a variety of actions, including deleting and changing the visibility of previous posts. These AI posters were more privacy conscious by design, but still focused on winning; they posted a majority of the time and ended up winning a small percentage of the games.

10. Future Direction

In the current version, individual players can run through the simulation with AI opponents, free to select the parameters of the simulation, including number of opponents, number of turns, and topics available. In the Scenario Management Module (see Fig 2), new topics can be added to the simulation by the educator in order to customize topics to user preferences. An open source copy of the simulation is available at www.ghazinour.com

Our immediate next step is to add stronger scripting elements to the simulation in order to create guided simulations that proceed according to a lesson plan, both for the purpose of educational curricula and for product demonstrations to prospective adapters. During these demos, users will be walked through the basics of the system, given chances to make posts and view reactions, then experience what it is like to have your identity stolen. Then they will be taken over to the other side and shown how the gameplay process of being a thief functions. This demonstration will serve as the sample material for a series of focus group tests aimed at our core demographics: high-school age students, middle-aged office professionals, and the elderly. Through exposing them to the prototype, we will be able to derive insights on the user experience and modify our system accordingly. We have already performed certain focus group tests, mainly focusing on high school students, but it would be informative to expand our testing to other age groups, including schoolteachers or office professionals.

We will proceed from there to creating a fully rounded unguided simulation, where the experience is not scripted in any shape or form. In this phase, the player will take the role of a poster or thief throughout the game interchangeably and will engage in the

simulation alongside other posters and thieves represented by automated processes.

We will begin a new round of focus group testing and private release to certain test customers to bug-test the game, as well as continually reviewing the results to update and improve the game as we make the transition from single-player to multiplayer. The automated players will be still kept to increase the size of a simulation of limited population), and real-world players will be added into the simulation.

We are fully aware that the face of social media could fundamentally change in the upcoming years. Therefore, we have a responsibility to be constantly taking the pulse of social media and observing what services and what features are in vogue for the purposes of implementing them into the simulation.

On top of this potential difficulty, the cultural climate of social media culture is in a state of constant change, and content we add at a point early in the process could be completely obsolete less than a year later. Therefore, rather than focusing on loading content on our end, functionality will be implemented to fully customize the simulation content in order to tailor it to the proper time and place. This comes in three particular areas: world events, the proportion and makeup of the different categories in the audience, and the available topics to post in.

As we progress, we begin the process of offering this simulation to high schools, higher education institutions, private companies, and public offices and departments, as well as to individual consumers. Once we reach this stage, we will likely have acquired a large body of data from running simulations with different groups and individuals. This can serve as a reviewable and analyzable database for further studies to come, both for ourselves and other academics who may find interest in the dataset. The results hold potential interest to a wide variety of fields, including psychology and education. Another part of our work in development will be devising a means of properly reviewing and analyzing the play records and creating statistical data based on simulation results specifically for research purposes.

11. Conclusion

Simulations have been proven to be an effective component of an educational curriculum, providing additional motivation, engagement, interest, and context to any lesson. Applying the principles to social media instruction has proven equally effective, and we have great confidence that in time our simulation model will have a noticeable impact on privacy education.

In this work, we discussed the challenges toward gamification of privacy behavior and privacy preferences and presented our initial thoughts for a proposed model of simulating social networks for educational purposes. We discussed the building blocks of such model and explained the way each part interacts with other pieces of our design.

Next, we discussed four case studies that we undertook, and scrutinized the results. Finally, we presented our plans for future research and indicated the directions for development that we would like to pursue and gave guidelines for interested researchers in this area. We believe that this is the starting point of an effective simulation-based educational system that would raise awareness on the potential privacy dangers that individuals on social media experience, and through gamified motivation schemes and a detailed curriculum would have a major impact on training individuals on how to protect their personal information online.

12. References

1. Barker K, Askari M, Banerjee M, Ghazinour K, Mackas B, Majedi M, Pun S, Williams A "A data privacy taxonomy." *Proceedings of the 26th British national conference on databases: dataspace: the final frontier (BNCOD 26)*, pp 42–54. Birmingham, 2009.
2. Borges, S., Durelli, V., Reis, H. & Isotani, S. "A systematic mapping on gamification applied to education." In *SAC '14 Conference* (p 216–222). 2014.
3. Caponetto, I., Earp, J., & Ott, M. "Gamification and education: a literature review". Genova: ITD-CNR. 2014
4. Chang R., Yang R. "Developing a mobile app for game based learning in middle school mathematics course", *International Conference on Applied System Innovation (ICASI)*, 2016.
5. Cone, B.D., Thompson, M.F, Irvine, C.E., Nguyen, T.D. "Cyber Security Training and Awareness through Game Play," *IFIP International Federation for Information Processing*, Vol. 201, 2006
6. De-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. "An empirical study comparing gamification and social networking on e-learning." *Computers & Education, 75*, 82–91. 2014.
7. Denny, P. "The effect of virtual achievements on student engagement." In *Proceedings of CHI 2013: Changing Perspectives* (p. 763–772), Paris. 2013
8. Deterding, S., Khaled, R., Nacke, N. & Dixon, D. "Gamification: Toward a definition." *In CHI 2011 Gamification Workshop Proceedings* (pp. 12–15) 2011.
9. Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. "Gamifying Learning experiences: practical implications and outcomes." *Computers & Education, 63*, 380–392.CrossRef 2013.
10. Ern, M. "The use of gamification and serious games within interventions for children with autism spectrum disorder. A systematic review." University of Twente: Master thesis. 2014.
11. Gibson, D., et al. "The role of gamification and game-based learning." *Research and Development in Higher Education: The Place of Learning and Teaching, 36*, 514–523. 2013.
12. Glover, I. "Play as you learn: gamification as a technique for motivating learners." *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2013*. 2013.
13. Gómez Álvarez, M. C., Piedad Gasca-Hurtado, G., Villalón, C. M., Antonio, J., San Feliu Guilabert, T. "Design of a pedagogic instrument for teaching software process improvement: Teaching instrument for university and business environments." *Information Systems and Technologies (CISTI), 2014 9th Iberian Conference* P. 1-7. 2014.
14. Gondree M, Flushman T, Peterson Z. "This Is Not A Game: Early Observations on Using Alternate Reality Games for Teaching Security Concepts to First-Year Undergraduates." *CSET'15 Proceedings of the 8th USENIX Conference on Cyber Security Experimentation and Test*, 2015

15. Hamari, J. & Koivisto, J. “Social motivations to use gamification: an empirical study of gamifying exercise.” *Proceedings of the European Conference on Information Systems*, Utrecht. 2013.
16. Hamari, J., Koivisto, J. & Sarsa, H. “Does gamification work? A literature review of empirical studies on gamification.” *System Sciences (HICSS), 47th Hawaii International Conference System Science* (p. 3025–3034), Hawaii. 2014.
17. Hsiao, I., Yang S., Chang T., Wei Y., Lan Y., “Creating a 3D game-based learning system in a virtual world for low achieving students in mathematics”, *IEEE 16th International Conference on Advanced Learning Technologies (ICALT)*, 2016.
18. Hsiao, P., Lo, C.S. "Use of a digital game based Tang Poetry learning system to improve learning for children in nursery school", *Proceedings of the IEEE International Conference on Advanced Materials for science and engineering (ICAMSE)*, p75-78, 2016.
19. Hsu, W., Lin, H.K. “Impact of Applying WebGL Technology to Develop a Web Digital Game-Based Learning System for Computer Programming Course in Flipped Classroom”, *International Conference on Educational Innovation through Technology*, p64-69, 2016.
20. Huang, Wendy Hsin-Yuan, Soman, Dilip, “A Practitioner's Guide to Gamification of Education,” *Behavioural Economics in Action Report Series.* Rotman School of Management, University of Toronto. Toronto. 2013
21. Kapp, K. M. “The gamification of learning and instruction: game-based methods and strategies for training and education: John Wiley & Sons.” doi: http://dx.doi.org/10.4018/jgcms.2012100106. 2012.
22. Kiryakova, G., Angelova, N., & Yordanova, N. “Gamification in education.” In *Proceedings of 9th International Balkan Education and Science Conference.* Trakya University, Edirne. 2014.
23. Lee, J. J., & Hammer, J. “Gamification in education: What, how, why bother?” *Academic Exchange Quarterly*, 15(2), 146. 2011.
24. Li, W., Grossman, T., Fitzmaurice, G. “GamiCAD: a gamified tutorial system for first time AutoCad users.” In *Proceedings of the 25th annual ACM symposium on User interface software and technology, October 7-10, 2012 (pp. 103-112).* Cambridge, Massachusetts, USA: ACM. 2012.
25. Morris, B., Croker, S., Zimmerman, C., Gill, D., Romig, C. “Gaming science: the gamification of scientific thinking.” *Frontiers in Psychology*, *4*(607), 1–17. 2013.
26. Muntean, C.I. “Raising engagement in e-learning through gamification.” *The 6th International Conference on Virtual Learning* (p. 323–329). 2011.
27. Roussos, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C., & Barnes, C. “Learning and building together in an immersive virtual world.” *Presence: Teleoperators and Virtual Environments, 8(3),* 247-263. doi: http://dx.doi.org/10.1162/105474699566215. 1999.
28. Surendeleg, G., Murwa, V., Yun, H. K., & Kim, Y. S. “The role of gamification in education a literature review.” *Contemporary Engineering Sciences, 7*,(2932), 1609–1616.CrossRef . 2014.
29. Thompson, M. F., Irvine, C. E. "CyberCIEGE Scenario Design and Implementation", *USENIX Summit on Gaming, Games, and Gamification in Security Education*, San Diego, 2014.
30. Wang, R. “Demystifying Enterprise Gamification for Business.” *Constellation Research*. 2011.
31. Yang, K., Chen, J., Lu, B. "Development of a digital game based learning system with graduated prompting strategy for math course", International Congress on Advanced Applied Informatics (IIAI-AAI), p423-426, 2016.
32. Zapata, C. y. A., G. “Requirements Game: Teaching Software Projects Management.” *CLEI Electronic Journal, 1.* doi: http://www.clei.org/cleiej/paper.php?id=133 . 2007.
33. Zichermann, G., & Cunningham, C. “Gamification by design: implementing game mechanics in web and mobile apps.” USA: *O’Reilly Media*. 2011.