Design and evaluation of a cybersecurity education game

Paul Gestwicki and Kaleb Stumbaugh
Ball State University

We present the background, design, and evaluation of *Social Startup Game*—an original cybersecurity education game for ages 10–16. The game is designed with two primary goals: first, to teach the player fundamentals of cybersecurity, and second, to show them possible careers and educational paths to careers in cybersecurity. *Social Startup Game* is a single-player strategy game in which the player takes the role of a security consultant at a fictional social media software development company, Social Jam. The player balances the tasks of their employees to maximize user acquisition while reducing security vulnerabilities; during the simulation, the player has to make several narrative-based decisions that are designed to foreground our design goals. We evaluated the game using a qualitative research methodology involving semi-structured interviews and recorded gameplay with thirteen minors in our target demographic. This led to four primary findings: the players have mixed views about the role of education and degree toward career goals; there are diverse opinions about professional developers' appearances and interests; players' background knowledge had a significant impact on their ability to learn from playing the game; and there were two distinct modes of character-based decision making, which we distinguish as *pragmatic* or *empathic*. Among our conclusions are the need for continued study of the role of characters, narrative, and player background in educational simulation games, especially with respect to classical theories of constructivist learning and more contemporary theories of situated learning.

Keywords: cybersecurity, computer security, education, game, social startup game, assessment, diversity

Introduction

The Cybersecurity Education Workshop Final Report (2014) identified six major research themes for immediate action, in order to address deficiencies in contemporary computer security education practice. *Concepts and Conceptual Understanding* is a call for improved epistemology and pedagogy for cybersecurity; *Assessment* deals with how we determine what someone knows about cybersecurity and whether our pedagogy is effective; and *Recruitment and Retention* deals with the "pipeline" problem—getting people into computer security careers and keeping them there. A game-based learning approach satisfies these three themes: a game is a formalization of concepts, a reification of the abstract that can be used to teach and to assess knowledge. By incorporating outreach into the game's narrative, it can also be used as a tool for recruitment and retention.

Inspired by these themes, we created a *Social Startup Game* (Figure 1), an educational game designed around two objectives: first, to teach players fundamentals of cybersecurity, and second, to show them careers and educational paths to careers in cybersecurity (Figure 1). The game has been

Figure 1. *Social Startup Game* logo

completed and can be played for free online at `http://play.socialstartupgame.info`.[1] We chose to focus our attention on middle school and early high school, ages 10–16, because children at this age make critical decisions about their futures, particularly with regard to whether or not they can succeed as scientists and engineers. Margolis & Fisher (2003) and Margolis et al. (2010), for example, describe how cultural factors disproportionately affect women and ethnic minorities, contributing to their absence in IT careers generally. *Social Startup Game* incorporates themes of diversity, both to inspire players and to help foreground such issues in our formal project evaluation.

Game Design

As part of our epistemology-building, we observe that when one models cybersecurity as a game, it has the following formal characteristics:

competitive attackers versus defenders

asynchronous the two sides may take their actions at any time, regardless of what the other is doing

distributed the two teams are not collocated, although allies may be

asymmetric the two sides take different actions

These characteristics create an interesting and significant design challenge. Like all creative and scientific endeavors, game design generally derives new ideas from previous ones, punctuated by occasional paradigmatic shifts and inventions. In our analysis, we found no existing entertainment or educational games with this combination of characteristics except for the genre of capture-the-flag already used to teach computer security. Even these games are not fully asynchronous since events and tournaments are scheduled and timeboxed.

We developed and playtested a series of paper prototypes based on these characteristics. Early prototypes relied upon the *de facto* standard taxonomy of *confidentiality*, *integrity*, and *availability* as key security concepts, but we found that these terms do not carry much meaning to those outside the field. That is, our designs featured authentic treatments of these terms, each being a defense against a different kind of attack, but this was arbitrary to the players: they could not see the metaphorical connection, and so they left without any viable declarative or operational knowledge of these concepts beyond their gameplay manifestations. It is also worth noting that our testers were traditional college-age students, and so we realized that this approach would be untenable for younger players.

Given that we were seeing little affordance for transfer in established cybersecurity abstractions, and given our team's previous experience regarding the technical complexity of multiplayer games, we decided to shift toward single-player simulation games. The short history of digital games has been dominated by single-player games, and this gave us a broad range of related work to draw from. Unfortunately, it also meant abandoning principles of competitive asymmetric play explored in some of our prototypes. However, this also eliminated a design problem that we uncovered in our multiplayer asymmetric designs: would potential audiences such as teachers, librarians, or parents reject a game that forced a young player to simulate illegal,

[1] Note to reviewers: The game includes an optional credits screen which essentially "unblinds" portions of this submission. The provided URL goes directly to the game, skirting an informational main page. We trust you to use your best judgment in this regard.

Figure 2. Social Jam company logo

"black hat" activities? Our consequent single player designs took a strictly "white hat" perspective.

Considering the importance of choosing metaphors familiar to the player—as opposed to terms like "integrity" that are abstruse to our target audience—we adopted a theme of social media. Informal polling showed that youth in our target demographic were familiar with social media, particularly through popular brands such as Facebook and Instagram. We also decided to keep the balance represented in our previous multiplayer designs, where security concerns are "defensive" moves necessary but not sufficient for victory.

One of the drawbacks of our earlier, multiplayer designs was that they cast each player abstractly as both a person and a company, and this elided the fact that software development is cooperative activity. We were inspired by the cooperative game theory, which observes that software development is a cooperative game of invention and communication (Cockburn, 2006). This led to our decision to frame the player as a security expert with a team of developers working under them. This theme allowed us to abstract away money as a directly manipulable resource and instead to focus on social media's fungible currency: users. This line of investigation led to the game we focus on for the rest of this report: *Social Startup Game*.

Social Startup Game is a single-player simulation game in which the player is cast as the chief security advisor for an up-and-coming social media startup, Social Jam (Figure 2). The player manages three employees over ten days and is given two competing goals: first, to increase the number of Social Jam users to 20,000, and second, to keep the estimated exposure (security risk) of the company at less than 15%. The player does this by assigning employees to either Development or Maintenance. Development contributes to the completion of features, which earn new users but also increase exposure: Social Jam is constantly under threat by hackers, and the higher the exposure, the higher the chance of an exploit. Maintenance is used both for fixing exploits and reducing exposure.

Figure 3 shows the nine main characters of *Social Startup Game*. At the start of the game, one is randomly chosen to be CEO, and three more are assigned as your employees. The main interface shows employee names, skills, and tasks. By tapping on a character, a player can also read about their background, interests, and credentials (Figure 4).

The buttons at the bottom of the screen are used to change between four different views. The first—selected by default—is the status view, which shows the users generated per hour, days remaining in the game, estimated exposure, and current progress toward goal. The sparkline chart fills as time elapses to show the player's performance, as shown in Figure 5. The feature view shows the currently completed features, the number of users they generate, and the progress toward completing the next feature; similarly, the exploits view shows the currently-identified exploits, how many users have left because of each, and progress toward fixing each (Figure 6). The fourth view—alerts—is used for interactive narrative events. Embedded into the simulation game are the eight narrative events summarized in Figure 7; these events serve as important decision points for contextualizing educational content. Figure 8 gives an example of a narrative event—one that features prominently in our evaluation. The introductory event explains the game's interface and always occurs as soon as the game starts; the others occur in a random order during gameplay.

Figure 3. The cast of *Social Startup Game*

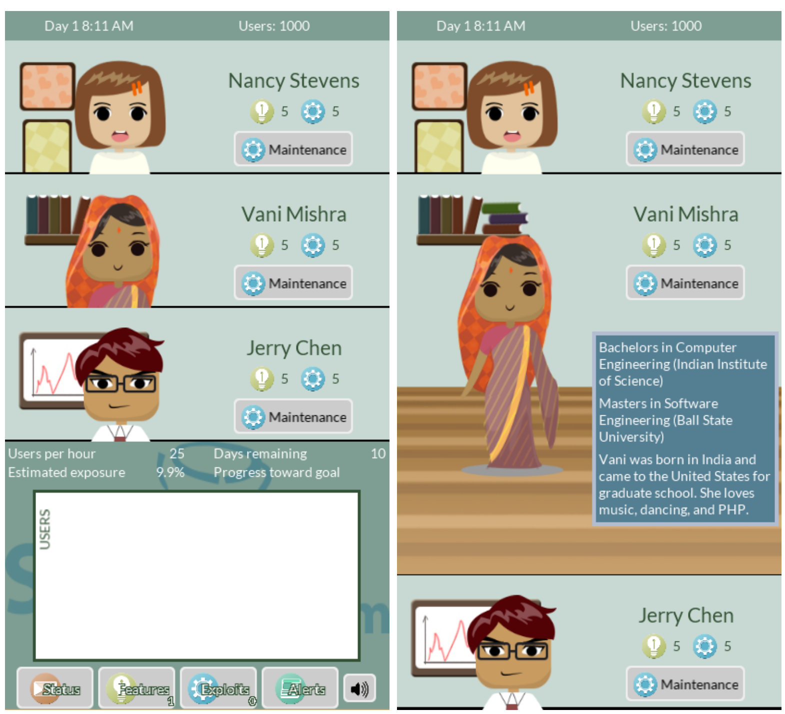

Figure 4. On the left, the screen at the start of the game. On the right, the screen after tapping on Vani to see her bio.

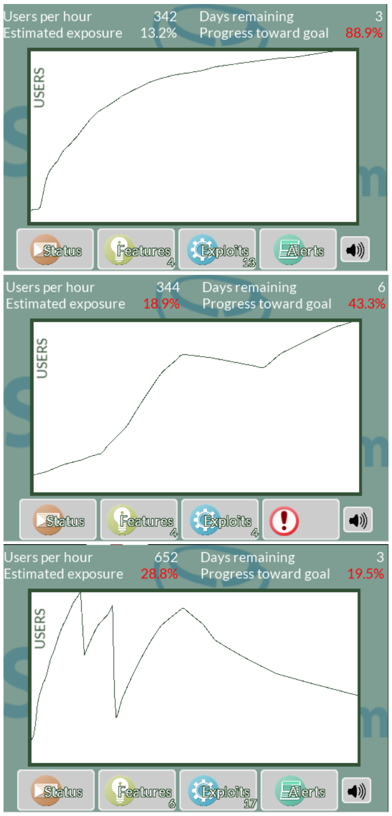

Figure 5. Three sample user sparkline charts, showing ideal play, moderate play, and poor play outcomes

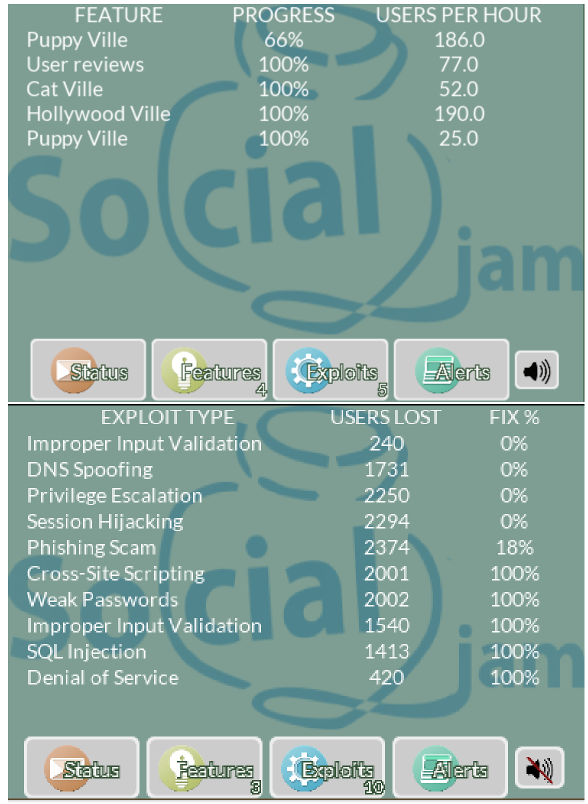

Figure 6. The features and exploits views

Evaluation

Methods

Social Startup Game is designed as an entertaining and educational intervention, but perhaps more importantly, it is also a catalyst for understanding our target demographic. We adopt qualitative research methods in order to gain a detailed, nuanced understanding of our subjects, recognizing that there are many overlapping and potentially-confounding factors at play. By applying a qualitative lens, we can build a sound and empirical local theory that gives us both an evaluation of the game and a better understanding of its cultural context. For an overview of qualitative research methods, we recommend Stake (2010).

The paucity of scholarly cybersecurity educational game assessments (citation removed for review) leaves almost no grounds for forming the reliable and valid instruments necessary for quantitative methods. Put another way, we could not know whether what we measured would actually contribute to answering our research questions. Qualitative research methods, by contrast, deal with observations and descriptions, meeting the subjects within their complex cultural contexts, and seeking understanding of phenomena that are not directly measurable.

Our research protocol was approved by our Institutional Review Board. The subject pool consisted of thirteen youth, ages ten to fifteen, with eight females and five males, as listed in Figure 9. All of the subjects were white American students living in Indiana. They came from families who responded to a call for participation that was sent to

Introduction Explains the game interface.

Take-your-child-to-work Day A coworker asks the player to recommend a major for his daughter to pursue in college.

Data stolen Choose responsible disclosure or secrecy when data is stolen by hackers.

DDOS Respond to an angry ex-user's initiation of a distributed denial-of-service attack.

Insecure password When a weak password causes a problem, choose whether to ignore it, fix it, or train the staff to prevent future problems.

Script kiddie Faced with amateurish attacks, choose to ignore it, report it, or fight back.

Security conference One employee may be sent to a two-day conference to improve their Maintenance Skill

Input sanitization When an employee expresses concern over input sanitization, you can either take action or tell them to ignore it.

Figure 7. Narrative events of *Social Startup Game*

Figure 8. An example narrative event. This is the "Take your child to work day" event, which features prominently in the qualitative evaluation.

Subject ID	Gender	Age
1	Female	14
2	Male	10
3	Female	12
4	Female	11
5	Male	12
6	Male	15
7	Female	12
8	Male	12
9	Female	13
10	Female	16
11	Male	15
12	Female	10
13	Female	10

Figure 9. Research Subjects

our institution's mailing lists and shared via social media. This group is a convenience sample of those who responded to the call, chosen for their ability to schedule sessions and with an emphasis on female participants. The subjects include a mixture of homeschool and public school attendees.

Each subject participated in a semi-structured interview that included pre-game questions, logged gameplay, and post-game questions. Interview audio was digitally recorded, and these were transcribed into 918 paragraph-separated units. The investigators also produced 27 paragraphs of analytic memos, consisting of 1378 words. Gameplay logging produced 16 logs consisting of 326 discrete game events.

Interview data were coded using techniques inspired by Saldaña (2009). The two researchers conducted three phases of coding before comparing results in order to ensure inter-rater reliability. The first round consisted of open coding, and these codes were refined into 96 second round codes. In the third round, we identified 18 categories, from which emerged the four themes described below.

Findings

Mixed views on the role of education and degree toward career goals. The subjects articulated a predictable mix of opinions about their school experience. Although the interview question about favorite school subjects was intended primarily as an ice-breaker, many students eagerly described what subjects they enjoyed and why. For example, subject 7 enjoyed English and science because they allow her to be "more creative than in other classes," and subject 5 likes mathematics because he identifies as "a problem-solver." In contrast, subject 6 gives his favorite subject as lunch, since he is "not a big fan of school… I know I have to do it, but it's not very fun." We observe some subjects making distinctions between "academic" and "non-academic" courses: subject 9 prefers art and band, but she identifies English has her favorite academic course. We lack data to indicate whether this distinction is from schooling structures, environmental impacts, or a personal epistemology.

Our subjects shared a general belief that mathematics is important to being a developer. According to subject 11, "There's probably a lot of mathematics and, you know, how you have to know how games work or what other people did to make the games." They also perceived Computer Science as being a default path into application development, exemplified by Subject 6's supposition, "It seems like making an app would be more computers and that stuff than anything else." During the in-game event in which players have to choose what major to recommend to a coworker's daughter, ten chose Computer Science, two chose Graphic Design, and one chose Engineering—a surprising consistency of choices given that there were nine options (Figure 10). Although almost all the students refer to Computer Science as the subject to study to become a developer, there is no other data to suggest that they understand what Computer Science is beyond a gateway to application development.

Option	Times Chosen
Computer Science	10
Engineering	1
English	0
Graphic Design	2
Mathematics	0
Marketing	0
Performing Arts	0
Physics	0
Psychology	0

Figure 10. Possible college major choices to suggest to a coworker's daughter in *Social Startup Game* and the number of times they were selected by unique subjects during interviews

Some of the subjects recognized that a college education is not the only path to becoming a successful application developer. Subject 5 identifies reputation as an asset and entrepreneurship as a possibility, saying that "… somebody would hear about you, and you probably need to get a degree, or be an entrepreneur and start your own app-making business." Subject 6 recognizes that coursework can help someone learn to make apps but that there also is a wealth of "random sources" from which one can learn:

> If somebody decides they want to make an app, and they know how, they sit down and do it. If they don't know how, and they still want to do it, then depending on how complex the app is they either go to try to figure it out at some random source, or if it's a bigger, more complex project, they go take classes for it, and then they make the app.

Subject 4 recognized that teachers hold significant sway in students' opportunities. When asked how developers got their jobs, she suggested that "their teachers saw their ability to do something like that in an earlier stage, so when they were able to get the classes in high school and through college to be able to get better at that." This empowerment of authority is an interesting contrast against the entrepreneurial view of subject 6, yet it reflects a real and common understanding of the trajectory from elementary school through to a career—a trajectory primarily mediated by teachers rather than autonomy.

Diverse opinions about developers' appearances and interests. Most of the players claimed to have never thought about the physical characteristics of software developers, but among those that did, there was great variance. Subjects 6 and 13 reported developers to be "geeky," which matched our expectations based on earlier work on perceptions of engineers (Margolis & Fisher (2003), for example). Neither subject elaborated on what characterized geekiness, and we recognize that there are potential misunderstandings in leaving the terms used by teenagers up to the interpretation of adults. By contrast, subject 10 stated that app developers look "like normal people... like anybody." She went on to say that the characters in the game "didn't look like smart people—they looked like normal people." Unfortunately, as with geekiness, we lack the data to qualify for certain what she and the other subjects believe a "smart person" looks like.

We did not expect the responses that painted app developers as something attractive and aspirational. Subject 7 describes them as "17 to sort of early 20s in age," which is only about seven years older than herself. Subject 9 was amused by Ivar's eccentric appearance, saying, "He looks kinda creepy, and I love it. I like his creepy long hair with a bald head." This was particularly amusing to us as designers and developers, given that Ivar's character design was inspired by a classic photograph of Bjarne Stoustrup.

Subject 4 brings up a distinction between what she calls "designers" and "producers," which we believe to correspond to interface design and programming, respectively. When asked what app developers look like, she puts it like this:

> Well, I'm pretty sure [they look] like people, they're like artists and stuff, ... I picture the people who *design* it would be kind of like artists, kind of, messy kind of regular people, and the people who publish it would be kind of like, *official*.

It is worth noting that subject 4 was one of several students who had completed a business information technology course in the local public school, which we know to have included an introduction to programming and which we believe to have informed students about some specializations in the field.

Subject 6 does not like school, but he does like *Minecraft*. When asked about the appearance of *Minecraft*'s developers, he responds, "Probably most of them are not American, probably a lot of them are from Japan. It's a Japanese company, isn't it?" While other subjects who brought up *Minecraft* seemed to know something about its Swedish creator—Markus Persson ("notch")—subject 6 held an incorrect belief that it was created by Japanese developers. We suspect that this belief may be due to the relatively high number of Japanese game development companies, although this does not come up in the interview.

There was only one instance of overt sexism in the data: when asked why he sent Esteban to the security conference over Janine or Melissa, subject 2 responded cheekily, "He's the *guy*." No other subject referenced developers' gender in their discussion of their perceptions, backgrounds, interests, or educational preparation. However, we recognize that gender and racial prejudices may be subconscious or intentionally hidden from subjects' responses.

Impact of background knowledge. We saw players' background knowledge impact their in-game behavior and gameplay outcomes in a variety of ways. In a theoretical sense, this is unsurprising, given the wealth of evidence to support constructivism—the educational theory that individuals build their own mental models based on their experiences and background knowledge (see Duffy & Cunningham (1996), for example). However, the broad and multifaceted impact that this can have on educational game design and evaluation makes this worth analyzing.

Most players expressed surprise at the scale and persistence of cyberattacks. They were less clear on what the hackers were after. Subject 2 expressed a concern that revealed a few layers of confusion about the game, saying that perhaps the hackers "wanted to call the game their own." This shows an understanding of software copyright infringement ("piracy"), which then shaped his rationalization for their behavior. Although some subjects showed an understanding of the business and economic forces around software development, none of our subjects reflected aloud on the value of private information.

In the hacker event, most of our subjects chose to report the crime to the police. Subject 5, however, had a different perspective:

> I could report it to the police, or I could task someone to try to figure it out. I'll probably just task somebody to figure it out for now because, I don't really feel like getting mixed up with the police, so I'll task [someone] to do it.

We did not follow up to inquire why he did not want to get "mixed up with the police," but it stands in sharp contrast to those subjects who, seeing this option, quickly decided it was the best course of action. Subject 10 chose revenge over notifying the police—an action that happens to result in the largest possible loss of users in the game. In her post-game reflection, she regretted it, saying "I probably shouldn't have. I probably should have let the police handle it instead of taking it into my own hands."

Several subjects thought that decision points had right and wrong answers, and that their task was to choose the right one. This matches their experience with an industrial-empiricism model of education—where knowledge is assessed with computer-based tests that have one right answer, the rest being wrong—and conventional educational games. Given that their interaction with us was framed in terms of "an educational game," it is not surprising that this mental model would be at the fore. Subject 11's explanation of how he approached decision-making in the game could just as well be about how he took a quiz in school: "I knew some of them would hurt me no matter what I did so [I chose] the best possible answers. And some of them, I really didn't know what to do, so I just guessed." However, we also observed some players adopting a more nuanced view, particularly those who played more than once. Subject 9 puts it this way:

> See, last time I was like, "Oh, I'll let them know," and everyone got mad. So I thought I'd just keep it a secret. But then when I was exposed, everyone was like, "Why didn't you tell us that?!" They all got really mad and left me, which I think is what made my exposure go up. Last time they didn't like that; this time, they didn't like it more.

The players did not question whether the game was an authentic portrayal of a social media company, although they recognized the game as being a simulation. Subject 5 says, "I actually kind of liked it. I liked the way you have to try and run your company—I just kind of like that type of thing." None of the students commented on the diversity of the employees, although several college students and adults who playtested the game commented on the disparity between our portrayed diversity and their experiences and assumptions.

Subjects with personal experience or vicarious experience in software development had more positive views of the profession. We see this explicitly in subject 11's commentary about his uncle: "My uncle has a job in computer software but he didn't graduate from college. So I don't think you have to."

This is important when we note the disproportionate number of white male software developers. If this indicates that those with a role model are more likely to be informed about IT as a profession, then underrepresented groups who have relatively few role models are more likely to be unaware of how to become an IT professional.

The most intriguing support for this theme is a cautionary tale about branching narratives, coming from the interview with subject 5. He was faced with the "weak password" event, in which an employee's account is compromised due to their using a weak password. The player is prompted to either train the staff or to simply change the password. Subject 5 laughed after reading the options, saying, "Encrypt it! ... that's not one of the options." Referencing the training option, he asks, "Would this be training to encrypt it?" This was one of very few times that pre-existing security knowledge came up in the interviews, and so clearly this 12-year-old knew something about encryption—but his was a partial knowledge. Encryption is not something the user needs to do, but something the server or the developers need to manage. His own incomplete knowledge led to him not being able to understand or learn from this narrative event like the other players did, most of whom simply chose to train their staff. Had a researcher not been present to talk with him, he could have built an even less viable understanding of passwords, encryption, and networks for having played the game. This speaks to the need for further research on the intersecting roles of branching narrative, constructivist learning, and interactive games.

Finally, we note that students' vocabulary and cultural literacy were major factors in their ability to understand the gameplay experience. None of the players younger than 14 understood the term "press release," although they were able to understand their options when we explained the term. Subject 5 demonstrated a general misunderstanding of core gameplay terms when, a few minutes into the game, he asked, "Are exploits and features good or bad?" Game literacy also played an important role, as it appears many of our subjects did not recognize the character skill levels for what they were. By contrast, upper-division high school players—who likely have higher literacy and game literacy—had no trouble recognizing these.

Different modes of character-based decision-making. The players often referenced their workers when making decisions. Several took advantage of the game's paused state to review the bios of their employees. However, we observed an interesting dichotomy in how this information was used. Those who engaged in *pragmatic* decision-making chose characters who were deemed to be best at these roles; that is, they focused on the employee as a tool for the company's good. This is exemplified in subject 4's approach to dealing with the hacker event: "They needed somebody to fix it, so I sent [Vani] because she was good at fixing that kind of thing." Similarly, subject 7 explains why she sent Jerry to investigate the injection attack, contrasting his background against Janine's, whose bio indicates she is especially skilled at requirements analysis:

> He seemed like the most well-trained out of all of them, and the girl—I forget her name—she seemed more like someone who should be with, like, social, and not developer, but Jerry seemed like the better person for the job of that kind of thing.

Sometimes this pragmatic approach included assumptions about educational background as well, as we can see in subject 11's explanation of why he referenced their bios so often:

> I thought there was probably a pretty balanced team, because you had someone who majored in three of the important categories, such as computer engineering. I can't remember what the one in the middle was, and then there was another one who was good at working with people and finding out

> what they want. That's why I chose here to do the—oh no, I chose the engineering one to teach people how to do the password thing because he know how to make a safe password, and then I would have chosen her to find out what people liked. . . .

We also observed players engaged in *empathic* character-based decision-making, where they justified their decisions based on what the fictional characters would prefer rather than perceived output maximization. Subject 9 is the best example of this—an enthusiastic player who quickly learned the characters' names, referenced them frequently, and seemed particularly interested in being a good leader to them. When we asked why she chose Janine for several events, she told us, "From her description, she seemed really on it. She seemed really cool. I liked her a lot." She also commented on how much she liked Nancy, explaining that she saw some of herself in Nancy's artwork:

> She looked really sassy. And she looked like me before I went insane, 'cause I used to have long blond hair and that was normal. It was gross. But she was really sassy, so I liked her. "Sassy," like, not really sassy but, like, "cool."

It is worth adding that, when subject 9 was first reading Nancy's educational credentials, she had other reasons to identify with this character: Nancy and the subject's mother shared an *alma mater*, and she was also impressed with Nancy's hobbies, saying, "Nancy has a popular podcast about being a woman in technology. Oh, that's cool!" Our data are not clear on whether sharing a gender contributed to this identification as well, but we note the possibility.

Discussion

About the game

Our findings regarding game vocabulary affirm our design decision to avoid the terms "confidentiality", "integrity", and "availability" in the game since they are security jargon, not common parlance. The subjects in our study were not a random sample of their age group, and in fact, we have reason to believe that these were children of privilege; we expect that players' confusion over vocabulary and cultural references would be exaggerated among players from less privileged backgrounds.

The Social Startup Game features high degrees of randomization in the spirit of increasing replayability. For example, the sequence of story events and the set of characters included in the game are randomized, so they will be unpredictable for any given gameplay experience. While this improves replayability, which is generally considered a virtue in games of any kind, it also potentially hindered our evaluation. Each player was faced with a random set of characters, for example, so we cannot directly compare two players' opinions of their workers to each other, and we also could not intentionally match or mismatch character gender and ethnicity with the player.

A significant category in our analysis included usability problems with the game interface. The two that dominated the list were players not realizing that they could change characters' tasks and that they could read character bios. This surprised us during the first few interviews, given that all of the players read this introduction given by Frieda, Social Jam's administrative coordinator:

> [Employee 1], [Employee 2], and [Employee 3] are currently maintaining our software. You can tap them to find out more about them. You may reassign any number of them to new feature development at any time. Go ahead and try that now, and let me know when you are ready!

All of the players read this and then tapped "Next" without ever following her instructions, except for those whom we interrupted to demonstrate how to do these actions. This usability problem could likely have been discovered with more playtesting prior to the formal evaluation. However, we do not believe that this had any invalidating effect upon our findings, since players were quick to understand the interface once we demonstrated how to use it. We suspect part of the confusion came from the fact that our interface was designed for touch-based mobile devices, yet our evaluation was being done on non-touch laptops with mouse input.[2] We observe that, with mouse-driven input, players scan the screen with the pointer, watching for responses that would indicate interactivity, whereas on touch-based devices, players are more apt to try tapping things to see what happens. We expect that these problems would be ameliorated by (a) playing on touch-only devices such as tablets or (b) having designed for mouse-based input instead of touch.

We used *The Social Startup Game* as part of our cybersecurity workshop at the 2016 Congressional Leadership Academy, which is a full-day event held at Ball State University for high-achieving high school juniors from Indiana's sixth district. The students at this event, who were just barely outside of our target demographic, showed none of the confusion over terms and jargon; in fact, they seemed to enjoy the game much more than our target demographic. Of course, these were high-achieving students who elected to attend a session on cybersecurity, so in a sense we were preaching to the choir: these students were already on their way to potential careers in cybersecurity.

About security and software development

Our design privileges authentic contextualization of relevant decision-making over fine-grained security topics. We saw variation in how subjects approached their initial choices, informed by their individual backgrounds, but we also saw a move toward more ethical and informed decision-making after playing the game. This was especially pronounced when considering the principle of responsible disclosure, which was featured in more than one interactive narrative event. In the absence of a longitudinal study, we do not know whether these students will later recall specific kinds of attacks introduced in *The Social Startup Game*, such as injection and DDOS attacks.

Our players who had previous exposure to programming believed that they could get a job in software development. This is an important victory for advocates of early computing education intervention. One homeschooled subject mentioned having a positive experience with Code.org, while several public-school subjects cited positive experience with Scratch (`https://scratch.mit.edu`) through a required business information technology course. Subjects with no exposure to programming showed more anxiety or unease at the thought of software and IT careers. It merits repeating here that ours was not a random sample: our subjects came from families who were willing to bring them to an educational game evaluation, and all our subjects identify as "white." Fortunately, there are ample other programs—including within our own geographic community—that are building local theories around ethnic minorities and specifically targeting families of low socio-economic status. Despite this caveat around ethnicity, we observed no significant difference in opinions or articulations between our male and female participants. We hope that this points to a continuation of the trend toward greater gender diversity in Computer Science programs and, by extension, into industry.

[2]The reason for using mouse-driven input was an untimely defect in one of our supporting libraries, PlayN. The 2.0-rc3 release contains a previously unidentified defect that, in the HTML build, broke popup menu support on touch-input devices; hence, employee task selection was impossible on touch-based interfaces. Although PlayN 2.0 has yet to be released, the snapshot build has fixed this defect, and our current version of the game is built upon that snapshot.

About education

The number and confidence of players who chose Computer Science as the default academic option for IT careers is problematic. While it is true that many students graduate from Computer Science programs do pursue such careers, we know that there are many other routes to becoming a successful IT professional. (Indeed, we have met recruiters who prefer liberally-educated, critical-thinking philosophy majors over technology-focused Computer Science majors!) If there is any standard of what constitutes Computer Science, it is the ACM/IEEE Curriculum Guide (ACM/IEEE-CS Joint Task Force on Computing Curricula, 2013)—a sprawling guide that speaks to the truth of the insider joke, "If you ask five computer scientists to define 'computer science' you get seven different answers." We are starting to see new curricula and academic programs being developed, including undergraduate Software Engineering programs and transdisciplinary programs like Georgia Tech's computational media major. However, the fact remains that to many students, it looks like Computer Science is *the* path because it has "computer" in the title—despite the fact that every field of study now uses computers.

Although this study is aimed at youth, our experiences with undergraduate computer science students revealed to us a connection between our subjects and our university students: understanding of security principles and practices is negligible, and both pools seem to lack the concepts—the vocabulary—about which to converse on these. This echoes the findings of the Cybersecurity Education Workshop (Cybersecurity Education Workshop, 2014). We suspect that there is a causal relationship between the prevailing structures of Computer Science education and the lack of good security practices in industry. The ACM/IEEE curriculum guidelines for Computer Science curricula treat cybersecurity differently from all other domains, explaining that this topic must be included throughout the curriculum and not simply in isolated courses. However, conventional higher educational structures privilege the separation of content into courses as well as the academic freedom of faculty in the classroom. This leads to a situation where there is no practical way to create interleaving topics, motivate their adoption, or evaluate their efficacy.

Many subjects were quick to characterize people as "geeky" or "smart"—properties of both their behavior and their appearance. We suspect that these same students may subscribe to an entity theory of intelligence, which is the belief that intelligence is a fixed attribute. Hence, one would look smart because one is smart. However, students with an incremental theory of intelligence—that one improves through practice—are known to have generally better academic outcomes (see Blackwell et al. (2007) for example). *Social Startup Game* is designed around the incremental theory, with workers who improve their skills by practicing them, including professional development through the conference narrative event. Given the advantages of the incremental theory, and the prevalent character-development tropes in contemporary video games that emerged from tabletop roleplaying games, we hope to see future work that looks at how gameplay can encourage the more beneficial theory.

We believe there is reason to be concerned about the perception that subjects in school are either "academic" or "not academic." Our study does not reveal from whence these youth developed such perspectives, although we can assume that it relates to the administrative structure of their school experience. The divisions of school are artificial, created by bureaucracy for the purpose of perceived efficiency. Students who do not recognize this as an artificial construction become, ironically, victims of it rather than liberally educated. This is relevant to computer security, again giving its cross-cutting nature: if learners believe that all knowledge belongs in discrete, non-overlapping boxes, then understanding of orthogonal or cross-cutting issues is lost.

About game design

Those who made pragmatic character-based decisions did so assuming that character backgrounds impacted their performance, while in fact, all of the characters are interchangeable: the differences between them are only skin deep. It is interesting that the same players who lacked the game literacy to recognize some of our design tropes (such as a character statistics block) were the same to assume that character backgrounds provided some kind of hint to maximizing decision outcomes. Considering two design options—one in which characters are all fundamentally the same regardless of their appearance, bio-sketch, and credentials, and one where their in-game behavior is different based on these factors—we lack data to determine whether one will better meet our intended learning outcomes.

The discovery of empathic character-based decisions opens more potential opportunities for future design research. Our prototyping process described in Section refsec:design included several different quantitative factors such as users, defects, and money, but none of our approaches considered emotional health, such as happiness or job satisfaction. Players who tend toward empathic decision-making may feel more engaged with the game by receiving such feedback, and this, in turn, may lead to better learning outcomes. Inspired by Bartle's classic essay (Bartle, 1996) in which he describes the kinds of players who play MUDs, perhaps a similar kind of taxonomy could be built that describes how players approach decision-making with respect to fictional, in-game characters. We know of no such theory, but our findings suggest that it could be applicable to game design generally and educational game design specifically.

Many players made decisions based on a search for the "right" answer, not just a good answer. We suspect that there are two factors that contribute to this mental model. First, the game was framed as an "educational game", which is something of a loaded genre. Klopfer et al. (2009) describe the history of educational games and their many problems. One common design mode is to rely upon multiple-choice, quiz-style assessments, where there is just one right answer. Such assessments ignore constructivist and situated theories of learning (see Hickey & Zuiker (2003), for example). More to our point, they put students into a school-culture mental model that multiple-choice scenarios imply that there is only one correct answer. The second factor we suspect to contribute here is the predominance of such dichotomous decision trees in computer games. From our experience, games aimed at youth tend not to feature nuanced decision-making, particularly in dialog options. Hence, a player whose mindset is one of conventional games would not expect, say, that the advice given to a coworker's daughter did not have any "wrong" answers.

As mentioned earlier, we had the opportunity to introduce several high school juniors to *Social Startup Game* through the Congressional Leadership Academy. While this was not part of our formal evaluation, we noticed an interesting phenomenon that is worth noting here. This portion of the academy was held in a computer science teaching lab that is designed for pair programming, and the participants formed pairs based on the furniture configuration. Each play through the game was therefore completed by a pair of participants—strangers who had only met that day. Although they were not instructed to follow a think-aloud protocol as were our actual research subjects, they played the game much slower and more deliberately, seeking consensus among the two players before making any move. We believe that this points to an important possible direction for future evaluations of such pro-social games. This kind of situative perspective—drawing upon the tradition of Wenger (1999)—could reveal important truths about the subjects and their play experience (Hickey & Zuiker, 2003).

Conclusions

The design goals of this project were to create a game that teaches fundamentals of cybersecurity and to expose careers and career paths to cybersecurity careers. Regarding the former, we found that the topic of computer security poses significant challenges to the design of educational interventions, particularly because security is woven into the tapestry of professionalism in software development. While it's true that cybersecurity can be a focused area of study—as evidenced by courses and certificate programs—to understand security requires an operational understanding of the rest of the system. This property of cybersecurity is the reason why the ACM/IEEE Computer Science Curriculum guidelines uniquely identify information assurance as needing to be integrated through other courses; this is especially surprising given the conventional, course-structured approach for the rest of the curriculum guide.

Despite these challenges, our rapid prototyping process identified several possible mappings from cybersecurity concepts to learning game mechanisms. *Social Startup Game* is a single-player simulation game with embedded narrative events, but this is only one of several design possibilities to meet our goals. Future work should more rigorously explore the asynchronous, asymmetric, multiplayer mechanisms that were out of scope for this project, since these properties are integral to computer security—we suspect they represent the best effort to "find the game in the content" (Klopfer et al., 2009). The context of deployment and evaluation is also worth investigating, particularly extending to more situated environments, adopting more wholeheartedly a design-based research approach (Barab & Squire, 2004; Brown, 1992). As discussed earlier, our research design incorporates a constructivist perspective; we believe there may be value in moving to a more situated one (see, for example, Greeno (1998); Hickey & Zuiker (2003)).

Regarding the goal to expose players to career and educational paths, our evaluation has served to confirm the value of early computing interventions toward positive outlook of youth toward computing careers. Our subjects recognized the economic viability of computing careers. However, this was coupled with a wide variety of perspectives of the developers themselves: some positive and aspirational, some negative and stereotypical. Taking *Social Startup Game* as a pilot project, future work could look more carefully at how both aesthetic and formal properties of in-game characters can change perspectives of players toward real-life software developers and IT careers. Our findings present a snapshot—a local theory on how youth perceive software development careers, and future work can compare this to other times, places, and contexts.

Future work should explore more carefully the relationship that background knowledge and experience has on the efficacy of educational games. This manifested in various ways in the evaluation of *Social Startup Game*, the two most prominent being: players with any background with programming or relationship to software developers self-evaluated as being capable of pursuing such a career; and using a conventional branching narrative structure led to confusion—or worse, learning the wrong outcome—to a 12-year-old subject who had some prior experience with encryption technology. Bartle (2009) points out that there is relatively little theory in game design, relying more on craft and anecdote, and that theory which does exist is often misunderstood and misrepresented. In educational game design, we can draw upon established educational theory such as constructivism (Duffy & Cunningham, 1996), constructionism (Papert & Harel, 1991), and situated learning (Wenger, 1999), but more work is required to determine where these align or conflict with nascent game design theory.

Social Startup Game puts the player into an authentic although simplified simulation of a software development start-up company. Players made decisions based on what was best for the company, aligning in-game victory with successful decision-making. They identified with the in-game characters in rich ways, making both pragmatic and empathic choices in their attempt to win. This produced a corpus of data that led to important, interesting, and sometimes surprising findings—findings that we believe would not have been possible in a Type 1 or Type 2 experience. This work therefore provides an example of and a justification for formal evaluation of cybersecurity education games.

Acknowledgments

This research is based upon work supported by the National Science Foundation under Grant No. 0968959.

References

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The journal of the learning sciences*, *13*(1), 1–14.

Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD research*, *1*(1), 19.

Bartle, R. (2009). Understanding the limits of theory. In C. Bateman (Ed.), *Beyond game design: Nine steps to creating better videogames.* Charles River Media.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, *78*(1), 246–263.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The journal of the learning sciences*, *2*(2), 141–178.

Cockburn, A. (2006). *Agile software development: The cooperative game* (second ed.). Boston, MA, USA: Addison-Wesley.

Computer science curricula 2013. (2013). ACM Press and IEEE Computer Society Press.

Cybersecurity education workshop: Final report. (2014, February). `https://research.gwu.edu/sites/research.gwu.edu/files/downloads/CEW_FinalReport_040714.pdf`.

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. *Handbook of research for educational communications and technology*, 170–198.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American psychologist*, *53*(1), 5–26.

Hickey, D. T., & Zuiker, S. J. (2003). A new perspective for evaluating innovative science programs. *Science Education*, *87*(4), 539–563.

Klopfer, E., Osterweil, S., & Salen, K. (2009). *Moving learning games forward: Obstacles, opportunities, & openness* (Tech. Rep.). Cambridge, MA: The Education Arcade, MIT. (`http://education.mit.edu/papers/MovingLearningGamesForward_EdArcade.pdf`)

Margolis, J., Estrella, R., Goode, J., Holme, J. J., & Nao, K. (2010). *Stuck in the shallow end: Education, race, and computing*. MIT Press.

Margolis, J., & Fisher, A. (2003). *Unlocking the clubhouse: Women in computing*. MIT Press.

Papert, S., & Harel, I. (1991). Situating constructionism. *Constructionism*, *36*, 1–11.

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Sage.

Stake, R. E. (2010). *Qualitative research: Studying how things work.* The Guilford Press.

Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge university press.