Authenticity, Ethicality, and Motivation: A Formal Evaluation of a 10-week Computer Security Alternate Reality Game for CS Undergraduates

John R. Morelock
Virginia Tech

Zachary Peterson
Cal Poly, San Luis Obispo

Abstract

Alternate reality games (ARGs) have been shown to have desirable characteristics for computer security education and student motivation. We implemented a 10-week-long ARG in an introductory undergraduate computer science course, and formally assessed the ARG's impact on students' course experience, as well as examined students' motivation-related experiences in the course by gender. Among other conclusions, we found that the ARG enabled an authentic and motivating problem-solving environment, but also raised ethical concerns among students that could lead to constructive discussions on ethical behavior in computer security. We also found that the ARG's use of several programming languages has detrimental effects on novice students—especially women—who felt at a disadvantage compared to their peers. We discuss connections to extant literature and implications for future implementations of the ARG.

1 Introduction

In order to address issues of retention and diversity, many universities have sought to redesign their introductory CS curricula, employing new methods and objectives to address these deficiencies. In particular, CPE123 is a reinvented introduction to computer science offered at Cal Poly, San Luis Obispo [20, 21, 40]. In this course, students engage with core CS principles through constructivist, open-ended assignments. Multiple sections of the course are offered simultaneously, but in a variety of "flavors" (including robotics, computational art, video game design, and digital music), encouraging students to explore CS concepts through preexisting interests. Results from CPE123 and courses like it at other universities have shown decreased attrition rates, increased student performance in subsequent CS courses, and demonstrate a higher retention rates among women and minorities in CS [11, 12, 17, 40].

In this paper, we focus on a unique "flavor" of CPE123; one that explores core computer science topics through the lens of computer security, merging game play, puzzle-based learning, and storytelling to provide an engaging, relevant, and contextually meaningful experience. We achieve this by using techniques drawn from alternate-reality games (ARGs). ARGs are story-driven, trans-media art, designed to encourage players to collaboratively uncover and interpret fragments of a story, distributed across multiple forms of media, using the "real world" as its platform and basis for game mechanics. First and often used for promotional and marketing purposes, ARGs have garnered attention in educational settings [3, 6, 14, 19, 22, 26], and have been shown to have valuable pedagogical characteristics, *e.g.*, to inspire counterfactual thinking [18], to maintain engagement [15, 27], and to provide an authentic context and purpose for presented material, both online and in the real world [3]—many of the same characteristics observed to be lacking in current computer security curricula. Indeed, these were our findings in our previous efforts to integrate ARG-style narratives to bridge class concepts to the real world, drawing students into fictional scenarios presenting authentic computing problems [14].

Here, we revisit and improve upon our initial efforts, both in deployment and assessment. We have expanded beyond a collection of narrative-driven, thematic puzzles, to a full-scale, 10-week alternate reality game. One in which a single narrative connects coursework throughout the quarter—students interact with fictional characters and fabricated web sites, perform "real world" attack and defense exercises, and attempt to reveal the identify of a mysterious hacker determined to bring down their university. Further, we look to revisit and improve upon our initial efforts in evaluation, both in refining our instruments and increase the scope of inquiry. Specifically, our assessment of the course addressed the following, coarse-grained research questions:

1. What aspects of the ARG affected student experience in the course?
2. How did students' motivation-related experiences in the course differ based on gender?

Using an action research approach, we answered the first question using qualitative data from two large student focus groups at the end of the course, and answered the second question using mixed methods data combining these focus groups and an end-of-term survey on student motivation.

In answering our second research question, we have elected to measure student motivation using Eccles' Expectancy-Value Theory (hereafter EVT). EVT was initially created to explain gender differences in children's math performance and persistence [39], and it has since been adapted for use in predicting student pursuit in a wide range of academic disciplines, including engineering and computer science [23, 35, 38, 39]. EVT posits that students' goals, self-beliefs, and experiences influence five motivation-related constructs (one expectancy and four task values), which proceed to influence students' choices and performance in a given activity. The five constructs were defined by Wigfield [38] as follows:

1. **Expectancy for success:** The perception of how well one will do on a given activity. Conceptually and operationally similar to the constructs of self-concept and competence beliefs.
2. **Attainment value:** The importance of doing well on a given activity or domain.
3. **Utility value:** The usefulness of an activity or domain in relation to future plans.
4. **Intrinsic interest value:** The enjoyment one gains from an activity or domain.
5. **Cost:** The opportunity cost, difficulty, and emotional cost of engaging in an activity. We have operationalized cost as Eccles did in an earlier work, as divided into (1) effort required to do well, and (2) difficulty of doing well [13].

The remainder of this paper identifies the pedagogical goals and game mechanics used in designing our ARG. We then discuss the creative process in developing the game and the mechanics of deploying it. We then present and reflect on our findings, which look to assess the enjoyment and efficacy our game, as well as provide guidance to those looking to offer a similar course.

2 Course Goals & Design

A complete list of the course's learning objectives can be found in Appendix A of this paper and the specific knowledge, skills, and abilities are detailed in our prior work [14]. However, we summarize expected outcomes of our course here. At a high-level, all sections of CPE123 aim to engage incoming freshmen, especially those that have no prior experience in computer science, with authentic problems that demonstrate the relevance of computing to the world around them. The course highlights the role of computers in both solving and constructing problems, and to challenge students with creative, constructivist challenges that help students develop an ability to think "computationally."

Our variant of CPE123 augments these goals with objectives specific to computer security, including building confidence, developing good intuitions, and integrating the best security practices and behaviors into the students' own lives. This includes developing an informed sense of skepticism around security claims as well as an ability to think counterfactually and adversarially—*i.e.*, to think like an attacker would. Ultimately, we hope to impart complex security concepts typically reserved for advanced computer science majors (despite their wide relevance to all), but in a way that is accessible to a wide audience, and which fosters curiosity in security as well as a lifelong love for CS and other STEM disciplines.

2.1 Game Design

Before we began development, we further identified some game mechanics specific to the design of a pedagogical ARG, including: that we develop an intriguing and realistic narrative, that the game mechanics and challenges balance accessibility and technically accurate phenomena, and the game provide opportunities for collaboration, communication, and community-building.

With these goals in mind, we contracted with Lee Sheldon, a renowned pedagogical alternate reality game designer and screenwriter, to develop the story, characters, and high-level mechanics for the game. We engaged in an iterative design process, which began by sharing our broad learning objectives as well as a coarse-grained breakdown of the expected knowledge, skills and abilities for the course, and continued with weekly refinements on the alignment between story, puzzle elements, and course content. The outcome of which was a design document, organized by a week-by-week breakdown of the game, identifying the active characters, the narrative to be advanced, any videos or other messages to be communicated, and the digital and physical puzzles to be solved, all aligned with the lecture and lab topics to presented that week.

2.2 The Story[1]

In 1987, a computer science student at Cal Poly was expelled from the university for allegedly hacking into the school's servers, and stealing sensitive research for a foreign government. Just before police were about to arrest him, he disappeared with his wife and newborn child.

Nearly twenty years later, a hacker organization, led by a single, shadowy figure emerges on the dark web, claiming credit for several extremely embarrassing attacks on US universities. They have been relatively silent since those initial attacks, but rumors suggest they are planning something major they are calling a "devilish disruption."

The shadowy figure is, in fact, the daughter of the expelled student. To avoid prosecution, her father fled to the country that paid him, leaving our story's antagonist behind in foster care. It appears she is now out to avenge her parents' exile and her abandonment on Cal Poly, the school she believes shattered her family. Her initial point attack is the security section of Cal Poly's CPE123 course.

Two weeks into the course, during a lab where students practice their password cracking skills against a Cal Poly server specially designated for this task, students will encounter something that is clearly not part of the assignment. Students will come across a page indicating the server has been compromised, on which the attacker has left an ominous warning to both the students of CPE123 and administration of the university of a "devilish disruption." Rumors of the dark web again begin to swirl that the nefarious hacker group has returned for their final, and most serious attack.

As the course continues, the hacker group and the woman who leads them, will continue to taunt the class through videos posted to a dark web site and disrupt the course assignments, leading the instructor to enlist the students in attempting to investigate these attacks, and help determine the group's motivation and ultimate goal.

Through a series of network, web, cryptographic, and digital forensic exercises, students discover the true identity of the hacker, learn the source of her animus for the university, and her intention: to find a program written by her father useful in gaining a backdoor into the control systems of the Diablo Canyon nuclear power plant, located a few miles from the university, giving her the ability to cause a major, and possibly dangerous, disruption.

In a dramatic climax, students must collaborate and use their newly learned computing skills to prevent this attack and identify the hacker's whereabouts. Ultimately, the hacker will be digitally brought before the class, where she will claim to never had intended to cause any real harm, but to show that it simply *could* happen. A poorly judged proof of concept from a well-intentioned but disillusioned youth, and a teaching moment for those who she feels destroyed her family. Students will ultimately decide whether to let her go, having learned her lesson and already been punished enough, perhaps reflecting on their own ethical and legal choices throughout the course, or to inform the authorities of her whereabouts, which will lead to her arrest. A final post on the dark web will announce her fate.

3 Assessment Methods

In accordance with best practices in engineering education assessment [30], we implemented the ARG and course assessment following an action research model. Case and Light [5] specify that action research consists of four phases: (1) planning learning activities, (2) implementing learning activities, (3) formally assessing the outcomes of learning activities in context, and (4) reflection upon implementation to inform future iterations of the learning activities. The methods and results of this paper expound the third and fourth phases of our action research process. The remainder of this section will detail our formal assessment methods that—in tandem with instructor reflection—informed our results and plans to improve the ARG in future academic terms.

We answered our research questions through a mix of qualitative and quantitative assessment. We addressed our first research question through two large, qualitative student focus groups and our second research question through a mixed methods concurrent triangulation [9], combining the focus group data with quantitative student survey data to examine patterns in students' motivation-related experiences. All of our data collection activities took place at the end of the Fall 2017 academic quarter, and were approved by the Institutional Review Board at both Virginia Tech and Cal Poly, San Luis Obispo.

3.1 Qualitative Assessment Methods: Data Collection

We used semi-structured focus group interviews to capture overall student experience in course and ARG, as well as changes in students' perceptions of cybersecurity and computer science as a result of the course. In doing so, we were able to identify specific aspects of the ARG narrative and related course activities that students considered important to their experience throughout the term, helping us answer both our research questions. Interested readers can find our focus group protocol in Appendix B.

[1] We have deliberately omitted some details from this description to avoid spoiling future offerings of the game.

We conducted two focus groups with students taking the course. The instructor provided students with the date and time of each focus group and requested volunteers; students could choose to show up at either session with no RSVP required. The first focus group consisted of 12 students (3 female and 9 male). This focus group featured students with a range of prior programming and cybersecurity experience, and the participants agreed they had a positive experience in the course overall. The second focus group consisted of 8 students (1 female and 7 male.) These students also had a range of prior programming and cybersecurity experience, but were more critical of the course than the first focus group. Both focus groups were video-recorded and transcribed for analysis.

3.2 Qualitative Assessment Methods: Data Analysis

We followed the coding process outlined by Miles *et al.* [24]. We began by descriptively open-coding the two focus group transcripts, noting the gender of each speaker and documenting student statements related to the following variables: (1) positive and negative aspects of the course, (2) positive and negative aspects of the ARG, (3) suggested changes to the course and ARG, and (4) changes in students' perceptions, behaviors, and understanding related to cybersecurity. We then grouped codes within each variable into a smaller number of pattern codes to guide our discussion of results. In answering our first research question, we focused on the positive and negative aspects of the course and ARG identified by students. Answering our second research question involved looking for differences in patterns of responses by gender across all variables.

To ensure the rigor of our qualitative work, we have used strategies suggested by Anfara *et al.* [1]. To ensure our work was credible, Morelock independently conducted the qualitative data analysis, and we used peer debriefing to confirm that the qualitative results matched the experiences of the Peterson (the instructor), who had prolonged engagement with students during the course. To enhance the transferability of our work, we have described our course and ARG so that readers may compare our context to theirs. To enhance the dependability of our work, we used descriptive codes that aligned closely with students' actual phrasing. Because we worked with a small qualitative data set, we did not use triangulation to further bolster dependability.

3.3 Quantitative Assessment Methods: Data Collection

We collected survey data on students' motivation-related beliefs about the course and ARG by adapting a validated EVT instrument used in mathematics education by Eccles and Wigfield [13]. Our adaptation involved changing the wording of the items to address computer science and cybersecurity (instead of mathematics) and to be answered using a standardized 6-point Likert scale (ranging from "strongly agree" to "strongly disagree.") We also added a questions related to the ARG for each EVT component, which we analyzed separately from the other EVT items. Our survey instrument can be found in Appendix C.

The survey was sent to students via email at the end of the academic term, along with a request to fill it out. Out of 59 students (45 male, 14 female) taking the course, we received 55 responses, one of which a student asked us to redact because he felt he misunderstood the Likert scale and did not believe his responses were accurate. Of the remaining responses, 40 students self-identified as male, and 14 self-identified as female in an open-ended response requesting student gender. No other genders were entered, and all students responded to the gender question.

3.4 Quantitative Assessment Methods: Data Analysis

Our quantitative analysis focused on gender comparisons to answers our second research question in tandem with the focus group data. Accordingly, we conducted a t-test to compare responses from men and women along each EVT scale, once for computer science and cybersecurity (the original scales) and once for student perceptions of the ARG (the new questions we added.) We conducted all analysis in JMP, using a significant level of $\alpha = 0.05$ to determine statistical significance.

4 Assessment Results

Qualitatively, our results indicate that students found the ARG to be a reasonably authentic context to motivate them to solve cybersecurity problems, noted an important tension between authenticity and ethicality in the ARG, and identified changes in their perceptions of the cybersecurity profession. Both quantitative and qualitatively, we found that female students came into the course with less programming experience than male students, and this affected both their expectancies for success and perceived costs of doing well in computer science and cybersecurity. The remainder of this section will expound these results in greater detail.

RQ1. What aspects of the ARG affected student experience in the course?

In looking for patterns of how students' ARG experiences affected their course experiences, the authenticity of the ARG—how "real" the narrative felt, despite its "fakeness"—was the thread that tied most discussions together in both focus group. Our first observation was that students connected the game's authenticity to their motivation to engage in learning activities. As one male student said:

> *Even though we all knew exactly when [the ARG] started and exactly what was going on, it was way more engaging than just getting assignments...when you're like, "Hey, we're doing an attack," you're like, "Oh, I know this is fake, but it's exciting."*

Here, the student expressed that the ARG added a sense of motivating authenticity to assignments, even when he acknowledged it was fake. Similarly, another male student expressed that the ARG motivated him to do coursework:

> *And then with the game, that made it more fun. And even though some things ended up taking multiple hours, it didn't feel like it was a task. It felt like it was just a fun thing to do.*

Students noted that some elements of the ARG—including the videos and the reveal of TA as an FBI agent—detracted from its authenticity, but were nonetheless enjoyable. As one female student proclaimed:

> *I think because, at least for me, I knew from the beginning [the ARG] wasn't real, that even though throughout the quarter [with the introduction of videos], it became like more obvious, I still really enjoyed the videos...I thought they were so cool. I showed my family members, and it was really cool.*

To this student, the inauthentic elements of the course better allowed her to share her ARG experience with important people in her life, a sentiment echoed by other (predominantly female) students. Another male student, however, expressed relief that these elements came at the tail end of the game's narrative:

> *I thought the videos were entertaining, I was laughing at them too. But I'm glad they were at the end instead of at the beginning when we were all a little more believing [the ARG].*

Students considered authenticity to be an exclusively positive factor in their experience with the ARG, but also noted a tension between authenticity and another factor they considered important: the ethicality of performing adversarial learning activities. For example, although the instructor told students on the first day of class that the ARG was not real and their activities would not harm anyone, one student participant was not present on that day, and he shared his experience:

> *I missed the first day of this class when you said that this was not a real thing we were doing. So I actually thought we were doing a phishing attack against a person... I was like, "This could so be real, and then we're just going to go and do the worst possible thing?" So I had that whole thing, like, "No, I'm not going to participate." I think it was interesting to see myself, like, "No, I literally cannot do this." I had an introspective moment there, like, "I can't do this."*

This student revealed that missing the announcement that the ARG was fake caused course activities to raise ethical red flags that frustrated his learning experience. Even for those students who were present on the first day, some appreciated that inauthentic elements of the game helped keep the ethical boundary obvious. One female student in the first focus group discussed:

> *I think on the other end though, making [the ARG] more fake and not as believable I think made the [ethical] line a lot clearer...seeing this in a game and a set setting, you're not harming someone else, this is completely within [Cal Poly's] system...It was nice to know that it wasn't real and we were doing it in practice, rather than—I think it would have been a huge issue if like it wasn't like that.*

Students in the second focus group in particular noted the need to communicate the game's "fakeness" if inauthentic elements of the game are removed in future implementations. One male student recommended having an easily recognizable marker on any materials related to the ARG, an idea considered by other ethically-concerned ARG designers [28]:

> *Little markers, that could just...if a person is consciously trying to suspend disbelief, then that marker won't mean anything to them because you're consciously suspending your disbelief. It will just do away with any worries that they might have that [the ARG] is real.*

Finally, several students noted that the course helped change their perceptions of the cybersecurity profession, in terms of either their understanding of professional activities or their intended professional trajectories, both of which can be attributed at least in part to playing the role of the attacker in the ARG. In terms of their understanding of professional activities, multiple students noted that the technological world is much more vulnerable to attack than they previously thought, perhaps summed up best by one male students whose biggest takeaway from the course was, “Everything is terrible.” More particularly, students discussed realizing that cyberdefense is much more difficult than cyberattack, that even experts may struggle to prevent all attacks, that cybersecurity is a necessary element of most computer science work, and that there is an important need for more cybersecurity work. One male student, for example, stated:

> *You have to be right on defense 100% of the time. And on attack, you only have to be right once. And so, I was thinking about that, and it makes [cybersecurity] a very important profession, given how it must be a perfect system in that way. So it seems like it’s actually way more necessary.*

In terms of professional trajectories, students said that the course helped them identify which aspects of cybersecurity did or did not interest them. Multiple students mentioned that they could now “see themselves in” particular cybersecurity roles, while other students mentioned already having an interest that was reinforced by the course. One female student also mentioned that the course helped her identify aspects of cybersecurity in which she had no interest:

> *I’m going to get into web security. I don’t know, it looks fun...I think that what this class did tell me is don’t get into cryptography. Okay, I’m not going to lie. I was falling asleep during the lecture because I was just, I couldn’t grasp it. But web security, I was into it.*

Here, the student asserted that the course helped her discover that she was interested in web security but not cryptography, an outcome that we consider positive given that other students said the course reinforced their interest in cryptography.

RQ2. How did students’ motivation-related experiences in the course differ based on gender?

Quantitatively, we found that there were no statistically significant differences between genders in terms of perceptions of the ARG, but there were significant differences in perceptions of computer science and cybersecurity as a whole. Particularly, as shown in Figure 1, female students have lower expectancy for success in computer science and cybersecurity than male students, and higher perceptions of the cost of doing well in terms of both difficulty and effort required.

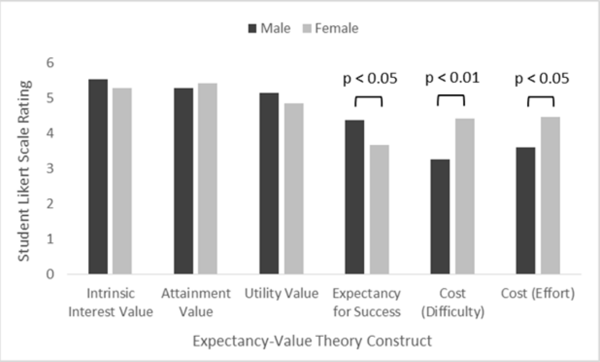

Figure 1: Gender Comparison of EVT Constructs (n = 54: 40 male, 14 female).

Qualitatively, we found that these differences likely stem from a lack of programming experience upon entering the course among women. Although women were underrepresented in our focus groups, they were far more vocal than men about the course’s requirement to be at least literate in multiple programming languages, most notably Python and C. They expressed that, compared to more experienced students, they felt “left behind,” required excessive use of search engines to make up the knowledge gap, and had trouble knowing what to do in some ARG-related labs. As one female student said:

> *I felt like with a lot of the labs, I knew conceptually what I had to do, but I did not have the skills in writing code to actually do it, because all the lectures were conceptual, and I seemed to understand them. But then I had no tools to go in and do that in the lab.*

Notably, several male students with prior programming experience expressed that they found that challenge level of the course gratifying, and one expressed that he would have been disappointed if the course were made easier to accommodate novice students.

5 Discussion

Overall, our results offer evidence that an ARG can be an effective format for conducting adversarial learning activities in an introductory cybersecurity course. In implementing our adversarial learning activities, having an ethical safe space to practice the role of the attacker is an important requirement that we found our ARG is well-suited to address. The ARG added a narrative that students viewed as authentic, motivating them to step into the attacker role and engage with a variety of coursework. Moreover, the ARG offers enough "fakeness" that most students did not need to worry about the ethicality or legality of their actions.

As noted by famed ARG designer Andrea Phillips [28], creating alternate realities raises new legal and ethical questions around the role of alternate reality game designers and their creations. Seemingly, in games where players are represented by digital or imaginary avatars (*e.g.* video games or traditional role playing games) those players can more easily make ethical and moral decision that may not align with their own belief systems. The Grand Theft Auto series exemplifies this phenomena well—where amoral decisions come with little consequence and indeed, may be required to advance in the game. Whereas in ARGs, players adopt a near identical version of themselves in a game environment that is intentionally indistinguishable from reality. As such, players bring their belief systems with them into the game, and may be less willing to engage in unethical or immoral behavior. Concerns lay on the other side as well, with examples of players gaining a false sense of ability and authority which have led players (in other games) to engage in unethical, illegal, and dangerous behavior.

Quandaries around students "hacking back," establishing limits of "self-defense" in cyberspace, and violating the privacy of fictional, though perhaps malicious, characters are equal parts compelling and troubling. However, these issues are nonetheless topical, and we take advantage of these moments in the game to break from a more traditional philosophical discussions and more constructively explore these issues, while being able to rely upon a student's suspension of disbelief and assurance of a safe space to explore, in ways other pedagogical approaches cannot.

It is worth noting that while our students generally did not express anxiety that they were acting unethically during the game, many did acknowledge that they would consider adversarial activities to be unethical in most other contexts. Accordingly, following a constructivist teaching approach [4], the ARG also presents an opportunity to challenge students' assumptions about adversarial activities, and prime them to discuss the kinds of scenarios in which "white hat" hacking is appropriate. Incorporating such a discussion into the ARG without sacrificing the ARG's authenticity is one improvement we aim to incorporate in the next iteration of the course.

Indeed, one of the most compelling aspects of an ARG, generally, is a participant's willingness to suspend disbelief; to more easily adhere to a "this is not a game" philosophy [33], where students engage with the game in their own reality, not in an contrived or artificial environment. Future iterations of our game may abandon some of the more "over the top" and inauthentic aspects of the game—for example, the jokiness in posts made on fictional web sites by fictional characters, or in the too-high production value of video messages from our antagonist—in favor of game elements that are more inline with the "look and feel" of a hacker narrative, while still identifiable as game elements. If we were to take measures to make the ARG feel more authentic in future academic terms, as students suggested, accompanying effort would be needed to ensure the game still feels "fake" enough to abate student worries about acting unethically.

In terms of student persistence and motivation-related beliefs, the ARG helped students "try on" roles in different aspects of cybersecurity and learn or reinforce what interested them and what didn't. According to expectancy-value theory [38], this knowledge of one's own interest can influence persistence by affecting the utility value of the classes students take or by making them more intrinsically interested in their classes. In a similar vein of literature, engineering identity researchers have found that the ability to "try on" different engineering roles can allow students to build a stronger engineering identity [8, 10, 25], which can improve retention in engineering programs [23, 29].

Quantitatively, we found that female students' intrinsic interest value, attainment value, and utility value for cybersecurity and computer science were not significantly different from those of men. These results are encouraging in terms of persistence, as task values taken together are strong predictors of student persistence [37, 38]. However, these results are not particularly surprising, as previous studies have found that although gender gaps exist in these task values for non-computer-science majors, gender gaps are less pronounced or not present for computer science majors [16, 34].

However, because the ARG asks students to act as attackers in a variety of ways, it requires a baseline level of programming knowledge. Our data suggested men were more likely to have this knowledge than women, women considered it more difficult and effortful to make up the difference, and women were less confident in their computer science and cybersecurity skills than men. These are all commonly cited issues in computer science [2, 7, 31]. EVT suggests that lower expectancy for success can lead to lower achievement, and higher cost value can lead to a choice not to persist [38].

Accordingly, action must be taken to reduce the programming skill gap in future iterations of the ARG. Several students in our focus groups suggested making it easier for advanced students to help novice student along the way, such as establishing experience-balanced "help groups" or having students share contact information at the beginning of the course. Another suggestion was to hold optional class sessions for students who struggled in lab could be walked through how the lab could be done. Students' suggestions align well with research suggesting collaborative approaches to addressing experience gaps in computer science such as pair programming [2, 36], though studies have found that these approaches work best when the skill levels of students working together are not too disparate [32].

Finally, we found that running a 10-week alternate reality game to be an involved and relatively complicated endeavor, with lots of opportunity for the unexpected to occur. We faced a number of challenges related to student actions (positive and negative) that required us to adapt the story or other game elements on the fly. For example: students progressing through game elements at different and unplanned rates (both too fast and too slow); students finding continuity errors in the story line or flaws in the digital assets (*e.g.* discovering that two unrelated web site were hosted on the same physical machine); and running into scheduling problems related to the academic calendar (*e.g.* we observed a small break in the game play due to a cancelled class caused some disengagement).

Overall, we found it very important to be responsive to students' involvement: changing the narrative as necessary to maintain engagement, increasing their stake in the game, and rewarding intrinsic motivations and curiosity. Further, we found the instructor's willingness to participate in the game themselves, increased student buy-in. However, this does require some acting skills and a willingness to dynamically go "off script" in real time, in order to maintain the game's illusion.

6 Conclusion

We evaluate an alternate reality game in an introductory computer science course to allow students to practice cybersecurity concepts in an adversarial setting. Using mixed methods data, we assessed how the ARG affected student experience in the course, and explored gender differences in students' motivation-related course experience. We found that the ARG was an effective vehicle to motivate course activities and reveal ethically ambiguous situations in cybersecurity work, while also helping to refine students' understanding and interests related to the cybersecurity profession. However, we also found that tasks required of students throughout the ARG were best suited to students with prior programming experience, which was more common among men than women. In accordance with our findings, our goals for the next iteration of the ARG are to (1) have more productive ethics discussions throughout the term while maintaining the ARG's sense of authenticity, and (2) reduce the gendered experience gap to provide a less frustrating experience for novice programmers, especially women.

Acknowledgements

We would like to thank Dr. Lee Sheldon for his work in designing the alternate reality game narrative, Nick Gonella for his dual role as TA and FBI agent in our initial offering, as well as Dr. Holly Matusovich and her research group for their feedback, which helped refine this paper.

This paper is based on research supported by the National Science Foundation under Grant No. 1419318 and by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1651272. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] ANFARA, V. A., BROWN, K. M., AND MANGIONE, T. L. Qualitative Analysis on Stage: Making the Research Process More Public. *Educational Researcher 31*, 7 (2002), 28–38.

[2] BARKER, L. J., MCDOWELL, C., AND KALAHAR, K. Exploring factors that influence computer science introductory course students to persist in the major, 2009.

[3] BONSIGNORE, E., HANSEN, D., KRAUS, K., VISCONTI, A., AHN, J., AND DRUIN, A. Playing for real: designing alternate reality games for teenagers in learning contexts. In *Proceedings of the International Conference on Interaction Design and Children* (2013).

[4] BROOKS, J. G., AND BROOKS, M. G. *In Search of Understanding: The Case for Constructivist Classrooms*. Association for Supervision and Curriculum Development, 1999.

[5] CASE, J. M., AND LIGHT, G. Emerging Research Methodologies in Engineering Education Research. *Journal of Engineering Education 100*, 1 (2011), 186–210.

[6] CHOTHIA, T., HOLDCROFT, S., RADU, A.-I., AND THOMAS, R. J. Jail, Hero or Drug Lord? Turning a Cyber Security Course Into an 11 Week Choose Your Own Adventure Story. In *2017 USENIX Workshop on Advances in Security Education (ASE 17)* (Vancouver, BC, 2017), USENIX Association.

[7] COHOON, J. M. G., AND ASPRAY, W. *Women and Information Technology: Research on Underrepresentation*. MIT Press, Cambridge, MA, 2006.

[8] CREDE, E., AND BORREGO, M. The effect of international diversity on graduate engineering education: A literature review. In *2010 ASEE Annual Conference and Exposition, June 20, 2010 - June 23, 2010* (Louisville, KY, United states, 2010), American Society for Engineering Education.

[9] CRESWELL, J. W. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications, 2009.

[10] DEHING, F., JOCHEMS, W., AND BAARTMAN, L. Development of an Engineering Identity in the Engineering Curriculum in Dutch Higher Education: an Exploratory Study from the Teaching Staff Perspective. *European Journal of Engineering Education 38*, 1 (2013), 1–10.

[11] DODDS, Z., AND ALVARADO, C. Women in CS: an evaluation of three promising practices. In *Proceedings of the ACM Technical Symposium on Computer Science Education* (2010).

[12] Dodds, Z., Libeskind-Hadas, R., Alvarado, C., and Kuenning, G. Evaluating a Breadth-First CS 1 for Scientists. In *Proceedings of the ACM Technical Symposium on Computer Science Education* (2008).

[13] Eccles, J. S., and Wigfield, A. In the Mind of the Actor: The Structure of Adolescents' Achievement Task Values and Expectancy-Related Beliefs. *Personality and Social Psychology Bulletin 21*, 3 (1995), 215–225.

[14] Flushman, T., Gondree, M., and Peterson, Z. This is not a game: Early observations on using alternate reality games for teaching security concepts to first-year undergraduates. In *Proceedings of the Workshop on Cyber Security Experimentation and Test* (2015).

[15] Grobstein, P. Revisiting science in culture: Science as story telling and story revising. *Journal of Research Practice 1*, 1 (2005).

[16] Haller, S., and Beyer, S. Gender Differences and Intragender Differences in Computer Science Students: Are Female CS Majors More Similar to Male CS Majors or Female Nonmajors? *Journal of Women and Minorities in Science and Engineering 12*, 4 (2006), 337–365.

[17] Hambrusch, S., Hoffmann, C., Korb, J. T., Haugan, M., and Hosking, A. L. A multidisciplinary approach towards computational thinking for science majors. In *Proceedings of the ACM technical symposium on Computer science education* (2009).

[18] Hansen, D., Bonsignore, E., Ruppel, M., Visconti, A., and Kraus, K. Game design for promoting counterfactual thinking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (May 2012).

[19] Hansen, D., Bonsignore, E., Ruppel, M., Visconti, A., and Kraus, K. Designing reusable alternate reality games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013).

[20] Haungs, M., Clark, C., Clements, J., and Janzen, D. Improving first-year success and retention through interest-based CS0 courses. In *Proceedings of the ACM Technical Symposium on Computer Science Education* (2012).

[21] Haungs, M., Clements, J., and Janzen, D. Improving Engineering Education Through Creativity, Collaboration, and Context in a First Year Course. In *Proceedings of the American Society for Engineering Education Annual Conference* (2008).

[22] Macvean, A., Hajarnis, S., Headrick, B., Ferguson, A., Barve, C., Karnik, D., and Riedl, M. O. WeQuest: scalable alternate reality games through end-user content authoring. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology* (2011).

[23] Matusovich, H. M., Streveler, R. A., and Miller, R. L. Why Do Students Choose Engineering? A Qualitative, Longitudinal Investigation of Students' Motivational Values. *Journal of Engineering Education 99*, 4 (2010), 289–303.

[24] Miles, M. B., Huberman, A. M., and Saldaña, J. *Qualitative data analysis: a methods sourcebook*, 3rd ed. SAGE Publications, Inc, Thousand Oaks, Ca, 2014.

[25] Morelock, J. R. A systematic literature review of engineering identity: Definitions, factors, and interventions affecting development, and means of measurement. *European Journal of Engineering Education* (2017), 1–23.

[26] Moseley, A. An Alternate Reality for Education? Lessons to be Learned from Online Immersive Games. *International Journal of Game-Based Learning* (2012).

[27] Niemeyer, G., Garcia, A., and Naima, R. Black cloud: patterns towards da future. In *Proceedings of the ACM International Conference on Multimedia* (2009).

[28] Phillips, A. `http://www.deusexmachinatio.com/blog/2010/8/17/the-ethics-of-transmedia-at-sxsw.html`, August 2010.

[29] Pierrakos, O., Beam, T. K., Constantz, J., Johri, A., and Anderson, R. On the development of a professional identity: engineering persisters vs engineering switchers. In *2009 39th IEEE Frontiers in Education Conference. Imagining and Engineering Future CSET Education (FIE 2009), 18-21 Oct. 2009* (Piscataway, NJ, USA, 2009), IEEE, p. M4F (6 pp.).

[30] Raubenheimer, C. D. Assessment for improving teaching and student learning within a course. In *Designing Better Engineering Education Through Assessment: A Practical Resource for Faculty and Department Chairs on Using Assessment and ABET Criteria to Improve Student Learning*, J. E. Spurlin, S. A. Rajala, and J. P. Lavelle, Eds. Stylus Pub., Sterling, Virginia, 2008, ch. 9, pp. 246–265.

[31] Rubio, M. A., Romero-Zaliz, R., Mañoso, C., and de Madrid, A. P. Closing the gender gap in an introductory programming course. *Computers & Education 82* (2015), 409–420.

[32] Salleh, N., Mendes, E., and Grundy, J. Empirical Studies of Pair Programming for CS/SE Teaching in Higher Education: A Systematic Literature Review. *IEEE Transactions on Software Engineering 37*, 4 (2011), 509–525.

[33] Szulborski, D. *This is not a game: a guide to alternate reality gaming.* New Fiction Publishing, 2005.

[34] Wang, J., Hong, H., Ravitz, J., and Ivory, M. Gender Differences in Factors Influencing Pursuit of Computer Science and Related Fields, 2015.

[35] Wang, M.-T., and Degol, J. Motivational pathways to STEM career choices: Using expectancyvalue perspective to understand individual and gender differences in STEM fields. *Developmental Review 33*, 4 (2013), 304–340.

[36] Werner, L. L., Hanks, B., and McDowell, C. Pair-programming helps female computer science students. *J. Educ. Resour. Comput. 4*, 1 (2004), 4.

[37] Wigfield, A., and Cambria, J. Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review 30*, 1 (2010), 1–35.

[38] Wigfield, A., and Eccles, J. S. Expectancyvalue theory of achievement motivation. *Contemporary educational psychology 25*, 1 (2000), 68–81.

[39] Wigfield, A., Tonks, S., and Klauda, S. L. Expectancy-value theory. In *Handbook of motivation at school*, K. Wentzel, A. Wigfield, and D. Miele, Eds. Routledge, 2009, pp. 55–75.

[40] Wood, Z. J., Clements, J., Peterson, Z., Janzen, D., Smith, H., Haungs, M., Workman, J., Bellardo, J., and DeBruhl, B. Mixed approaches to CS0: Exploring topic and pedagogy variance after six years of CS0. In *Proceedings of the ACM Technical Symposium on Computing Science Education* (2018).

A Learning Objectives

Computer Science Learning Objectives

1. Be exposed to the interdisciplinary nature of computer science, and its myriad professional opportunities.
2. Begin to see how computer science, as a discipline, can be meaningful, socially relevant, and change the world for the better.
3. Develop an ability to think “computationally,” allowing students to analyze and create solutions to computational problems.
4. Be provided with the core programming and problem solving skills to be successful in follow-on CSC coursework, build a community of peers, and have fun!

Computer Security Learning Objectives

1. Demonstrate a general knowledge of, and be able to differentiate between, core computer security principles.
2. Approach claims of security with an informed sense of skepticism, as well as develop an ability to intelligently critique those claims.
3. Recognize that information we store digitally can have real value, but can also be valued differently, warranting different protections.
4. When considering the value of digital assets, identify the common and emerging threats to those assets, and identify the appropriate techniques for protecting those assets.
5. Gain confidence, develop good intuitions, and integrate into their own lives the best practices and behaviors to keep secure in daily life.
6. Establish, in addition to computational thinking skills, an ability to think “counterfactually” and “adversarially” about systems, allowing students to protect these systems against adversaries thinking about them in the same way.

B Focus Group Protocol

1. Now that you’re finished with your first quarter, how have your perceptions about cybersecurity as a profession changed?
   (a) What about your first quarter contributed to those changes in perception?
   (b) What about this course led to those changes in perception?
   (c) How do you foresee these changes affecting your trajectory in the computer science program as you move forward?
2. Now that the course is over, how do you feel about your overall experience in the course?
   (a) What about the course did you like? What didn’t you like?
   (b) How did your experience in this course compare to your experience in other courses you took this quarter?
3. What kinds of things do you notice in your daily life as a result of what you learned in the course? How has your typical behavior changed as a result?
   (a) Can you think of an example where something cybersecurity-related happened in your life outside of class? What happened? How did you handle that when it happened?
4. How did the ARG you played throughout the quarter affect your experience in the course?
   (a) How do you feel the things you did in this courseand for the ARG in particularrelate to things you’d like to do here at Cal Poly or after you graduate?
5. Given everything we’ve talked about today, if you could change anything about the course to have a more meaningful experience, what would you change?

C Expectancy-Value Theory Survey Instrument

Table 1: Items related to computer science and cybersecurity as a whole.

EVT Component	Items
Intrinsic Interest Value	1. In general, I find working on cybersecurity problems interesting. 2. I enjoyed this cybersecurity class. 3. I enjoy computer science as a major.
Attainment Value	1. The amount of effort it will take to do well in computer science is worth it to me. 2. The amount of effort it took to do well in this cybersecurity class was worth it to me. 3. I feel that, for me, being good at solving cybersecurity problems is important to me. 4. It was important to me that I did well in this cybersecurity class. 5. It is important to me to do well in my computer science classes.
Utility Value	1. Learning about cybersecurity is useful for what I want to do after I graduate. 2. Learning about cybersecurity is useful for my daily life outside of classes.
Expectancy for Success	1. Compared to other students, I expect to do well in my computer science classes this year. 2. I think I will do well in my computer science courses this year. 3. Compared to other students, I believe I did well in this cybersecurity class. 4. I think I did well in this cybersecurity class. 5. I am good at solving cybersecurity problems. 6. If I were to order all the students in this class from the worst to the best at solving cybersecurity problems, I would put myself among the best.
Cost (Difficulty)	1. In general, computer science is a hard major subject for me. 2. In general, cybersecurity is a hard topic for me. 3. Cybersecurity is more difficult for me than for other students in my classes. 4. Computer science is more difficult for me than for other students in my classes. 5. Compared to other courses I am taking, this course was among the most difficult for me.
Cost (Effort)	1. I would will have to try very hard to do well in future computer science courses. 2. I had to try very hard to get good grades in this course. 3. I have to study very hard,for cybersecurity tests to get a good grade. 4. To do well in this course, I had to work much hard than in my other courses this quarter.

Table 2: Items related to the alternate reality game.

EVT Component	Items
Intrinsic Interest Value	1. The alternate reality game contributed to my interest in cybersecurity.
Attainment Value	1. It was important to me that I contributed to solving problems in the alternate reality game.
Utility Value	1. The alternate reality game changed my perception of how useful it is to learn about cybersecurity.
Expectancy for Success	1. The alternate reality game helped me feel more confident in my ability to do well in this course. 2. The alternate reality game helped me feel more confident in my ability to do well my computer science courses this year.
Cost (Difficulty)	1. I found the tasks required of me during the alternative reality game to be too difficult.
Cost (Effort)	1. I had to try very hard to successfully complete the tasks required of me during the alternate reality game.