PenQuest Reloaded: A Digital Cyber Defense Game for Technical Education

Robert Luh
St. Pölten UAS &
University of Vienna
Austria
rluh@fhstp.ac.at

Sebastian Eresheim
St. Pölten UAS &
University of Vienna
Austria
seresheim@fhstp.ac.at

Stefanie Größbacher
St. Pölten University of
Applied Sciences
Austria
sgroessbacher@fhstp.ac.at

Thomas Petelin
ICS Information Systems
Vienna, Austria
thomas.petelin@gmail.com

Florian Mayr
St. Pölten University of
Applied Sciences
Austria
it201505@fhstp.ac.at

Paul Tavolato
University of Vienna
Austria
paul.tavolato@
univie.ac.at

Sebastian Schrittwieser
University of Vienna
Austria
sebastian.schrittwieser@
univie.ac.at

***Abstract*—Today's IT and OT infrastructure is threatened by a plethora of cyber-attacks conducted by actors with different motivations and means. Furthermore, the complexity of these exposed systems as well as the adversaries' sophisticated technical arsenal makes it increasingly difficult to plan and implement an organization's defense. Understanding the link between specific attacks and effective mitigating measures is particularly challenging – as is understanding the underlying information security concepts.**

To support the training of current, and more importantly, nascent security engineers, we propose PenQuest, a digital attack and defense game where an attacker attempts to compromise an abstracted IT infrastructure and the defender works to prevent or mitigate the threat. The game is based on MITRE ATT&CK, D3FEND, and the NIST SP 800-53 security standard and incorporates a multitude of concepts such as cyber kill chains, attack vectors, network segmentation, and more.

PenQuest is built to support security education and risk assessment and was evaluated with a class of engineering students as well as independent security experts. Initial results show a significant increase in knowledge retention and attest to the game's feasibility for educational use.

Keywords—cyber-attack, security, awareness, gamification

I. Introduction

Cyber-attacks on IT and OT systems have become a common occurrence. Next to sheer volume, the economic impact of such threats is rising constantly: the center for Strategic and International Studies (CSIS) estimates worldwide losses close to $600 billion [1], which translates to roughly 1% of global GDP. The complexity of targeted information systems and the asymmetric nature of digital threats pose an immense challenge for organizations that have to consider countless attack vectors. Understanding the link between specific attacks likely to occur and effective mitigating measures is particularly challenging – as is internalizing the underlying information security concepts, which is key to implementing effective defense strategies. This puts security education into the spotlight.

While creating a curriculum is far from trivial in itself [2], conventional cyber-security lectures often compartmentalize topics in a way that prioritizes memorization, comprehension, and application [3] within its own context alone. Cross-topical analysis and synthesis can be lacking, particularly when considering information security management (i.e., risk assessment and organizational controls) and technical education (i.e., applied networking or OS specifics).

Serious games have been identified as a feasible means to educating students [4]. Different research [5] as well as security guidelines [6] emphasize that such games are not only well-suited to teach, but also to model information security principles for different audiences.

With this in mind, we propose PenQuest, a digital attack and defense game where an attacker attempts to compromise an abstracted IT infrastructure and the defender works to prevent or mitigate the damage. PenQuest is built upon a realistic model [7] [8] that was reworked from the ground up to accommodate gamified education and risk assessment in the digital domain. PenQuest is a virtual two-player board game that offers students a means to play through and dissect complex attacks without having to physically conduct them. More importantly, it enables learners to discover appropriate countermeasures on a technical, organizational, and human level. With PenQuest, we aim to a) make it easier to understand vulnerabilities and threats derived from adversary behavior; b) match attacks to appropriate security controls; and c) make it entertaining to explore such a complex topic, thereby motivating students to engage in the field. At the same time, we seek to retain a maximum of realism that will ultimately allow the game to be used in awareness and risk assessment scenarios. In short, the purpose of PenQuest is to offer a means to better understand practical IT security on a technical, strategic, and tactical level without having to work on a live system.

This paper details the design of the game on four different levels, which are discussed in Section IV: The foundational model of how different IT security concepts are translated to a non-cooperative, imperfect- & incomplete-information game, the data used to populate said model, game design principles related to visual arrangement and usability, and general game mechanics. In the evaluation part (Section V) we assess how well these aspects play together to create a suitable learning environment for IT and engineering students and how PenQuest impacts knowledge retention and learner reflection. Furthermore, we take a look at the underlying model's accuracy in terms of realism – in particular the attack-defense mapping of the game's actions that teaches students appropriate countermeasures to specific cyber-attacks.

This research was funded by the Austrian Science Fund (FWF): P 33656

II. Related Work

In the literature, we see three distinct approaches to games in the security context: works that focus on game-theoretic aspects and formal modeling, serious games used in e.g., awareness workshops, and the emerging field of AI-enabled adversarial games that use machine learning (ML) to facilitate malware detection or specific threat responses.

In the following review we focus on the second category. Refer to [7] for a more game-theoretic look at current strategy games. Future plans for PenQuest relating to artificial intelligence and strategy optimization are discussed in the concluding section of this paper.

One of the aforementioned serious games is "Elevation of Privilege" [9], a physical print-at-home card game which has been built to help people get started with threat modeling and aid aspiring analysts identify general threats to their IT and software systems. It is based on Microsoft's STRIDE mnemonic [10], which links threats to desired security properties such as confidentiality or availability. The game includes no mechanics for cyber-defense, and topological factors are not considered.

Also based on STRIDE, "Riskio" [11] offers attack and defense gameplay facilitated by a human game master, who decides which actions are successful by listening to the players' arguments. The game uses three distinct diagrams representing the game boards: an office map, a network diagram, and a data flow diagram.

"Operation Digital Chameleon" [12] is a red-team exercise in the form of a board game. Players are asked to collaborate to build an attack and defense strategy for a given scenario. The proposed solutions are again assessed by a game master. While this workshop approach offers flexibility and is suitable for dedicated events comprising large groups, "Operation Digital Chameleon" does not provide a distinct security model or ways to resolve scenarios computationally. There are similar solutions that combine physical game components and team-based decision-making. For example, "Backdoors & Breaches" [13] uses cards to provide attack tactics, tools, and methods. The goal of the defender is to reveal the attacker's cards within a turn limit. While the game considers a kill chain similar to PenQuest, it requires human facilitation akin to a role-playing game and does not include advanced mechanics like modeling different assets and their interconnections, threat prevention, or specific types of compromise.

"OWASP Cornucopia" [14] follows a different premise: The game offers cards to assist software development teams in identifying the security requirements of their projects by expediting discussion. The game links their technology-agnostic concepts to exemplary weaknesses (CWE) and attack patterns (CAPEC).

Moreover, there are educational and commercial games created to raise user awareness. Most put an emphasis on entertainment and do not encompass a more complex security model. Some applications, such as "Keep Tradition Secure" [15], the US Department of Defense's Cyber Awareness Challenge [16], "Targeted Attack: The Game" [17], or "The Weakest Link" [18] can be considered quizzes or interactive decision-making games where the user needs to pick the option that is least likely to lead to a compromise (e.g., not to publicize certain information or not to click a suspicious link) or spot potential security violations in a virtual room [19]. Other games, e.g., "CyberEscape Online" [20] add team building to the mix. These games incorporate a number of security best practices without focusing on the (technical) background. Phishing-specific solution such as "Craft that Phish" [21], "What.Hack" [22] or "Jigsaw" [23] take a closer look at specific threats and how they can be spotted. While all these games promise to increase awareness in novice and intermediate users by addressing human and some organizational vulnerabilities, none of them aim to teach IT security wholistically.

Additional games, many of them intended for younger audiences, can be found in Adam Shostack's repository [24].

III. Background

This section briefly summarizes the security concepts and principles directly or indirectly incorporated into PenQuest. Details on the integration into the game – the conversion to game rules and mechanics, respectively – are discussed in Section IV.

A. *IT Infrastructure*

PenQuest aims to impart a *Defense in Depth* approach to designing secure IT infrastructures. The term originally comes from military jargon and describes the strategy of slowing down and eventually halting an attack through a multi-layered defense system [25]. It further encompasses the use of access control systems, diversity in terms of utilized hard- and software, and the reduction of information disclosure by e.g., hiding publicly viewable error messages.

Our educational game focusses on *assets*, which generally describe tangible or intangible entities that need to be protected in in the context of an organization or computer system [26]. This includes storage systems and the data on them, databases, programs, memory content, and more. In risk assessment – for which PenQuest can be used in addition to training – the required security properties for each asset are evaluated and improved upon.

Another concept the game implements is that of *attack vectors*, which describes a possible attack path to, for example, gain unauthorized access to a computer system [27]. Attack vectors can be caused by security vulnerabilities in the software or by a misconfiguration of a system. It may also be inherent to the infrastructure since some systems have to be exposed by design in order to provide their services to the public. In PenQuest, the concept is applied to actions (see also Subsection C) as well as the possible paths an attacker can take to traverse a modeled infrastructure.

B. *Security Objectives*

The three most important objectives in information security and their relationship to each other are often summarized in the so-called CIA triad [28]. Confidentiality is achieved when no unauthorized gain of information is possible in a system. Integrity describes the correctness of a system and/or of its data. Availability describes the degree to which the functionality of a system is correctly made available to users. Other protection goals include non-repudiation (communication cannot be denied after the fact) and authenticity (ensuring that data originates from a specific entity).

PenQuest uses the CIA triad to model damage and mitigating effects on assets and assigns the respective game actions a number representing a net loss or gain in security.

C. Attack Stages

Similar to *Defense in Depth*, the term *kill chain* [29] is also rooted in a military concept. It describes the structure and sequence of actions of an attack. In information security, *cyber kill chains* [30] have become a popular concept. They describe the general stages of a cyber-attack ranging from pure observation and selection of potential targets (reconnaissance) to system access as well as exploitation, exfiltration or destruction of data. Supporting tasks like gaining persistence or evasion techniques are sometimes covered as well.

Some models go one step further: The *MITRE ATT&CK framework* is a full taxonomy of tactics and techniques for attacking an IT system [31]. While cyber kill chains represent a high-level view, the ATT&CK framework exists much closer to the actual technical implementation of an attack. Its tactics and techniques are more flexible and do not necessarily happen in a set order. Examples of MITRE ATT&CK tactics are 'Elevation of Privilege' (exploitation of a software vulnerability or configuration error with the purpose of gaining more privileges on a computer system) and 'Credential Gathering' (unauthorized access to e.g. a user's password hashes).

PenQuest uses both a simplified kill chain as well as most of the techniques found in MITRE ATT&CK to teach how such attacks typically play out. Examples for ATT&CK techniques can be found in the following section.

D. Threats, Techniques, and Vulnerabilities

In information security, a *threat* is a potential security violation which is often caused by human actions – either intentional (e.g., hacking a system) or unintentional (e.g., human error). Natural disasters (e.g., floods or earthquakes) can also be understood as threats to computer systems.

Techniques are concrete actions to achieve tactical goals. An example of a technique in the MITRE ATT&CK framework is 'process injection', where the attacker's code is injected into an existing process (typically a known program) to bypass process-based defenses or to achieve elevation of privilege.

A *vulnerability* is a concrete weakness of a computer system that can be exploited by an adversary. The Common Vulnerabilities and Exposures (CVE) system [32] provides an inventory of all publicly known vulnerabilities.

PenQuest aims to teach all of these concepts and implements numerous techniques as actions usable in the game. Exploits to vulnerabilities are available in the form of purchasable equipment.

E. Security Controls

On the defense side, a number of measures and techniques are available to mitigate or delay attacks. These include technical countermeasures as well as organizational controls and policies outlined in various security standards.

MITRE D3FEND [33] classifies technical cybersecurity countermeasures into the categories *harden* (e.g., message encryption), *detect* (e.g., operating system monitoring), *isolate* (e.g., network traffic filtering), *deceive* (e.g., decoy files), and *evict* (e.g., account locking). On the organizational side, several widely accepted security standards provide guidelines for preventative and mitigating controls. These standards include the ISO/IEC 27000 series [34], Common Criteria [35], and NIST SP 800-53 [36].

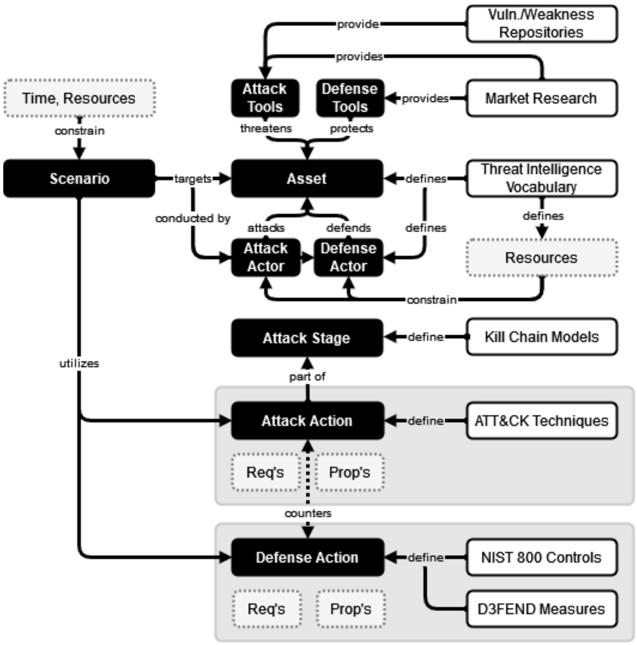


Fig. 1: PenQuest game model, simplified representation of components and data sources.

Our two game variants incorporate both approaches: MITRE D3FEND as well as NIST SP 800-53 is used as template for technical and more organizational actions that counter the various attack techniques. This teaches students which mitigating measures exist and how they can be used to protect IT systems from (intentional) threats.

IV. The Game

The original base of PenQuest is grounded in a security model [7] [8] that has seen major revisions since the initial publication. In this section, we focus on these updates and summarize the primary aspects of the game's formal foundation. Furthermore, we examine the actual mechanics translating the security concepts discussed in Section III into game elements. Various data sources and vocabularies used to provide an optimum of realism are listed as well. Finally, we examine the visual design concepts of the newly created digital app and provide a look at the graphical user interface.

A. Model

Formally, PenQuest models adversarial behavior as part of an asymmetric, non-cooperative two-player game on the basis of imperfect, incomplete information [37]. This is due to the fact that the actors use opposed strategies and are not necessarily aware of the other's actions. It is actually a key element of the game to optimize detection in order to gain insight into the opponent's activities. The payoff associated to certain actions is not always known, either, but familiarity generally improves as part of the learning process.

It is not straightforward to classify PenQuest as zero- or non-zero-sum. While the core mechanics of system compromise (see Subsection B) is zero-sum – the attacker's gain in damage points equals the defender's loss – other mechanics are more complex. This includes the gaining of insight, various positive and negative effects on certain assets, and the placement of exploits. Refer to [7] for a more detailed view on the formal model.

Practically speaking, the PenQuest model depicts scenarios that seek to compromise or protect certain assets

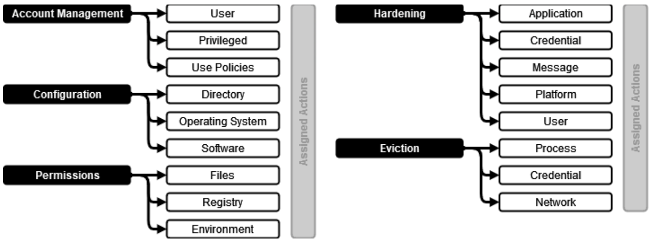


Fig. 2: Examples of the categories used for mapping attack to the defense actions. Left: Based on ATT&CK mitigations (46 total). Right: Based on D3FEND tactics and techniques (20 total).

through a set of attack and defense actions, which are associated to an attack stage. Tools utilized by both sides support the effort and can impact the assets as well. Several resources (e.g., time, financial means, skill bars, etc.) act as constraints for certain actions, tools, and the game scenario itself. Fig. 1 provides a high-level overview of a scenario's components and summarizes the data sources used. See also Subsection C for more information on the data-driven aspects of the game.

B. *Core Mechanics*

Each game of PenQuest follows a specific scenario that defines the attacker's and defender's (i.e., actors') goals. Scenarios are generally constrained by a timeframe in which the attacker must achieve their malicious goal in order to be victorious. This time limit is determined by an actor's *attributes*, namely Skill and Motivation, which we measure on a scale from 1 (low) to 5 (high). These attributes also contribute to how successful an actor is in attacking or defending and how likely it is that they remain undetected. Additional attributes include Wealth and Insight. The latter tracks how much knowledge an actor has gained about their opponent, which provides gameplay advantages in certain situations.

In the current version of PenQuest the aforementioned goals which characterize a scenario require an attacker to compromise an asset in a certain way. *Assets* can be construed as IT systems such as servers, workstations, hardware governing individual network segments, cloud services, or even mobile devices. The attacker attempts to compromise one or several of these assets while the defender seeks to prevent just that. Since cyber-attacks can have a wide range of consequences, we model the game's objectives and the incurred "damage" to systems as numeric values on a 'C-I-A' scale [28], which signifies the main security objectives mentioned in the previous section: theft, manipulation, and availability attacks. Any actions used by the defender either reduce that damage (response) or prevent it from happening in the first place.

To cause or mitigate damage both actors utilize *actions* that represent attack techniques or controls that are technical, organizational, or human in nature. The attacker's actions correspond to a so-called *attack stage* as defined by the cyber kill chain discussed in Section III. To streamline this concept for the game, we distinguish Reconnaissance, Initial Access, and Execution actions. It is not possible to use actions that require a stage that has not yet been unlocked for the respective asset. For example, you cannot make your attack persistent or move laterally to another asset (Execution stage) before gaining access to the system first (Initial Access).

Actions are supported by various *tools*, be that a vulnerability scanner looking for flaws in the configuration, malware, or a security appliance that scans network traffic for suspicious patterns. Tools typically increase or decrease various success chances and either provide permanent or temporary modifiers to certain actions. Some also "stick" to an asset until removed (e.g., exploits). Tools generally require monetary resources to procure, which are governed by the actors' Wealth.

Both tools and actions are constrained by attributes as well as a multitude of other factors; insufficient skill may prevent an attacker from using a sophisticated action, and certain equipment may only work in concert with an asset of the 'web server' type. Other prerequisites the game considers include administrative privileges required by some techniques, and the asset's operating system.

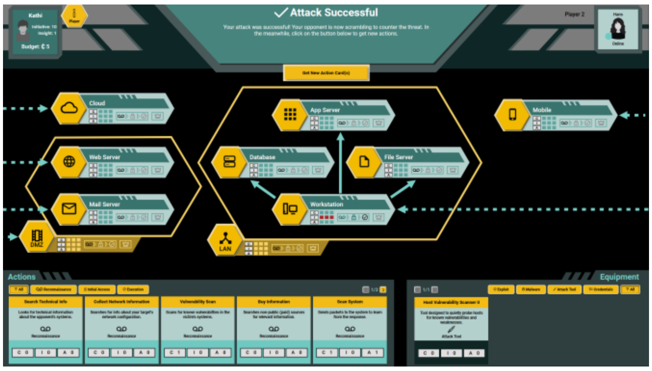
C. *Data Sources*

Under the hood, PenQuest translates version 8 of the MITRE ATT&CK framework [31] into game actions and pits these techniques against security and privacy controls for information systems and organizations derived from NIST SP 800-53 [36] as well as countermeasures listed in the MITRE D3FEND knowledge base [33], version 0.9.3. This results in two distinct game variants that focus on organizational and technical aspects, respectively.

For example, the ATT&CK-based action 'Phishing' may be prevented by organizational, NIST-based defense actions such as 'Security Awareness Training' (user training) or 'Trusted Path Enforcement', which defines a policy that disallows non-trusted communication channels such as certain external e-mail services. On the technical, D3FEND-based side, valid controls include 'Message Authentication' (basically the use of digital signatures for electronic messages) and 'Identifier Analysis' (i.e., checking URL strings for suspicious components). PenQuest's flexible nature makes it possible to easily focus on one or the other aspect – or to mix and match as desired.

A key element of the game is the mapping mechanism that connects the various attack and defense actions. Here, we have developed two alternative approaches: the first, primarily built for the aforementioned organizational variant of the game, is loosely based on ATT&CK's 'mitigations', which range from 'Account Use Policies' to 'Vulnerability Scanning'. Each attack action in the framework is already linked to such mitigations. To determine which category corresponds to a certain defense action we have translated it to the control families found in NIST SP 800-53 and further refined them. For example, the mitigations 'Privileged Account Management' (M1018) and 'Caution with Device Administrator Access' (M1007) in ATT&CK now translate to the NIST control family 'Account Management'. Each control within was then checked for references to administrative accounts and assigned the respective category number. During play, the game checks if at least one category assigned to both actions is identical to determine if it is a valid prevention our response measure. See Fig. 2 (left) for a number of example categories.

The second mapping approach takes a different perspective and expands not on adversary behavior but on MITRE D3FEND's defensive tactics and techniques, e.g., 'Credential Hardening' and 'Process Analysis', which are pre-assigned to the defense actions derived from that knowledge base. To establish a link to ATT&CK's techniques, we assigned the same categories to all of the game's attack



✓ Attack Successful

Your attack was successful! Your opponent is now scrambling to counter the threat. In the meanwhile, click on the button below to get new actions.

Search Technical Info: Looks for technical information about the opponent's systems.

Collect Network Information: Searches for info about your target's network configuration.

Vulnerability Scan: Scans for known vulnerabilities in the victim's systems.

Buy Information: Searches non-public (paid) sources for relevant information.

Scan System: Sends packets to the system to learn from the response.

Host Vulnerability Scanner II: Tool designed to quietly probe hosts for known vulnerabilities and weaknesses.

Fig. 3: PenQuest game board; attacker view after successfully compromising a user workstation.

actions, thereby defining which attacks can be mitigated by which existing control. See Fig. 2 (right) for examples and refer to Section V for an evaluation of the two approaches done by security experts.

There are additional data sources that helped us create an accurate representation of today's threat landscape. For example, we based the game's actors – i.e., designations of attackers and defenders that hint at their general motivation, such as "Hacktivist" or "Organized Crime" – on the STIX threat information language [38]. Vulnerabilities that can be exploited as part of the corresponding game mechanic are taken on a sample basis from the Common Vulnerability Enumeration (CVE) Database [32]. Other models such as cyber kill chains that influenced the game's design are introduced in Section III.

D. Visual Design

Creating the visual structure of PenQuest used an interdisciplinary, design-oriented research approach based on user (students as well as teachers) and stakeholder requirements (project owner, university). The most promising ideas were iteratively selected and turned into an interactive prototype, which was subsequently evaluated and improved based on feedback.

Specifically, we used greyscale mockups that were combined into an interactive prototype using Figma[1]. This helped to string together the attacker's and the defender's game turns and visualize dependencies between PenQuest's many stages. The issue of which information to present the user at a given time was addressed as well; since our educational game uses imperfect, incomplete information (see Subsection A), a workable compromise between realism and accessibility had to be developed.

This user-centered design process to creative problem-solving ensured optimal outcomes in terms of user experience and acceptance. Typically, this process comprises five phases: research, design, prototyping, evaluation, and implementation [39] [40]. For the design phase, an iterative design funnel [41] [42] was used. In the beginning, as many ideas as possible were generated. The focus here was on the quantity of the ideas and not their depth, i.e., their degree of implementation. This was in part accomplished by inviting students of a design master class to provide ideas for the digital version of the gameboard. As the project progressed, the ideas were reduced and elaborated in more detail until functional high-fidelity prototypes could be implemented. Within this process, the target audience is placed at the center of the design, which has proven to be an effective approach [43] [44].

Fig. 3 shows PenQuest's game board with its many interactive elements. After excessive testing, we opted for a cyan-and-yellow color scheme that highlights interactive and informational elements. The gameboard can be found in the vertical center of the screen with information elements on top and the player's actions and equipment below.

E. Terminology

While visual design is vital to motivate learners and better bring across teaching points, the simplification of technical terms is equally important [45]. Since the game is based on established vocabularies that typically require medium to high technical or managerial understanding on at least an IT Bachelor's level, we have conceptually reworded the names and descriptions of over 1,200 ATT&CK, D3FEND, and NIST-derived actions. The resulting data set is suitable for students of the fifth form (upper school) and above, as well as equivalent vocational education in the area of information technology. Further simplification is planned for future iterations; ultimately, we want to create additional game variants that aim at secondary school students.

V. Evaluation

A. Methodology

The efficacy and feasibility of employing PenQuest was evaluated using a three-pronged approach: educational effects and entertainment value were determined through a blinded study, where one group of students was given a conventional research task (Group "Research"; control) and the other was asked to play the game for 45 minutes or one match (Group "PenQuest"; test), whichever eventuated first.

For the topical research task, students were asked to research the NIST SP 800-53 control family "Incident response" [36], translate the technical terms to everyday language, and create a simple incident response plan for a fictitious company that included at least 5 processes or systems they would implement to meet the standard's requirements.

In the end, both groups were handed a test with questions regarding subject matter that was only implicitly touched on during their respective tasks – no verbatim content, terms or specific techniques were discussed. The goal was to measure the degree of reflection, general understanding, and knowledge gain in addition to a change in confidence with and without PenQuest in comparison to the research task.

Each block of multiple-choice knowledge questions was scored from -5 (all incorrect) to +5 (all correct), whereas confidence was rated from 1 (low) to 5 (high). Group "PenQuest" was also asked to score the game from 1 to 5 in respect to several categories ranging from learning curve and abstraction level to visual design and entertainment value. All resulting scores were compared using a two-sample t-test [46].

The following example shows one of the knowledge questions with all possible answers. Keep in mind that the referred principle was not explicitly mentioned in the game or the research exercise – the various aspects, however, were

[1] https://www.figma.com

indirectly referenced in both through e.g., game actions or incident handling controls listed in the NIST document. To answer this and other questions correctly required students to reflect on the matter and establish relevant connections on their own.

Question: “Which aspects does the principle of ‘defense in depth’ encompass?”

- Have users provide their username and password when using a system (correct)
- Buy computer equipment and software from different manufacturers (correct)
- Segment your network into different zones (correct)
- Ask users to change their password as often as possible (incorrect)
- Return detailed error messages to users to make bug-fixing easier (incorrect)
- Focus your defense efforts on securing the boundary between Internet and DMZ (incorrect)

To assess confidence, we asked students to grade statements like “I feel ready to plan the defense of a company against cyber-attacks” or “I consider the field of IT security easy to understand” on a scale of 1 to 5. No further proof of knowledge was required.

Part 2 of the evaluation focused on measuring the quality of the underlying model, in particular the two different versions of mapping algorithms linking attacks to defensive measures. Here, we used a Python script to randomly generate pairings from the hundreds of actions that are at the core of PenQuest (support-type actions that accompany other attacks were exempt) and asked security experts to grade their relationship, i.e., how well a response measure would help to restore a compromised asset or work towards the goal of evicting or isolating an attacker. Random samples of prevention and detection measures were similarly scored on a scale of 1 to 5. This enabled us to assess the game’s inherent realism and collect data to improve its mechanics.

Lastly, we conducted expert interviews and hands-on playtests with company and higher education representatives to collect feedback and evaluate the game’s accessibility and visual components. To this end, we presented participants with an 18-point questionnaire asking them to score from 1 to 5 the various gameplay and visual elements such as the process of selecting and executing an attack on a certain asset, or how comprehensible they consider the game’s kill chain tracker. A total of 5 categories were assessed: the game’s lobby, game board, information box, asset details screen, attack/defense window, and equipment shop. Each category was rated for overall layout & visual appeal, comprehensibility, and usability. Testers were also asked to provide verbal or written feedback explaining their score and were encouraged to suggest improvements to the game’s visual and gameplay elements.

Please note that there was no overlap of participants of the three parts of the evaluation. None of the expert interviews were conducted with in-house personnel or paper authors and no tester has had the opportunity to play PenQuest before. Communication to other participants was actively discouraged to prevent in-group bias.

B. *Results*

For part 1, initial results with a class of 14 pre-bachelor engineering (non-IT) students showed an average increase of 25.5% in terms of knowledge retention/reflection in comparison to students completing the conventional study exercise. The strongest result of 3.1 points on a -5..5 range was achieved for the question “What does an intrusion detection system do?” which sought to determine the understanding of the benefits of common security appliances. At 1.0 points, the “Research” group scored significantly lower. On the other hand, PenQuest players struggled (-0.7) with the question “Which of the 4 attacks require the least preparation by the attacker?” – here, students were asked to estimate required attacker efforts to complement risk assessment. The research task, which dealt with these aspects more explicitly, yielded slightly better results (0.4).

Interestingly, the students’ confidence in their own skills decreased by 6.6%. This is likely owed to the fact that PenQuest introduces intricate concepts that are otherwise not fully grasped and where a lack of understanding might make the subject matter appear less complex than it actually is [47]. Specifically, test group students claimed to best understand attacker motivation (4.0 out of 5) but felt overwhelmed by the prospect of planning an organization’s defense against cyber-attacks in general (1.9). Learners that did not play the game had a slightly higher confidence in their ability to take responsibility (Δ=0.28). The question with the most notable increase in terms of self-confidence thanks to PenQuest revolved around the understanding of IT system vulnerabilities: Here, players claimed to have a better understanding of the concept (Δ=0.57).

In terms of suitability to education, usability, and entertainment value the test group awarded PenQuest 3.9 points out of 5, with the strongest aspects being an increase in security awareness as well as straightforward fun (4.4 out of 5 each), and the lowest score pertaining to the game’s steep learning curve (3.3). Refer to Table I for all mean (x̅) and median (x̃) values in the categories ‘knowledge’, ‘confidence’, and ‘game score’.

TABLE I. LEARNING EVALUATION (PART 1)

Categories	**Educational effects, blinded study (n=14)**				
	PenQuest		***Research***		**Δ**[d]
	x̅	x̃	x̅	x̃	
Knowledge[a]	1.38	1.36	1.10	1.14	+0.28 (+25.5%)
Confidence[b]	3.24	3.43	3.47	3.57	-0.23 (-6.6%)
Game[c]	3.89	3.71	-	-	-

[a] Knowledge scored from -5 (all incorrect) to +5 (all correct), 6 multiple-choice questions

[b] Confidence in own skills scored from 1 (low) to 5 (high), 10 questions

[c] Game scored from 1 (low) to 5 (high), 14 questions

[d] Observed change in percent when employing PenQuest (references average)

Overall, we achieved a net promoter score (NPS) of 57%, which is computed as follows:

$$(n_p + n_d) \,/\, n = \text{NPS} \qquad (1)$$

…where the n_p is the number of promoters that awarded scores of 9 and 10 (out of 10) points and n_d are detractors that gave a score of 6 points or lower.

TABLE II. MAPPING ACCURACY EVALUATION (PART 2)

Algorithm		Random action pair scoring (n=304)				
		x	x_{min}	x_{max}	$\bar{x}$	$\tilde{x}$
ATT&CK vs. NIST[a]	Prevention	113	16	44	3.16	3
	Response	20	0	25	4.05	4
	Overall	133	16	69	3.29	4
ATT&CK vs. D3FEND[b]	Prevention	118	16	23	3.47	4
	Response	53	5	8	3.81	4
	Overall	171	21	31	3.57	4

[a] Random selection from 106 attack and 154 defense actions

[b] Random selection from 140 attack and 108 defense actions

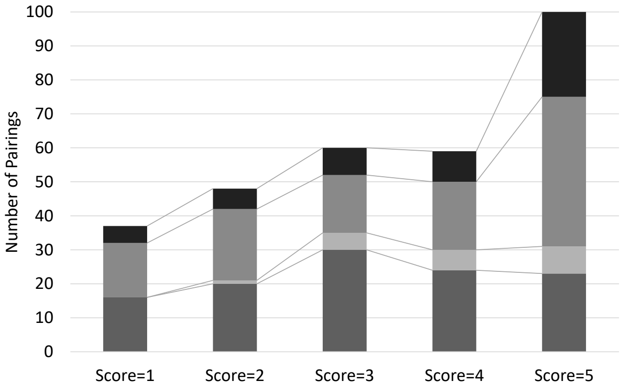


Fig. 4: Score distribution of the evaluated pairings. From top to bottom, the categories are: ATT&CK ↔ D3FEND response, prevention, ATT&CK ↔ NIST response, prevention.

In part 2 of the evaluation, we asked 6 security experts to score attack and defense action pairings. At an average of 3.57 out of 5 points, the algorithm linking MITRE ATT&CK and D3FEND actions was generally perceived as more accurate. This is opposed by a mean score of 3.29 for ATT&CK ↔ NIST actions. Apart from NIST-based prevention measures, all awarded median scores equated to 4. See Table II for more information on mean and median values as well as pairings with the highest (x_{max}) and lowest scores (x_{min}).

As depicted in Fig. 4, close to a third of the evaluated sample set of attack-defense pairs were given a perfect 5 out of 5 score. Approximately 11% of the pairings were seen as entirely inappropriate countermeasures with a score of 1. Upon assessing the lower end scores, we did not find a systemic flaw in the mapping algorithm but rather individual deficiencies in need of further finetuning. For example, some NIST-derived defense actions like 'OPSEC' and 'Encrypt Data' were simply too generic in their description to be seen as appropriate controls for certain technical attacks. At other times, defense measures were too detailed in their description and addressed an issue that was not specifically mentioned in the attack, even though it would constitute a valid adversarial approach.

Distinctive to the ATT&CK ↔ D3FEND algorithm, the 'decoy' and 'analysis' categories of defense measures were scored lowest at only 3.17 and 3.19 points, respectively. Here, including further distinction between the affected entities is likely to improve the mapping in the eyes of the users. For example, 'files' and 'e-mails' were so far considered to be synonymous, which was not well received.

TABLE III. USABILITY EVALUATION (PART 3)

Categories	User interface questionnaire (n=9)			
	$\bar{x}$	$\tilde{x}$	min	max
Lobby[a]	4.42	4.50	3	5
Game board[b]	4.18	4.00	2	5
Infobox[c]	4.46	5.00	3	5
Asset details[d]	4.24	4.00	3	5
Attack window[e]	4.04	4.50	2	5
Shop[f]	4.36	4.25	4	5
Overall	4.26	4.25	2	5

[a] Game lobby & creation screen, game objectives, 5 questions

[b] Game board elements, 5 questions

[c] Hints and information box, 4 questions

[d] Asset details screen presenting additional information, 2 questions

[e] Attack/defense window used to play actions, 5 questions

[f] Shopping overlay for procuring equipment, 2 questions

Expert interviews conducted in part 3 of the evaluation largely mirrored the student's assessment of the game and highlighted its visual appeal. Several company representatives expressed their interest to use PenQuest as part of their security awareness programs while providing invaluable development feedback.

Taking a closer look at the results we saw the highest average score of 4.7 out of 5 awarded to the overall gameboard and information box layout. The comprehensibility of attack detection information and the overall shopping process were scored highly as well (4.6 and 4.5 points, respectively). Testers saw most room for improvement in the presentation of error messages (3.7) and the damage tracker visualizing an asset's compromise level (3.9). Category-wise, the info-box was received best (4.4) while the attack window received the lowest average score (4.0). The categories 'game board' and 'attack window' saw the biggest discrepancy in tester scores. Refer to Table II for an overview.

Individual testers provided feedback on which game elements should be better highlighted while others wished for more detailed information. Numerous features were implemented in response to this feedback, including but not limited to an event log tracking all actions played, additional visual elements, and tooltips.

C. Discussion

We have seen that PenQuest helps students to reflect on learned content and that certain aspects and concepts of IT security were indeed better understood after playing the game. Nevertheless, future iterations will have to consider that some players may be intimidated by the complexity of the game and, by extension, the subject matter. We have seen this most often with people unfamiliar with (offline) strategy games or who exhibit a shorter attention span.

While the complexity of PenQuest could motivate many to delve deeper into the topic, the opposite is possible as well. To lessen the risk of discouraging students, a lecture's curriculum would have to be adapted to accommodate PenQuest and give students unfamiliar with the game ample time to discuss rules, individual in-game situations, and session outcomes. Only more experienced users should play unattended. Furthermore, we recommend to slowly expand the game's available action set to stay in sync with current lectures. Successfully completing a class or exercise could award new "cards" for use in PenQuest matches, thereby helping learners reinforce knowledge gained without overwhelming them.

It stands to mention that most pre-bachelor engineering students needed between 15 and 30 minutes to get the gist of the game's rules; mastery took significantly longer and was directly linked to a player's IT security knowledge and strategy game experience. Expert-level users were able to play the game within minutes and required little explanation.

In regard to realism, the second part of the evaluation has shown us that linking attacks to countermeasures using mapping functions based on categorization is largely feasible but will require expert review and occasional manual correction. Future work will explore crowdsourcing the process to build a comprehensive knowledge graph that governs mappings and can even be used independently of PenQuest.

VI. Conclusion

Despite its beta stage, experiments have shown that PenQuest can improve security (awareness) education through its gamified approach when presenting a complex technical topic to an interested audience with a basal technical aptitude. Thanks to our model's flexibility we are able to apply the developed mechanics to a wide range of scenarios, including – but not limited to – IT system attacks targeting an abstracted network topology, web application threats (e.g., OWASP Top Ten [48]), industrial systems, and physical security.

Future research will add new action vocabularies and encompass further risk assessment applications including strategy optimization through reinforcement learning as well as model checking, which will allow students to play against an intelligent AI opponent and enable security practitioners to use the game to identify the most severe threats to their own assets. Depending on the use case, this will enable PenQuest users to focus specifically on factors such as probability of success, detectability, impact, or cost.

Acknowledgments

The financial support of the Austrian Science Fund (FWF), Digital Innovation Hub OST, as well as St. Pölten University of Applied Sciences is gratefully acknowledged.

References

[1] J. A. Lewis, "Economic Impact of Cybercrime," Center for Strategic and International Studies, 2018.

[2] J. Van den Berg, J. Van Zoggel, M. Snels, M. Van Leeuwen, S. Boecke, L. Van de Koppen, J. Van der Lubbe, B. Van den Berg and T. De Bos, "On (the Emergence of) Cyber Security Science and its Challenges for Cyber Security Education," NATO STO/IST, 2014.

[3] B. S. Bloom, Taxonomy of Educational Objectives: Vol. 1: Cognitive Domain, New York: McKay, 1956.

[4] J. Locke, Some Thoughts concerning Education, London: Black Swan, 1693.

[5] I. Caponetto, J. Earp and M. Ott, "Gamification and Education: A Literature Review," in *European Conference on Games Based Learning*, 2014.

[6] Federal Office for Information Security, "Durchführung von Planspielen zur Informationssicherheit," BSI, 2014.

[7] R. Luh, M. Temper, S. Tjoa and S. Schrittwieser, "PenQuest: A Gamified Attacker/Defender Meta Model for Cyber Security Assessment and Education," *Journal of Computer Virology and Hacking Techniques*, 2020.

[8] R. Luh, H. Janicke and S. Schrittwieser, "AIDIS: Detecting and Classifying Anomalous Behavior in Ubiquitous Kernel Processes," *Computers & Security*, vol. 84, pp. 120-147, 2019.

[9] A. Shostack, "Elevation of Privilege: Drawing Developers into Threat Modeling," in *USENIX Summit on Gaming, Games, and Gamification in Security Education*, 2014.

[10] L. Kohnfelder and P. Garg, "The Threats to our Products," Microsoft Corporation, 1999.

[11] S. Hart, A. Margheri, F. Paci and V. Sassone, "Riskio: A Serious Game for Cyber Security Awareness and Education," *Computers & Security,* vol. 95, 2020.

[12] A. Rieb and U. Lechner, "Operation Digital Chameleon: Towards an Open Cybersecurity Method," in *Proceedings of the 12th International Symposium on Open Collaboration*, 2016.

[13] Black Hills Information Security, "Backdoors & Breaches," [Online]. Available: https://www.blackhillsinfosec.com/projects/backdoorsandbreaches/. [Accessed 17 11 2021].

[14] OWASP Foundation, "OWASP Cornucopia," 2020. [Online]. Available: https://owasp.org/www-project-cornucopia/. [Accessed 17 11 2021].

[15] Texas A&M University, "Keep Tradition Secure," Division of Information Technology, [Online]. Available: https://keeptraditionsecure.tamu.edu/. [Accessed 17 11 2021].

[16] Defense Information Systems Agency, "Cyber Awareness Challenge 2022," 2021. [Online]. Available: https://public.cyber.mil/training/cyber-awareness-challenge/. [Accessed 17 11 2021].

[17] Trend Micro, "Targeted Attack: The Game," 2015. [Online]. Available: http://targetedattacks.trendmicro.com/. [Accessed 17 11 2021].

[18] IS Decisions, "The Weakest Link," [Online]. Available: https://www.isdecisions.com/user-security-awareness-game/. [Accessed 17 11 2021].

[19] LivingSecurity, "Hotspot," [Online]. Available: https://hotspot.livingsecurity.com/. [Accessed 17 11 2021].

[20] LivingSecurity, "Living Security Teams: CyberEscape Online," [Online]. Available: https://www.livingsecurity.com/cyberescape-online. [Accessed 17 11 2021].

[21] LivingSecurity, "Craft the Phish," [Online]. Available: https://phishing.livingsecurity.com/. [Accessed 17 11 2021].

[22] Z. A. Wen, Y. Li, R. Wade, J. Huang and A. Wang, "What.Hack: Learn Phishing Email Defence the Fun Way," in *CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017.

[23] Google, "Jigsaw Phishing Quiz," [Online]. Available: https://phishingquiz.withgoogle.com/. [Accessed 17 11 2021].

[24] A. Shostack, "Tabletop Security Games & Cards," 2021. [Online]. Available: https://adam.shostack.org/games.html. [Accessed 17 11 2021].

[25] M. Stytz, "Considering Defense in Depth for Software Applications," *IEEE Security & Privacy,* vol. 2, no. 1, pp. 72-75, 2004.

[26] T. R. Peltier, Information Security Risk Analysis, Second Edition, CRC Press, 2005.

[27] C. B. Simmons, S. G. Shiva, H. Bedi and D. Dasgupta, "AVOIDIT: A Cyber Attack Taxonomy," in *Symposium on Information Assurance*, 2014.

[28] D. Bell and L. La Padula, "Secure Computer System: Unified Exposition and MULTICS Interpretation," MITRE Corporation, 1976.

[29] J. W. Greenert, "Kill Chain Approach," Chief of Naval Operations, 2013.

[30] E. Hutchins, M. Cloppert and R. Amin, "Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion KIll Chains," in *Leading Issues in Information Warfare and Security Research*, Good News DIgital Books, 2021.

[31] MITRE Corporation, "MITRE ATT&CK Framework," 2021.

[32] MITRE Corporation, "CVE," 2021. [Online]. Available: https://www.cve.org/. [Accessed 1 12 2021].

[33] P. E. Kaloroumakis and M. J. Smith, "Toward a Knowledge Graph of Cybersecurity Countermeasures," 2021.

[34] ISO/IEC Information Technology Task Force, "ISO/IEC 27000," 2018.

[35] Common Criteria Recognition Arrangement, "Common Criteria," 2021. [Online]. Available: https://www.commoncriteriaportal.org/. [Accessed 1 12 2021].

[36] NIST Information Security Laboratory, "Security and Privacy Controls for Information Systems and Organizations, SP 800-53 Rev. 5," 2020.

[37] R. Sankardas, E. Charles, S. Sajjan, D. Dipankar, S. Vivek and W. Qishi, "A Survey of Game Theory as Applied to Network Security," in *Hawaii International Conference on System Sciences*, 2010.

[38] OASIS Open, "Cyber Threat Intelligence Technical Committee," 2021. [Online]. Available: https://oasis-open.github.io/cti-documentation/. [Accessed 1 12 2021].

[39] D. A. Norman, The Design of Everyday Things, Basic Books, 2002.

[40] C. Bowles, "A List Apart," 2013. [Online]. Available: https://alistapart.com/column/looking-beyond-user-centered-design/. [Accessed 1 12 2021].

[41] B. Buxton, Sketching User Experiences, Morgan Kaufmann, 2007.

[42] S. Pugh, Total Design: Integrated Methods for Successful Product Engineering, Addison-Wesley, 1991.

[43] K. Goodwin, Designing for the Digital Age: How to Create Human-Centered Products and Services, John Wiley & Sons, 2009.

[44] H. Sharp, J. Preece and Y. Rogers, Interaction Design: Beyond Human-Computer Interaction, Wiley, 2002.

[45] S. Vicente, J. Orrantia and L. Verschaffel, "Influence of Situational and Conceptual Rewording on Word Problem Solving," *British Journal of Educational Psychology,* vol. 77, no. 4, pp. 829-848, 2010.

[46] R. Wilcox, Statistics for Social Sciences, Academic Press Inc., 1996.

[47] J. Kruger and D. Dunning, "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments," *Journal of Personality and Social Psychology,* 1999.

[48] Open Web Application Security Project, "OWASP Top Ten," 2021.