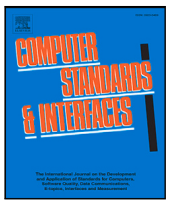
Computer Standards & Interfaces 92 (2025) 103924

Contents lists available at ScienceDirect

Computer Standards & Interfaces

journal homepage: www.elsevier.com/locate/csi



Evaluating and validating the Serious Slow Game Jam methodology as a mechanism for co-designing serious games to improve understanding of cybersecurity for different demographics

Shenando Stals [a,*], Lynne Baillie [a], Ryan Shah [a], Jamie Iona Ferguson [b], Manuel Maarek [a]

[a] School of Mathematical and Computer Sciences, Heriot-Watt University, Earl Mountbatten Building, Edinburgh, EH14 4AS, Midlothian, United Kingdom
[b] School of Simulation and Visualisation, Glasgow School of Art, The Hub, Pacific Quay, Glasgow, G51 1EA, Lanarkshire, United Kingdom

ARTICLE INFO

Keywords:
Cybersecurity
Serious Slow Game Jam
Serious games
Evaluation
Human-computer interaction
Secure code citizens

ABSTRACT

We present an evaluation of a Serious Slow Game Jam (SSGJ) methodology as a mechanism for co-designing serious games in the application domain of cybersecurity, to evaluate how the SSGJ methodology contributed to improving the understanding of cybersecurity for different demographics. The aim of this study was to evaluate how the SSGJ contributed to improving the understanding of cybersecurity for young persons between the ages of 11 and 16 years old who had no formal training or education in cybersecurity, and to validate and compare these results to previous work where the same SSGJ methodology was used with a different target demographic (i.e.,M.Sc. students with no formal training or education in secure coding). To this end, we engaged 23 participants between the ages of 11 and 16 years old for 5 consecutive days over a one-week period, in a multidisciplinary SSGJ involving domain-specific, pedagogical, and game design knowledge, and encouraged engagement in-between scheduled events of the SSGJ. Findings show improved confidence of participants in their knowledge of cybersecurity, for both demographics, after undertaking the Serious Slow Game Jam (from 41.2% to 76.5% for young persons, and from 12.5% to 62.5% for M.Sc. students). Free-text answers specifically indicate an improved understanding of cybersecurity in general, and one specific security vulnerability, attack or defence for a quarter of young persons, and the trichotomy of security vulnerabilities, attacks, and defences for three quarters of the M.Sc. students. Also, confidence in knowledge of game design improved for both demographics (from 47.1% to 82.4% for young persons and from 12.5% to 75% for M.Sc. students). The SSGJ methodology also successfully engaged both demographics of participants in-between scheduled days. Finally, two serious games in the application domain of cybersecurity are presented that were co-designed during the SSGJ with participants and produced as an output of the SSGJs.

1. Introduction

Within the software industry, there exists a subset of individuals who are code-literate and are able to build and deploy software code. However, they often do not have any formal training or education in software engineering or secure coding and may not adhere to best practices. Because of this, it is important that they understand the implications of not adhering to best practices and deploying insecure code [1]. For example, deployment of insecure code can come from code they have written themselves due to lack of formal training, or the reuse of insecure code snippets which have been shown to make up part of over a million published Android applications [2,3]. Various cybersecurity best-practices are employed in organizations, such as password-security policies, and have shown some success in mitigating the potential liabilities that may arise as a result of cybersecurity failures [4,5]. However, given the identification of the majority of cybersecurity failures (e.g., data breaches) occurring as a result of human error [6,7], one key question that remains concerns whether users who follow these practices truly understand the impact of not adhering to them in a constantly evolving threat landscape.

The use of game-based approaches have shown to be successful for training and education of cybersecurity and software engineering [8–11], however many of these are oriented around entertainment as their primary purpose, with some also requiring varying levels of prior knowledge of cybersecurity topics [12,13]. Serious games are defined as those "that do not have entertainment, enjoyment and fun as their primary purpose" [14] and bespoke serious games have been shown to be a promising approach for learning and training for domain-specific knowledge [15–17]. For designing serious games, serious game

* Corresponding author.
E-mail address: S.Stals@hw.ac.uk (S. Stals).

https://doi.org/10.1016/j.csi.2024.103924
Received 31 May 2024; Received in revised form 5 August 2024; Accepted 2 September 2024
Available online 11 September 2024

jams have been recommended for facilitating serious game design and research [18,19]. In addition to this, serious game design also needs to be well understood with regard to application domain and pedagogical objectives. Unfortunately, the conventional game jam format has shown to be less adequate for serious game design and research [20] due to the fast-paced nature [21,22] and little time for refinement and/or reflection [23].

Prior work has shown that the use of a Serious Slow Game Jam (SSGJ) methodology as a mechanism for co-designing serious games is very successful in providing support and mentorship from domain experts throughout [20,24]. Due to their inherent characteristics of being multi-disciplinary, involving domain-specific, pedagogical and game design knowledge, SSGJs are an ideal candidate for the design of serious games. Furthermore, results have shown that confidence in domain-specific knowledge for cybersecurity and game design increased significantly, engaging participants throughout the entire game jam, including in-between scheduled days.

Despite successful results, a potential limitation of the work in [25] relates to demographics. The participants recruited in that study were masters-level conversion students in computer science and engineering (but not cybersecurity), with most participants describing themselves as having little cybersecurity or secure coding expertise. But they would have had some familiarity with software development at some stage in higher education. However, the target group to create secure code citizens is wider, looking at people with different levels of coding expertise, different age ranges, different genders, and different professional and educational backgrounds. In addition, the way participants experience game jams and gain knowledge through participation in game jams has been shown to differ depending on their age [22], educational background [26], technical knowledge [22,27], and prior game jam experience [22,28].

In this study, we aimed to contribute to further investigations of the SSGJ methodology by exploring a different demographic. Specifically, we recruited participants with ages ranging from 11 to 16 years who have mixed coding experience and limited experience with game jams. We found that the SSGJ methodology was successful in engaging these young persons in the co-design of serious games to improve their understanding of cybersecurity. Furthermore, the participants in the workshop contributed to the output of a serious game, No-Entry, which aims to provide an easy and enjoyable way for players to get a better understanding of (the breadth of) cybersecurity in general, and the trichotomy of vulnerabilities, attacks, and defences in particular.

1.1. Research questions

The overall aim of the research presented in this paper is to evaluate how our SSGJ methodology contributed to improving the understanding of cybersecurity for this new demographic, and present a serious game in the application domain of cybersecurity that was co-designed with young persons between the ages of 11 and 16 years old, and was produced as an output of this SSGJ. It investigates how different aspects of the SSGJ may have contributed to this goal. To investigate this, we were guided by the following research questions:

- RQ1: How has the SSGJ affected young participants' understanding of cybersecurity?
- RQ2: How can the cards for the application domain (in our case the Cybersecurity cards), Learning Mechanics cards, and Game Mechanics cards that are part of the SSGJ toolkit, assist young participants in serious game design?
- RQ3: What are the workload and motivation levels of young participants during the SSGJ?
- RQ4: How has the "slow" format of the SSGJ affected engagement of young participants?

The remainder of this paper is organized as follows. Section 2 provides background knowledge regarding game jams, serious game design, the rationale for the proposed SSGJ methodology. Section 3 presents the SSGJ methodology and the procedure for evaluating a SSGJ we conducted in the application domain of cybersecurity. The results of this evaluation are presented in Section 4, and the serious games produced as an output of the SSGJ are presented in Section 5. The findings are discussed in Section 6, and the conclusions and future work are presented in Section 7.

2. Background

2.1. Serious games

Serious games are designed in a manner to promote education, training or promote behavioural changes in those who play them, by combining the traditional elements of games with pedagogical content to improve engagement and interactivity of learning experiences. They allow learners to experience real-world situations that may otherwise be difficult to experience due to reasons such as economic cost, safety or time. Further, serious game design is inherently multidisciplinary as it encapsulates domain-specific and pedagogical knowledge, as well as knowledge of game design. In serious game design, the gameplay design needs to be understood with regards to the application domain and pedagogical objectives, conceptualized in Triadic Game Design (TGD) [29] as Reality (i.e., the application domain), Meaning (i.e., pedagogical value), and Play (i.e., gameplay). Therefore, serious games design places much more emphasis on mapping these three aspects for effective learning outcomes [25].

Serious games have seen an increase in utilization in a wide array of domains, including education, business and healthcare. The primary goals of applying serious game-based approaches in these domains stems from the benefits to training, motivation and education in domain-specific areas. In the domain of cybersecurity, several game-based approaches have been developed, ranging from those that raise awareness about threats to providing training and skills development for defending against cyber attacks in real-world scenarios. Examples include Darknet [30] — a Virtual Reality (VR) serious game that simulates hacking and cyber espionage in the real world — and Cyber Threat Defender (CTD) [31] — a serious collectible card game developed by the Center for Infrastructure Assurance and Security to teach users how to defend against cyber threats and practice security skills in a gamified environment.

2.2. Serious game jams

Traditional game jams are inherently fast-paced and aim to create prototypes with design and time constraints (usually between 24 and 48 h). In the context of education and learning, game jams can provide an opportunity for creative thinking and project management practice in a short timeframe, which can supplement formal educational practices. Because of this, game jams have been shown to improve development practices employed by participants, as well as the impact on soft- and hard-skills and developer practices.

Most game jams are aimed at creating entertainment prototypes. However, the traditional, conventional entertainment-oriented game jam format is not best suited for the needs of serious game design and serious game design research [20]. The use of serious game jams has been recommended for serious game design and research, such as the use of academic game jams. It has been noted that the conventional entertainment focus that resides with the outputs of conventional game jams possess characteristics that may be less adequate for serious game design. Specifically, a key trait of serious game design is multidisciplinarity, reflected by TGD [29,36]. The mapping of the competencies of domain-specific, pedagogical and game design knowledge requires a refocus on the game jam lifecycle, where the emphasis is now placed

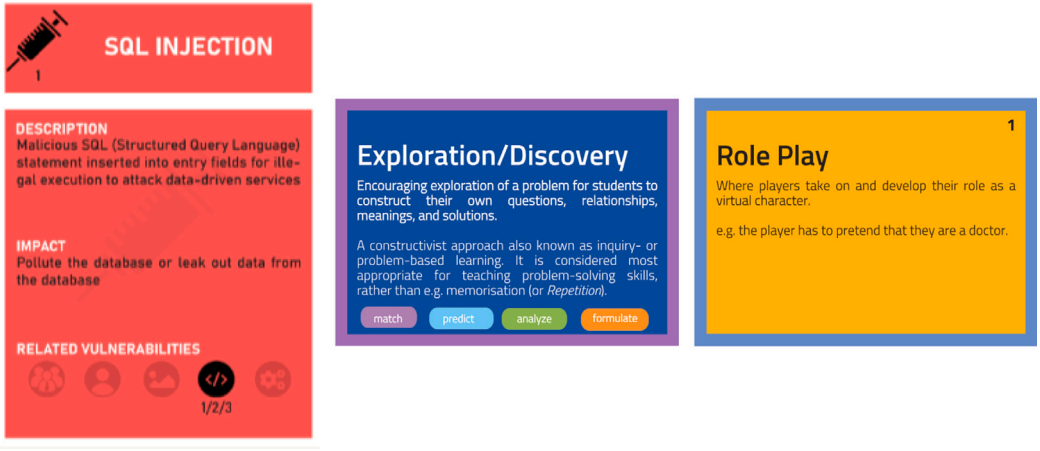


Fig. 1. Examples of the cybersecurity cards (left) by [32] based on CyBOK [33], and LM cards (including verbs and color coding from Bloom's Taxonomy [34]) (middle) and GM cards (right) both based on the LM-GM framework by [35].

on the serious outcome of the prototype game output, as opposed to polishing the prototype. As well as this, conventional game jams are typically time- and skill-constrained, which is a barrier for incorporating inclusivity of multi-disciplinary knowledge. Their generally fast and intensive pace can make them inaccessible [21,22,37,38] and leaves little room for participants to refine or reflect on their work [23]. Furthermore, given the collaborative and multi-disciplinary nature of serious game jams, those involved in the game jam are likely to be different compared to those in conventional, entertainment-oriented game jams and presents a challenge that is required to be addressed by serious game jams.

2.3. Serious Slow Game Jam (SSGJ) methodology

To address the challenges above in relation to serious game design parameters, Abbott et al. [23] propose the SSGJ methodology. It provides a multidisciplinary collaborative framework for serious game design, putting participants and experts at the centre of the design. It provides mentorship by application domain and game design experts to support participants, to support the value and validity of outputs, and to provide a structured, accessible, and educational experience. Although the SSGJ method is applied here to the application domain of cybersecurity, it is intended to be flexible and generic so that it can be used irrespective of application domain [23].

Game jams can have diverse formats depending on their aims and contexts. However, an analysis of the literature on game jams identified several shared design parameters [28,39–41]. These are the theme of the game jam, time constraints (i.e., typically ranging from 8–72 h), the location (i.e., physical, online, or hybrid), participation and team requirements (i.e., prior experience and skills), technology use (i.e., technology-agnostic or dedicated platforms for games production), participant support (i.e., keynote talks, workshops, presentations, mentoring), and deliverables (i.e., game prototypes, supporting multimedia, and documentation) [23,25,28,39–41].

The design of the SSGJ methodology is presented in terms of these game jam design parameters [23]. The **Theme** is guided by domain expert mentors and structured educational materials. These are included in the SSGJ toolkit to facilitate the SSGJ (see the Support parameter) [23]. The cybersecurity theme for that specific SSGJ, was that of secure software development lifecycles [25].

For the **Time** design parameter, emphasis is on accessibility and inclusion, aiming for a non-crunch working environment with non-exhausting session durations for each day of the SSGJ [23,38]. Based on lessons learned from previous serious game jams [20], the SSGJ is structured into three phases in the serious game creation lifecycle (i.e., design, development, and pre-release), consisting of two work days each, resulting in six days in total. The first phase consists of an introduction to the application domain and TGD. During this introduction phase presentations by domain experts, a deck of Cybersecurity, Learning Mechanics (LM), and Game Mechanics (GM) cards (describing concepts within each topic, and are used throughout all days of the SSGJ), and a small provoking game are introduced to kick off discussions about the application domain. Phase two covers the design of the serious game loop and prototype design, and phase three covers the development of the serious game prototype and other deliverables. There is sufficient time in between each of these phases for reflection, feedback, and refinement (2 weeks), resulting in an overall duration for the SSGJ of 5 weeks [23,25].

For the **Location** parameter, due to the timing of this SSGJ which took place in spring 2022, the COVID-19 pandemic at the time, and the diversity of the target group, the SSGJ was run in a hybrid format, where participants may be in-person or synchronously online to prioritize inclusivity and accessibility [23].

For **Participation and Teams**, recommendations from the literature [42] were followed, with the organizers of the SSGJ creating teams based on self-identified roles collected during participant registration. Where serious game research is an intended outcome, explicit inclusion of both domain experts and serious game designers as participant-mentors allows for the delivery of high-quality support materials (see [20,43,44]), guidance in framing the SSGJ theme [44], supporting and contextualizing domain related material, and validating its inclusion in the serious game [23,25]. Application domain experts in game design and cybersecurity rove between teams, to enhance contact and knowledge exchange between experts and participants [20,44].

In addition to in-person communication and collaboration, **Technology** in the form of Discord [45] was used for online communication and Miro [46] with structured activity worksheets for online collaboration. Due to the wide and diverse skill set of the target audience, there were no limitations imposed for game platforms [23,25].

Regarding the **Support** of participants, the SSGJ methodology includes guided educational group activities, supported by physical and digital materials in the SSGJ toolkit [23]. The TGD method [29] is used to inform and guide participants with respect to serious game design. It is proposed this will result in strong learning outcomes for participants as well as serious game prototypes that have high rigour and domain validity [23,25]. The SSGJ toolkit to support the SSGJ includes presentations of domain experts, a provocative game [47], Miro boards [46], and three decks of cards: Cybersecurity [32] (i.e., for the application domain), LM, and GM (see Fig. 1) [25].

The Cybersecurity cards [32] have been developed based on the Cyber Security Body Of Knowledge (CyBOK) [33], a comprehensive body of knowledge to inform and underpin educational and professional

training for the cybersecurity sector [33], while the LM and GM cards were created based on the LM-GM framework [35] to support the TGD method to inform and guide participants with respect to serious game design. The Cybersecurity cards consist of general and more detailed Attack cards, Defence cards, and Vulnerability cards, which are related but not necessarily in a one-to-one relationship. As our focus is on the domain of software engineering and secure coding, the knowledge areas of CyBOK selected for the design of the cybersecurity cards are Human Factors, Malware and Attack Technologies, Software security, Web and Mobile Security, and the Secure Software Lifecycle [32]. The full decks of Cybersecurity cards, LM cards, and GM cards, which are part of the SSGJ toolkit, are freely available from our Secrious project website[1] [25].

Informed by recommendations from the literature [28], **Deliverables** for each team include a serious game prototype, and a Serious Game Design Document (SGDD) which lays out the serious game design according to a provided template, which all SSGJ activities feed into [23,25].

2.4. Results of the first SSGJ with MSc students

This SSGJ methodology has previously successfully been used as a mechanism to co-design serious games to improve the understanding of cybersecurity among MSc conversion students in Computer Science (but not cybersecurity) between the ages of 22–35 years old [25]. Findings showed improved confidence of participants in their knowledge of cybersecurity (from 12.5% to 62.5%, $Z = -2.041$, $p = 0.041$) after undertaking the SSGJ, with free-text answers specifically indicating an improved understanding in terms of vulnerabilities, attacks, and defences for three quarters of the participants. Also, confidence in knowledge of (serious) game design improved (from 12.5% to 75%, $Z = -2.112$, $p = .035$) [25].

In comparison to traditional, fast-paced game jams, which have a high workload and temporal demand [21,22,37], the SSGJ successfully reduced the time pressure. Workload and temporal demand, measured using the NASA-TLX [25,48,49], was never considered very high (18–21) by those participants. Furthermore temporal demand was only considered high (11–17) for 5 out of 20 activities, namely for the Serious Game Loop Design (11.8) and during development activities (11.9–12.7) [25].

Motivation levels, measured using the IMI [25,50,51], indicated the SSGJ managed to engage participants in software security concepts. Average levels of the subscale Interest/Enjoyment (5.00–5.60), and Perceived Value/Usefulness (5.40–6.17) were positive for all days and phases of the SSGJ. Interest/Enjoyment and Perceived Value/Usefulness was even very high for three out of six days of the SSGJ [25].

The SSGJ also successfully engaged participants in-between scheduled days. The "slow" format encouraged engagement in-between scheduled days of the SSGJ. All participants (100%) engaged, by actively creating content for their serious game (86% of participants), reflecting on things learned during the SSGJ (29%), and conducting further research in cybersecurity (57%), learning context (43%), or games (14%) [25].

3. Methodology

In this work, we carry out a further investigation into the use of SSGJ as a methodology for co-designing serious games in cybersecurity, with a new demographic of young persons between the ages of 11 and 16 years old. The SSGJ methodology was designed such that it remains flexible and generic [23], irrespective of the application domain and design, so long as the core values of the framework are still adhered to (i.e., supporting the value and validity of outputs). We adopt the SSGJ methodology proposed by Abbott et al. [23], an applied game jam framework which involves four aspects: the problem space, game jam design, delivery and outcomes, and follow-on opportunities.

While the design and implementation of the methodology we follow is the same as that proposed by Abbott et al. [23] and evaluated by Stals et al. [25], some design parameters in the game jam design are slightly adjusted in this study to make the implementation of the SSGJ more suitable for the target demographic. This will be discussed in the next Section 3.1.

3.1. SSGJ design

The shared game design parameters used in this study that remain unchanged are the **Theme** (which remained cybersecurity, but with a focus on code security), **Participation and Teams**, and **Support** [23,25].

The only specific parameters that have been slightly adapted to suit the target demographic are: **Time**, **Location** and **Technology**:

- **Time**: The emphasis is still placed on accessibility and inclusion, with the primary aim being to alleviate time-criticality and exhaustion during SSGJ sessions. The SSGJ conducted by Abbott et al. [23] and Stals et al. [25] was split into three phases consisting of two workdays each (6 total) spread over 5 weeks. In this study, the SSGJ structure still consists of three phases, but is split over 5 consecutive days in a single week to suit the availability of the target demographic (see Table 1). The first phase consists of an introduction to the cybersecurity application domain and an introduction to TGD. Participants then played a small provoking game and used the cybersecurity cards to identify cybersecurity problems. The second phase covers the design of the serious game loop and prototype design. Phase three covers further development of the serious game prototype and other deliverables. In contrast to the previous SSGJ, there is less time in-between phases for reflection, feedback and refinement. Each day consisted of a morning and afternoon session, with a 15 min break within each session, separated by a 45 min lunch break.
- **Location**: The previous SSGJ was run in a hybrid format such that the participants could be in-person or online to prioritize the goals of inclusivity and accessibility during the Covid-19 pandemic [25]. For this study, the SSGJ was run fully in-person, as during the summer 2022 when this SSGJ took place, local government restrictions regarding Covid-19 had been lifted, and several studies have reported lower turnout rates, lower commitment effort of participants, and less interaction among participants who participated in studies online (e.g., [1]) and in game jams in particular (e.g., [25,52]).
- **Technology**: As the SSGJ was all in-person, Discord [45] was no longer used. All communication and collaboration was done in-person, with Miro [46] boards being used for structured activity worksheets used for in-person collaboration.

3.2. SSGJ evaluation procedure

For the evaluation of this SSGJ, the SSGJ evaluation procedure by Stals et al. [25] for the evaluation of the SSGJ methodology as a mechanism for co-designing serious games to improve understanding of cybersecurity, was used. It identified several aspects of SSGJs that should be evaluated and the methods and tools to do so. These are knowledge, workload, motivation, and engagement. In this work, we follow the same evaluation guided by the research questions in Section 1, and informed by the specific design parameters of our SSGJ outlined in Section 3.1. The structure of the SSGJ, including evaluation activities, can be seen in Table 1.

[1] https://secrious.github.io/.

Table 1
SSGJ structure showing duration (in minutes), milestones, and evaluation activities.

Phase	Activity	Rationale	Procedure	Time (min)
Prep	Participant information sheet	Inform participants	Online reg.	5
	Informed consent form	Inform consent	Online reg.	5
	Demographic quest.	Participant profiles	Online reg.	5
	Adjusted IMI	Motivation pre-SSGJ	Online reg.	5
1	**Day 1**: Introduction to cybersecurity through provocative game and expert presentations. Introduction to TGD. Cybersecurity problems are identified using the cybersecurity cards.			
	Pre-Test quest.	Assess understanding	Discord	10
	TLX: Day 1 (See Table 2)	Measure workload	Paper/Discord	4 × 5
	IMI: Day 1	Measure motivation	Discord	5
	Day 2: TGD Meaning session: suitable learning mechanics selected using cards. TGD Play session: suitable game mechanics selected using cards and matched with learning mechanics.			
	TLX: Day 2 (See Table 2)	Measure workload	Paper/Discord	3 × 5
	IMI: Day 2	Measure motivation	Discord	5
2	**Day 3**: Game loops and prototyping.			
	TLX: Day 3 (See Table 2)	Measure workload	Paper/Discord	4 × 5
	IMI: Day 3	Measure motivation	Discord	5
3	**Day 4**: Second round of prototyping, elevator pitch and further development. Questionnaires for 3 decks of cards			
	TLX: Day 4 (See Table 2)	Measure workload	Paper/Discord	3 × 5
	Cybersecurity Cards quest.	Evaluate cards	Discord	10
	Learning Cards quest.	Evaluate cards	Discord	10
	Game Cards quest.	Evaluate cards	Discord	10
	IMI: Day 4	Measure motivation	Discord	5
	Day 5: Rule development and playtesting serious game prototypes with peer development.			
	TLX: Day 5 (See Table 2)	Measure workload	Paper/Discord	2 × 5
	IMI: Day 5	Measure motivation	Discord	5
Post	Post-Test quest.	Assess understanding	Discord	10
	SSGJ Experience quest.	Evaluate SSGJ format	Discord	30

With reg.: registration, quest.: questionnaire.

3.2.1. Measuring knowledge, workload, motivation and engagement

Assessing learning outcomes from game jams is difficult [53,54] and the learning experience is a very private experience [54]. Therefore, reports of learning are typically self-assessed by game jam participants [53,55]. In order to determine how the SSGJ has affected participants' understanding of the application domain of cybersecurity, pre-/post-tests, self-assessment, and peer-assessment are used [25]. Pre-/post-tests are one of the most used experimental designs in educational research to assess the effect of new teaching methods [56,57]. Confidence in key skills is also an important measure of the learning experience [58]. An adapted version of the Student Instrument for measuring Confidence in Key Skills (SICKS) [58] is used for a quantitative measurement of confidence regarding cybersecurity and game design knowledge and skills [25]. At the start of Phase 1 and after the end of Phase 3, participants' level of knowledge and understanding of cybersecurity and game design and development and their confidence in key skills in those areas, was collected using a one-group pre-/post-test questionnaires [25] (see Table 1). It consisted of seven 7-point Likert questions, with the key skills in the SICKS questionnaire [59] being replaced by the key skills in cybersecurity [1] and game design/development [25]. It was administered individually online using Microsoft Forms [60] at the start of Day 1 and at the end of Day 5 of the SSGJ (see Table 1). Self-assessment and peer assessment at team level using the Cybersecurity cards, LM cards, and GM cards of the SSGJ toolkit, combined with feedback from experts, has been used for the qualitative assessment of participants' understanding of cybersecurity and serious game design and development during the SSGJ [25]. This was done on Day 4 and 5 of the SSGJ (see Table 1).

Unlike traditional, fast-paced game jams [22], the SSGJ model re-evaluates the time pressure based on serious game design needs, and to reinforce accessibility and inclusivity [23]. It aims to be a 'no-crunch' working environment by having session durations that are non-exhausting. To evaluate this, the workload of each of the activities during the SSGJ needs to be assessed. During all three phases, workload was assessed using the NASA Task Load Index (NASA-TLX) [48,49]. It was administered individually at the end of each activity, using pen/paper for in-person (see Table 1) [25].

Another important measure of the learning experience in game jams is (intrinsic) motivation, as it can drive future learning and have a positive impact on self-efficacy [53,61]. Motivation also plays an important role in participating in game jams [27] and, due to the "slow" aspect of the SSGJ, participants are asked to commit over a longer overall duration than a traditional game jam [22]. Therefore, during all three phases of the SSGJ, motivation to adhere to and complete the activities of the SSGJ was measured using the Intrinsic Motivation Index (IMI) [50,51] at the end of each day (see Table 1). The IMI was administered individually and online by sharing a Microsoft Forms link to the questionnaire [25].

In the Post-Phase at the end of Day 5 (see Table 1), an open-ended questionnaire was used to encourage participants to reflect on the innovative "slow" format of the SSGJ and to evaluate their engagement between scheduled days [25]. It was administered individually by sharing a Microsoft Forms link on the final day of the SSGJ.

3.2.2. Evaluating the cybersecurity, LM and GM cards

In the third phase on Day 4, for each of the deck of cards in the SSGJ toolkit (Cybersecurity, LM, and GM cards), the usefulness, usage, and design of each deck of cards was evaluated. This was done using a questionnaire consisting of 12 items using a 7-point Likert scale, three questions using checkboxes with predefined options, and four free-text questions [25]. These three questionnaires (one for each deck of cards) were administered online by sharing Microsoft Forms link on Day 4 of the SSGJ (see Table 1).

4. Results

In this section, the results of the evaluation of the SSGJ with schoolchildren between the ages of 11 and 16 years old are presented. These findings will be discussed in Section 6.

4.1. Participants

Ethical approval was obtained from the researchers' university ethics board. In total, 27 participants were recruited via local youth organizations who run Science, Technology, Engineering and Mathematics (STEM) based learning programs. Only the data of participants that have actively participated in at least 3 out of 5 days of the SSGJ have been included, resulting in data of 23 participants in 6 teams being analysed.

Participants were aged between 11–16 years (mean 12.8 years, 7 female, 16 male) who had no formal training or education in cybersecurity. All participants were in late primary or early secondary school, with very limited prior game jam experience (only 2 have participated in one before the SSGJ). Eight participants had stated they had some coding experience, noting skills in either Scratch, Python or GDevelop, with only one participant indicating they have secure coding expertise, identifying themselves as a hacker. With regard to gaming experience, all participants have played digital games with only 2 participants indicating they had very limited experience in gaming in general. Three participants stated they do not play board games. During the SSGJ, experts in cybersecurity and game design would always be present with at most 7 experts present, periodically roving between different teams to check if there were any questions or they wanted to have a discussion with an expert. The serious games were co-designed among participants (i.e., experts were not part of a team).

4.2. Pre- vs. Post-questionnaire

The 7-point Likert-scale data was categorized by the percentage of SSGJ participants who reported whether they were confident (score $\geq$ 5), neutral (score = 4) or not confident (score $\leq$ 3). These percentages were then compared pre- and post-SSGJ [59,62]. In addition to this, a Wilcoxon signed-rank sum test [63,64] was done to determine whether there exists a statistical difference between pre- and post-test scores. Free-text answers were coded using the constant comparative method over three passes (coder 1, coder 2, resolve/combine) and grouped into themes. The entire data set was initially coded by two postdoctoral researchers with expertise in HCI and cybersecurity, and experienced in free-text coding qualitative data from questionnaires. Open discussion was used to systematically discuss and resolve the codes to reach consensus for the final coding [17,65,66].

4.2.1. Cybersecurity (Pre- vs. Post-SSGJ)

The Wilcoxon signed-rank sum test [63,64] showed participants' confidence in their knowledge and understanding of cybersecurity in the post-test scores had improved significantly compared to the pre test scores ($Z = -2.392$, $p = 0.016$). Responses relating to *Code Practices* indicated confidence in current level of knowledge and understanding of cybersecurity shifted positively from 41.2% to 76.5%. Confidence with reviewing and updating existing code regarding cybersecurity dropped from 41.2% to 35.3%. Responses relating to *Resources* indicated increased confidence to ask for more time to improve code security from 58.8% to 64.7%, while confidence to ask help from other people to improve code security dropped from 64.7% to 54.9%. Responses relating to *Communication* showed confidence in raising a security issue with their non-expert teacher increased from 35.3% to 47.1%, as did their confidence to raise a security issue with a cybersecurity expert (from 35.3% to 58.8%). Responses for *Morality* showed increased confidence to go against their teacher when the teacher finds finishing the code/programming assignment more important than creating secure code (from 17.6% to 35.3%), but confidence to bring up a security issue that they know will have an impact on the user of the software or app decreased 64.7% to 47.1%. Other items, such as asking for more focus on improving code security, showed little change.

4.2.2. Game design and development (Pre- vs. Post-SSGJ)

The Wilcoxon signed rank sum test also showed participants' confidence in their knowledge and understanding of game design and development in the post-test scores had improved significantly compared to the pre-test scores ($Z = -2.792$, $p = 0.005$). Regarding game design, confidence in current level of knowledge and understanding of game design increased (from 47.1% to 82.4%) and to teach others about game design (35.3%–58.8%) and the ability to design a game (58.8%–82.4%) shifted positively. Regarding game development, confidence in current level of knowledge of game development (41.2%–76.5%), in sharing their knowledge and understanding about game creation with others (29.4%–47.1%), and in ability to implement a game (41.2%–52.9%) all shifted positively. Other items, such as teaching others about game development, showed little change.

4.2.3. Free-text responses

The free-text responses provided some additional insights in the quantitative data presented above. A third of the participants self-reported that through participation in the SSGJ they had obtained a greater general awareness of cybersecurity: *"How vulnerable you can be by doing certain things like downloading things from illegal websites"* (P15). In addition, a quarter of the participants reported learning more about

Table 2
Overview of the workload for each of the NASA-TLX subscales and the average per activity of the SSGJ, with the top-5 highest values in the "High" and "Very High" classifications per subscale highlighted in bold underlining.

Day	Activity	MD	PD	TD	EFF	FRU	PER
1	Playing provocative game	11.5	6	10.2	10.9	**11.9**	8.7
	Introduce cybersecurity cards via game	**13.7**	6.7	9.7	12.0	9.3	8.1
	Cybersecurity metaphors in game	9.9	8.1	9.1	12.0	8.2	7.8
	Reality phase	**13.8**	8.5	8.9	13.8	8.8	8.1
2	Meaning phase	**14.3**	9.0	9.3	7.1	**15.2**	7.1
	Play phase	**14.7**	8.8	10.5	**15.0**	9.9	8.1
	Bringing it all together (SGDD)	12.9	9.4	10.7	13.8	8.9	7.9
3	Game loop design	**13.8**	8.1	10.3	6.6	**12.7**	6.6
	Serious game loops	12.5	10.5	10.2	13.8	9.6	7.6
	Prototyping session	13.5	**11.4**	**11.8**	**15.3**	10.6	6.5
4	Prototype sharing	12.0	10.7	**11.9**	**14.4**	9.9	7.2
	Elevator pitch and name of game	12.9	10.8	**12.3**	**14.3**	9.1	5.9
	Development – Part 1	12.4	**11.0**	**16.7**	12.9	9.0	6.6
	Assess own game with cards	12.2	7.5	9.4	12.1	6.6	4.4
5	Rules development	10.3	9.9	10.4	12.3	8.6	5.8
	Playtesting	12.7	9.8	9.4	12.3	10.0	5.6
	Peer assessment	12.6	9.5	10.8	13.6	9.6	5.5

a specific attack, defence, or vulnerability: *"I have learnt about specific types of attack like brute force"* (P3). A little over half of the participants explicitly indicated that the SSGJ matched their expectations regarding learning about cybersecurity: *"Game jam school was a really cool and fun way to learn about cybersecurity"*. (P10), with none of the participants indicating it did not match their expectations in this regard. Almost half of the participants also explicitly indicated the SSGJ matched their expectations regarding serious game design, with only one in ten indicating it did not: *"The game Jam kind of met my expectations except the fact that it wasn't going to be digital"*. (P15).

4.3. Workload

The raw NASA-TLX workload data was analysed by taking the mean of all responses for each activity and workload subscale [48] (see Table 2). The average workload values for each of the subscales correspond to mental demand (MD), physical demand (PD), temporal demand (TD), effort (EFF), frustration (FRU) and performance (PER). The scores were classified as Low (1–3), Medium (3–7), Bit High (7–11), High (11–17) and Very High (17–21) [48,67].

Overall, the workload was never *"Very High"* on average. The overall workload was the highest for the Prototyping session, and the Development activity, with four of the six NASA-TLX subscales being *"High"* for those activities. Mental demand (MD) was *"High"* throughout the SSGJ, specifically for TGD (MD = 13.8–14.7), Serious Game Loop Design (13.8), and the Introduction of the cybersecurity cards using a game (13.7). Effort was the highest for the Prototyping session (EFF = 15.3) and Play phase (15.0) in TGD respectively. The lowest mental demand came from identifying the cybersecurity metaphors in the provocative game (MD = 9.9) and lowest effort was in the serious game loop design (EFF = 6.6). Even though mental demand and effort where high for most of the activities during the SSGJ, temporal demand (TD) however is only *"High"* for the prototyping and development activities (TD = 11.8–16.7), particularly on Day 4.

With regard to performance, participants felt they performed well for all activities of the SSGJ, with performance always falling in the *"Medium"* or *"Bit High"* category (PER = 4.4–8.7), with a low score for perceived performance indicting participants having a positive perception of their performance. In addition, participants felt they performed increasingly well as the SSGJ progressed. Participants perceived their performance to be the least good on the first day of the SSGJ (while playing the provocative game (PER = 8.7)), with the best perceived performance noted during the self-assessment of their own serious game with the cards (PER = 4.4) on Day 4. Frustration (FR) was only *"High"* for playing the provocative game (FR = 11.9), game loop design (FR = 12.7) and during the meaning phase of TGD (FR = 15.2). Finally, with regard to physical demand, this was only *"High"* during the prototyping session (PD = 11.4) and development (PD = 11.0).

4.4. Motivation

Motivation was measured using the Intrinsic Motivation Inventory (IMI) and was analysed by averaging each subscale for each day of the SSGJ. It was then further refined by looking at the percentage of participants who scored *"Very High"* one each of the subscales (as per [68–70]) (see Table 3 and Fig. 2).

Results in Table 3 shows that Interest/Enjoyment and Perceived Value/Usefulness are all positive (≥4.5/7) on average and stay positive throughout the SSGJ, indicating the SSGJ had successfully engaged participants on the topic of cybersecurity. In particular, 2 out of the 5 days were even rated *"Very High"* on average (Days 2 and 3). Perceived Competence is neutral on the first day of the SSGJ, but is positive for all the other days of the SSGJ.

The IMI sub-scale scores of each individual participant support the findings in Table 3, with Fig. 2 visualizing the percentage of participants who scored *"Very High"* on the sub-scales of the IMI per day. The SSGJ was rated very highly on Interest/Enjoyment by 50%–54.5% of participants on Day 2 and Day 3 of the SSGJ, as well as Value/Usefulness (54.5%–59.1%). Perceived Competence also peaked on Day 3 at 68.2%.

Table 3
Average IMI scores for each subscale per day of the SSGJ, with very positive scores of 5.50 and over highlighted in bold underlining.

Day	Interest/enjoyment	Competence	Choice	Value/usefulness
Pre	5.33	4.68	5.40	**5.87**
1	4.75	4.17	5.38	5.01
2	**5.50**	5.27	5.35	**5.62**
3	**5.69**	**5.74**	5.35	**5.73**
4	5.18	5.09	5.29	5.23
5	5.23	5.28	5.14	5.32

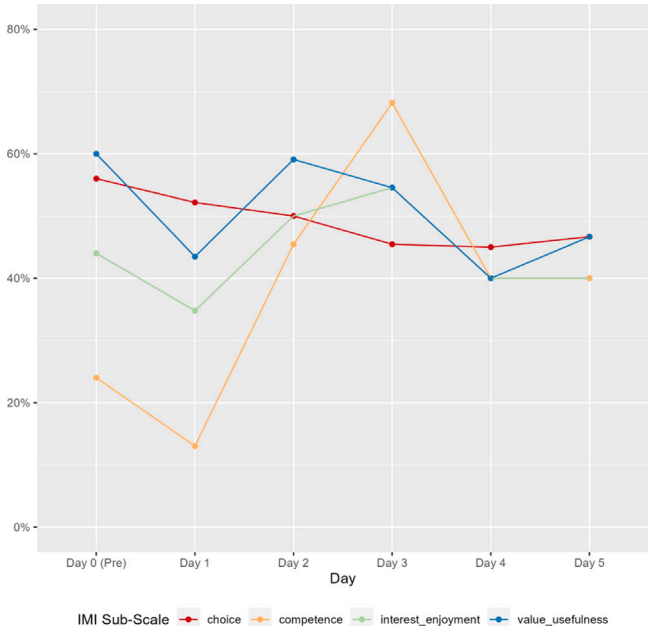


Fig. 2. Percentage of participants who scored "Very High" (6–7 out of 7) per IMI sub-scale per day.

4.5. Cybersecurity, LM and GM cards

During the SSGJ, online questionnaires to evaluate each deck of cards were distributed among participants, consisting of twelve 7-point Likert-scale questions, three questions with tick boxes allowing for a selection of multiple pre-defined options, and four free-text responses to open questions for further clarification. The response rates were 100% for the Cybersecurity Cards Questionnaire, 78% for the LM Cards Questionnaire, and 91% for the GM Cards Questionnaire.

4.5.1. Cybersecurity cards

The cybersecurity cards aimed to provide participants with a knowledgebase for cybersecurity (see Section 3.2.2). In relation to the usefulness of the cybersecurity cards (red and white cards in Fig. 3), participants reported that they provided knowledge about individual cybersecurity concepts (74%), the wider scope of cybersecurity concepts (74%), the relationship between vulnerabilities, attacks and defences (70%) and terminology (61%). Participants reported they also provided a means for independent learning (48%) and self-efficacy by providing access to cybersecurity knowledge when experts were not present (57%). Furthermore, they improved accessibility by acting as an interface to discuss cybersecurity topics with experts (61%) and other peers in their group (61%) throughout the SSGJ. There were 70%

Fig. 3. Picture of team 4 analysis of serious game prototype using the cybersecurity cards, selecting cards linked to web-based attacks (i.e., Injection) and Malware. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of participants who indicated that the cybersecurity cards provided them with inspiration for the design of the serious game. Throughout the serious game design and serious game creation and development, 65% of the participants noted that the cards were used as a reference point for ensuring the game links to real cybersecurity issues. While the cybersecurity cards were used throughout the entire SSGJ, they were most used in the TGD phases (57%–65%), during game creation and development (65%) and when analysing a serious game (65%) (see Fig. 3).

4.5.2. LM cards

For the LM cards (the blue cards in Fig. 3), 72% participants agreed that they provided knowledge about individual LM concepts, the range of LM concepts, LM terminology. Furthermore, participants also agreed that the LM cards provided access to LM knowledge when serious game design experts were not present (72%), as well as providing both inspiration for their serious game design (72%) and linking LM concepts to their game during creation (72%). Interestingly, participants were split regarding whether the LM cards limited their creativity for the design of their serious game, with 39% agreeing or disagreeing and the remainder neutral to this. During the activities of the SSGJ, the LM cards were mostly used during serious game creation and development (67%), the creation of the serious game loop (61%), and during both the meaning phase of TGD and analysis of their serious game (both 56%).

4.5.3. GM cards

Participants agreed the GM cards (yellow cards in Fig. 3) provided knowledge of individual GM concepts (71%), the scope (71%) and the relationship between GM concepts (81%) and terminology (76%). More than half agreed the GM cards were useful for independent learning (67%) and self-efficacy by providing access to knowledge of game mechanics when the serious game design experts were not present (67%). They improved accessibility through provision of an interface for discussion with the experts (67%) and with other participants (67%) throughout the SSGJ. Over half the participants (67%) agreed the GM cards provided them with inspiration and acted as a reminder to link and ground the serious game in the mechanics they chose during game design (62%) and development (67%). The GM cards were consulted for different phases and activities throughout the entire SSGJ. Specifically, serious game creation and development (67%), the Play phase of TGD (62%) and Creation of the serious game loop (52%).

Table 4
Overview of the percentages of participant-selected limitations in the design of the Cybersecurity, LM, and GM cards.

Limitation	% for		
	Cybersecurity cards	LM cards	GM cards
Total number of cards too high	35	44	38
Colour coding not clear	39	33	24

4.5.4. Feedback on the design of the cybersecurity, LM and GM cards

Limitations of the design of each deck of cards are summarized in Table 4, indicating the percentage of participants who ticked that selection box with predefined options. Across the three different decks, the main redesign suggestions were that the total number of cards was considered too high and that the colour coding was not clear. For the cybersecurity cards specifically, the main redesign suggestions were that the total number of cards was too high (35%), the number of cards within each category was too high (39%), and the colour coding was not clear (39%). For the LM cards, the total number of cards was considered too high by 44% of participants, the relationship between LM cards and the learning verbs on those cards was not clear (33%), the numbering and colour coding of the LM cards was not clear (33%), and the information on the LM cards was considered to be too abstract (33%). For the GM cards, the total number of cards was also considered too high (38%) and the relationship between the LM and GM cards was considered not to be clear by 31%.

4.6. Assessment of serious games using cards

Self-assessment and peer-assessment at team level were used to assess participants' serious game prototypes, using the cybersecurity cards, LM cards, and GM cards from the SSGJ toolkit.

4.6.1. Self-assessment

Firstly, each team assessed their own game using the three decks of cards and discussed and explained their selection to at least one of the experts (see Fig. 3). These discussions with the experts reflected an evolving understanding of both cybersecurity as well as serious game design, as participants could quickly, confidently, and adequately explain why certain cards were (or were not) part of their serious game design.

It also illustrated how many cards have been used during the design of the serious game, which has been summarized in Table 5. It should be noted that a higher number of cards does not imply that the serious game design is better (or worse), but it does illustrate that the Cybersecurity cards played a prominent role in the serious game design.

Fig. 4. Playtesting of the serious game by another team in the SSGJ.

Table 5
Number of cards per serious game per team.

Team	Cybersecurity	LM	GM	Total used
1	1	9	14	24
2	70	8	12	90
3	15	5	7	27
4	9	8	12	29
5	11	8	5	24
6	1	5	7	13
Range	1–70	5–8	5–14	13–90

Table 6
Number of cards in the peer-assessment of the serious game that matched the team's self-assessment.

Team	Cybersecurity	LM	GM	Total matched
1	N/A	N/A	N/A	N/A
2	70/70	1/8	4/13	75/89
3	5/15	2/5	3/7	10/27
4	2/9	4/8	5/12	11/29
5	0/11	3/8	0/5	3/24
6	0/1	4/5	5/7	9/13
Total	77/106	14/34	17/44	3–75

4.6.2. Peer-assessment

The game was then played by another team (see Fig. 4), who subsequently also assessed it as a team using the three decks of cards (see Table 6). A member of the team that created the game would afterwards explain which cards matched in their opinion and which ones did not. This indicated a high level of successful mapping of learning outcomes into the teams' games. It should be noted though that since one of the teams (Team 1) did not manage to produce a playable serious game, it could not be play-tested or assessed by another team in the SSGJ using the three decks of cards.

In addition, it should be noted that the peer assessment is not about selecting a sub-selection of cards that is "correct" or "matched" or "incorrect" or "Did not match", but about the discussion that occurred around this matching exercise, which provided valuable feedback from peers and experts on the design of the serious game, and how it was interpreted by others.

4.7. Engagement between scheduled days of the SSGJ

The response rate for this questionnaire was 78%. Half of the participants indicated they had engaged with their serious game project in between scheduled days, as a: team (11%), subsection of a team (11%), and individually (22%) (and 56% did not specify their response). Responses indicated activities such as reflection on things they had learned (50%), and content creation for their game (28%). Furthermore, further research was conducted in: cybersecurity (22%), learning context (28%) and on games (38%). The main reason the other half of participants gave for not engaging with their serious game project in between scheduled events was that they had managed to finish their work during the scheduled events of the SSGJ: *"There was nothing to do at home cuz we finished it in the time period"*. (P9).

5. Serious game output from SSGJ

From the SSGJ, five out of six teams managed to deliver a playable prototype of their serious game as an output (see Fig. 5), as well as any supporting documentation in the form of both a rule book for their game and a Serious Game Design Document (SGDD).

In this paper, we will discuss two of these serious games that were further developed by serious game design experts from the prototypes delivered through the SSGJ methodology. These two games were decided to be further developed by consensus decision, chosen as winners by experts in serious game design, cybersecurity (one external expert), and software engineering (one external expert) in the SSGJ. The only team that did not produce a playable prototype still engaged with the application domain and serious game design and managed to create characters and a story line around a specific cybersecurity topic, but did not manage to develop that into a playable prototype during the SSGJ.

5.1. ScareCity

ScareCity is a 2–4 player serious board game that is underpinned by the theme of secure software development lifecycles. This theme involves the integration of security into the software development

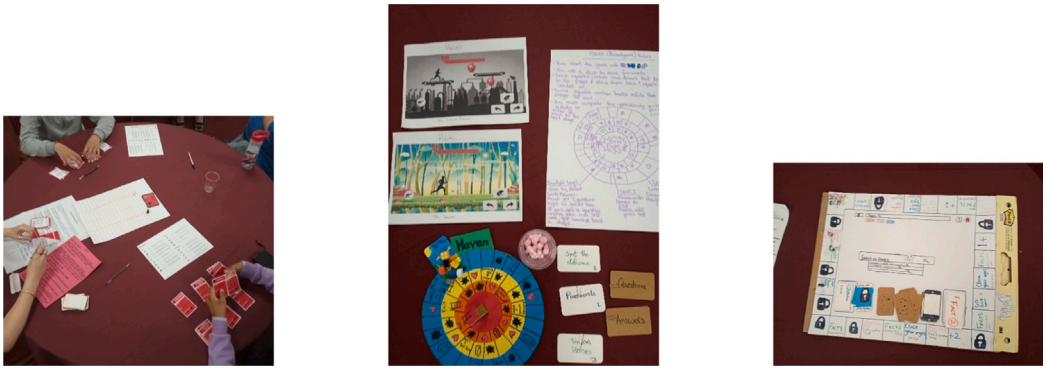


Fig. 5. Examples of the playable serious game prototypes co-designed with participants during the SSGJ.

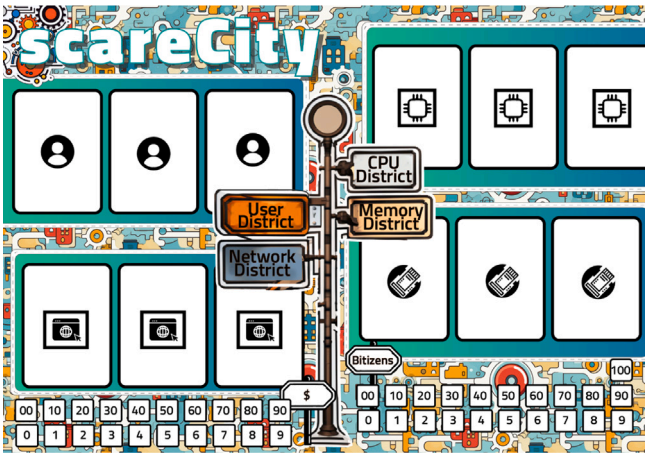


Fig. 6. The ScareCity game board showcasing the four program districts.

lifecycle, coupled with additional processes such as risk analysis and security testing. In ScareCity, each player has their own game board (see Fig. 6) and acts as the *mayor* of their city.

The goal is to expand the city's features through *Program Cards* to manage both the security of the city and improve the trust of its *Bitizens*. Program cards can be abstracted to strategic security decision making (i.e., implementing defences), and fall into one of four *Districts*: CPU, Network, User or Memory. During the game, before the end of each round, *Impact Cards* are drawn which are akin to security attacks or defences, which either cause havoc to the player's program cards (causing them to degrade) or may bring benefits (i.e., increasing the number of bitizens). Examples of both *Program* and *Impact* cards can be seen in Fig. 7. The first player to reach a total of 200 *Bitizens* wins the game.

Linking back to the theme of secure software development lifecycles, the goal of the game is to implement security into the player's city (programs), with aspects such as risk analysis and security testing done through analysing the available program cards to minimize the potential for serious impacts to a player's programs. The game also highlights the trade-offs that are often required to be made by cybersecurity experts due to limited resources. From this game design, it is clear that participant's in the SSGJ also successfully met the jam's intended learning outcomes regarding serious game design — understanding the importance of the three aspects of TGD [29] evidenced by appropriate usage of reality, meaning and play elements during design.

5.2. No-entry

No-Entry is the second serious game output from the SSGJ. It is a serious card game underpinned by CyBOK [33] and the cybersecurity cards from the SSGJ toolkit (see Section 3 Support-parameter and [32]). It is designed to be played by 2 players, with an optional third player taking a turn to act as a game master. One of the players will act as an attacker, holding the attack cards from the cybersecurity card deck (see Fig. 8), with the other as a defender holding the defence cards from the deck. The game master will then draw a vulnerability card at random, with the attacker and defender tasked with finding an appropriate attack or defence, respectively, that targets the drawn vulnerability. Ultimately, this game uses the full deck of cybersecurity cards as game elements. In addition to this, a playable surface (i.e., a flat table), a form of timer (i.e., stopwatch/egg timer or smartphone timer) and a method of deciding who plays as an attacker or defender (i.e., coin flip) is also required.

The objective of the game is to reach 20 points and the first player, either the attacker or defender, to do so wins the game. Both the attacker and defender will find and select a single card they think targets the vulnerability card. Before the end of the round and a winner of the round is chosen, both the attacker and defender must verbally describe why they believe their chosen card is correct. Points are awarded to the player(s) who successfully attack or defend the vulnerability. If only one of the players (either the attacker or defender) is the one to find a matching attack/defence, they are awarded 2 points and in the case both are correct, both players are awarded a single point. The primary role the game master plays in this game is deciding whether the card selected by the attacker and defender are correct, which is influenced based on how well the attacker and defender describe why their chosen card matches the vulnerability. Secondary responsibilities is time-keeping and keeping score of the game.

The main purpose of this game is to provide a gamified approach to interacting with the cybersecurity cards, which has been shown to

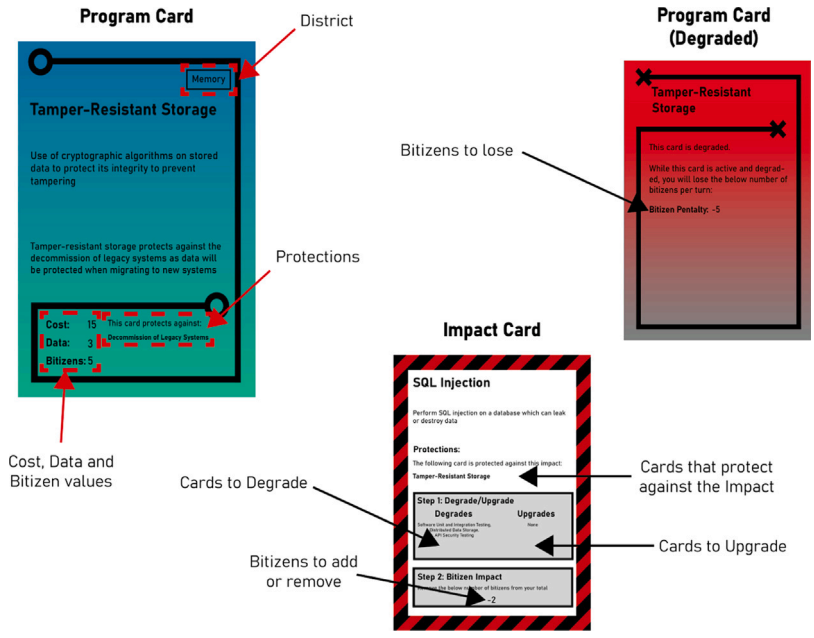


Fig. 7. Example of ScareCity program and impact cards.

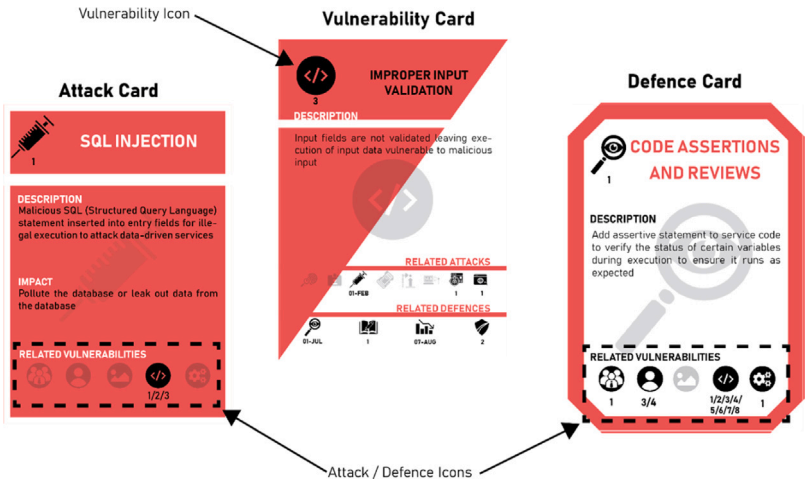


Fig. 8. "No-Entry" serious game being played, with, counterclockwise, the defender, game master, and attacker.

be an effective method for improving learning and understanding of pedagogical content [71]. During the SSGJ, the cybersecurity cards were introduced using a basic introductory game, in which each team was given a vulnerability card from the cybersecurity cards deck, and they had to find a matching attack and defence card held by the researchers present in the room. Although participants were given more time to explore the full cybersecurity card deck afterwards, this only allowed a very limited exploration of the cybersecurity cards as part of the game. Building on the existing cybersecurity cards, No Entry further addresses the issues around accessibility of CyBOK [32], and the difficulties with navigating this comprehensive information base, as well as the ability to express different cybersecurity scenarios with it. From a game design perspective, the No Entry game improves the balance of the three aspects of TGD [29] evidenced by appropriate usage of reality, meaning and play elements during design.

6. Discussion

This paper presented an evaluation of a SSGJ methodology as a mechanism for co-designing serious games in the application domain of cybersecurity, to evaluate how the SSGJ methodology contributed to improving the understanding of cybersecurity for different demographics. The aim was to evaluate how the SSGJ contributed to improving the understanding of cybersecurity for young people between the ages of 11 and 16 years old who have little or no knowledge of cybersecurity, and use this to validate and compare to the results from earlier work where the SSGJ methodology was used with a different target demographic (i.e., MSc students with no formal training or education in cybersecurity). To investigate this, we were guided by the research questions in Section 1.1. Finally, we presented two SSGJ outputs in the form of serious games in the application domain of cybersecurity that were co-designed with participants. We will now use the results to reflect on each research question in turn.

6.1. RQ1: How has the SSGJ affected young participants' understanding of cybersecurity?

The SSGJ has contributed to improving young participants' confidence in, and understanding of, cybersecurity, and insight into where their skills may be lacking.

The assessment of the serious games using the three decks of cards presented in Section 4.6 showed that participants were able to assess their own serious game design using the cybersecurity cards. Discussion of their card selection with experts showed that they could effectively explain why certain cybersecurity cards were (or were not) part of their game design. Secondly, participants were able to communicate their knowledge of cybersecurity to others by relating cybersecurity concepts to elements in their game. For example, the No-Entry serious game presented in Section 5 incorporated the trichotomy of vulnerabilities, attacks and defences and communicated this by having players take on the role of attacker or defender of vulnerabilities in the game. Thirdly, participants were also able to match cybersecurity cards to the serious game design of another team of participants, as results in Section 4.6 show. Matching all cybersecurity cards would be difficult though, as it also depends on how abstract the in-game metaphor is, the interpretation of the participant, and metaphors can be interpreted in multiple ways, as interpretive flexibility is an element of serious game design [47,72].

Pre-/Post-test results in Section 4.2 indicated that regarding key cybersecurity skills related to *Code Practices*, for young participants their confidence in their knowledge and understanding of cybersecurity shifted positively from 41.2% to 76.5%. This is in line with findings by [25], who found that for MSc students in an SSGJ their confidence in their knowledge and understanding shifted positively from 12.5% to 62.5%. So after participation in the SSGJ, more young participants felt confident, but their confidence in key cybersecurity skills prior to participation in the SSGJ was also higher as compared to the participants who were MSc students. Key cybersecurity skills regarding reviewing and updating existing dropped slightly (i.e., from 41.2% to 35.3%). This may have been due to the fact that actual coding or reviewing existing code was not part of the serious game design or SSGJ in general. In addition, a decrease in confidence may indicate that this is an area where confidence may initially have been inflated or perhaps unrealistic, and the process of the SSGJ has allowed participants to reassess areas in which they may need improvement. This is supported by the literature, showing that participants with less experience in cybersecurity are more willing to acknowledge their mistakes and lack of expertise in their skills and decision-making processes compared to experienced cybersecurity experts [16,25].

Free text answers showed a difference between what both demographics had learned about cybersecurity through the SSGJ. For the SSGJ with MSc students, three quarters of participants self-reported the main thing they had learned were the different types of security vulnerabilities, attacks, and defences, half of participants reported factors influencing cybersecurity (e.g., human factors), and almost a third mentioned the relationships between vulnerabilities, attacks, and defences as well as terminology [25]. Young persons however self-reported that the main thing they had learned through participation in the SSGJ was a greater general awareness of cybersecurity, while a quarter of the participants reported learning more about one specific attack, defence, or vulnerability.

In addition to improved confidence regarding participants' knowledge and understanding of cybersecurity, also young participants' confidence in their knowledge and understanding of (serious) game design increased significantly from (47.1% to 82.4%). This is line with findings by [25] who found that for MSc students their confidence in their knowledge and understanding of (serious game design) shifted positively from 12.5% to 75.5%. Again, after participation in the SSGJ a higher percentage of young persons reported feeling confident in this area, but their confidence in (serious) game design prior to participation in the SSGJ was also higher.

6.2. RQ2: How can the cards for the application domain (in our case the cybersecurity cards), learning mechanics cards, and game mechanics cards that are part of the SSGJ toolkit, assist young participants in serious game design?

Regarding the design of the serious games, it was observed that the cybersecurity cards assisted in the design of the serious game in all three phases of the SSGJ in similar ways for both demographics. As shown by the results in Section 4, the cards contributed to the SSGJ by providing a knowledge base for individual cybersecurity, LM, and GM concepts and terminology, enabled self-efficacy for when the experts were not present, and improved accessibility by acting as an interface for discussion. They also acted as a reminder to link and ground the serious game design in cybersecurity, effectively mapped to LM and GM mechanics, which has been shown in the literature to be important in order to create an effective serious game [29,35,36,47,73]. A difference between both demographics though was that the young persons did not feel the cybersecurity cards limited their creativity for serious game design, only the LM cards did. This is unsurprising though, as a serious game needs to achieve the learning outcomes using learning mechanics, while an entertainment game does not have such restrictions.

6.3. RQ3: What are the workload and motivation levels of young participants during the SSGJ?

Traditional, fast-paced game jams tend to have a high workload and temporal demand [21,22,37]. The SSGJ successfully reduced time pressure for both demographics. Based on the NASA-TLX data presented in Table 2 in Section 4.3 it can be concluded the workload was never considered "Very High" (17–21) by young participants and temporal demand was only considered "High" (11–17) for 4 out of 17 activities during the SSGJ. However, there were some differences between the two demographics regarding the various workload subscales. Where MSc students reported the highest mental demand for development activities (MD = 15.3–17.1), young persons reported the highest mental demand for the three TGD activities (13.8–14.7), serious game loop design (13.8), and the game played to introduce them to the cybersecurity cards (13.7). In addition, young persons reported the highest effort for activities involving prototyping and presenting or sharing their prototype (14.3–15.3). This was supported by free-text answers, in which young persons indicated they struggled with having to present their work in front of others: *"Presentation [was the most difficult], I don't really like talking in front of lots of people"* (P1). For both demographics, temporal demand was the highest for development activities. This may partially be related to participants' inexperience with those activities. In addition, another noticeable difference is that young persons reported a higher Physical Demand (PD) during the SSGJ. For the MSc students, this ranged from 2.4–7.0 which falls in the Low to Medium range [25]. For young persons this ranged from (6.0–11.4) which falls in the Medium to High range, with prototyping (11.0) and development activities (11.4) being the highest. This finding indicates that different demographics find different activities during the SSGJ challenging. This is a useful insight from a planning perspective, as it indicates that participants may need additional breaks and support during specific activities to offset the fatigue from increased workload based on their demographic.

Motivation levels reported in Section 4.4 indicate the SSGJ managed to engage young people in software security concepts, which is in line with findings by [25]. Average levels of the sub-scale Interest/Enjoyment (4.75–5.69), and Perceived Value/Usefulness (5.01–5.73) were positive on average for all days and phases of the SSGJ and even considered very high for the majority of participants for 2 out of 5 days (i.e., Day 2 and 3) which included the TGD, serious game loop design and prototyping activities. An interesting difference between the two demographics though is that the percentage of young persons who scored very high on perceived competence, shoots up after the first day to between 40% and 68% for the remaining days of the SSGJ, while for the MSc students this was below 40% throughout the SSGJ. This is supported by the NASA-TLX data, which shows a steady increase of perceived performance (from 8.7–5.5), indicating that young persons feel they perform better as the SSGJ progresses. This shows that for both demographics, but in particular for young persons, the SSGJ contributes to increasing their confidence in their own abilities regarding cybersecurity and serious game design.

6.4. RQ4: How has the "slow" format of the SSGJ affected engagement of young participants?

Results in Section 4.7 show the "slow" format encouraged engagement of young persons in-between scheduled days of the SSGJ. Half of the young participants engaged with their serious game project outside the SSGJ. This is particularly impressive, as the SSGJ for young persons was scheduled for 5 consecutive days as outlined in Section 3. In contrast, for the SSGJ with MSc students all participants engaged with their serious game project in between scheduled events, but their SSGJ was structured into three phases of two workdays each (so 6 days in total), spread over 5 weeks. That allowed more time for reflection and refinement in-between. But the results of the SSGJ with young persons shows that, even when there is less time in-between scheduled days of the SSGJ, it still encourages engagement from participants. Young participants engaged in-between scheduled days of the SSGJ by actively creating content for their serious game (28% versus 86% for MSc students) and for reflection on things they had learned (50% versus 29% for MSc students). Furthermore, further research was conducted in: cybersecurity (22% versus 57%), learning context (28% versus 43%) and on games (38% versus 14%). This indicates that, compared to older participants, younger participants reflected on things they had learned and they undertook more research on games in-between scheduled events.

7. Conclusion

This paper has presented an evaluation of a SSGJ methodology as a mechanism for co-designing serious games in the domain of cybersecurity. The aim was to evaluate how the SSGJ contributed to improving the understanding of cybersecurity for young persons between the ages of 11 and 16 years old, and to validate and compare this to earlier work where the same SSGJ methodology was used with a different target demographic (i.e., MSc conversion students in Computer Science with no formal training or education in secure coding). To this end, we engaged 23 participants between the ages of 11 and 16 years old for 5 consecutive days over a one-week period, into a multidisciplinary SSGJ involving domain-specific, pedagogical, and game design knowledge, and encouraged engagement in-between scheduled events of the SSGJ.

The confidence of participants from both demographics regarding their knowledge and understanding of cybersecurity improved, from 41.2% to 76.5% ($Z = -2.392$, $p = 0.017$) for young persons and from 12.5% to 62.5% ($Z = -2.041$, $p = 0.041$) for MSc students. However, younger participants reported that the SSGJ experience mainly improved their general understanding of cybersecurity, specifically

in terms of one specific vulnerability, attack, or defence, while MSc students reported the SSGJ experience mainly improved their understanding of cybersecurity in terms of the trichotomy of vulnerabilities, attacks, and defences. Also the confidence of participants for serious game design and development improved for both demographics, from 47.1% to 82.4% ($Z = -2.792$, $p = 0.005$) for young persons and from 12.5% to 75% ($Z = -2.112$, $p = 0.035$) for MSc students. Furthermore, the SSGJ format worked well in engaging participants in between scheduled events of the SSGJ, with half of the young persons engaging in between scheduled events even though they had less time for refinement and reflection in comparison, and all MSc students engaging in between scheduled events.

The findings and resulting discussion in this paper provide useful insights into how the different phases and activities of the SSGJ have contributed to enhancing understanding of cybersecurity and game design for different demographics. As the SSGJ is intended to be flexible and applicable across multiple domains, these insights are not only useful for research, education, and training in cybersecurity, but can also be useful to researchers in the wider HCI community and game research community who are interested in using SSGJs to co-create serious games to improve understanding in other application domains and with various target demographics. The findings regarding the workload and motivation levels of the SSGJ also provided a baseline for workload and motivation levels for other game jams to be compared against, and show how those can differ for various target demographics regarding.

For future work, we will continue to use SSGJs as a mechanism to co-design serious games to improve the understanding of different themes within cybersecurity, focusing on code security, API security, and the security lifecycle. In addition, exhibitions for different audiences (i.e., cybersecurity, HCI, and game design experts, as well as the general public) will be organized to showcase the SSGJ toolkit, and to further playtest and evaluate the two serious games that have been co-designed during our SSGJs and produced as outputs of the SSGJ with their target audiences.

CRediT authorship contribution statement

Shenando Stals: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Lynne Baillie:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Ryan Shah:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Formal analysis, Data curation. **Jamie Iona Ferguson:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Investigation, Formal analysis, Data curation. **Manuel Maarek:** Writing – review & editing, Software, Resources, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This work contributes to the SECRIOUS project and is supported by the Engineering and Physical Sciences Research Council, United Kingdom (Grant number: EP/T017511/1).

References

[1] T. Georgiou, L. Baillie, O. Chatzifoti, S.C. Chan, Future forums: A methodology for exploring, gamifying, and raising security awareness of code-citizens, Int. J. Hum.-Comput. Stud. 169 (2023) 102930, http://dx.doi.org/10.1016/j.ijhcs.2022.102930.

[2] F. Fischer, K. Böttinger, H. Xiao, C. Stransky, Y. Acar, M. Backes, S. Fahl, Stack overflow considered harmful? the impact of copy&paste on android application security, in: 2017 IEEE Symposium on Security and Privacy, SP, IEEE, 2017, pp. 121–136, http://dx.doi.org/10.1109/SP.2017.31.

[3] H. Hong, S. Woo, H. Lee, Dicos: Discovering insecure code snippets from stack overflow posts by leveraging user discussions, in: Annual Computer Security Applications Conference, ACSAC '21, Association for Computing Machinery, New York, NY, USA, 2021, pp. 194–206, http://dx.doi.org/10.1145/3485832.3488026.

[4] C. Van't Wout, Develop and maintain a cybersecurity organisational culture, in: ICCWS 2019 14th International Conference on Cyber Warfare and Security: ICCWS, Vol. 457, 2019.

[5] A. Onumo, I. Ullah-Awan, A. Cullen, Assessing the moderating effect of security technologies on employees compliance with cybersecurity control procedures, ACM Trans. Manag. Inf. Syst. (TMIS) 12 (2) (2021) 1–29, http://dx.doi.org/10.1145/3424282.

[6] R. Kelly, Almost 90% of cyber attacks are caused by human error or behavior, 2017, ChiefExecutive. net. URL: https://chiefexecutive.net/almost-90-cyber-attacks-caused-human-error-behavior/.

[7] L. Li, W. He, L. Xu, I. Ash, M. Anwar, X. Yuan, Investigating the impact of cybersecurity policy awareness on employees' cybersecurity behavior, Int. J. Inf. Manage. 45 (2019) 13–24, http://dx.doi.org/10.1016/j.ijinfomgt.2018.10.017.

[8] T. Denning, A. Lerner, A. Shostack, T. Kohno, Control-Alt-Hack: the design and evaluation of a card game for computer security awareness and education, in: Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security, 2013, pp. 915–928, http://dx.doi.org/10.1145/2508859.2516753.

[9] T. Denning, A. Shostack, T. Kohno, Practical lessons from creating the Control-Alt-Hack card game and research challenges for games in education and research, in: 3GSE USENIX Summit on Gaming, Games and Gamification in Security Education, 2014, URL: https://www.usenix.org/conference/3gse14/summit-program/presentation/denning.

[10] M.K. Thomas, A. Shyjka, S. Kumm, R. Gjomemo, Educational design research for the development of a collectible card game for cybersecurity learning, J. Form. Des. Learn. 3 (1) (2019) 27–38, http://dx.doi.org/10.1007/s41686-019-00027-0.

[11] J. Anvik, V. Cote, J. Riehl, Program wars: a card game for learning programming and cybersecurity concepts, in: Proceedings of the 50th ACM Technical Symposium on Computer Science Education, 2019, pp. 393–399, http://dx.doi.org/10.1145/3287324.3287496.

[12] M. Swann, J. Rose, G. Bendiab, S. Shiaeles, F. Li, Open source and commercial capture the flag cyber security learning platforms-A case study, in: 2021 IEEE International Conference on Cyber Security and Resilience, CSR, IEEE, 2021, pp. 198–205, http://dx.doi.org/10.1109/CSR51186.2021.9527941.

[13] L. McDaniel, E. Talvi, B. Hay, Capture the flag as cyber security introduction, in: 2016 49th Hawaii International Conference on System Sciences, HICSS, IEEE, 2016, pp. 5479–5486, http://dx.doi.org/10.1109/HICSS.2016.677.

[14] D.R. Michael, S.L. Chen, Serious Games: Games That Educate, Train, and Inform, Muska & Lipman/Premier-Trade, 2005.

[15] M.R.d.A. Souza, L. Veado, R.T. Moreira, E. Figueiredo, H. Costa, A systematic mapping study on game-related methods for software engineering education, Inf. Softw. Technol. 95 (2018) 201–218, http://dx.doi.org/10.1016/j.infsof.2017.09.014.

[16] S. Frey, A. Rashid, P. Anthonysamy, M. Pinto-Albuquerque, S.A. Naqvi, The good, the bad and the ugly: a study of security decisions in a cyber-physical systems game, IEEE Trans. Softw. Eng. 45 (5) (2017) 521–536, http://dx.doi.org/10.1109/TSE.2017.2782813.

[17] J.M. Rojas, T.D. White, B.S. Clegg, G. Fraser, Code defenders: crowdsourcing effective tests and subtle mutants with a mutation testing game, in: 2017 IEEE/ACM 39th International Conference on Software Engineering, ICSE, IEEE, 2017, pp. 677–688, http://dx.doi.org/10.1109/ICSE.2017.68.

[18] M. Deen, R. Cercos, A. Chatman, A. Naseem, R. Bernhaupt, A. Fowler, B. Schouten, F. Mueller, Game jam: [4 research], in: CHI'14 Extended Abstracts on Human Factors in Computing Systems, 2014, pp. 25–28, http://dx.doi.org/10.1145/2559206.2559225.

[19] R. Ramzan, A. Reid, The importance of game jams in serious games, in: 10th European Conference on Games Based Learning: ECGBL 2016, Academic Conferences and Publishing International Ltd.(ACPI), 2016, pp. 538–546.

[20] M. Aibara, S. Kawakami, M. Furuichi, Lessons learned from serious game jams organized by DiGRA Japan, in: Abstract Proceedings of DiGRA 2020 Conference: Play Everywhere, 2020.

[21] L. Grace, Deciphering hackathons and game jams through play, in: Proceedings of the International Conference on Game Jams, Hackathons, and Game Creation Events, 2016, pp. 42–45, http://dx.doi.org/10.1145/2897167.2897175.

[22] G. Lai, A. Kultima, F. Khosmood, J. Pirker, A. Fowler, I. Vecchi, W. Latham, F. Fol Leymarie, Two decades of game jams, in: Sixth Annual International Conference on Game Jams, Hackathons, and Game Creation Events, 2021, pp. 1–11, http://dx.doi.org/10.1145/3472688.3472689.

[23] D. Abbott, O. Chatzifoti, S. Louchart, J. Ferguson, S. Stals, Serious 'slow' game jam: A game jam model for serious game design, in: Proceedings of the International Conference on Game Jams, Hackathons and Game Creation Events, ICGJ'23, 2023, http://dx.doi.org/10.1145/3610602.3610604.

[24] D. Abbott, O. Chatzifoti, J. Craven, Serious game rapid online co-design to facilitate change within education, in: Games and Learning Alliance: 10th International Conference, GALA 2021, La Spezia, Italy, December 1–2, 2021, Proceedings 10, Springer, 2021, pp. 233–238, http://dx.doi.org/10.1007/978-3-030-92182-8_22.

[25] S. Stals, L. Baillie, J. Ferguson, D. Abbott, M. Maarek, R. Shah, S. Louchart, Evaluating serious slow game jams as a mechanism for co-designing serious games to improve understanding of cybersecurity, ACM Games: Research and Practice. (2024) submitted for publication.

[26] A. Zook, M.O. Riedl, Game conceptualization and development processes in the global game jam, in: Workshop Proceedings of the 8th International Conference on the Foundations of Digital Games, Vol. 5, 2013.

[27] L. de Almeida Melo, T.H. Leite, F. Freire, M.G. Perin, F. Figueira Filho, C.R. de Souza, A.C.D. Batista, What motivates different people to participate in game jams? in: Anais Estendidos Do XV Simpósio Brasileiro de Sistemas Colaborativos, SBC, 2019, pp. 135–140.

[28] R. Eberhardt, No one way to jam: game jams for creativity, learning, entertainment, and research, in: Proceedings of the International Conference on Game Jams, Hackathons, and Game Creation Events, 2016, pp. 34–37, http://dx.doi.org/10.1145/2897167.2897181.

[29] C. Harteveld, Triadic Game Design: Balancing Reality, Meaning and Play, Springer Science & Business Media, 2011.

[30] E. McNeill, Darknet, 2014, URL: http://emcneill.com/press/sheet.php?p=Darknet. (Accessed 18 October 2023).

[31] Center for Infrastructure Assurance and Security, Cyber threat defender (CTD), 2024, URL: https://cias.utsa.edu/ctd/. (Accessed 18 October 2023).

[32] R. Shah, M. Maarek, S. Stals, L. Baillie, S.C. Chan, R. Stewart, H.-W. Loidl, O. Chatzifoti, Introducing and interfacing cybersecurity: A cards approach, 2024, doi:arXiv:2307.16535. under review.

[33] CyBOK, The cyber security body of knowledge (CyBOK), 2020, URL: https://www.cybok.org/. (Accessed 17 Febuary 2022).

[34] L.W. Anderson, D.R. Krathwohl (Eds.), A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives, second ed., Allyn & Bacon, New York, 2001.

[35] T. Lim, S. Louchart, N. Suttie, J. Ritchie, R. Aylett, I. Stanescu, I. Roceanu, I. Martinez-Ortiz, P. Moreno-Ger, Strategies for effective digital games development and implementation, in: Cases on Digital Game-Based Learning: Methods, Models, and Strategies, IGI Global, 2013, pp. 168–198.

[36] C. Harteveld, Triadic game evaluation: A framework for assessing games with a serious purpose, in: Workshop of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems, 2010.

[37] A. Fowler, F. Khosmood, A. Arya, The evolution and significance of the global game jam, in: Proceedings of the Foundations of Digital Games Conference, Vol. 2013, 2013.

[38] A. Kultima, Game jam natives? The rise of the game jam era in game development cultures, in: Sixth Annual International Conference on Game Jams, Hackathons, and Game Creation Events, 2021, pp. 22–28, http://dx.doi.org/10.1145/3472688.3472691.

[39] A. Fowler, G. Lai, F. Khosmood, R. Hill, Trends in organizing philosophies of game jams and game hackathons, in: GJ Workshop. FDG2015, 2015.

[40] J. Musil, A. Schweda, D. Winkler, S. Biffl, Synthesized essence: what game jams teach about prototyping of new software products, in: Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering-Volume 2, 2010, pp. 183–186, http://dx.doi.org/10.1145/1810295.1810325.

[41] G. Lai, F. Khosmood, What is a game jam? in: Game Jams–History, Technology, and Organisation, Springer, 2022, pp. 1–20.

[42] A. Fowler, F. Khosmood, A. Arya, G. Lai, The global game jam for teaching and learning, in: Proccedings of the 4th Annual Conference on Computing and Information Technology Research and Education New Zealand, sn, 2013, pp. 28–34.

[43] A. Fowler, Informal stem learning in game jams, hackathons and game creation events, in: Proceedings of the International Conference on Game Jams, Hackathons, and Game Creation Events, 2016, pp. 38–41, http://dx.doi.org/10.1145/2897167.2897179.

[44] J.A. Preston, Serious game development: Case study of the 2013 CDC games for health game jam, in: Proceedings of the 2014 ACM International Workshop on Serious Games, 2014, pp. 39–43, http://dx.doi.org/10.1145/2656719.2656721.

[45] Discord Inc., Imagine a place, 2023, URL: http://discord.com. (Accessed 13 January 2023).

[46] Miro, Take ideas from better to best, 2023, URL: https://miro.com. (Accessed 13 January 2023).

[47] D. Abbott, O. Chatzifoti, S. Louchart, Provocative games to encourage critical reflection, in: ECGBL 2022 16th European Conference on Game-Based Learning, Academic Conferences and Publishing Limited, 2022.

[48] S. Hart, NASA-task load index (NASA-TLX); 20 years later, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 50, Sage publications Sage CA: Los Angeles, CA, 2006, pp. 904–908, http://dx.doi.org/10.1177/154193120605000909.

[49] S. Hart, L. Staveland, Development of NASA-TLX (task load index): Results of empirical and theoretical research, in: Advances in Psychology, Vol. 52, Elsevier, 1988, pp. 139–183, http://dx.doi.org/10.1016/S0166-4115(08)62386-9.

[50] E. McAuley, T. Duncan, V.V. Tammen, Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis, Res. Q. Exerc. Sport 60 (1) (1989) 48–58, http://dx.doi.org/10.1080/02701367.1989.10607413.

[51] R. Ryan, Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory, J. Pers. Soc. Psychol. 43 (3) (1982) 450, http://dx.doi.org/10.1037/0022-3514.43.3.450.

[52] T. Faas, I.-c. Liu, L. Dombrowski, A.D. Miller, Jam today, jam tomorrow: Learning in online game jams, Proc. ACM Hum.-Comput. Interact. 3 (GROUP) (2019) 1–27, http://dx.doi.org/10.1145/3361121.

[53] R. Aurava, M. Meriläinen, V. Kankainen, J. Stenros, Game jams in general formal education, Int. J. Child-Comput. Interact. 28 (2021) 100274, http://dx.doi.org/10.1016/j.ijcci.2021.100274.

[54] A. Fowler, X. Ni, J. Preston, The pedagogical potential of game jams, in: Proceedings of the 19th Annual SIG Conference on Information Technology Education, 2018, pp. 112–116, http://dx.doi.org/10.1145/3241815.3241862.

[55] A. Arya, J. Chastine, J. Preston, A. Fowler, An international study on learning and process choices in the global game jam, Int. J. Game-Based Learn. (IJGBL) 3 (4) (2013) 27–46.

[56] P. Dugard, J. Todman, Analysis of pre-test-post-test control group designs in educational research, Educ. Psychol. 15 (2) (1995) 181–198, http://dx.doi.org/10.1080/0144341950150207.

[57] F. Bellotti, B. Kapralos, K. Lee, P. Moreno-Ger, R. Berta, Assessment in and of serious games: an overview, Adv. Hum.-Comput. Interact. 2013 (2013) http://dx.doi.org/10.1155/2013/136864.

[58] A. Bray, P. Byrne, M. O'Kelly, A short instrument for measuring students' confidence with 'key skills'(sicks): Development, validation and initial results, Think. Skills Creat. 37 (2020) http://dx.doi.org/10.1016/j.tsc.2020.100700.

[59] K.K. Silveira, S. Musse, I.H. Manssour, R. Vieira, R. Prikladnicki, Confidence in programming skills: gender insights from StackOverflow developers survey, in: 2019 IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings, ICSE-Companion, IEEE, 2019, pp. 234–235, http://dx.doi.org/10.1109/ICSE-Companion.2019.00091.

[60] Microsoft, Microsoft forms - easily create surveys, quizzes and polls, 2022, URL: http://forms.office.com. (Accessed 25 January 2023).

[61] M. Meriläinen, R. Aurava, A. Kultima, J. Stenros, Game jams for learning and teaching: a review, Int. J. Game-Based Learn. (IJGBL) 10 (2) (2020) 54–71.

[62] D.A. Plecher, C. Eichhorn, J. Kindl, S. Kreisig, M. Wintergerst, G. Klinker, Dragon tale-a serious game for learning japanese kanji, in: Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts, 2018, pp. 577–583, http://dx.doi.org/10.1145/3270316.3271536.

[63] R.F. Woolson, Wilcoxon signed-rank test, in: Wiley Encyclopedia of Clinical Trials, Wiley Online Library, 2007, pp. 1–3.

[64] S. Arnab, T. Lim, M.B. Carvalho, F. Bellotti, S. de Freitas, S. Louchart, N. Suttie, R. Berta, A. De Gloria, Mapping learning and game mechanics for serious games analysis, Br. J. Educ. Technol. 46 (2) (2015) 391–411, http://dx.doi.org/10.1111/bjet.12113.

[65] I. Mols, E. van den Hoven, B. Eggen, Informing design for reflection: An overview of current everyday practices, in: Proceedings of the 9th Nordic Conference on Human-Computer Interaction, NordiCHI '16, Association for Computing Machinery, New York, NY, USA, 2016, http://dx.doi.org/10.1145/2971485.2971494.

[66] B. Chinh, H. Zade, A. Ganji, C. Aragon, Ways of qualitative coding: A case study of four strategies for resolving disagreements, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, pp. 1–6, http://dx.doi.org/10.1145/3290607.3312879.

[67] A.D. Prabaswari, C. Basumerda, B.W. Utomo, The mental workload analysis of staff in study program of private educational organization, in: IOP Conference Series: Materials Science and Engineering, Vol. 528, IOP Publishing, 2019, 012018, http://dx.doi.org/10.1088/1757-899X/528/1/012018.

[68] D. Gundry, S. Deterding, Trading accuracy for enjoyment? Data quality and player experience in data collection games, in: CHI Conference on Human Factors in Computing Systems, 2022, pp. 1–14, http://dx.doi.org/10.1145/3491102.3502025.

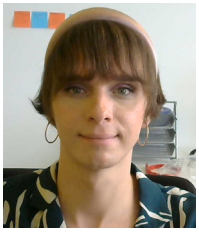
[69] D. Kao, R. Ratan, C. Mousas, A. Joshi, E.F. Melcer, Audio matters too: How audial avatar customization enhances visual avatar customization, in: CHI Conference on Human Factors in Computing Systems, 2022, pp. 1–27, http://dx.doi.org/10.1145/3491102.3501848.

[70] S. Uzor, L. Baillie, Recov-R: evaluation of a home-based tailored exergame system to reduce fall risk in seniors, ACM Trans. Comput.-Hum. Interact. 26 (4) (2019) 1–38, http://dx.doi.org/10.1145/3325280.
[71] M. Kordaki, A. Gousiou, Digital card games in education: A ten year systematic review, Comput. Educ. 109 (2017) 122–161, http://dx.doi.org/10.1016/j.compedu.2017.02.011.
[72] P. Sengers, K. Boehner, S. David, J. Kaye, Reflective design, in: Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility, 2005, pp. 49–58, http://dx.doi.org/10.1145/1094562.1094569.
[73] G.M. Troiano, D. Schouten, M. Cassidy, E. Tucker-Raymond, G. Puttick, C. Harteveld, All good things come in threes: Assessing student-designed games via triadic game design, ACM Int. Conf. Proc. Ser. (1) (2020) http://dx.doi.org/10.1145/3402942.3403010.

Dr. Shenando Stals is a Postdoctoral Research Associate in Human-Computer Interaction at Heriot-Watt University. He obtained a B.Sc. (1st Class Hons) in Computer Science & Software Engineering, and a M.Sc. (with distinction) in Human-Technology Interaction from the Eindhoven University of Technology, before moving to Edinburgh Napier University to obtain a Ph.D. in Computing. As a Research Associate on the SECRIOUS project, his research focuses on the use of Serious Slow Game Jams and serious games to engage new code-citizens on the topic of cybersecurity in software engineering. He co-organized the Deconstructing Gamified Approaches to Security and Privacy (DGASP) workshop at SOUPS 2023.

Prof. Lynne Baillie is a Professor at Heriot-Watt University in Human Computer Interaction (HCI). Her research focuses on the development of novel user-centred methods to facilitate the design and evaluation of new applications and technologies for pervasive and ubiquitous contexts. This work has been published over 100 times in peer-reviewed, high level conferences and journals, and she has a strong track record of over 15 years in research leadership at a senior level in two countries (UK and Austria) and has led or been part of successful teams that have been awarded funding of over £9.5M. She co-organized the Deconstructing Gamified Approaches to Security and Privacy (DGASP) workshop at SOUPS 2023.

Dr. Ryan Shah was a Research Associate on the SECRIOUS project. His expertise lies in cybersecurity. He obtained a B.Sc. (1st Class Hons) in Computer science at Heriot-Watt University before moving to the University of Strathclyde to obtain a Ph.D. in Cybersecurity, investigating the security of teleoperated robotics systems. His research interests include network security and privacy, useable security, robotics security and side channels. He co-organized the Deconstructing Gamified Approaches to Security and Privacy (DGASP) workshop at SOUPS 2023. He is now a security consultant at Sapphire.



Dr. Jamie Iona Ferguson is a Postdoctoral Research Associate in serious games design in Glasgow School of Art's School of Simulation and Visualisation. Their research interests are serious games (particularly applied to software engineering) and audio/haptic interaction.

Dr. Manuel Maarek is an Associate Professor at Heriot-Watt University, whose research focuses on advanced methods in software engineering and programming language design for building secure and safe systems. He has industrial experience in safety analysis and secure software development. His research has more recently focused on serious games approaches for developer-centred security. He co-organized a SICSA Cyber Nexus funded workshop on serious games for cybersecurity and co-organized the Deconstructing Gamified Approaches to Security and Privacy (DGASP) workshop at SOUPS 2023.