THE QUEST GAME-FRAME: BALANCING SERIOUS GAMES FOR INVESTIGATING PRIVACY DECISIONS

Patrick Jost

Follow this and additional works at: https://aisel.aisnet.org/scis2020
THE QUEST GAME-FRAME:
BALANCING SERIOUS GAMES FOR INVESTIGATING
PRIVACY DECISIONS

Research paper

Jost, Patrick, Norwegian University of Science and Technology, Trondheim, Norway, patrick.jost@ntnu.no

Abstract
Digitalisation permeates all areas of social life. The use of digital games in research settings to analyse social phenomena is thereby no exception. However, games that can successfully achieve research objectives and at the same time create an engaging experience require thoughtful balancing. When investigating decision-making, for example, asking players directly about their reasoning in the game is breaking the game flow and prone to distorting influences from the game experience.

This paper presents the design science (DS) process of a quest-based game-frame (QGF) oriented on the investigation of privacy decision-making. The design-empirical cycle of the QGF is outlined and applied to design two privacy decision scenarios for investigating reflection tendencies.

The conducted binational experiment reflects the behaviour of 78 educators, university students and high-school students from Austria and Norway in online ordering security and fake news sharing while monitoring the game flow.

Results demonstrate the potential of the QGF for unobtrusively investigating privacy decisions while maintaining high fluency of performance. Significant differences between educators and high-school students are found in time spent for reflection before making online security decisions. Additionally, Norwegian high-school students show a low awareness when deciding on real/fake news sharing.

Keywords: serious games, game-based assessment, decision making, privacy behaviour, design science research

1 Introduction and Serious Games Challenges

Although a universally accepted definition does not exist, a serious game (SG) is broadly viewed in the scientific community as a digital game created with the intention to entertain and to achieve at least one additional – characterising (or domain) – goal. Knowledge acquisition, skill/health improvement, behaviour change or raising awareness regarding a specific context are such frequently addressed domain goals (Dörner et al., 2016). However, to successfully engage and inform, the balance between game goal and domain goal is crucial. In general, there are multiple configurations of SG possible to address each of the goals, but when unevenly balanced the resulting SG experience is either tedious or fails to transmit the domain objectives (Iten & Petko, 2016). As Iten and Petko (2016) point out, playing and learning must be integrally connected and not just alternate during a game. Emergent emotions that are originating from the hedonic playing experience can thereby either support or impair the successful mediation of the pragmatic intentions. On the other hand, disruption from presenting learning content or reflection activities can break concentration and thus, the intrinsically motivating state of flow (Nakamura &
Csikszentmihalyi, 2009). Since reflection is considered indispensable for successful knowledge acquisition (Boud et al., 2013), Kiili et al. (2011) propose to align it in-game with conflicts, competition, performance/peer feedback and challenging comments from game characters. However, the balancing task between game goal and transmitting knowledge or skill is only the first component to consider for a balanced SG.

The second component is, nonetheless, inextricably linked with the first as it influences the level of challenge and originates from interacting with a game. A game is inherently interactive and must be controlled by actions of players and react on these inputs with changes in the presentation of the game world. However, both control and presentation influence the cognitive load of players (Sweller et al., 2011). While, for example, an immersive Virtual Reality game with natural hand-gesture control can help training motor skills (Piedra et al., 2016; Ren & Wu, 2019), it would introduce extraneous cognitive load when training mental arithmetic in a game for school children (Jost et al., 2019). Nevertheless, SG interaction design is affecting the balance of a SG configuration in more ways than cognitive load. User interaction is contributing not only to Ease-of-Use but particularly in games also a part of Joy-of-Use. It is interwoven with the game experience and triggers emotions. In fact, using intricate controlling mechanisms creates a challenge that can contribute to engagement in a SG when considering arcade-style games and player motivation types that perceive enjoyment from mastery (i.e. achievers) (Bourke et al., 2018; Yee, 2016). Motivation and engagement, on the other hand, are the key concepts in the prominent SG research field addressing improvements of behaviour and people’s choice-making (Chow et al., 2020). This points to the third component of balance in a SG configuration when applied in research environments – scientific evaluation.

Importantly, cognitive and motivational influences from interaction/presentation design are thereby only one part to regard for balancing when applying SG in a human-centred research environment. The other part originates from balancing an engaging game with scientific data collection without disrupting the game flow but just as well without confounding the collected data. Before applying behaviour adapting strategies such as nudging (Acquisti, 2009), the scientific process entails to better understand a phenomenon either by observing/protocolling behaviour or directly questioning persons. An excellent example of this is the research in privacy and data sharing behaviour which is of great concern, especially with children and younger adults (Wisniewski et al., 2015). Several aspects, such as the privacy paradox (i.e. people acting against better privacy knowledge) (Kokolakis, 2017) impact making decisions and risk-taking when deciding on personal data sharing. To better understand how privacy awareness and decisions can be improved, a SG represents a genuine approach to create decision scenarios and investigate choice behaviour/influences. At the same time, it can reach the young target group with an enjoyable experience that raises awareness about the essential issues.

However, to utilise SG as a research instrument, it is imperative to integrate strategies of scientific assessment that are flexible and adaptable to research design variations. For investigating effects of different/altered game characteristics, it is unavoidable to assess progress/outcomes in-game where often it also is critical to evaluate with the right timing or stage of game progress. A good example of this is character-based interventions in the game flow where different game characters (e.g. advice of a mentor that is female or male respectively) are representing distinct interventions in behaviour analysis. Consequently, this requires balancing of assessment instruments with game experience. Not least, it requires the integration of an adaptable research design in the SG while maintaining valid measurements that correspondingly measure the right effects and control for confounding influences.
Seamlessly integrating user experience, game experience, learning/reflection (i.e. domain experience) and scientific evaluation (research goal) is, therefore, the non-trivial task that research groups are required to accomplish when applying SG as a research approach. Figure 1 synthesises the aforementioned factors of balancing a SG for research application.

This paper addresses the outlined problem of balancing a SG for investigating privacy decision making while maintaining an engaging game flow. After defining the research objectives (2.1) current privacy literature and privacy-oriented games are reviewed (2.2). Subsequently, the cyclic design science approach (DS) proposed by Hevner et al. (2007) is followed by developing an open-world game environment – the Quest Game Frame (QGF) with two iterative design cycles and playtests. The developed quest-based structure is based on a 3D game engine that allows researchers to flexibly create and adapt real-world oriented 2D or 3D scenarios (3.1). In the following relevance cycle, the QGF is used to set up two privacy mini-quests, including a city-based environment, avatars, and narrative dialogues (3.2). Next, reflection tendencies in privacy choices between educators, university and high school students are investigated by an experiment with the QGF quests while monitoring the quality of flow (4). Finally, the analytical outcome (5.1), as well as the practical implications (5.2) and conclusions (6), are presented.

2 Research Objectives and Related Work

2.1 Research objectives

The created QGF quests were designed and iteratively improved to raise awareness about privacy concerns that pose a growing concern with ubiquitous digitalisation (Crabtree et al., 2017). Data is being collected everywhere, leaving people unaware of what data they share, with whom and what it is used for. Studying choices involving the use of personal data can be approached with game experiences that visualise or simulate what will happen with personal data. Such scenarios can be mediated with narratives that are familiar to the user. As outlined in the section above, not only the play/privacy objectives are central in this respect, but also how people reflect on their decisions and what influences them. Therefore, well-designed quest-based game scenarios addressing privacy should evoke system two modes of thinking (Evans, 2003), i.e. helping players to move from an automatic reaction when sharing data to consciously reflecting on the current problem, and deliberately anticipate potential future consequences of their behaviours. The design of a game quest should, in this regard, implement cues on the consequences of sharing private information or provoke moral evaluation of the conduct of apps and tech-companies (Böhm & Pfister, 2017). As outlined in the introductory section, reliably investigating the specifics of decisions but maintaining at the same an enjoyable narrative demands unobtrusive
strategies of evaluation without interrupting fluency and absorption of the game experience. Thus, our research objectives in this study are twofold:

1. Establish a quest game-frame in a DS approach for flexible construction of quest scenarios that enables researchers the unobtrusive investigation of privacy decision behaviour.
2. Create two quest scenarios that raise awareness about privacy decisions and allow the investigation of reflection tendency regarding privacy choices while maintaining an enjoyable game flow. Respectively, apply the designed quest scenarios in an experiment to test applicability.

2.2 Related work: privacy-focused serious games

The term privacy is not definable as a solitary concept. In fact, as Solove (2008) has outlined, it could rather be understood in the view of “family resemblances” than having a single common characteristic. Privacy entails many different but related aspects. A good way to approach this fact to raise awareness is, therefore, to shift focus away from the broad term to a problem-centric viewpoint. Raise awareness about activities that cause privacy problems resembles a promising practice-oriented approach. A taxonomy of activities creating privacy problems is mapped out by Solove (2008). Dissemination and processing are thereby the activities which present the most privacy concern. Both activities are related to control of flow and use of personal data as an essential aspect of privacy (Westin, 1968). Another aspect focused by Nissenbaum (2004, 2009) and connected to dissemination (i.e. sharing) is contextual integrity. Privacy concern might arise delayed in this sense, namely when shared information, for example, on social media, leaves a defined context. A matter that complicates the decision-making process of data sharing since weighing of trade-offs (i.e. consequences) becomes boundless. In effect, this complication of privacy calculus (Laufer & Wolfe, 1977) could lead to (automatic) system one decision-making resulting from overgrowing cognitive load. A SG could address this by demonstrating clear consequences in scenarios of this kind and create problem solving quests involving decisions along the same patterns.

Contemporary user studies and game-oriented research addressed some of these privacy aspects and aimed to raise awareness of involved concerns (Baxter et al., 2016; D’Apice et al., 2015; Malheiros et al., 2011; Vanderhoven et al., 2015). Findings from the studies provide good evidence on the ability of SG for raising awareness on privacy. Notably, raising awareness is not necessarily resulting in a change of behaviour in privacy decisions. Several studies refer to this with the term privacy paradox (Barth & de Jong, 2017; Gerber et al., 2018; Kokolakis, 2017). However, before changing behaviour with concepts such as, for example, nudging (Acquisti, 2009) it is a required to understand (i.e. evaluate) how people are deciding and what factors do contribute to the decisions which are the focus of this study and the QGF. Looking at specific game experiences for raising awareness, examples such as Friend Inspector (Cetto et al., 2014) deal with the social network scenario by showing discrepancies between perceived and actual visibility of shared information. Taking privacy awareness into the Smart City context is PrivaCity (Berger et al., 2019). A game based on chatbot conversations that includes approaches such as dialogue-oriented quizzes to raise players attention on privacy issues. Several data security topics are taken up by Google’s Interland (Seale & Schoenberger, 2018). The SG is composed of mini quests presented on islands that each present challenge that educates about internet security.

Several other SG, however, represent a single effort with many addressing phishing as the characterising goal and either aimed at enterprise (CJ et al., 2018) or individual use (Misra et al., 2017; Weanquoi et al., 2019). Examples given are featuring a story-based approach with either artificial aquatic theme or bird/worm theme. If a SG should be incorporating a fish theme in this regard is, however, questionable. More likely, it should focus on real-world decision-making, where players are actually having conversations. The findings of Berger et al. (2019) support real-life orientation, as real examples of data protection issues have been perceived as helpful by stakeholders. An approach that was also pursued by Wen et al. (2019) by creating roleplay/simulation scenarios. The phishing topic is understood as cybersecurity activity by all these authors and might be popular because of the fun name “phishing”. However, additionally, it has an inherent training character that allows for an obvious game-loop that presents players iteratively with different items for improving the recognition of deceptive information. A recent SG on mobile security awareness that addresses decision-making – “What could go wrong” – is as well
incorporating this training character (Zargham et al., 2019) and tries to incorporate humour as a game element. By integrating a “time machine”, players can revert to a decision node and decide differently. A feature that is promoting training but on the other hand, might reduce suspense and reflection at the point of decision. Nonetheless, SG mechanisms for increasing and solidify domain knowledge are to be encouraged since they lead to improved privacy decision-making when regarding the privacy calculus.

The outlined research on SG in privacy awareness is suggesting a strong focus on narrative experiences. In fact, when turning now to story-driven games in current literature it shows that adventure games are generally a popular SG type as systematic reviews observed (Boyle et al., 2016; Connolly et al., 2012). Recent studies have employed quest like SG in such diverse domains as engineering education within a 3D environment (Morsi & Mull, 2015), teaching about information literacy (Kwak et al., 2018), training of computational thinking (Gossen et al., 2018), drug discovery and development (Chang et al., 2018) but also computer assembly (Hou & Li, 2014) and security (D’Apice et al., 2015). When analysing the mentioned studies, it can be noted that they are closely related to real-world workflow simulations. A non-surprising fact since systematic mappings confirm educational simulation games and roleplaying games are an even more commonly applied game type (Boyle et al., 2016; Connolly et al., 2012) specifically when training is involved (Graafland et al., 2012). The reasons for this seem obvious since SG are aligned on real-life scenarios and problems and aim to transmit knowledge on them, train them or raise awareness on issues surrounding them. Therefore, it seems obvious to simulate the circumstances in the game world and take on the role that should deal with the matter. Even more so when interested in the decision tree that evolves from progressing through a game. Adding to the usefulness of such games of progression (Juul, 2005) is that the creators usually have control over the events as compared to games of emergence where rules (or physics) and chance are influencing the action unpredictably. But these pragmatic advantages are to be balanced with the hedonic qualities such as curiosity and surprise (Grodal, 2000; Malone, 1981; Provenzo Jr, 1991) that contribute to the game experience and subsequently engagement and flow state to avoid the classic pitfall of a dreary learning/evaluation “game” (Berger et al., 2019; Van der Spek et al., 2013). If narration is contributing beneficially to learning in a SG is subject to discussion (Adams et al., 2012). As Ravyse et al. (2017) note in this regard that aligning narration and learning goal is a success factor which equally applies to the research goal when evaluating with SG. This is particularly important since the matter of distraction could influence cognitive load and thereby a valid scientific assessment.

Creating authentic in-game scenarios with decision-making quests can be considered as an approach to keep a close relationship to a real-world decision process. A good example of this is the study of D’Apice et al. (2015) who created a game to teach principles of information security with a playable avatar that solves problems and makes decisions with the help of a mentoring character. Characters could thereby not only serve as a story and educational ingredient but as well be utilised as an investigative entity. When balanced to the story and not breaking the game experience, a character could pose questions related to decision-making or influences thereof. Additionally, D’Apice et al. (2015) integrated other elements that potentially support reflecting and learning, such as a journal that logs all quests, cutscenes, interactions with characters and in-game objects. All these ingredients can as well be applied when using a quest-based SG as a research apparatus. The proposed QGF provides an environment that can integrate and test corresponding hypotheses and conditions concerning decision situations and reveal influences or potentially conflicting configurations between SG components. However, in terms of achieving and not breaking the game flow, the evaluation of decision-making must be aligned with the narrative when linked to game entities. Furthermore, unnoticeable strategies of data collection need to be integrated when looking at an objective assessment of different decisions. D’Apice et al. (2015) implemented in-process assessment (Michael & Chen, 2005) which involves event triggers on in-game objects or variables. Real-time adaption can then inform the player for example, through the mentor character and the game can be balanced to the players progress. How this approach can be instead configured to unobtrusively investigate privacy decision-making presents a further research challenge worth investigating with QGF quests.
Similar strategies of stealth assessment are proposed by Shute (2009) and Ke and Shute (2015) by utilising data mining and analytics or Bayesian Networks, respectively. While these studies investigated the utilisation on learning evaluation, they can be valuable strategies when utilising the SG as a decision-making research instrument. Equally, experience sampling (e.g. Berkel et al., 2017) which is linked to flow state would be suitable as a “stealth” approach to test in decision research with the QGF. Analysis of contemporary related work suggests that a quest-based approach is promising when designing SG for raising awareness or learning on a specific domain. Related studies focus on the link between motivation and learning, assess the impact of both and the flow experience. However, the utilisation of game entities and SG configurations that simultaneously allow for research on decision-making in, for example, privacy-related topics is basically unaddressed at this point.

3 Research Design

3.1 The Quest Game-Frame

Based on the contemporary SG literature outlined above, a DS research approach is pursued to develop a modular, quest-based SG framework for studying privacy decision-making in relevant, real-world scenarios. Theoretical foundations of DS were laid out by Simon already in 1969 and progressed by Cross (2006). Further development has been revolving around concepts from Peffers et al. (2006), Hevner et al. (2007; 2004) and various other contributions (Baskerville, 2008; Baskerville et al., 2016, 2018; Gregor & Hevner, 2013). DS is inherently a practical approach since design activity is connected to users and context to create applied artefacts in mostly case-driven studies.

In the presented “design-empirical” quest cycle, the QGF is defined as a new approach for investigating decision-making with enjoyable game experiences based on identified requirements from the discussed SG literature. Subsequently, two decision-making scenarios are created to demonstrate applicability in empirical evaluation. Figure 2 illustrates the iterative DS approach adapted to QGF requirements.

![Figure 2](Exploring, explaining, and prescribing through quest cycles (design | rigor | relevance) – adapted from Hevner (2007))

The QGF structure is designed to represent a decision-making research setting informed by privacy awareness/game studies literature and flow theory (sections 1 and 2). The initial configuration thus comprises the following identified requirements oriented at the application in research:

a. **Scalable open-world frame** that can host and be extended with modular quest-oriented decision scenarios modelled on real-world decision processes.

b. **Component-based evaluation approach** that enables unobtrusive or story-aligned evaluation of decisions through in-game entities (e.g. characters, dialogues) and data collection through stealth assessment procedures such as data logging/analytics, experience sampling, Bayesian networks or digital questionnaires/scoring interwoven with the game narrative.
Jost, P. / The Quest Game-Frame: Investigating Privacy Decisions

c. **Flexible configuration and placement of mini-quests/evaluations** with event triggers and adaptability of game-assets (e.g. change of characters, change of narrative or timing)

d. **Easily extendible/adaptable game elements and visuals** that encompass a variety of possibilities for situation design, rewarding, inducing enjoyment/flow and balancing game/domain experience

e. **Multi-platform distribution of SG quests** to all relevant game platforms addressing desktop, mobile and game consoles with varied interaction paradigms (e.g. touch, keyboard, game controller)

f. **Secure DB communication, in-game informed consent and display of data privacy statement**

In order to meet the technical requirements and allow for fast development in a research project, known game development engines have been screened and tested. The Unity engine, which has been used in a research context for more than a decade (Craighead et al., 2008) and is therefore familiar to researchers, offered several additional advantages. Compared to non-game engine choices Unity contains a physics engine, an editor for rapid prototyping/collaboration in research teams, bundled assets and pre-scripted 3D objects. Advantages over other game engines are primarily the extensive asset store with free assets, resource optimised mobile rendering and wide-ranging cross-platform support (Gregory, 2018). Unity projects thus provide a basis for the listed requirements with the component-based structure (a, b, c), a pool of free and customisable assets (d) and multi-platform build options (e). Additionally, secure DB communication (f) is facilitated with prepared libraries for secured web requests and asset store components. The modular system architecture for the QGF is subsequently patterned for empirical application on the research of privacy decisions and illustrated in Figure 3.

![Figure 3. Component architecture of the QGF for investigating privacy decisions](image)

In order to address the different areas of privacy issues and create relevant choice scenarios, a segmented approach is proposed by arranging the open world in cities. Each city represents a key problem area with challenges that address privacy problem-solving. For example, the “City of Knowledge” where privacy calculus and the knowledge about data regulation and rights such as GDPR is the focused area. Other cities would similarly represent the “City of Deception” (Bösch et al., 2016) or the “City of Data Security” (Kumar et al., 2018). Contributing to the modular and extensible nature of the QGF this creates numerous opportunities for linking privacy to everyday experiences when, for example, addressing privacy decisions involved in Smart Cities (Berger et al., 2019; Farahat et al., 2019), Big Data/Cloud Computing or IoT (Benjelloun & Lahcen, 2019; Dabbagh & Rayes, 2019; Kumar et al., 2018).

However, most importantly, it connects several cognition and motivation specific considerations. First, it structures the information about privacy problems in essential areas. Second, it allows for mini-quests that link privacy decisions to the overarching topic and at the same time meets suggestions of Miller (1956) and Sweller (2011) regarding mental load when segmenting (i.e. chunking) and limiting an awareness cycle to, for example, five mini-quests for each city. Third, it contributes to rewarding and achievement (e.g. exploration and mastery) (Bjork & Holopainen, 2005; Heeter et al., 2011) when a city
(i.e. topic) can be claimed as conquered. Finally, it enables a level-oriented increase in challenges and thus the maintenance of the flow state with a structuring unit known from real life. Besides, it creates opportunities for taking breaks in the gameplay and not least saving game progress also regarding research goals (e.g. saving decision outcomes or data logging to the database). Even finer structuring in Point of Interest (POI) triggers that are located in the city can thereby put people in everyday, demanding or exciting decision situations. Correspondingly, saving of each mini-quest or saving of sub-data points (e.g. experience sampling) is recommended to optimise data collection for analysis. This approach of collecting mini-quests in the city is providing opportunities for engagement and rewarding while creating appeal through collecting (Bjork & Holopainen, 2005). In this first design/empirical cycle, the focus was set on data sharing decisions. Accordingly, the SG artefact was named “Conquest of Shareadise” (CoS) and the created mini-quests were located in the “City of Data Sharing” (Figure 4).

![Figure 4. “City of Data Sharing” – POI decision quest trigger](image)

### 3.2 Empirical Research Approach

For empirically investigating privacy decisions, two privacy quests were created with the QGF. First, a data-sharing scenario when meeting a friend at the bus station. It involves online ordering and deciding about several important privacy behaviours during the task. Second, a news headline scenario with fake and real news where players decide on their likeliness of sharing the headline with peers.

![Figure 5. Quest A: Deciding about sharing of personal data / Quest B: Deciding about sharing of news](image)

The first quest [A] was triggered when the player arrived at the bus station and met a friend there. A dialogue-driven scenario brought the player in a decision situation about private data and personal data dissemination. The decisions were thereby ranked from low to high with respect to disseminating personal data or enlisting your name into promotion databases. The scenarios featured a dialogue with a female friend to drive the story and decision process. Three choice situations were presented with each encompassing three options to decide. Table 1 lists example decisions with weighted awareness.
Jost, P. / The Quest Game-Frame: Investigating Privacy Decisions

Table 1. Evaluation model of personal data decisions in Quest A

<table>
<thead>
<tr>
<th>Private data decision example</th>
<th>Awareness</th>
<th>Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create online account for ordering food with your personal e-mail address</td>
<td>Low</td>
<td>1</td>
</tr>
<tr>
<td>Create online account with another mail address not used for private communication</td>
<td>Medium</td>
<td>3</td>
</tr>
<tr>
<td>Create a temporary e-mail account that is deleted after ordering and register with it</td>
<td>High</td>
<td>6</td>
</tr>
</tbody>
</table>

The second quest [B] was triggered when the player passed a newspaper rack in the city. Six news headlines were presented that could be shared with friends. The player was informed by a short text to evaluate how likely s/he would share the news headline with her friends. A hint was shown that not all news might be real. The fake news headlines represented the most shared fake crime news on Facebook from 2016 (according to buzzfeednews.com), and the real headlines were taken from the independent.co.uk at the time of the study. Range of the likeliness for sharing could be set between 1 and 10 on a sliding bar with guiding text-marks: don’t share (1), likely share (5) and definitely share (10). Figure 5 shows both decision scenarios, dialogue, and news headlines examples. The QGF was used to place POI triggers in the city, create the scenarios/dialogues plus user-interface elements and set-up for data-logging which provided an evaluation of reflection time between decisions and total decision time for news sharing. For evaluating flow experience, a digital rating possibility of the 10-item flow short scale (FSS) (Rheinberg et al., 2002) with a slider rating from 1 to 7 was presented to the players following the last quest and could optionally be provided as feedback. The SG “Conquest of Shareadise” was built for mobile use on the Android platform and web browser with WebGL 2.0. No personal data was assessed; all players were prompted with information on data collection and only advanced to playing after informed consent.

Using the QGF quests, privacy decision-making differences between researchers/educators, university students and high-school students were analysed. Additionally, it was assessed if the game experience can achieve/maintain flow while it is used as a research instrument. The QGF is addressing players starting from the age of 15 and counts researchers and educators as primary stakeholders and users. In order to involve most of this target group and to get an international perspective, actors were recruited at a Norwegian and an Austrian university, either by e-mail invitation (educators/researchers), in a workshop or a lecture (students). For high-school students, informed consent was also obtained from their parents before taking part in the workshops. The hypothesis was put forward that teachers/researchers should be more experienced in data sharing decisions and more careful in exchanging fake messages than students and therefore differences between the three groups were expected. Ambient music, overarching or curious story, direct feedback, collectables or rewarding mechanics were not yet implemented, and both quests together had a playing time of only about 5 minutes. Therefore, an achievable average flow value of 3.6, which is just above mid-scale, was anticipated for general flow as well as the two sub-factors “fluency of performance” and “absorption by activity”. Reference values of the flow short scale show that the maximal flow experience was found in graffiti spraying with an average of 5.16. Thus, the null hypotheses established for the empirical investigation were:

\[ H_{0A}: \text{There are no significant differences between the groups educators/researchers, university students and high-school students in making decisions about private data sharing.} \]

\[ H_{0B}: \text{There are no significant differences between the three groups in making decisions about sharing real and fake news headlines.} \]

\[ H_{0C}: \text{The created QGF SG is not achieving a flow experience significantly above an average of 3.6.} \]

A total of 78 participants from two research groups, two university classes and two high-school workshops were taking part in the study and played the created quest game. Participating researchers were also teaching classes, so they were grouped researchers/educators. They played the game in the web...
browser just as the university students and mostly on notebooks. High school students played the game as part of a SG workshop where some used the touch controls/tablet computers or smartphones and others used notebooks/keyboard controls. All participants played individually. Four participants were excluded from the study since the data logging revealed inconsistent data such as being a high school student and being over 60 years old. For specific comparisons, single cases were excluded since, for example, a decision time of over 3 hours for a decision was indicating taking a break during the game quest. Table 2 lists the distribution and demographic attributes of participants related to the two decision quests and the FSS response. The FSS number of participants was considerably lower since it was not collected from high school students and not obligatory at the end of playing the game.

<table>
<thead>
<tr>
<th>Age in years</th>
<th>High School</th>
<th>Educators</th>
<th>University</th>
<th>female</th>
<th>male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>15.96</td>
<td>36.50</td>
<td>24.29</td>
<td>24.85</td>
</tr>
<tr>
<td>Sharing private data</td>
<td>valid N</td>
<td>24</td>
<td>24</td>
<td>26</td>
<td>38</td>
</tr>
<tr>
<td>Sharing fake news</td>
<td>valid N</td>
<td>23</td>
<td>26</td>
<td>26</td>
<td>39</td>
</tr>
<tr>
<td>Flow experience (FSS)</td>
<td>valid N</td>
<td>-</td>
<td>18</td>
<td>12</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 2. Participant distribution and demographics

4 Findings

Since the data collected on the two decision quests was not consistently normally distributed, a conservative non-parametric analysis was performed (α=0.05). A Kruskal-Wallis (1952) test for independent samples with subsequent pairwise testing (Dunn-Bonferroni) thereby revealed (adjusted) significant differences in the decision-making on private data security [A] and sharing of real/fake news [B] (Dunn, 1961). When first turning to the private data quest [A] weighted awareness scoring was not different over the three groups, \( H(2) = 0.11, p = .947 \). However, the time spent on deciding how to configure the password for the online account differed significantly, \( H(2) = 16.63, p < .001 \) between educators and high school students \((z = 3.96, p < .001, r = .57)\) with a large effect size as well as between university and high school students \((z = 2.86, p = .013, r = .40)\) with a medium effect size (Cohen, 1988; Field, 2009). The same applies to the decision to register to receive promotional material, \( H(2) = 15.11, p = .001 \) where the difference \((z = 3.72, p = .001)\) shows again larger effect \((r = .54)\) between educators and high school students than between the latter and university students \((z = 2.88, p = .012, r = .41)\). Concerning password security behaviour, high school students were averagely reflecting 17.7 seconds more than educators and 15.7 seconds more than university students (Table 3). Analysing decision-making in the second quest [B] exposes interesting distinctions between educators, high school, and university students. Combined time spent on deciding for sharing real/fake news headlines, \( H(2) = 9.91, p = .007 \) and actual sharing of fake news, \( H(2) = 6.36, p = .045 \) differed both significantly. High school students did on average reflect almost double the time when deciding about how likely they would share the six news headlines compared to educators \((z = 2.87, p = .012, r = .41)\) and university students \((z = 2.62, p = .027, r = .37)\).

<table>
<thead>
<tr>
<th>Reflection time before deciding on password security</th>
<th>Mean time mm:ss</th>
<th>High School</th>
<th>Educators</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Error of Mean</td>
<td>00:04:44</td>
<td>00:03:4</td>
<td>00:02:2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reflection time before deciding on subscription for promotion</th>
<th>Mean time mm:ss</th>
<th>High School</th>
<th>Educators</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Error of Mean</td>
<td>00:04:9</td>
<td>00:02:8</td>
<td>00:03:1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Combined reflection time deciding on sharing real/fake news headlines</th>
<th>Mean time mm:ss</th>
<th>High School</th>
<th>Educators</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Error of Mean</td>
<td>01:52:5</td>
<td>01:01:6</td>
<td>01:03:2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tendency to share fake news headlines [min 3; max 30]</th>
<th>Mean likelihood to share</th>
<th>High School</th>
<th>Educators</th>
<th>University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Error of Mean</td>
<td>1.59</td>
<td>1.54</td>
<td>1.52</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Significant mean differences between groups when reflecting about sharing of private data and real/fake news headlines
Moreover, they were significantly more susceptible to decide on sharing fake news headlines ($z = 2.44, p = .045, r = .35$) than educators despite the longer reflection time. As illustrated by Figure 6 high school students shared on average 62% ($M = 10.35$) more likely fake news headlines with their peers than educators ($M = 6.42$) although spending almost 84% ($M = 1:52$ minutes) more time reflecting before deciding on sharing compared to educators ($M = 1:01$ minutes).

![Figure 6](mean_reflection_time_tendency_to_decide_sharing_fake_news.png)

Inferential examination regarding the third hypothesis about players achieving a flow perception above a 3.6 rating was conducted with a one-sample $t$-test since data was normally distributed. Analysis showed the quest-based game is significantly $t (30) = 2.39, p = .012, r = .43$ achieving a general flow perception above 3.6 average. Players thereby experienced fluency of performance also significantly above a mean of 3.6, $t (30) = 3.99, p < .001$ with a large effect size ($r = .72$) but did not feel absorbed of the activity, $t (30) = -.569, p = .287$ as this sub-component of flow was not significantly rated above 3.6 (Table 4). Consequently, analysis suggests rejecting all three null hypotheses (H0A, B, C) as testing revealed significant differences among groups in decision-making and an above-average flow experience is stimulated by the created quests.

<table>
<thead>
<tr>
<th></th>
<th>valid $N$</th>
<th>$M$</th>
<th>SE</th>
<th>95,0% CI</th>
<th>95,0% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total flow</td>
<td>31</td>
<td>4.16</td>
<td>.23</td>
<td>3.68</td>
<td>4.64</td>
</tr>
<tr>
<td>Fluency of performance</td>
<td>31</td>
<td>4.64</td>
<td>.26</td>
<td>4.11</td>
<td>5.17</td>
</tr>
<tr>
<td>Absorption of activity</td>
<td>31</td>
<td>3.44</td>
<td>.29</td>
<td>2.84</td>
<td>4.03</td>
</tr>
</tbody>
</table>

Table 4. Players mean experience of flow and fluency/absorption when playing Conquest of Shareadise

5 Discussion

5.1 Interpretation of results

Empirical testing of the established QGF with two quest-based evaluation scenarios successfully demonstrated applicability for decision investigation. Results regarding decision specifics and differences between educators and high school/university students were showing the expected experience advantage of educators when making decisions. Specifically, the differences found in reflection time before deciding on privacy data sharing and sharing of fake news between students in high school age and researchers/educators with a mean age of 36.5 in this study were striking. However, in the case of the private data sharing quest, it did not affect the weighted awareness score. This seems obvious since educators can draw from experience and need less time to reflect before deciding. Even though these time differences have to be put into perspective, because news headlines have to be read and interpreted first before
a reflection can start, it can be assumed that the difference in reflection is substantial since headlines consist only of one short sentence. Students did reflect nearly twice as long on whether to share specific headlines, yet they were still almost twice as susceptible to share fake news. This finding should encourage researchers and educators to explore the factors involved in educating students with appropriate measures and tools to detect false information and data security pitfalls.

Although the utilised QGF quests did manage to achieve an above general average flow state score, the mini-quests are still a long way from being engaging experiences. This becomes obvious when looking at the subcomponents of flow. Fluency of performance was reported as high by the participants, which can be explained through having met a good level of challenge but must be confirmed with the next relevance cycle and improved privacy quests. Obviously, the time has a crucial influence in this regard since the whole “adventure” was completed between 5 and 10 minutes, and it could be that players attributed fluency to ease of completion. In any case, the low score in absorption of activities which was below average is related to this shortness of experience. To get into a flow state, it needs not only the right balance of game components but also an experience long enough that it allows for time passing unnoticed.

5.2 Implications for design and further research

The application of the QGF in this empirical study has shown the applicability of the approach. The expected results suggest that there was no substantial interference or extraneous cognitive load (Sweller et al., 2011) that would have distorted reflection time during decision making. Interaction and presentation were, however, kept very simple for this relevance cycle. Future scenarios should include, for example, a progressing story, interacting characters, and more advanced audio-visual presentation to examine how reflection results are impacted by engaging SG elements. More elaborated quest scenarios can then be applied for researching influential factors on the decision-making outcome. In this regard, cognitive biases (Waldman, 2019) or nudging strategies (Acquisti, 2009) can be modelled in QGF scenarios for investigating their influences on decision-making results.

When reflecting the experiences from this study with SG literature (section 2.2) several implications for the improvement of quest-oriented SG applied in decision research can be identified:

- An open-world environment supports the investigation of multiple decision scenarios while keeping a reasonable fluency in events
- Structuring in mini-quests helps grouping into research topics and allows variations in research design as well as progressive feedback to participants and stepwise storage of evaluation data
- The mini-quest structure facilitates inclusion, exclusion, and extension of research questions and thus the balancing of game/domain and research goals
- Using a game engine that supports effortless switching between 2D and 3D presentation helps implement unobtrusive questioning and clear consent forms
- Multi-platform support makes it possible to reach the target population efficiently while investigating or controlling the influences of different interaction paradigms

However, as pointed out in the introduction and discussed in the related work section, there are several more SG balancing factors than structure, presentation, and interaction. Most important are evaluation strategies that do not interfere with the investigated decision situation. The designed QGF scenarios did apply event-triggered data logging of time variables as main unobtrusive assessment of reflection. This concept can be extended to quest-based logging of character interaction (D’Apice et al., 2015; Michael & Chen, 2005) or combined with experience sampling (Berkel et al., 2017) for more advanced decision analysis. When aligned to the quest narrative, questioning by a non-playing character, for example, can provide explanatory insights to measured reflection times.

Further investigation concepts could look at team decision-making to explore shared mental models (Kocher et al., 2020; Langan-Fox et al., 2001) or employ in-game interrogations at the time of decision
making with the researcher taking part as game character. Another essential aspect to address is to what extent making decisions in quest-based games relates to real-world decision making. Naturally, some risks (e.g. those involving life-threatening outcomes) can hardly be investigated rigorously in a game world. Besides that, more studies on SG for researching decisions are needed for a better insight in:

- The influences between game/domain experience and scientific assessment
- Discovery and recording of emerging effects in the long-term play of a SG
- Further development of unobtrusive in-game assessment on decision-making
- Concepts and guidelines for aligning assessment with game entities (e.g. characters)
- Artificial intelligence with respect to in-game scientific evaluation
- Influences of interaction paradigms and device usage on SG scientific evaluation
- Creating flexible tools and frameworks for researchers to facilitate the development of research games without game expertise

6 Conclusion

The QGF, as a suggested approach for SG decision research, was applied in an empirical trial to investigate privacy decision-making and raise privacy awareness. In the design science cycle, after identifying requirements from the literature, an architecture was developed to create modular, question-based games for the exploration of decisions or other behavioural scenarios. The presented empirical trial demonstrated applicability of the QGF with creating expectable results of decision-making differences between educators and students while successfully maintaining a sense of fluency. The empirical study revealed the susceptibility of high school students to sharing fake news headlines despite reflecting on the decision almost twice as much than educators/researchers. This result should be a starting signal – if solutions are not already on the road – for teachers, parents, and authorities to involve privacy awareness and data security behaviour as a teaching subject and activity. Although phenomena such as the privacy paradox and mechanisms as the privacy calculus are known to the scientific community and various SG are produced on the topic, there is much more research and engagement needed on revealing how awareness and decisions in privacy matters can effectively be improved among children and teenagers. One approach is utilising the proposed QGF to create experiences for researching and teaching while at the same time creating enjoyable experiences for the target audience. This study identified the relevant requirements, demonstrated the applicability of the generated artefacts, and initiated the relevance cycle with implications for improving SG for decision research purposes. Future work should extend and refine these suggestions but just as well address the discussed SG challenges.

Acknowledgements. The research is funded by the NFR IKTPLUSS project ALerT, #270969. We thank all players who joined the evaluation and the supervisors of the research project with distinct thanks to Monica Divitini.

References


