12-7-2022

How to increase ethical awareness in cybersecurity decision-making

Bakhtiar Sadeghi
Macquarie university, bakhtiar.sadeghi@hdr.mq.edu.au

Deborah Richards
Macquarie university, deborah.richards@mq.edu.au

Paul Formosa
Macquarie university, paul.formosa@mq.edu.au

Mitchell McEwan
Macquarie university, mitchell.mcewan@mq.edu.au

Michael Hitchens
Macquarie university, michael.hitchens@mq.edu.au

How to increase ethical awareness in cybersecurity decision-making

Full research paper

Bakhtiar Sadeghi
School of Computing
Macquarie University, Australia
Email: bakhtiar.sadeghi@hdr.mq.edu.au

Deborah Richards
School of Computing
Macquarie University, Australia
Email: deborah.richards@mq.edu.au

Paul Formosa
Dept of Philosophy
Macquarie University, Australia
Email: paul.formosa@mq.edu.au

Mitchell McEwan
School of Computing
Macquarie University, Australia
Email: michael.mcwean@mq.edu.au

Michael Hitchens
School of Computing
Macquarie University, Australia
Email: michael.hitchens@mq.edu.au

Abstract

Cybersecurity technologies offer secure channels to enable the confidentiality, integrity, and availability of data and services. Human factors; e.g. demographics, personality traits, and human values, which are linked with greater cybersecurity vulnerabilities, have drawn less attention. It is important to understand how to increase ethical awareness for cybersecurity professionals via training. This ethical awareness helps professionals make better moral judgments prior to final decisions and reduces the risk of unexpected human implications. To sensitise players to five cybersecurity ethical principles (beneficence, non-maleficence, justice, autonomy, and explicability), we created a serious game. This game allows players to explore multiple cybersecurity scenarios based on these five cybersecurity ethical principles. Although the analysis does not support the claim that the game increased ethical awareness in general, it did help promote better ethical understanding in some cases where players advanced from providing non-ethical to ethical justifications in a cybersecurity scenario after playing the game.

Keywords: ethical training, serious games, cybersecurity ethical decision-making, ethical fading

1. Introduction

The purpose of cybersecurity systems is to provide a secure channel to transmit and protect users' data and services against any unsolicited access or breach. On the one hand, there are massive research and development resources invested in improving cybersecurity technology and systems. On the other hand, there is a vulnerability gap related to ethics in cybersecurity (e.g., around ethical fading (Bazerman 2011) that has received far less attention and effort than the technical side of computer systems. This vulnerability gap relates to various stakeholders, from behaviours of the end-users of a system to the cybersecurity domain experts who make decisions about system design and implementation. The professionals who make cybersecurity decisions need to understand the human aspects (e.g., social engineering (Mann 2008)) of the solutions they design. Otherwise, the solutions may have unexpected impacts that professionals are unaware of. For example, when system administrators push the deployment of two-factor authentication (2FA) on software to be authenticated only via a smartphone, they can cause issues and even harm to some users, such as those with a disability or those who lack access to smartphones. Cybersecurity professionals also need to be aware that despite their implementation of policies, procedures and technology to protect IT infrastructure and data, personal values can override norms, resulting in end-user decisions to breach policies and circumvent procedures (Christen et al. 2017; Schwartz 2012). To address the vulnerability gap due to human behaviour, the system administrator should receive adequate ethical cybersecurity education to raise their awareness of possible ethical implications in designing cybersecurity solutions.

There are ethical training solutions intended to close the value-action gap (Narvaez 2005), such as live role playing games, which can be online or in person. These solutions, however, rely on players/trainees being available at the same time and location, which is costly to organise, hard to scale up, doesn't allow for anonymity, and can require players to have expert domain knowledge which is difficult to obtain. Thus traditional solutions make it difficult to expose many trainees to real situations with significant ethical implications (e.g. human factors, roles and skills) (Gee 2007). Serious games for ethical training is one of the promising approaches that can offer cybersecurity professionals the opportunity to privately and conveniently explore ethical decision-making dilemmas (Lapsley 2005). It provides professionals a playground to make decisions and experience ethical implications without fear of a risk of serious consequence (Staines et al. 2017). While a number of such games have been developed for cybersecurity training, limited attention has been paid to their effectiveness and ethical principles.

To train players about the influence of human values on cyber-ethical decision-making, we created a game (Ryan et al. 2022) that first sensitises the player to five ethical principles (detailed below). The game also enables the player to navigate multiple cybersecurity ethical scenarios to understand the application of these five ethical principles in different contexts. The aim of this serious game is to sensitise players to the factors that influence their own and others' ethical decision making related to cybersecurity (Ryan et al. 2017). We used Schwartz's theory of basic human values (Schwartz 1994) to study human values and its implications on cybersecurity decision-making. In addition, in order to connect the study of human factors with cybersecurity ethical decision-making, we leveraged a principlist framework for cybersecurity ethical decision-making (Formosa et al. 2021). This framework consists of five ethical principles: beneficence (cybersecurity technologies should improve human lives), non-maleficence (cybersecurity technologies should not harm individuals' lives), justice (cybersecurity technologies should improve fairness and provide impartial access for all), autonomy (cybersecurity technologies should not limit users' choices) and explicability (cybersecurity technologies should be both understandable and accountable clearly), which are described in the following literature review section.

2. Literature review

A robust cybersecurity system relies on both the cybersecurity technologies (being built based on the CIA triad: Confidentiality, Integrity, and Availability) and the ethical values and behaviours of human decision-makers and end-users of the system. While cybersecurity technologies might cause various ethical issues, these ethical implications are often ignored (Formosa et al. 2021; Vallor 2018). Therefore, it is crucial for cybersecurity professionals to understand how their decisions may affect users. For example, consider the possible introduction of 2FA described in the introduction. Moral issues raised by this case include those related to the ethical principles of non-maleficence (harming some users who lose access), beneficence (better security for others), and justice (unfair that more vulnerable users may not benefit as much as others from the technology). Hence, professionals and other system users (e.g. system architects, administrators) need training about ethical conflicts and dilemmas that might arise in cybersecurity (Blanken-Webb et al. 2018). This kind of training is helpful for acquiring a clearer understanding of ethical issues in a domain (Jamal et al. 2016), but we need a suitable ethical framework

to develop a proper ethical education. For this, we utilised a principlist approach and a framework that has been proposed for the cybersecurity domain by Formosa et al. (2021) as it connects cybersecurity to basic ethical concerns. The framework consists of five ethical principles applied to the cybersecurity domain: beneficence, non-maleficence, autonomy, justice, and explicability. This framework is modelled on the five AI4People's principles (Floridi et al. 2018) for ethical AI. By adding explicability, the AI4People's principles extends the four basic ethical principles developed in a bioethics context by (Beauchamp and Childress 2001).

Beneficence describes the situation where cybersecurity technologies should enhance individuals' lives. This applies to day-to-day life activities, such as e-commerce and the private sharing of data, to promoting human well-being, protecting privacy, and strengthening trust. The main purpose of beneficence is to secure a safe cybersecurity environment that benefits all. **Non-maleficence** describes the situation where cybersecurity technologies should not be used to harm individuals. A poor cybersecurity practice (such as using an outdated security patch), for instance, can harm users of a system by exposing them to vulnerabilities and threats that could compromise their data. It can also be financially harmful, and so reduce their emotional health and well-being. **Autonomy** describes the situation where leveraging cybersecurity technologies by users should not limit their informed choices about how they use that technology. To some reasonable extent, users should be given freedom to manage their own cybersecurity options. Acquiring user consent for accessing users' data also respects their autonomy. **Justice** describes the situation where cybersecurity technologies should improve fairness and provide equitable access for all. This applies to avoiding bias, exploiting the vulnerable, and undermining solidarity. For example, consider designing a cybersecurity solution. If the solution is designed so that it is not accessible or useable by members of disadvantaged social groups, it raises important justice concerns. As another example, deploying machine learning algorithms trained on deeply biased data sets for cybersecurity might also treat users inequitably. Lastly, **explicability** describes the situation where cybersecurity technologies should be clearly understandable and accountable for their functioning. In addition, cybersecurity professionals are responsible to keep their professional skills and knowledge up-to-date, which includes understanding the ethical ramifications of the technology.

Buchan (2005) spoke about a gap between people's ethical attitudes and their ethical intentions, and thus people's ethical values in a cybersecurity context may not align with their ethical choices. There is extensive research that demonstrates ethical education is a promising way to improve ethical decision-making and help to close this gap (Cagle and Baucus 2006; Geiger and O'Connell 1998; Luthar and Karri 2005; Stead et al. 1990). Therefore, it is important to understand how to improve ethical cybersecurity decision-making via a training tool, such as a serious game, that provides trainees with interactions that enable them to reflect upon different human factors. To that end, we need to understand which human factors influence ethical decision-making to help tailor the training.

According to Gino et al. (2009), reminding people about moral behaviour may decrease their dishonesty. They studied the effect of a group member's unethical behaviour on the group. The result of Gino et al.'s (2009) research asserts that people's reaction to unethical behaviour depends on the active social norms related to observed dishonesty. The role played by individuals and the wider social norms and patterns of behaviour are thus important to consider when studying what factors influence ethical decision-making. Whitty et al. (2015) focused on the different behaviours by individuals when it comes to ethical decision-making in cybersecurity domain, such as with sharing passwords. They examined locus of control (an individual's belief about their control over their environment), perseverance (the ability to remain on a task until completion and avoid becoming bored), and self-monitoring (people who are sensitive to social and situational cues and change their behaviour) as human factors among participants to verify if they share passwords. The result of their analysis shows that the group of people with a considerable lack of perseverance as well as a high level of self-monitoring are more likely to share passwords (i.e., engage in unethical behaviour). This study also highlights the limited resources in the literature on this topic and the need to do more research to understand better which people are more likely to engage in risky behaviours (Whitty et al. 2015). Therefore, it is important to study human values and how those values drive ethical decision-making in a cybersecurity context, and how to help cybersecurity professionals to be more aware of the ethical implications of their decisions. There is also a novel study by (Ferro et al. 2022) which shows the importance of human factors in cybersecurity. The study asserts that utilising a game can improve users' awareness for good cybersecurity practices. However, the study did not focus on ethics and ethical training for cybersecurity.

Gratian et al. (2018) studied the correlation of four parameters: demographic factors, personality traits, risk-taking preferences, and decision-making styles to cybersecurity behaviours of users in a university environment. The result of their analysis suggests that extraversion is a significant predictor of good

device securement behaviour intentions. This means that people with outgoing personalities are more likely to be careful to lock their devices than those with introverted personalities. Therefore, we aim not only to study the role of the game on ethical awareness in cybersecurity contexts, but also to collect data on human factors (e.g., demographic and human values) in order to understand different individual responses to ethical dilemmas when playing the game. We can also use this data to improve the game with more realistic features in future studies.

3. Methodology

The overall aim of this paper which reports on a study is to improve the ethical awareness of cybersecurity professionals. As a starting point, we have created a game designed based on the three distinct stages of Rest-Model (Jones 1991) (awareness, orientation, and intention) to help cybersecurity students understand possible ethical implications when making cybersecurity decisions. We also seek to understand whether people who make different cybersecurity decisions in the game (e.g. counterattack/disclosure) have distinct demographic factors and/or personality traits. To that end, in this study we aim to address the following research questions.

Research Question 1: Does our designed game influence players to acquire better understanding of different ethical principles in a cybersecurity context?

Research Question 2: Can we differentiate players who take different cybersecurity response actions based on their demographic factors /personality traits /moral stance/human values?

To answer the above questions, we conducted an online study approved by our university's Human Ethics Research Committee. This study aims to examine how our serious game sensitises cybersecurity students to ethical principles and impacts their cyberethical decision-making (i.e. decision process). This training opportunity has been designed to increase the students' (participants') awareness of ethics in a cybersecurity context. The participants were given a total of forty minutes to complete the study. The study starts by providing the participants with an information and consent form that is followed by demographic information collection. The next steps are to answer pre-game questions, play the game, complete a player experience survey, and answer other post-game questions. The study is concluded by answering instruments to measure personality, moral foundations and values, as detailed below.

3.1 Sample

We recruited 366 students as our participants who were studying a first-year cybersecurity unit at our university. The study was conducted in the final week of the first semester of 2021 during a scheduled class. Students voluntarily consented to allow their data to be used for research purposes. Overall, 318 of the participants gave consent to use their data. We removed the remaining 48 participants from further analysis. However, only 250 played the game and gave valid responses with a majority of males (190) to females (54) and other gender options (6). To exclude bias/random answers, we only considered a response valid if participants did not choose the same option to all questions in each section.

3.2 Study Design

The designed game (Ryan et al. 2022) sought to emulate some cybersecurity decisions faced by a cybersecurity professional as well as other issues that might arise online in an office environment to add realism and complexity. The main purpose of the game is to increase the participants' awareness of the ethical situation by letting them explore the ethical decision-making dilemmas. The ethical dilemmas are designed to sensitise the participants to the ethical issues that arise in real-world cybersecurity scenarios (ethical awareness, orientation, and intention) (Jones 1991). An example of such a dilemma is where a system admin needs to decide whether to force installation of an outstanding security update organisation-wide or prioritise current operability if the release has detrimental impacts on accessibility functionality relied upon by vulnerable end-users. The game also provides the participants insights on the possible implications their decisions might have for end-users. This is compared with so-called nudging (Sunstein 2014). Nudging, generally, is a different way of presenting options to help users make a better decision in their routines. (e.g. products include labels showing energy consumption rates).

Our study included two different perspectives to answer the research question. The first perspective focused on whether the game could improve ethical decision-making. The second perspective focussed on the player's experience of the game. Addressing the first perspective, we asked participants to respond to the same text-based ethical decision-making dilemma in a cybersecurity context before and after the game was played. To determine whether the game had any impact on participants' ethical decision-making, we compared participant responses to the external prompt outside the game and how

the responses were justified before and after playing the game. Our goal in doing this was to understand if participants identify and apply ethical choices differently after playing the game. That is, does the game help participants make more ethical choices?

To analyse players' cybersecurity action response in-game, we created two main ethical dilemmas (disclosure and counterattack) in the game based on cases reported in the literature (Formosa et al. 2021). Players must decide what action to take in each scenario, i.e. whether or not to counterattack. The scenarios were also enriched with narrative to provide players with experiences that include a range of ethical conflicts between the five ethical principles. After players had made a major decision in the game (i.e. whether to disclose or counterattack), they needed to complete an in-game reflection report outlining the main ethical considerations regarding their decision by drawing on the relevant five ethical principles. To analyse the game experience, we asked participants to answer the Player Experience of Need Satisfaction (PENS) questionnaire. This helps us to assess the players' level of competence, autonomy, relatedness, presence, and mastery of control controls that they experienced when they played the game.

3.3 The game

The game's interface appears as a desktop computer with various communication and project management apps, which is meant to simulate communications and activities at a medium sized organisation based on real world scenarios. The players were given a mid-level role in the organisation to enable them to make decisions autonomously. The participants had the chance to explore the environment of the simulated platform which includes project management, email, news portals, real-time status updates, and text messaging. The designed game enabled players to communicate with non-player characters (NPC) via direct messages, emails, or in a group chat. The players were given various pre-written choices as actions to take in response to communications. There are two main systems within the game. The first system is a scripted narrative system to liaise with NPCs. The second system is a resource management system to manage the players' time and other resources (such as bandwidth). More information about the philosophy underlying the design and screenshots of the game are provided in (Ryan et al. 2022; Ryan et al. 2017).

3.4 Pre and Post game test scenario

A text-based scenario was designed to present an ethical decision-making dilemma in a cybersecurity domain. The participants were asked to make a decision on the exact same dilemma both prior to and after playing the game. Typical of ethical dilemmas, there was no obvious correct decision, but the choice would depend on the ethical stance taken. The dilemma is presented in Figure 1 and the decision involved is whether or not to release an ethical worm that will install updates automatically. Participants were given an option to justify their decision. This provides an opportunity to learn whether, and why, participants made different decisions before and after playing the game.

You notice that many staff in your organisation have been failing to follow company procedure by installing security and operating system updates. You have been sending lots of reminder emails about the importance of updating to staff. Current policy leaves the installing of software updates to individual employees, but leaving updates uninstalled could cause major security problems for the organisation. Should you release an ethical worm that will automatically install all outstanding updates on the computers of all staff in your organisation?

- Release the ethical worm that will install updates automatically
- Do not release the ethical worm that will install updates automatically

Please describe the important considerations and issues in responding to the above scenario and why they are important?

Figure 1 - An ethical decision-making dilemma

3.5 Data collection and analysis

To collect game data, we utilised bespoke designed online in-game analytics systems to capture data about player choices. Via an online survey we collected demographic data and participants' knowledge about ethics. Other data collected included player responses to a pre-game and post-game test scenario (as outlined in 3.4) and the following validated instruments; the Ten-Item Personality Inventory (TIPI) (Gosling et al. 2003), Moral Foundation Questionnaire (MFQ) (Graham et al. 2011), and Portrait Values Questionnaire (PVQ) (Schwartz et al. 2012) to collect personality, moral decision-making, and human values (Schwartz values), respectively. The reason we selected Schwartz's framework among others is because it is widely accepted in the research community due to its

comprehensive level of study for 60 cultures that lead to recognise 10 universal human values, later refined to 19 values. Qualitative analysis of the open-ended responses to the pre and post-game test scenario is described in the next subsection.

After playing the game participants also completed the Player Experience of Need Satisfaction (PENS) survey (Ryan et al. 2006), which assesses players satisfaction of the basic psychological competence, autonomy and relatedness by the game, as well as their experience of presence/immersion and intuitive controls. The competence scale measures a player's satisfaction of their need for challenge and feelings of effectiveness. The autonomy scale assesses a player's feeling of freedom in the game and the presence of choices of interest to them. The relatedness scale describes when players feel connected with others in a game, such as during a multiplayer game. We included that scale to see if players felt some satisfaction of relatedness via their sense of connection to the virtual characters (NPCs) they interacted with in the game. The fourth scale, presence, describes players' deep involvement and includes physical, narrative, and emotional presence. Finally, intuitive controls assesses the extent to which players feel they have control over the game's actions (Ryan et al. 2006). The 18 items are measured using a 7-point Likert scale (from 1=Do not agree to 7=Strongly agree). We calculated the score of each of the five components of the PENS based on the average value of its items, in accordance with the authors instructions.

3.6 Pre-game and post-game scenario qualitative analyses.

After preliminary investigation from the collected data, we devised a coding scheme for analysing participants' justifications to the pre/post game scenario. Nine codes were created for analysing pre-game and post-game responses, as presented in Figure 2. These codes were created after a joint discussion of the team (group of academics and research students). After trialling on about 30 entries by two coders, the set was finalised by the team. The purpose of the coding scheme was to categorise and analyse the given text justification that the participants made to support their decision regarding whether to release the ethical worm or not. We grouped the codes 2, 3, 4, 5, and 8 as "ethical choices" and grouped codes 1, 6, 7, and 9 as "other" or non-ethical. We received 230 valid responses to the pre-game justifications. Overall, we only received 190 responses to the post-game justifications in which we excluded 13 records of participants who left irrelevant answers to the justifications. In addition, the blinded peer coding has been used by two different persons separately to code the justifications. To assess reliability and consistency we compared both above coding by Cohen Kappa's coefficient. The results of Cohen Kappa for the two independent persons who analysed the justifications for pre-game and post-game were 0.93 and 0.96 respectively, showing high consistency and reliability. Also, as presented in Table 1, we compared (post-game) and (pre-game) by using a T-test to determine changes before and after playing the game.

1) No answer or irrelevant answer given
2) Did they use at least one of our terms (Autonomy, Justice, Beneficence, non-maleficence, explicability)
3) Did they use the word "ethics" or derivative (e.g. ethically)
4) Did they discuss ethics generally without using particular words or terms
5) Staff should be informed
6) Don't do it due to technical issues
7) Don't do it due to contravention of company policy
8) Consent should be obtained from staff
9) Relevant answer but not referring to ethics

Figure 2 - Codes for analysing the justification of pre-game and post-game questions

4. Results and findings

Participants included 54 females and 190 males, with six persons selecting other gender options, such as non-binary. All records (i.e. 250) were included in our analyses. Participants were aged from 17 to 34 years old, with an average age of 19.83 and standard deviation of 2.97. The participants' self-reported knowledge about ethics ranged from 1 to 5 (1=Terrible, 2=Poor, 3=Average, 4=Good, 5=Excellent), with a mean of 3.68. Computing is the main area of study for 70.4% (176/250) of participants, followed by business 12.4% (31/250). The other main area of study 11.60% (29/250) included people doing double degrees combining both computing and business. Other areas of study included Security Studies, Engineering, Accounting, Science, Software Engineering, Criminology, and Information technology. Also, in the following, we present the results (i.e. total and parentages) concerning the cultural group to which participants identified. Oceania (including Australian) 126/250 (50.40%), South-East Asian 55/250 (22.00%), Southern and Central Asian 15/250 (6.00%), North-East Asian 9/250(3.60%),

Northern-Western European 8/250(3.20%), Southern-Eastern European 7/250 (2.80%), North African and Middle 5/250 (2.00%), Sub-Saharan African 4/250 (1.60%), and People of the Americas 1/250 (0.40%) No answer or do not identify 20/250 (8.00%).

Table 1 compares the results for responses to out of the game prompts scenarios (pre and post game) about decisions and how the decisions were justified. 67 individual changed their decision after playing the game. 16 individuals changed their pre-game and post-game justifications from "other" to "ethical", and another 18 persons changed their justifications from "ethical" to "other". Descriptive statistics for the PENS survey results are also presented in Table 2. We also are focusing on two in-game choices involving whether to counterattack or disclose. Due to programming errors and technical errors where logfiles became truncated, much of our planned in-game data analysis was not possible. Lost data includes the player responses to the reflection reports that were featured in the game. From the data we did capture, we could only tell if a player had chosen "1" (i.e. yes) in the counterattack and disclosure scenarios. We were unable to differentiate between players who had chosen "0" (i.e. no) or who had not answered at all. Thus, we cannot report comparisons between those who chose to counterattack/disclose and those who did not. Of the participants who played the game and gave consent, only 23 persons were identified as having chosen disclosure, of which only 19 records have valid data (non-empty responses to the constructs' questions), and 12 persons as having chosen counterattack. Table 3 presents the average scores of two groups of participants who selected to counterattack or disclosure for their MFQ, TIPI, and PVQ constructs in comparison to the overall average of the complete dataset (250 records).

	Pre-game			Post-game			*P-value
Decisions	Release	Don't	Total	Release	Don't	Total	0.373
	135	112	247	124	95	219	
Justifications	Ethical	Other	Total	Ethical	Other	Total	0.067
	118	112	230	86	91	177	

Table 1: Pre-game and post-game decisions/justifications. Paired t-test significant at p<0.05

PENS Construct	Mean	SD
Competence	4.94	1.52
Autonomy	4.53	1.37
Relatedness	3.66	1.41
Presence	4.07	1.18
Intuitive controls	5.00	1.25

Table 2. PENS Analysis Results (1=Do not agree - 7=Strongly agree)

5. Discussion

This paper seeks to determine the value of a serious game for improving ethical choices and understanding of ethical principles in a cybersecurity context. To understand the value of the game we measure changes in ethical reasoning before and after playing the game. To understand if the game had any impact on players learning on ethical principles in a cybersecurity context, we were less interested in whether players changed their decision regarding the scenario (i.e. release the ethical worm or not) compared to whether they changed the justification for the choice indicating a change in their reasoning. We found no significant differences (p=0.373) in their choice after playing the game. However, Table 1 reports that 41.30% (118/230) of participants gave justifications that involved "ethical" reasoning prior to playing the game and this percentage increased to 48.84% (86/177) after playing the game. As described earlier, we only had 230 valid responses to the pre-game justifications. Also, we only received 190 responses to the post-game justifications, as in 13 records of the 190 responses who left irrelevant answers were excluded. While this change was not statistically significant (p=0.067), the change is in the right direction, as it considered marginally significant. The largest increase in ethical justifications was from participants who gave no answer or an answer unrelated to the scenario. This indicates that these students did not know how to evaluate the situation before playing the game, yet the game equipped them with an appropriate vocabulary to properly justify their post-game scenario response. This raises the issue of whether participants have increased their ethical awareness (that is, the ability to identify ethical issues and ethical considerations), their knowledge of ethical approaches and decision processes (such as the use of the ethical principles outlined to determine correct ethical action) or simply their ability to provide ethical responses (such using ethical language to describe the reasons for one's actions or responses).

Constructs	Scales	Counterattack (n=12)		Disclosure (n = 19)		Total (n=250)	
		Mean	SD	Mean	SD	Mean	SD
MFQ (Range 0-20)	Harm	13.91	3.11	11.84	7.16	13.81	3.62
	Fairness	15.24	3.41	12.05	7.17	14.94	3.24
	Ingroup	10.83	3.73	8.15	7.34	10.77	4.06
	Authority	10.91	2.77	8.42	5.85	11.2	4.25
	Purity	11.08	4.01	9.26	6.38	12.32	5.93
TIPI (Range 1-7)	Extraversion	3.71	1.72	4.05	1.37	3.79	1.25
	Agreeableness	4.54	0.75	4.29	1.2	4.32	0.93
	Conscientiousness	4.46	0.89	4.55	1.17	4.58	1.19
	Emotional Stability	4.38	0.97	4.39	1.55	4.52	1.16
	Openness to Experiences	4.71	1.05	4.84	1.06	4.73	1.02
PVQ (Range 1-6)	Self-direction Thought	4.94	0.71	5.02	0.87	4.59	0.91
	Self-direction Action	4.86	0.81	4.74	0.84	4.4	0.92
	Stimulation	4.25	0.98	4.29	1.25	4.11	1.1
	Hedonism	5.08	0.88	4.74	1.03	4.45	0.98
	Achievement	4.67	0.84	4.76	0.9	4.27	1.02
	Power Dominance	3.69	1.42	3.86	1.22	3.73	1.28
	Power Resources	3.92	1	3.81	1.04	3.51	1.15
	Face	4.22	0.88	4.17	1.39	3.93	1.12
	Security Personal	4.92	1.24	4.81	0.68	4.31	0.98
	Security Societal	4.78	1.02	4.52	0.98	4.07	1.19
	Tradition	4.11	1.29	3.02	1.52	3.47	1.44
	Conformity-Rules	4.67	1.05	4.07	1.48	4.04	1.28
	Conformity-Interpersonal	4.03	0.87	4.17	1.38	3.93	1.22
	Humility	4.64	0.74	4.4	1.15	4.18	1.05
	Benevolence-Dependability	4.97	0.7	5.17	0.8	4.99	0.99
	Benevolence–Caring	5.06	0.76	5.1	0.73	4.54	0.94
	Universalism-Concern	5.25	0.94	4.9	0.97	4.45	1.02
	Universalism-Nature	4.44	0.97	4.26	1.25	3.99	1.13
	Universalism-Tolerance	5	0.65	4.98	0.93	4.49	1.01

Table 3: Analysis results for counterattack and disclosure options

To understand how and why participants had changed their reasoning and if the game has had any influence on their decisions, we further differentiated how many individuals changed their justifications from "other" to "ethical" group and vice versa (i.e. changed their justifications that were coded differently). This analysis did not include participants who did not answer or who provided a relevant answer that was either useless or unrelated to the scenario. Our analyses revealed that there were 16 persons who shifted their justifications from "other" (here only coded as 6, 7, and 9)" to "ethical" and 18 other persons whose justifications changed from "ethical" to "other" (here only coded as 6, 7, and 9) category. Following Gino et al. (2009) who assert that warning people about morality can reduce dishonest behaviours, the 16 participants who changed from providing non-ethical to ethical justifications after playing the game may have been moved toward ethical thinking after being exposed to the ethics related discussions and situations provided by the game. The majority of the 16 records that shifted from other to ethical codes were mainly labelled with code 9 (relevant but not referring to ethics) in their pre-game response. For instance, a person shifted their justifications from "updates need to be done regularly, failure to do so [can] result in cyberattack" (code 9) to "it's not right to release a worm without consent" (code 8 and relevant to the ethical principle of autonomy/justice). Furthermore, the second category of 18 records that swapped from "ethical" to "other" codes were mainly changed from codes 8 (consent) and 4 (general discussion implying ethics) to code 9 (relevant answer but not related to ethics). For instance, a person changed their justifications from "Doing so ensures the security of the system though can be seen as an invasion into the employee's privacy" (Code 4) to "Do what need[s] to be done" (Code 9). Although their justification for this category after playing the game should not be coded as "ethical", it is unclear whether they still agreed with their previous justification and now were

simply adding further responses that were triggered either by the game or extended consideration of the same scenario. Thus, these 18 responses that moved from "ethical" to "other" may have had a negative and incorrect impact on our analyses. Due to this lack of clarity, it is recommended that future studies either explicitly elicit whether players still agree with what they said before or that a second different, but similarly complex and structured scenario, is provided for review after playing the game.

To answer the first research question, based on the comparison of pre-game and post-game reasoning, we are unable to conclude that the game increased consideration of the ethical issues underlying participants' cybersecurity decision-making. However, as described above, we observed records where participants acquired better ethical understanding after playing the game and promoted from providing non-ethical to ethical justifications.

Concerning the player experience, as presented in Table 2, these descriptive statistics give us some initial indication of the game's satisfaction of psychological needs and the experience it created for players. For intuitive controls, the average score for the whole cohort was 5 out 7 (the highest mean score for any component) which, since it is on average above a neutral score of 4, indicates in general participants felt the game's interface was fairly intuitive and that it didn't interfere with their experience of the game. Scores for autonomy and competence are also close to 5, indicating some good satisfaction of these psychological needs. This is a positive because having a sense of control or agency to make ethical decisions is essential. In particular ethical decision making is affected by one's sense of moral agency and ability to act according to one's moral values (Bandura 2006). This is further supported by the Theory of Planned Behaviour (TPB) (Fishbein and Ajzen 2011) which notes that intention to take an action is impacted by one's sense of control within a given context. If an individual does not feel they have self-efficacy in the given context they will not be motivated to act (Gerber and Rogers 2009). Future work may accordingly look to measure players' self-efficacy and sense of agency or control in cybersecurity contexts before and after playing such serious games. This also could be employed in the context of pre and post scenarios, as provided in our study. The PENS results and change from no/irrelevant answer to the pre-game scenario to an ethical response after the game, together suggest that the game may have had a positive on self-efficacy and sense of control concerning ethical decision-making. However, these results may also simply indicate that players felt competent in playing the game and had control over making decisions in the game. Future work can explore these interpretations. Finally, participants scored fairly neutral results for the relatedness scale (the lowest average component score) and the presence scale. The former result may mean participants did not feel highly connected to others in the game, which is not surprising since this was not a multi-player game. However, this close to neutral average score also suggests some sense of connection, although not a strong one, to the NPCs in the game. The latter result also indicates players aren't reporting strong presence or disengagement with the game on a physical, narrative or emotional level, but are at least reporting some sense of transportation into the game's world. Together, these results indicate that further design and development work on the game could seek to increase extent to which it can support a sense of engagement with its world and connection with its characters since empathy plays an important role in ethical reasoning. For example, caring about beneficence or justice requires caring about treating others kindly and fairly (Graham et al. 2011). The PENS results also provide a benchmark for player experience and need satisfaction against which further iterations of the game, or other cybersecurity serious games, can be measured against.

To answer the second research question, we analysed a number of characteristics of participants to understand if these characteristics can explain their choices in the game in response to the counterattack or disclosure scenarios. Due to technical issues with data collection discovered after the study had ended, we did not have the players' reflective reports and participants that did not select 'disclosure' or 'counterattack' was recorded as a zero. Thus, we were unable to separate participants who chose not to disclose or counterattack from those who never answered this question or who did not reach this part of the game. Given that we only have a limited number of positive cases and no negative cases, we sought to determine if those who made the choice to counterattack or disclose exhibited certain profiles based on their individual data. Therefore, we compared the average scores for the MFQ, TIPI, and PVQ constructs with the overall average of the complete dataset in Table 3. Although the numbers are too small to perform detailed statistical analyses such as T-tests, we can see differences that could be insightful for the disclosure group compared to the overall dataset. While the MFQ scales average score for the counterattack group is close to the overall average score for all participants, in the disclosure group the participants reported a lower average score for all MFQ scales in comparison to the overall average scores for all participants. For instance, the average score of 9.26 for purity is lower than the overall average of 12.32. As stated previously, the numbers of both groups are too small to generate statistical significance, however these results can be a useful basis for future studies. The MFQ, TIPI, and PVQ data collected from players could be used to analyse and potentially predict whether these

individual factors impact on cybersecurity decision-making, as is suggested by (Gratian et al. 2018) who highlight that demographic factors and personality traits, especially extraversion, are important in influencing cybersecurity behaviours.

6. Conclusions and future work

This paper reports a study to investigate the effectiveness of a serious game on improving ethical awareness and decision-making for cybersecurity students towards addressing a gap in ethical knowledge and sensitivity in cybersecurity professionals. We created a serious game that offers training on five ethical principles relevant to cybersecurity decision making. The game provides players with a series of cyberethical dilemmas requiring them to apply those five principles. However, our study did not show that the game increased the use of ethical justifications in a related cybersecurity ethical scenario after playing the game. We acknowledge a few limitations of our study and plan to address them in future studies. Firstly, the loss of valuable data due to technical issues has significantly impacted our study and our ability to assess the value of the game. This is being rectified through a new study. Participants were all university students studying introductory cybersecurity, and there was a gender imbalance with a large majority of males. We aim to address this by recruiting a wider range of target audiences in the next study, such as cybersecurity professionals and more females for a better gender balance. This study aims not only to understand whether the designed game increased ethical awareness for cybersecurity practitioners or students, but also how to improve it by leveraging human profiles that could be used to create realistic NPC agents in the serious game. Learning human profiles enables us to add more features to the designed game by leveraging artificially intelligent (AI) agents to interact with human players in cybersecurity within the context of organisational policies and norms (Dignum and Dignum 2009). We therefore plan to expand this work to accommodate organisational policies and norms via a multi-agent systems framework to integrate AI agents into gameplay (Jensen et al. 2014). Finally, we plan to enrich future studies with different contexts and scenarios in accordance with specific ethical principles, which will help us to study which ethical principles are relevant and important for cybersecurity decision making.

7. References

Bandura, A. 2006. "Toward a Psychology of Human Agency," *Perspectives on psychological science* (1:2), pp. 164-180.

Bazerman, M. H. 2011. *Blind Spots: Why We Fail to Do What's Right and What to Do About It*. Brilliance Audio; Unabridged edition (August 12, 2014).

Beauchamp, T. L., and Childress, J. F. 2001. *Principles of Biomedical Ethics*. Oxford University Press, USA.

Blanken-Webb, J., Palmer, I., Deshaies, S.-E., Burbules, N. C., Campbell, R. H., and Bashir, M. 2018. "A Case Study-Based Cybersecurity Ethics Curriculum," *2018 (USENIX) Workshop on Advances in Security Education (ASE18)*.

Buchan, H. F. 2005. "Ethical Decision Making in the Public Accounting Profession: An Extension of Ajzen's Theory of Planned Behavior," *Journal of Business Ethics* (61:2), pp. 165-181.

Cagle, J. A. B., and Baucus, M. S. 2006. "Case Studies of Ethics Scandals: Effects on Ethical Perceptions of Finance Students," *Journal of Business Ethics* (64:3), pp. 213-229.

Christen, M., Gordijn, B., Weber, K., van de Poel, I., and Yaghmaei, E. 2017. "A Review of Value-Conflicts in Cybersecurity," *The ORBIT Journal* (1:1), pp. 1-19.

Dignum, F., and Dignum, V. 2009. "Emergence and Enforcement of Social Behavior," *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*: Citeseer, pp. 2942-2948.

Ferro, L. S., Marrella, A., Catarci, T., Sapio, F., Parenti, A., and De Santis, M. 2022. "Awato: A Serious Game To improve Cybersecurity Awareness," *HCI in Games,* X. Fang (ed.), Cham: Springer International Publishing, pp. 508-529.

Fishbein, M., and Ajzen, I. 2011. *Predicting and Changing Behavior: The Reasoned Action Approach*. Taylor & Francis.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., and Vayena, E. 2018. "Ai4people-an Ethical Framework for a Good Ai Society: Opportunities, Risks, Principles, and Recommendations," *Minds Mach (Dordr)* (28:4), pp. 689-707.

Formosa, P., Wilson, M., and Richards, D. 2021. "A Principlist Framework for Cybersecurity Ethics" *Computers & Security* (109).

Gee, J. P. 2007. *What Video Games Have to Teach Us About Learning and Literacy*. Palgrave Macmillan.
Geiger, M. A., and O'Connell, B. T. 1998. "Accounting Student Ethical Perceptions: An Analysis of Training and Gender Effects," *Teaching Business Ethics*), pp. 371–388.
Gerber, A. S., and Rogers, T. 2009. "Descriptive Social Norms and Motivation to Vote: Everybody's Voting and So Should You," *The Journal of Politics* (71:1), pp. 178-191.
Gino, F., Ayal, S., and Ariely, D. 2009. "Contagion and Differentiation in Unethical Behavior: The Effect of One Bad Apple on the Barrel," *Psychol Sci* (20:3), pp. 393-398.
Gosling, S. D., Rentfrow, P. J., and Swann, W. B. 2003. "A Very Brief Measure of the Big-Five Personality Domains," *Journal of Research in personality* (37:6), pp. 504-528.
Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., and Ditto, P. H. 2011. "Mapping the Moral Domain," *J Pers Soc Psychol* (101:2), pp. 366-385.
Gratian, M., Bandi, S., Cukier, M., Dykstra, J., and Ginther, A. 2018. "Correlating Human Traits and Cyber Security Behavior Intentions," *Computers & Security* (73), pp. 345-358.
Jamal, A., Ferdoos, A., Zaman, M., and Hussain, M. 2016. "Cyber-Ethics and the Perceptions of Internet Users: A Case Study of University Students of Islamabad," *Pakistan Journal of Information Management and Libraries* (16).
Jensen, A. S., Dignum, V., and Villadsen, J. 2014. "The Aorta Architecture: Integrating Organizational Reasoning in Jason," *International Workshop on Engineering Multi-Agent Systems*: Springer, pp. 127-145.
Jones, T. M. 1991. "Ethical Decision Making by Individuals in Organizations: An Issue-Contingent Model," *Academy of Management Review* (16:2), pp. 366-395.
Lapsley, D. N. a. D. K. 2005. "The Psychological Foundations of Everyday Morality and Moral Expertise," Character Psychology and Character Education,,"), pp. 140-165.
Luthar, H. K., and Karri, R. 2005. "Exposure to Ethics Education and the Perception of Linkage between Organizational Ethical Behavior and Business Outcomes," *Journal of Business Ethics* (61:4), pp. 353-368.
Mann, I. 2008. *Hacking the Human: Social Engineering Techniques and Security Countermeasures* (1st Edition ed.).
Narvaez, D. 2005. "Integrative Ethical Education," in *Handbook of Moral Development*. Taylor and Francis.
Ryan, M., McEwan, M., Sansare, V., Formosa, P., Richards, D., and Hitchens, M. 2022. "Design of a Serious Game for Cybersecurity Ethics Training," in: *DIGRA conference* Poland.
Ryan, M., Staines, D., and Formosa, P. 2017. "Focus, Sensitivity, Judgement, Action: Four Lenses for Designing Morally Engaging Games," *Trans of the Digital Games Research Association* (2:3).
Ryan, R. M., Rigby, C. S., and Przybylski, A. 2006. "The Motivational Pull of Video Games: A Self-Determination Theory Approach," *Motivation and Emotion* (30:4), pp. 344-360.
Schwartz, S. H. 1994. "Are There Universal Aspects in the Structure and Contents of Human Values?," *Journal of Social Issues* (50:4), pp. 19-45.
Schwartz, S. H. 2012. "An Overview of the Schwartz Theory of Basic Values," *Online Readings in Psychology and Culture* (2:1).
Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lonnqvist, J. E., Demirutku, K., Dirilen-Gumus, O., and Konty, M. 2012. "Refining the Theory of Basic Individual Values," *J Pers Soc Psychol* (103:4), pp. 663-688.
Staines, D., Formosa, P., and Ryan, M. 2017. "Morality Play: A Model for Developing Games of Moral Expertise," *Games and Culture* (14:4), pp. 410-429.
Stead, W. E., Worrel, D. L., and Stead, J. G. 1990. "An Integrative Model for Understanding and Managing Ethical Behavior in Business Organizations," *Jrnl of Business Ethics*:9, pp.233–242.
Sunstein, C. R. 2014. "Nudging: A Very Short Guide," *37 J. Consumer Pol'y 583 (2014)*.
Vallor, S. 2018. "An Introduction to Cybersecurity Ethics.," *Markkula Center for Applied Ethics*.
Whitty, M., Doodson, J., Creese, S., and Hodges, D. 2015. "Individual Differences in Cyber Security Behaviors: An Examination of Who Is Sharing Passwords," *Cyberpsychol Behav Soc Netw* (18:1), pp. 3-7.

Acknowledgements

This work is funded by an Australian Research Council Discovery Grant: DP200102131 - Cybersecurity ethics training simulations for values-based decision-making.