Design of a Serious Game for Cybersecurity Ethics Training

Malcolm Ryan
School of Computing
Macquarie University
Australia
malcolm.ryan@
mq.edu.au

Mitchell McEwan
School of Computing
Macquarie University
Australia
mitchell.mcewan@
mq.edu.au

Vedant Sansare
School of Computing
Macquarie University
Australia
vedant.sansare@
mq.edu.au

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

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

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

ABSTRACT
Serious moral games offer a tool for moral development that can help players translate 'head knowledge' of ethical principles into habits of everyday practice. In this paper, we present the design process behind one such game: *Prescott & Krueger*, a serious game for training information technology students in cybersecurity ethics. Our design draws on the Four Component Model of moral intelligence and the Morality Play model for serious moral game design. We reflect on how these models influenced our design process. The Four Component Model proved a useful set of lenses for developing learning outcomes and game narrative and mechanics, however the more prescriptive Morality Play model was more difficult to apply as the development of a sophisticated 'moral toy' required modelling both low-level cybersecurity systems and high-level ethical interpretations. We reflect on the broader implications of this problem for serious moral game design.

Keywords
serious games, game design, ethics, moral psychology, cybersecurity, education

INTRODUCTION
'Serious moral games' (Christen, Faller, Götz, & Müller, 2012) have been proposed as a vehicle for ethics training which can help address the "judgement-action" gap in which an individual's moral judgement of the right thing to do in a given scenario fails to coincide with their moral behaviour (Francis et al., 2016; Williams & Gantt, 2012). One possible explanation for this failure is 'ethical fading' in which the immediate, in-the-moment needs and constraints of a problem crowd out ethical concerns we might recognise with a more detached perspective (Tenbrunsel & Messick, 2004). By providing an immersive environment in which we can practice morality *in situ*, games can potentially help us turn our explicit, analytical knowledge of morality into implicit awareness of moral concerns and habits of moral practice.

Over the past decade, a body of ethical game design theory has been developed, drawing on theories of moral psychology and serious game design, as well as reflective analysis of existing moral games (Katsarov, Christen, Mauerhofer, Schmocker, &

Proceedings of DiGRA 2022

Tanner, 2019; Schrier, 2015; Sicart, 2013; Staines, Consalvo, Stangeby, & Pedraça, 2019). However, there have been relatively few examples of this theory being put into practice (Consalvo & Staines, 2021; Hilliard et al., 2018; Katsarov, Biller-Andorno, Eichinger, Schmocker, & Christen, 2020), and so little evidence of their practical value in the development of new games. In this paper, we address this gap by describing the application of two models derived from the field of moral psychology: the Four Component Model (4CM) of moral intelligence (Rest, Narvaez, Thoma, & Bebeau, 1999) and Integrative Ethical Education (IEE) framework for moral development this model inspired (Narvaez, 2006). In particular, we draw on the elaboration of the 4CM to the context of game design by Christen et al. (2012) and Ryan et al. (2017) and the Morality Play framework of Staines, et al. (2019) which integrates the IEE with serious game design theory. We present the application of these ideas for the design of a serious game for training information technology students in the principles and practice of cybersecurity ethics.

Our aim for this game was for students to be able to recognise and apply five key ethical principles (Formosa et al., 2021) that are relevant in cybersecurity in their decision making – *beneficence*, *non-maleficence*, *autonomy*, *justice*, and *explicability* – and to do so in the face of other competing concerns that might otherwise prompt ethical fading. We present a discussion of how the 4CM informed our development of learning goals, game narrative and mechanics, leading to development of the first playable prototype of our game *Prescott & Krueger*. Empirical evaluation of the game is still underway, so we cannot yet report on its effectiveness from a player's perspective. Instead, in this paper, we want to consider the project from a designer's perspective and evaluate these theoretical design frameworks in terms of their usefulness in actual design practice.

In the following, we first outline the principles of cybersecurity ethics that form the ethical basis for our game. We also elaborate on the Four Component Model and the ways it has already been applied to serious moral game design theory and practice, as well as other work in this area. We then describe the design process undertaken, including identification of learning outcomes and target audience, and the design of appropriate narrative and mechanical elements. Finally, we reflect on the strengths and weaknesses of the 4CM and Morality Play model in informing this process, before concluding with discussion on areas of design theory in need of further development, and our future plans for evaluating the game.

BACKGROUND

Cybersecurity ethics

Cybersecurity technologies and practices raise many significant ethical issues (Christen et al., 2020; Formosa et al., 2021; Manjikian, 2018; Vallor, 2018), and it is important that students and practitioners as well as general users of information communication technologies (ICT) are educated about these issues. Cybersecurity technologies aim to provide for the *availability*, *integrity* and *confidentiality* of data and computer systems (Brey, 2007). However, ethical issues arise because of the competing ethical demands that these technologies create. For example, for data to be available to all users of a system, including those with limited ICT skills, login procedures and password requirements must be broadly accessible. This highlights a tension between maintaining accessibility through simple login procedures and ensuring the highest levels of security to protect the confidentiality and integrity of data by, for example, requiring more complicated two factor authentication (2FA) for login.

The dominant approach to outlining cybersecurity ethics has been to list a series of relevant ethical principles (Formosa et al., 2021), which is an approach known as

ethical principlism commonly used in other areas of applied ethics (Beauchamp & DeGrazia, 2004). Principlist approaches are useful in a pedagogical context as they emphasize the competing ethical concerns at stake in a given domain, and make it straightforward to generate ethical conflicts and dilemmas as material for ethical reflection and training. For this reason, we have adopted a principlist approach to cybersecurity ethics as the framework around which to base our serious game. While there are several competing frameworks based on different principles, we have adopted the framework defended in Formosa, Wilson and Richards (2021) as it is based on five widely used ethical principles which translate directly to the context of developing a serious game for cybersecurity ethics.

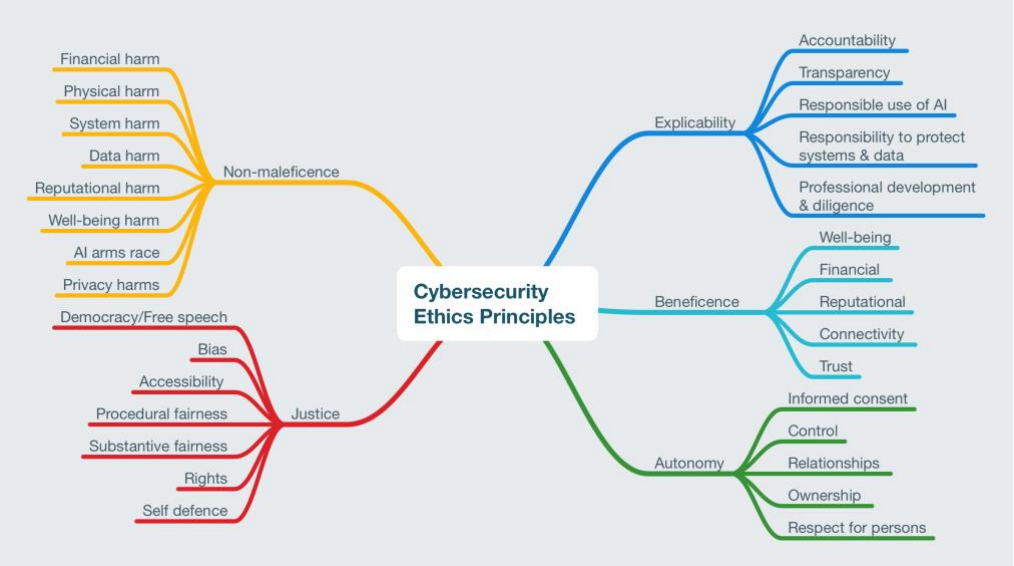

Figure 1: Reproduced figure of the five cybersecurity ethics principles from Formosa, Wilson and Richards (2021).

The five ethical principles in this framework are modelled on the AI4People's principles for ethical AI (Floridi et al., 2018), which in turn builds on the four principles for bioethics developed by Beauchamp and Childress (2001). The principles in the framework used here are *beneficence*, *non-maleficence*, *autonomy*, *justice*, and *explicability*, as illustrated in Figure 1. *Beneficence* is the principle that cybersecurity technologies should be used in ways that benefit and improve our lives. It can do this by making possible various interactions, such as e-commerce, that rely on good cybersecurity, and which have various benefits for us. *Non-maleficence* is the principle that cybersecurity technologies should be used in ways that do not harm us. Poor cybersecurity can result in many harms, such as the financial and psychological harms from a lack of data confidentiality caused by poor security settings. *Autonomy* is the principle that cybersecurity technologies should allow people to make their own choices in how they use that technology. This could include allowing people to make their own decisions about what level of security is appropriate for them. *Justice* is the principle that cybersecurity technologies should promote fairness and equality and not undermine solidarity. This includes ensuring data and systems have high useability and accessibility for even technically challenged users and avoiding bias and discrimination against minority users. *Explicability* is the principle that cybersecurity technologies should be intelligible and used in accountable and responsible ways. This requires, for example, transparency around cybersecurity and related privacy policies and ensuring clear lines of accountability for policy violations. These five principles form the ethical framework for our design.

Designing Serious Moral Games

While ethically significant games continue to be prominent in the entertainment market (Staines et al., 2019) and their value as vehicles for moral education has been discussed at length, there have been relatively few published titles that address this opportunity. Some notable examples include *Quandary* (FableVision Studios, 2012), *uFin: The Challenge* (Kobold Games, 2018) and *uMed: Your Choice* (Kobold Games, 2019). The development of *Quandary* was led by a team at Tufts University, who evaluated the game in terms of its engagement and educational value (Hilliard et al., 2018; Ilten-gee & Hilliard, 2019), but little has been revealed about the design thinking behind the game's systems. *uFin* and *uMed* are both part of a project by the University Zürich (Katsarov et al., 2020; Tamner et al. 2022). Their design philosophy draws on the same cognitive models of moral intelligence for the design of serious moral games (Christen et al., 2012) as our own work.

Our approach to serious moral game design is founded in the Four Component Model of moral intelligence, developed by psychologist James Rest and colleagues (Rest et al., 1999). The model is comprised of four broad categories of cognitive/affective capabilities: *moral focus* (prioritizing morality), *moral sensitivity* (recognizing morality in the world), *moral judgement* (judging what morality requires), and *moral action* (doing the right thing). This breakdown of moral intelligence into separate components can inform the design of serious moral games by allowing us to specify design goals for each component and target the design of formal game elements (such as narrative and mechanics) towards achieving these goals. It invites designers to consider the questions:

- **Moral Focus**: Why is morality a priority? What motivates the player to treat moral decisions as moral decisions, and not instrumentally?

- **Moral Sensitivity**: How is moral content presented to the player? Are the important factors clearly signposted or is the player expected to recognise them without prompting?

- **Moral Judgement**: What are the issues at stake? What are the different moral codes or norms that might drive players' choices? What factors complicate their choices?

- **Moral Action**: Is a moral problem solved once a choice is made, or does the player have to put it into action? How difficult is it to put your choices into action? What skill is required? (Staines, Formosa, et al., 2019)

The 4CM is a dual process theory recognising the impact of both implicit/automatic and explicit/deliberate mental processes at play in moral decision making (Lapsley & Hill, 2008). Explicit reasoning 'guides the individual in determining action' and provides 'objective rationale that can be challenged, revised, reputed, or accepted' (Narvaez, 2006, pp. 718-19), but morality also needs to be put into practice. Our ethical codes need to be automatized into unconscious ways of reading and interacting with the world. Real-world problems are rarely explicitly ethical. Rather, they arise when our conscious reasoning is engaged on more obvious pragmatic issues, leading to ethical fading unless habits of ethical thinking are implicitly established.

This model of moral intelligence connects well with the social constructivist approach to learning through discovery and problem-solving, championed by James Paul Gee and others, which remains one of the most popular approaches to serious games design (Gee, 2003; Ryan, Costello, & Stapleton, 2012). The Morality Play model of Staines, Formosa and Ryan (Staines, Formosa, et al., 2019) sets out how such a game can be

designed, starting with the development of a 'moral toy' that provides 'a simulation of a morally significant domain'. Players' experience of this domain is scaffolded through stages of discovery and mastery of increasingly complex ideas and skills. Through practice, deliberate ethical reasoning becomes automatic. Conversely, we can invite the player to make their implicit priorities explicit through reflection and discussion, either in-game with non-player characters (NPCs) or out-of-game with a community of fellow players (e.g., through moderated discussion forums).

REQUIREMENTS

Learning Outcomes

We began the design process by establishing a clear statement of our intended learning outcomes for the game. Our primary goal was to build greater awareness of the five principles of cybersecurity ethics described above, and to build habits of moral reflection in our players when they are in the midst of solving complicated technical or personnel problems, to counteract the problem of ethical fading. Using the 4CM, we broke this down into specific goals for each component.

Our goal for moral focus (*LO-MF*) was to establish and maintain the player's moral engagement in the work in a dynamic environment with other priorities and distractions. It is easy to have moral focus when presented with a clearly signposted ethical dilemma, but it is harder to maintain that focus in a complex environment with other competing priorities and distractions. We want our game to exercise and improve the player's moral focus by reminding them to keep moral questions in mind even when there are many other things to think about.

For moral sensitivity, we had two goals. The first (*LO-MS1*) was to demonstrate the ability to identify the five different categories of ethical issues relevant to cybersecurity. We want players to not only know the definitions of the five principles, but also to be able to recognise them in action and use them to analyse ethical problems as they arise.

The second goal for moral sensitivity (*LO-MS2*) was the ability to recognise competing perspectives in ethical decision making and have empathy for different points of view. We wanted the player to be able to recognise the different stakeholders in a decision and understand what outcome they might prefer and why.

With regard to moral judgement, our first goal (*LO-MJ1*) was for the player to be able to make decisions based on ethical principles and relevant cybersecurity factors within time and resource constraints. While we wanted players to make judgements based on ethical principles, it is also important to recognise that pragmatic circumstances may limit the available options, and trade-offs need to be made. We also wanted to encourage players to reconsider and re-evaluate moral judgements at multiple decision points based on an emerging scenario as new information comes to light (*LO-MJ2*), while maintaining ethical consistency across multiple decision points (*LO-MJ3*).

Finally, we wanted to provide the player with scope for moral action by allowing for a variety of ways to put moral judgements into effect. Our aim was to avoid the 'make it so' approach of many ethical narrative games, in which a moral choice is implemented as soon as it is decided. Instead, we aimed for the player to learn to strategise ways to achieve their ethical goals (*LO-MA*).

As can be seen from this discussion, the Four Component Model provided a useful set of lenses to consider the distinct ethical skills involved in decision-making and allowed us to refine our broad overall goal into a selection of specific learning outcomes to inform ongoing design practice.

Target Audience

While general, and specialised cybersecurity, IT professionals were a clear target audience for our game, we sought to narrow this focus for our initial game prototype to a more accessible sample for which principlist cybersecurity ethics knowledge and practice also represented an important foundational base. The initial target audience was thus defined as 'University-level cybersecurity (or other IT) students for whom the themes of cybersecurity ethics are relevant'. Pertinent player demographics for this group were not immediately clear, so we had to gather user requirements and establish representative player personas for a user-centred agile game development approach. At this point, the settings for game genre and mechanics that would effectively motivate and engage this demographic group were also unclear. As such, we designed a study to capture these requirements and establish a clear picture of our target audience to guide subsequent design and development.

The target audience study was deployed as an online survey with a range of relevant demographics questions, such as age, gender, year of enrolment in their degree, game platform and genre preferences, and exposure to serious moral games. To get a clearer picture of player preferences, a version of Yee and Quantic Foundry's *Gamer Motivation Model* (Yee, 2016) was included, where participants were asked to rate the importance of 12 distinct motivations for play on a 5-point Likert scale from 'not important at all' to 'extremely important'. Further to this, a set of questions was included to sample previous gaming experience (across all video games of any kind) in line with McEwan et al.'s Game Technology Familiarity (GTF) instrument (McEwan, Blackler, Wyeth, & Johnson, 2020).

An undergraduate introductory cybersecurity unit at our Sydney-based university was identified as an appropriate testbed for the initial game prototype in an upcoming semester, and so students in an earlier offering of this unit were invited to take part in our requirements study. In all, 125 students took part in the online questionnaire, with 110 completing enough of the survey to be included in subsequent analysis.

Analysis of the captured demographics data revealed important and overlapping patterns. These broadly coalesced into three distinct player profiles, including trends related to previous gaming experience (and a more hardcore male audience that prefers certain play styles), gender (and a more casual female audience that prefers certain interactions), and age (with an older, more moderate audience possessing distinct player motivations). To facilitate factoring these different target audience characteristics into the design and development of the game prototype, three personas were composed that encapsulated the key findings of the target audience study.

1. **Wei:** Wei is 18 and identifies as male and plays video games (mostly FPS and some MOBA) on PC. He self identifies as a hardcore gamer, playing games for more than 20 hours per week (and often claiming he plays video games more than he does anything else – including sleep!). The most important motivational aspects of video games for him are those related to challenge and story.
2. **Kat:** Kat is 19 and identifies as female and predominantly plays games on her tablet PC or smartphone. She is more of a casual player, playing for less than 10 hours per week, but still enjoys story driven RPGs, interactive narratives and puzzle games. She's open to different types of games but knows that she doesn't enjoy destruction, violence or the competition that she's seen in online communities, preferring to take her time with the games that she plays by herself before bed or on public transport.
3. **Sam:** Sam is 24 and identifies as non-binary and is a moderate gamer that plays games across current-gen consoles, smartphone and PC. They play for between 10-20 hours per week. Their favourite genre is action adventure, and they particularly

like the Uncharted and Tomb Raider games. They are motivated to play by game elements that support fantasy fulfilment and destruction.

These personas formed an important role in early design workshops and were also used to communicate the user requirements to the development team, who subsequently posted them on their design requirements board as a reference point to guide prototype production.

DESIGN

Following the Morality Play model, our next task was to design the 'moral toy' i.e., 'a simulation of a morally significant domain which implicitly represents important moral concepts and allows for sophisticated moral judgement and action' (Staines, Formosa, et al., 2019, p8). In this case, we aimed to simulate the kinds of day-to-day decisions that might be made in corporate cybersecurity, responding to a variety of security problems and threats. The design of the game focused around two major concerns: developing the narrative (setting, characters, and story) and developing the mechanical systems of the game.

Narrative

Setting and Characters

The narrative setting was designed to facilitate the game's learning outcomes. Since the aim was to teach cybersecurity ethics in an organic context, we situated the player as a Lead Security Analyst in a fictional cybersecurity firm called 'Prescott and Kreuger'. We intentionally chose a commercial firm rather than a military/intelligence setting, to avoid dealing with complex issues around just war theory that would arise in the context of national intelligence cybersecurity (Nissenbaum, 2005). Since the company provides cybersecurity services to other companies, we were able to explore ethical issues that arise within the company itself and with its relationship to other parties, including clients for its cybersecurity services and the broader public.

Prescott and Kreuger's narrative design established it as a medium-size firm offering a range of IT solutions, including network infrastructure provision, cloud data management and cyber security. Their consultants work with clients to develop integrated, tailored services. In the cybersecurity area, Prescott and Kreuger provides security audits, threat identification and general consultancy, as well as technology solutions. The company description, structure and job descriptions were fleshed out through review of: 1) current recruitment advertisements for security positions (which include job title, levels, skills, responsibilities, organization description) in a range of private and public sector organizations; 2) the Skills Framework for the Information Age (SFIA) which is used globally by IT professional associations and by employers such as government organisations to define the skills and competencies of a specific position (SFIA Foundation, 2018); and 3) organisational charts for medium-sized security companies.

We chose a medium sized organization with the player occupying a mid-level role to give the player sufficient scope and autonomy for decision making while still having to act under organisational constraints. A larger organisation would have risked the player considering themselves unable to affect organisational priorities and directions. Conversely, for a smaller organisation, the player would have been placed in a position of too much power without having to consider the constraints of company policy. Similar considerations applied to the player role. Narratively placing the player in charge of a team, while still being subordinate to higher level employees, gave them scope for independent decision making while managing employees within their team,

but still having to justify their decisions to those higher up in the context of the organisation.

Based on our review, we defined the player character as 'Alex', a Lead Security Analyst, who is both responsible for a small team of people and also reports to someone more senior (CTO/CIO), while also being part of a senior team with other managers (including HR, Marketing, and Legal). The player character's name, and subsequently designed visual representation, was intentionally gender neutral to support a more inclusive and accessible narrative entry point for the player and to appeal to our three target personas[1].

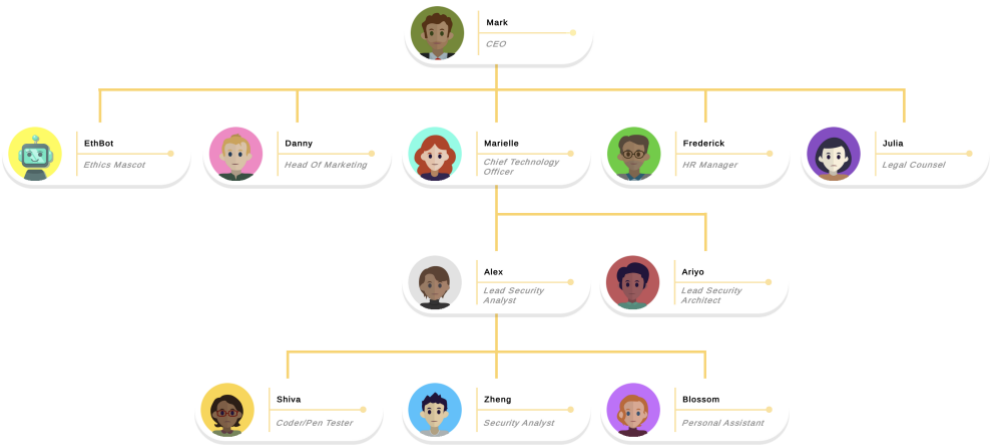

Figure 2: The Prescott & Kreuger org chart showing the different characters the player (as Alex) interacts with in the game.

This broader set of non-player characters (NPCs) were also introduced both to make the context interesting and plausible, but also to help facilitate the various learning outcomes. Figure 2 shows the complete org chart of characters. Several subordinate roles were created to give the player the opportunity to make leadership decisions and respond to various forms of ethical and unethical conduct for which they are, as a leader, partly responsible. This helps to challenge their moral sensitivity (*LO-MS2*), judgement (*LO-MJ1*) and leadership (*LO-MA*) skills. A peer NPC was created to allow for both social interactions and peer-level reflection (*LO-MF* and *LO-MJ1*). Higher level NPCs were designed to allow for issues around accountability and transparency to be explored, while NPCs external to the company support the depiction of a wide variety of cybersecurity contexts, allowing for a robust library of ethical scenarios to be developed and providing a range of contexts for players to demonstrate and deploy various moral skills. NPCs were also chosen to be racially diverse, including characters from a variety of Asian and European nationalities, to coincide with the nationalities of our target audience.

In terms of story development, we began by reviewing the cybersecurity ethics literature and identified key cases that were the most widely discussed (for an overview, see Formosa et al. 2021). This led us to develop a range of scenarios around DOS (denial of service) and DDOS (distributed denial of service) attacks, ransomware, penetration testing (including white, black and grey hat hacking, and bug bounties) and

[1] However, our previous design experience has shown that players of all genders tend to assume a player character is male in the absence of other information, so playtesting will be needed to validate this design choice.

system administration (including managing security and network settings and formulating and policing ICT policies). We then developed narratives around each of these types of scenarios that would allow players to experience the full range of ethical conflicts between our five ethical principles. For example, a ransomware scenario allows us to consider the benefits of paying a ransom in terms of getting quick access to encrypted data (*beneficence*), the potential costs this poses to others by incentivising more ransomware attacks (*non-maleficence*), issues around whether the client should be free to make whatever decision they wish in this regard (*autonomy*), any legal obligations that might be at play in terms of disclosure of a data breach (*justice*), and the appropriate level of transparency and accountability for any cybersecurity failures that occur (*explicability*).

Game mechanics

The mechanics of the game are divided into two interacting systems. First there is a scripted branching narrative system, told through a series of conversations with NPCs via email and a workplace chatroom interface. Alongside this there is a resource management system, involving management of both the player's and NPC's time, as well as NPC morale.

To challenge the player's moral focus (*LO-MF)*, we designed the narrative system to implement multiple simultaneous narrative threads playing out through email, chat, and other applications, as shown in Figures Figure 3 and Figure 4. Some of these threads are ethically important, while others are distractions. The player must deal with communications from a variety of other employees and external stakeholders and is under pressure to do their job well, with competing moral and pragmatic concerns.

Figure 3: The email interface in the game introducing an urgent DDoS attack which the player needs to address.

This setting then raises problems of moral sensitivity, which is the ability to read moral situations and recognise the competing moral concerns. This is where we introduce the five ethical principles previously outlined – *beneficence*, *non-maleficence*, *autonomy*, *justice* and *explicability* (*LO-MS1*). When a major issue arises the NPCs present different points of view on what should be done and why, and the player then needs to evaluate these alternatives and choose a path of action (Figure 4).

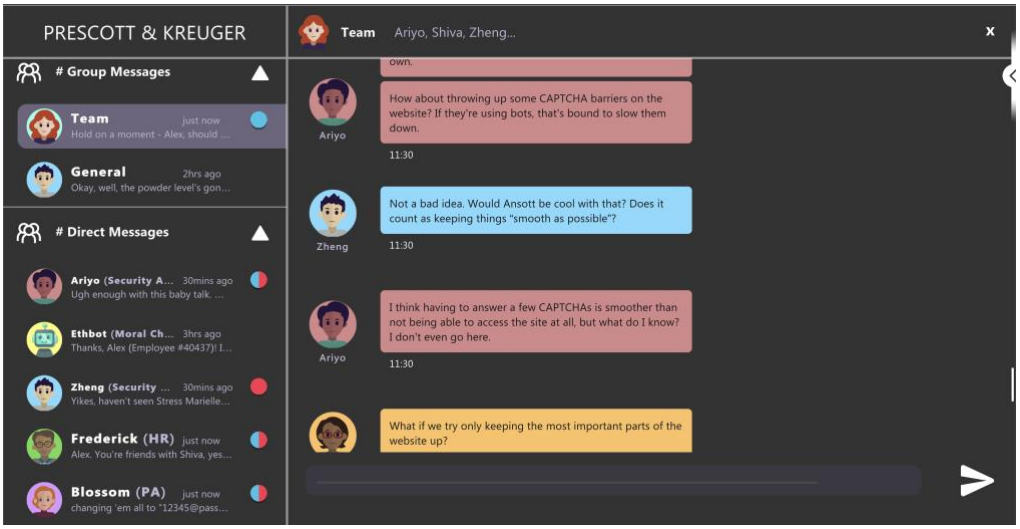

Figure 4: The chat interface in which NPCs provide advice on possible ways to deal with the DDoS attack.

Complicating this decision is a model of NPC morale, which provides each character with a 'moral compass' that gives different priorities to the five ethical principles. Decisions which favour a character's ethical priorities can improve their morale, while decisions that oppose those priorities can decrease morale. Morale is a variable that can affect story outcomes for better or worse, so the player needs to have empathy for NPC's points of view even when they act in ways that oppose them (*LO-MS2*).

To scaffold the players' evaluation skills, early in the game a mentor character, depicted as an automated AI ethics advisor called EthBot, prompts the player to consider the particular ethical principles that are at play in each given scenario (Figure 5). As the game progresses, this scaffolding is gradually taken away, and the player is expected to recognise these principles on their own initiative.

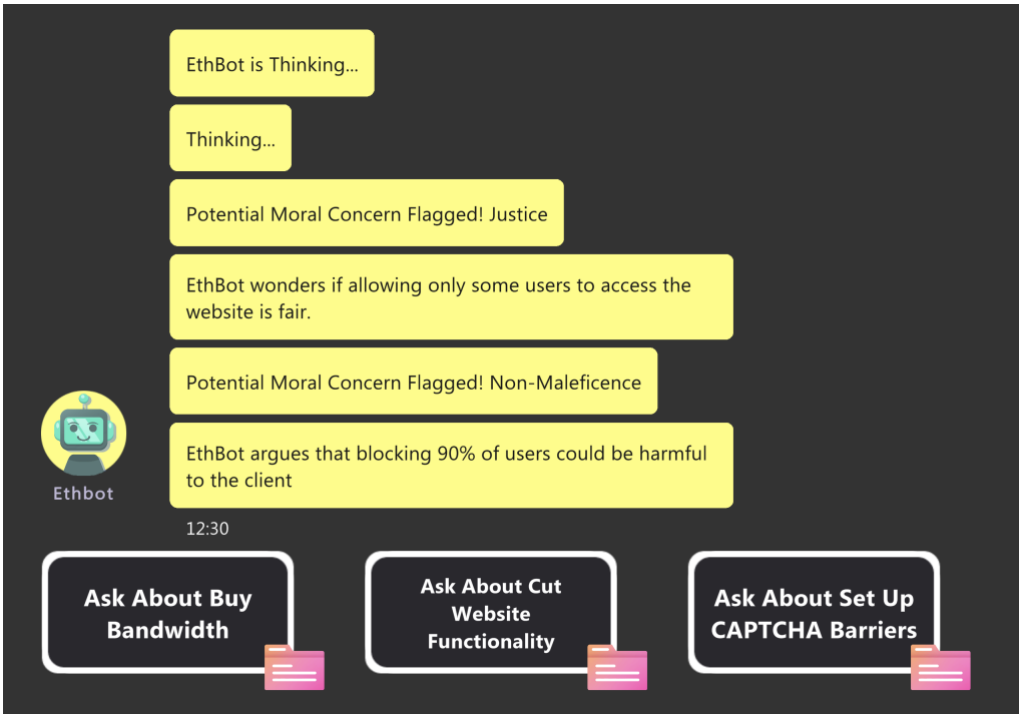

Figure 5: Asking for EthBot's advice about the ethical implications of dropping website traffic to counter the DDoS attack.

Making a moral judgement and putting it into action is complicated by the underlying resource economy of the game. The player has limited time and staff resources to spend on a problem. Assigning a staff member to implement a solution occupies them for a period of time, making them unavailable for other tasks (Figure 6). Discussing the problem at length also takes up time which might be better spent on other actions to address emerging narrative and ethical problems. In cases where there are multiple incidents at play simultaneously, the player needs to prioritize how resources will be assigned to each (*LO-MJ1*). This time economy also creates opportunities for the player to reconsider their decisions as the situation evolves, and new information comes to light (*LO-MJ2*).

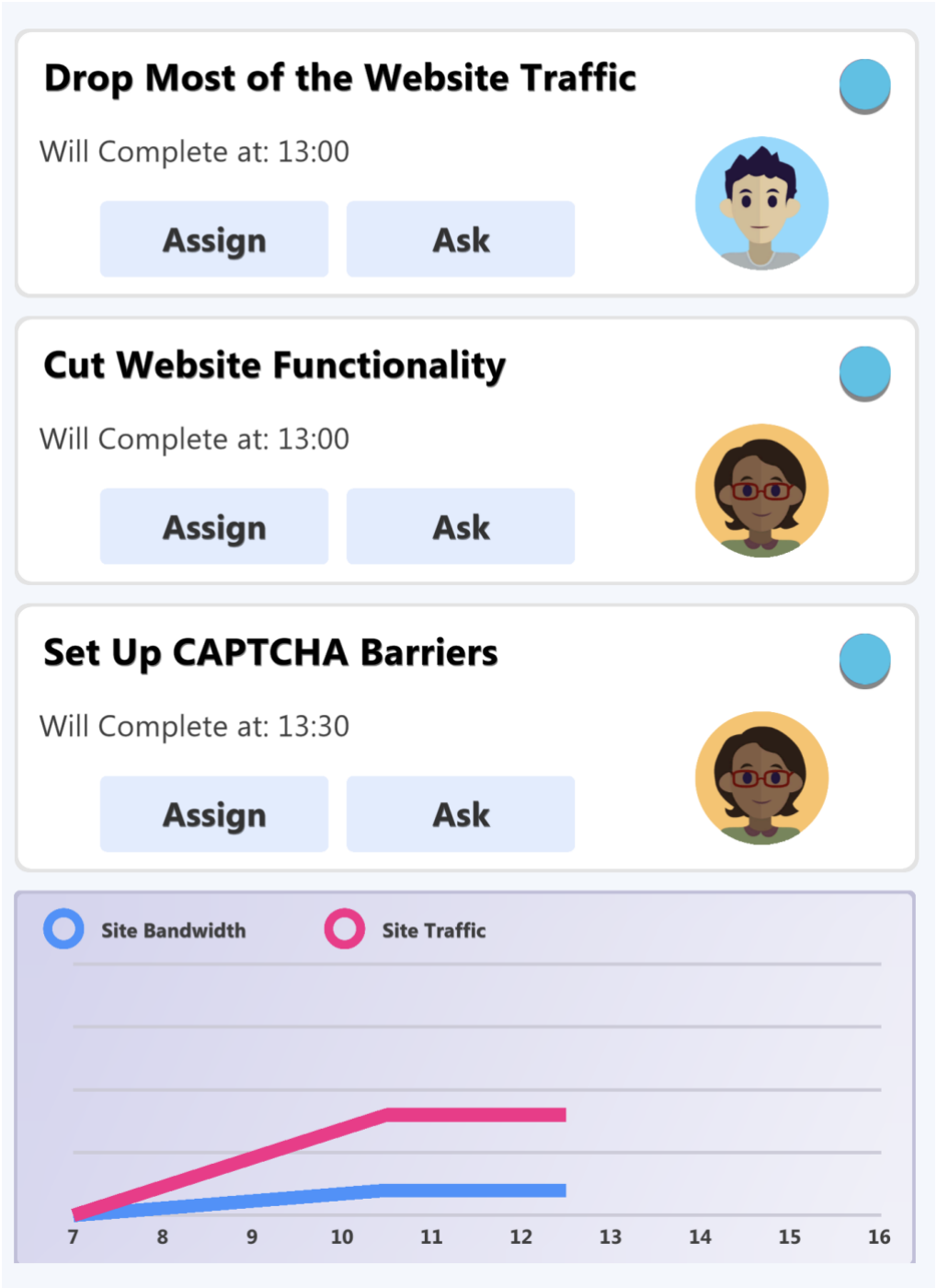

Figure 6: The resource economy. Assigning NPCs to complete tasks to address the DDoS takes time, during which the problem escalates.

In the design of these systems, we have attempted to increase the *atomicity* of choice (Formosa, Ryan, & Staines, 2016; Sicart, 2013), presenting each major decision as a collection of smaller choices that can interact and combine in different ways. While choices have specific scripted outcomes, they also interact through the underlying economy of the game, allowing a richer set of possibilities and outcomes. This increases the strategic space of the game and is intended to provide a sense of actual problem solving, rather than picking from a list of designer approved alternatives (*LO-MA*).

Finally, the game includes a reflection mechanic which appears periodically after major decisions in the game are complete. The designed narrative presents this as part of an 'HR-led initiative' to improve ethical judgement, in which the player is expected to complete a report outlining the main ethical considerations supporting and opposing the major decision they made, labelling each according to the relevant ethical principle it represents (Figure 7). This reflection is intended to remind the player of the main ethical principles and connect them to the choices they made (LO-MS1).

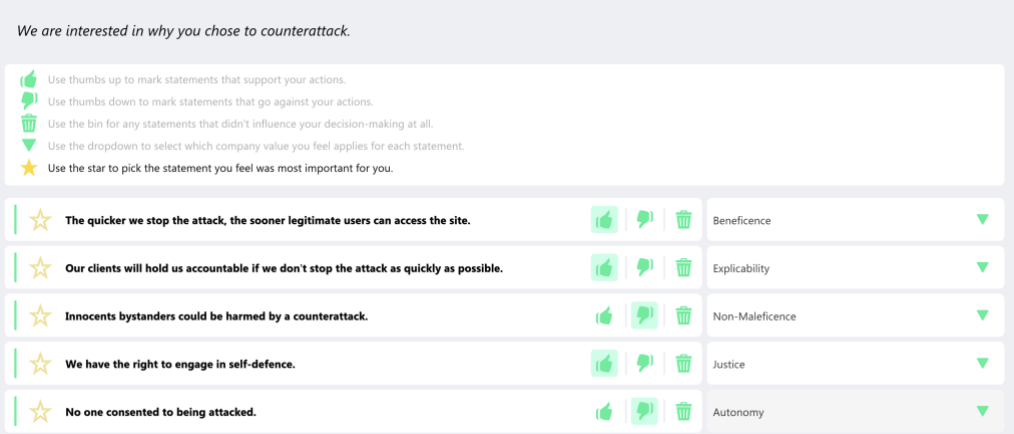

Figure 7: The reflection report invites the player to support their decision with ethical arguments for and against.

The current game prototype comprises a single 30-minute episode of what is planned as a multi-episode arc dealing with a variety of different ethical scenarios. In the longer version of the game, we plan to include modified scenarios in which a similar decision must be made under varying circumstances. For example, a DDoS scenario might reappear, but with a hospital as the target rather than an airline. Our plan is for such a scenario to include call-backs to the player's previous decisions and reflections, inviting them to reflect on their ethical consistency between episodes (*LO-MJ3*).

REFLECTIONS ON THE DESIGN

Overall, we found the Four Component Model very useful as a tool for designing this game. It provided a variety of perspectives on what a player is doing when they are engaging in moral decision making, which allowed us to formulate specific learning outcomes and design elements to address each. The 4CM doesn't provide immediate design solutions, rather it helped us to ask the right questions.

We found strong connections between the components of the 4CM and the common game design conceptions of *agency* and *engagement* (Phillips, Horstman, Vye, & Bransford, 2014; Wardrip-Fruin, Mateas, Dow, & Sali, 2009). Moral sensitivity, judgement and action together form the elements of agency, in which the player 'reads' the world, decides what to do, and then enacts their decision. These skills can be developed through a familiar process of discovery and mastery, in which explicit moral knowledge is translated into implicit moral skill. Moral focus, on the other hand, is more akin to engagement, as a force that maintains the players' motivation to engage with the game in moral terms rather than in purely strategic terms. Designing for moral focus parallels the more general problems of maintaining player engagement.

While the 4CM proved a useful tool, the more specific design recommendations of the Morality Play model were difficult to implement. Morality Play is based on the common constructivist model of serious game design which sees the players learn through scaffolded interaction with a "toy" model of the learning domain (Ryan et al.,

2012; Staines et al., 2019). Important domain concepts are implemented as dynamics and strategies within the system, allowing the player to discover them through concrete experience rather than as labelled abstractions. The Morality Play model also recommends this approach to serious moral game design, however we found this difficult to achieve in practice in our domain of interest given our limited resources.

This difficulty lay in the disconnect between the scripted and systemic elements of the game. Early in the design process, we investigated more detailed systemic models of cybersecurity management, such as those exhibited by the (non-moral) serious game *CyberCIEGE* (Thomps & Irvine, 2011). This approach presented two problems. First, designing such a systemic model that would be able to encompass all the different ethical scenarios we wanted to include would be a monumental task, given the diversity of underlying technical and social systems behind a DDoS attack, a ransomware demand, penetration testing and other system administration problems. Building such a comprehensive system was well beyond the scope of our project.

However, even if we focused on just one of these scenarios, we still found great difficulty in making the leap in abstraction between the player's choices in administering the system and the ethical ramifications of their choices. We knew that in order to scaffold the player's learning of the ethical principles, we would need to have other in-game characters present ethical perspectives on the player's choices, rather than always leaving it up to the player to notice the ethical outcomes for themselves. But the more fine-grained the player's actions were, and the broader the resulting space of strategies, the harder it would be to program NPCs that were able to interpret and comment on the result. Another human being might be able to look at the player's decision and give moral commentary, but it would require advanced artificial intelligence (AI) for an NPC to do the same.

Instead, we resorted to a simple office-management simulation with a largely scripted interaction. This allowed us to represent a much wider variety of ethical scenarios but sacrificed a large element of realism and limited the space of strategies the player could explore. To achieve some of our goals with regard to providing a rich simulation of moral play, we designed the narrative to provide more choices along the path to each major moral decision, so decisions did not appear as a monolithic moral dilemma with only designer-prescribed solutions, however this disguise is relatively thin, as the player's options for action are still heavily circumscribed. The systemic elements of time and morale management help provide a more dynamic sense of play, but do not contribute as much space for 'sophisticated moral judgement and action' as we would like.

This would appear to be a broader issue in the design of serious moral games in general. Morality is not a domain in and of itself, rather it is an interpretation we place over other more concrete domains of learning, such as cybersecurity, medicine, or economic policy. To design a serious moral game in such domains, we must not only model ethics but also the domain itself. With what fidelity should we do this? If we want to seriously address the problem of ethical fading, we need to allow for the possibility for the player to get lost in the technical detail of domain-specific problems, but at the same time we need to be able to provide ethical feedback in order to engage the player's moral focus and scaffold the development of their moral sensitivity and judgement. This is a very challenging design problem with significant resource implications.

CONCLUSIONS & FUTURE WORK

We have presented a case study of how the Four Component Model of moral intelligence and the Morality Play model of serious moral game design can be used as design tools in the design of a game to train cybersecurity ethics. We found the 4CM a

particularly useful tool in our design process, prompting the development of specific learning outcomes for each component of moral focus, sensitivity, judgement, and action, with narrative and mechanical elements targeted to achieve these goals. So far, we have only completed design of the first major prototype. External playtesting of this prototype is now underway to verify that it meets our expectations.

Once the game is complete, we plan to conduct experimental evaluation to measure how well it achieves our target learning outcomes. This may also require the development of novel instruments for the evaluation of moral engagement and situational moral development, as existing measures from moral psychology tend to focus either on intrinsic moral character or long-term development of moral perspective (Ryan et al., 2019).

Reflecting on this design experience, we have noted the usefulness of the 4CM but also realised some difficulties with the more prescriptive Morality Play model and with serious moral game design in general, when it comes to situating ethical concerns within a specific domain of action such as cybersecurity. How detailed should our simulation of the domain itself be? Ethical principles are largely abstract, but ethical decision-making is grounded in the specific details of the domain. To present the variety of ethical problems associated with any realistic domain, we either need a very broad and detailed simulation of many aspects of the domain, or else we need to discard detail and focus on decision making at a more abstract level. Working at the abstract level can make moral choices too obvious and constrained and fails to address the problem of ethical fading. Working at the detailed level, on the other hand, presents us with a difficult interpretive problem when we want to give players moral feedback on their actions. Addressing this problem remains a major challenge for the field of serious moral game design.

ACKNOWLEDGMENTS

We would like to acknowledge Meredith Porte and the team at Chaos Theory Games who made significant contributions to the design of this game and implemented the current prototype.

BIBLIOGRAPHY

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

Brey, P. “Ethical Aspects of Information Security and Privacy.” In *Security, Privacy, and Trust in Modern Data Management*, edited by Milan Petković and Willem Jonker, 21–36. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007. https://doi.org/10.1007/978-3-540-69861-6_3.

Christen, M., Faller, F., Götz, U., & Müller, C. (2012). *Serious Moral Games. Analizing and Engageing Through Video Games.* Insitute for Design Research, Zürich University of Arts, http://www.zhdk.ch/index.php?id=39457

Christen, M., Gordijn, B., and Loi, M., eds. *The Ethics of Cybersecurity*. Vol. 21. The International Library of Ethics, Law and Technology. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-29053-5.

Consalvo, M., & Staines, D. (2021). Reading Ren’Py: Game Engine Affordances and Design Possibilities. *Games and Culture*, *16*(6), 762–778. https://doi.org/10.1177/1555412020973823

FableVision Studios. (2012). *Quandary* (p. Game [Web]. Learning Games Network).

Learning Games Network. https://www.quandarygame.org

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., et al. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28, no. 4 (December 2018): 689–707. https://doi.org/10.1007/s11023-018-9482-5

Formosa, P., Ryan, M., Staines, D. (2016). Papers, Please and the systemic approach to engaging ethical expertise in videogames. *Ethics and Information Technology*, *18*(3). https://doi.org/10.1007/s10676-016-9407-z

Formosa, P., Wilson, M., Richards, D.. "A Principlist Framework for Cybersecurity Ethics." *Computers & Security* 109 (October 2021): 102382. https://doi.org/10.1016/j.cose.2021.102382.

Francis, K. B., Howard, C., Howard, I. S., Gummerum, M., Ganis, G., Anderson, G., & Terbeck, S. (2016). Virtual morality: Transitioning from moral judgment to moral action? *PLoS ONE*, *11*(10), 1–22. https://doi.org/10.1371/journal.pone.0164374

Gee, J. P. (2003). *What video games have to teach us about learning and literacy* (1st ed.). New York: New York : Palgrave Macmillan.

Hilliard, L. J., Buckingham, M. H., Geldhof, G. J., Gansert, P., Stack, C., Gelgoot, E. S., Bers, M. U., Lerner, R. M. (2018). Perspective taking and decision-making in educational game play: A mixed-methods study. *Applied Developmental Science*, *22*(1), 1–13. https://doi.org/10.1080/10888691.2016.1204918

Ilten-gee, R., & Hilliard, L. J. (2019). Moral reasoning in peer conversations during game-based learning : An exploratory study. *Journal of Moral Education*, *50*(2), 140-165. https://doi.org/10.1080/03057240.2019.1662775

Katsarov, J., Biller-Andorno, N., Eichinger, T., Schmocker, D., Christen, M. (2020). uMed: Your Choice---Conception of a Digital Game to Enhance Medical Ethics Training. In M. Groen, N. Kiel, A. Tillmann, & A. Weßel (Eds.), *Games and Ethics: Theoretical and Empirical Approaches to Ethical Questions in Digital Game Cultures* (pp. 197–212). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-28175-5_13

Katsarov, J., Christen, M., Mauerhofer, R., Schmocker, D., Tanner, C. (2019). Training Moral Sensitivity Through Video Games: A Review of Suitable Game Mechanisms. *Games and Culture*, *14*(4), 344–366. https://doi.org/10.1177/1555412017719344

Kobold Games. (2018). *uFin: The Challenge,* Game [Android], https://www.koboldgames.ch/project/ufin

Kobold Games. (2019). *uMed: Your Choice* Game [Windows], Unpublished. https://www.koboldgames.ch/project/umed

Lapsley, D. K., & Hill, P. L. (2008). On dual processing and heuristic approaches to moral cognition. *Journal of Moral Education*, *37*(3), 313–332. https://doi.org/10.1080/03057240802227486

Manjikian, Mary. *Cybersecurity Ethics: An Introduction*. London ; New York:

Routledge, Taylor & Francis Group, 2018.

McEwan, M., Blackler, A., Wyeth, P., & Johnson, D. (2020). Intuitive Interaction with Motion Controls in a Tennis Video Game. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (pp. 321–333). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3410404.3414242

Narvaez, D. (2006). Integrative Ethical Education. In M. Killen & J. Smetana (Eds.), *Handbook of Moral Development* (pp. 703–732). Mahwah, NJ: Erlbaum.

Nissenbaum, H. (2005) "Where Computer Security Meets National Security." *Ethics and Information Technology* 7, no. 2 (June 2005): 61–73. https://doi.org/10.1007/s10676-005-4582-3

Phillips, R. S., Horstman, T., Vye, N., Bransford, J. (2014). Engagement and Games for Learning: Expanding Definitions and Methodologies. *Simulation \& Gaming*, *45*(4–5), 548–568. https://doi.org/10.1177/1046878114553576

Rest, J. R., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (1999). *Postconventional Moral Thinking: A Neo-kohlbergian Approach*. Lawrence Erlbaum Associates.

Ryan, M., Costello, B., & Stapleton, A. (2012). Deep learning games through the lens of the toy. In *Meaningful Play 2012 Conference Proceedings*. United States: Michigan State University.

Ryan, M., Formosa, P., Howarth, S., & Staines, D. (2019). Measuring morality in videogames research. *Ethics and Information Technology*. https://doi.org/10.1007/s10676-019-09515-0

Ryan, M., Staines, D., & Formosa, P. (2017). Focus, sensitivity, judgement, action: Four lenses for designing morally engaging games. *Transactions of the Digital Games Research Association*, *3*(2).

Schrier, K. (2015). EPIC: A framework for using video games in ethics education. *Journal of Moral Education*, *44*(4), 393–424. https://doi.org/10.1080/03057240.2015.1095168

SFIA Foundation. (2018). SFIA 7. Retrieved from https://sfia-online.org/en/sfia-7

Sicart, M. (2013). *Beyond choices: The design of ethical gameplay*. MIT Press.

Staines, D., Consalvo, M., Stangeby, A., & Pedraça, S. (2019). State of play: Video games and moral engagement. *Journal of Gaming and Virtual Worlds*, *11*(3), 271–288. https://doi.org/10.1386/jgvw.11.3.271_1

Staines, D., Formosa, P., & Ryan, M. (2019). Morality Play: A Model for Developing Games of Moral Expertise. *Games and Culture*, *14*(4). https://doi.org/10.1177/1555412017729596

Tanner, C., Schmocker, D., Katsarov, J., & Christen, M. (2022). Educating moral sensitivity in business: An experimental study to evaluate the effectiveness of a serious moral game. *Computers & Education*, *178*, 104381. https://doi.org/10.1016/j.compedu.2021.104381

Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research*, *17*(2), 223–236. https://doi.org/10.1023/B:SORE.0000027411.35832.53

Thomps, M., & Irvine, C. (2011). Active learning with the CyberCIEGE video game. *4th Workshop on Cyber Security Experimentation and Test, CSET 2011*, 1–8.

Vallor, Shannon. "An Introduction to Cybersecurity Ethics." Markkula Center for Applied Ethics, 2018. https://www.scu.edu/media/ethics-center/technology-ethics/IntroToCybersecurityEthics.pdf.

Wardrip-Fruin, N., Mateas, M., Dow, S., & Sali, S. (2009). Agency Reconsidered, (1985), in Proceedings of DiGRA 2009

Williams, R. N., & Gantt, E. E. (2012). Felt moral obligation and the moral judgement–moral action gap: toward a phenomenology of moral life. *Journal of Moral Education*, *41*(4), 417–435. https://doi.org/10.1080/ 03057240.2012.665587

Yee, N. (2016). The Gamer Motivation Profile: What We Learned From 250,000 Gamers. *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, 2.