UNLOCKING ETHICS: DESIGNING AN ETHICS ESCAPE ROOM

VINESH KANNAN & MARIA GARGIULO

1. INTRODUCTION

Escape rooms, in which participants are locked in a themed room and solve puzzles to escape, have become increasingly popular. As of August 2019, there are more than 2,350 escape room establishments in the United States alone (Spira, 2019). More than just entertainment, however, escape rooms have made their way into classrooms and occupational settings, providing another medium to present educational content and other training materials (see Adams et al. (2018), Eukel et al. (2017), and Nicholson (2018) for examples). This paper provides a designer's postmortem of our attempt to create an escape room game about ethics.

Ethical questions are often high-stakes and serious. People may feel that discussions of ethics are too theoretical to have relevant practical implications, defensive of their own moral worth when given ethical advice, or frustrated that others are not held accountable for unethical conduct. At face value, this does not seem suitable for an engaging game. Yet, we see ethics and escape rooms as structurally complementary: both involve finding

information, solving puzzles, and separating the details that matter from those that are just part of the environment.

The ethics escape room brings to life the philosophical debate about the permissibility of torturing someone involved in an act of terror. Instead of physically locking players in a room, we created a "moral lock" for the room that could only be opened through thoughtful deliberation of ethical concepts. Players have 60 minutes to discover clues about the situation and solve puzzles that introduce and exercise ethical concepts, culminating in a morally challenging decision followed by a 30 minute reflective discussion.

We refined the game over three distinct prototype versions, with 18 total players playing six sessions. Overall, players found our moral lock compelling: they used the information on ethical concepts we presented to them to solve puzzles and used their full allotted time to have complex discussions about actions and potential ramifications. Each group's discussions were distinct and they yielded diverse gameplay outcomes, even in situations where groups were presented with the same information.

The escape room format provided a useful for creating an engaging ethics-oriented game prototype that has potential for further refinement in physical play environment, puzzle design, and presentation of player results. Additionally, we are excited by the possibility of expanding this work to explore other ethical questions.

2. DESIGN

2.1. Target Audience

Escape rooms are marketed as games that players solve with wit and ingenuity. Still, in the spirit of universal design, escape room designers are often advised to make sure that all their puzzles can be completed by people with high school level reading and

math proficiency, possibly less (Chen, 2016). We wanted to work with content that is often considered dense, so we relaxed this constraint and chose a more specific target audience:

- Age: 18

- Education: Some college

- Comfortable using mobile phone applications

- Some interest in philosophy, law, or politics

- Competitive, but also enjoy collaborative games

- Have friends also in the target audience, or comfortable playing with acquaintances

2.2. Choosing a Scenario

We set three requirements for selecting a scenario to base the escape room on:

- Accessible: Players in the target audience understand the vocabulary, context, and consequences of the scenario.

- Plausible: Players believe that the scenario could occur in real life, that important facets of the scenario can exist in the game, and that their actions can affect the outcome.

- Divisive: Players disagree over how to act in the scenario, preferably the scenario already has a body of work with competing viewpoints.

The trolley problem is perhaps the most well-known ethics thought experiment. In this scenario, the player must decide whether or not to divert a trolley: if they do nothing, the trolley will run over five people, if they divert it, those five will be saved, but a different person on the other track will die. Trolley problems tend to be accessible and divisive, but not plausible.

One incarnation of the trolley problem satisfied all three: the

ticking time bomb scenario. In this thought experiment, authorities captured a subject with knowledge that could prevent an impending terrorism act that would kill many people. Players debate whether it is permissible to torture the subject to extract their information.

2.3. Ethical Framing

The ethical grounding of our game consists of three perspectives: a pedagogical framework, a philosophical perspective, and a body of literature.

2.3.1. Pedagogical Framework

Our pedagogical framework defines what it means for the players to engage in "ethics" and what kinds of ideas we consider as "ethical." We use the Framework for Ethical Decision-Making from the Markkula Center for Applied Ethics at Santa Clara University (2009). We selected it because we have previously used it to teach students in the target audience. The framework lists ten reflective questions separated into five main steps:

1. Recognize an ethical issue
2. Get the facts
3. Evaluate alternative actions
4. Test decisions
5. Act and reflect on the outcome

The framework states that ethics is distinct from feelings, religion, law, cultural norms, and science. It also identifies five sources of ethical standards: utility, rights, fairness/justice, virtue, and the common good. We represent these sources in our game with ideas from philosophers that our target audience may already have exposure to, such as Aristotle (virtue), Jeremy Bentham (utility), and W.D. Ross (rights).

2.3.2. Philosophical Framework

We designed our game around moral particularism, which argues that there are no universal ethical principles that can apply in the same way in all scenarios (Dancy, 2004, 2017; Hare, 1963). Moral particularists must practice two habits: (1) learning many ethical principles and (2) identifying how those principles interact with the important details of a scenario to determine how to act. This philosophy complements the escape room format because it emphasizes discovery and incorporates ideas from many sources of ethics.

2.3.3. Body of Literature

Finally, we draw on the literature of torture ethics for content and situational details. We feature philosophers such as Henry Shue, who argues torture is never permissible (1978); Uwe Steinhoff, who argues torture is sometimes permissible (2015); and Carl Klockars, who writes about members of law enforcement can lose their sense of morality (1980). As designers, we expect that our target audience will not be familiar with these philosophers, providing an element of surprise and challenge for skilled players.

3. ESCAPE ROOM COMPONENTS

Our game has five components that liken it to a traditional escape room: a game master, a room, a lock, a way to escape, and puzzles.

In September 2019, Merriam-Webster added "escape room" to the dictionary, defined as: "a game in which participants confined to a room or other enclosed setting (such as a prison cell) are given a set amount of time to find a way to escape (as by discovering hidden clues and solving a series of riddles or puzzles)" ("We Added New Words to the Dictionary in September 2019", 2019"). Our game has all of these components,

as well as a "game master," a staff role that is common in commercial escape rooms.

3.1. The Game Master

Traditional escape rooms often have a "game master," usually an escape room employee, who performs the following tasks:

1. Room setup
2. Timekeeping
3. Answering clarifying questions
4. Providing hints

In the ethics escape room, our game masters also handle:

1. Assessing the players' performance according to various ethical standards.
2. Discussing the players' reflections after the game and their ethical evaluations.

To avoid overwhelming the game master, future versions of the ethics escape room could have a second role, the arbiter, who evaluates the players and leads the discussion.

3.2. The Room

The room provides the environment and boundaries: it tells players who they are, why they are trapped, and why they want to escape. When the players enter the room, the game master explains that they are staff in the Moral Defense Oversight Service (MDOS), a (fictional) organization in the Department of Homeland Security (DHS) that provides ethical guidance on matters of national security. Players read the mission briefing and learn that DHS has apprehended a man named Anthony Haven who claimed responsibility on social media for an impending bombing.
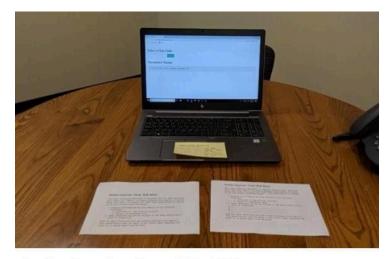
Image 1. Players searching the room for clues. Left: in the second prototype of the game, a player tries to check behind a painting on the wall. Right: in the third prototype, a player looks inside the couch.

3.3. The Lock

The lock represents the players' objective: it is the only thing keeping them in the room. Rather than a traditional physical lock, we use a "moral lock."

Players are given two proposed interrogation authorization forms: one permits "enhanced interrogation" and the other allows a psychologist to question the subject. The players have 60 minutes to decide which authorization form should be used and place it in a submission box.

Authorization Form DLN-0414

This order authorizes Department of Homeland Security (DHS) contractor David Snell (the "Agent") to engage in a one-hour session with Anthony Haven (the "Actor") on August 6th, 2019. This authorization is subject to the restrictions enumerated below.

1. Enhanced interrogation is authorized, but may only consist of the following techniques:
   a. Contortion in "the jetliner position"
   b. Exposure to deafening noise
2. These techniques may not be utilized if the Actor enters into a state of cooperation.

Only the Agent authorized here may conduct enhanced interrogation. This authorization does not apply to any future cases regarding the Actor or to anyone apart from the Actor.

Authorization Form DLN-0614

This order authorizes Department of Homeland Security (DHS) psychologist Shannon Harp (the "Agent") to engage in a one-hour session with Anthony Haven (the "Actor") on August 6th, 2019. This authorization is subject to the restrictions enumerated below.

1. The Agent may not use any measures of physical force.
2. The Agent may not lie to the Actor.

Only the Agent authorized here may conduct the session. This authorization does not apply to any future cases regarding the Actor or to anyone apart from the Actor.

Image 2. Photo from the second prototype. Players enter the room to find the two authorization forms and a note that shows them how to access the situation briefing. This first, simple note teaches them how to use the web application to find digital clues.

Escape rooms need not involve a literal "escape." According to the database compiled by escaperoomplayer.com, commercial escape rooms might challenge players to "investigate a crime," "prevent a crime," "find a cure," or even "find letters with numbers and put them in order"("Escape Room Themes", n.d.). In our escape room, choosing an authorization form is akin to finding a cure or preventing a crime.

Since there are no actual human lives at stake, for the moral lock to be effective, players must feel that the decision is of moral consequence, otherwise, they would simply choose a form and exit. As moral particularism suggests, the moral lock should also be resistant to simple maxims such as "it is okay to harm one person in order to save many." The ethical choice should depend on the specifics of the situation.

Through completing certain puzzles, players can also unlock a custom authorization form that allows them to specify all of the constraints on how the department should treat the subject. This is meant to provide advanced players with greater room for expression and to confront the fallacy of false choice. By design, the emphasis on two forms implies to players that one choice is good and one choice is bad, which can limit their ability to perceive alternatives. When players unlock the custom form, it becomes clear that nothing in the game restricts their recommendations, reminding them that they have full agency over how to crack the lock.

If the players choose an authorization form with more than 20 minutes left in the game, this triggers a storyline where they find out the results of their chosen action and must decide how to proceed by answering an open-ended follow-up prompt. This is analogous to how several escape rooms have a surprise room: players think they have successfully cleared the challenge, but there is actually another vestibule they have to escape.

3.4. The Escape

In traditional escape rooms, participants may share a photo with the time it took them to complete the game. In the ethics escape room, time to complete is a perverse metric: the faster players finish, the less likely they had a rigorous discussion.

We decided there should be no "correct" authorization form because providing a moral "answer" risks diminishing the merits of the other perspectives. Instead, we evaluate the players' escape plan based on their actions and justifications, not just the form they select.

We designed three evaluation instruments, each based on a primary branch of traditional western ethics: deontology, virtues, or utilitarianism. To reduce the length of the introduction and to see how players would behave, we chose not to reveal these instruments until the end of the game, when the game masters led a 30 minute reflective discussion about the players' results.

3.4.1. Deontology

Deontology is the study of the rights and responsibilities people must adhere to in a good society. We created a deontological definition of good and bad for this scenario by writing four basic rules.

Players can exercise two rights:

- Self-Defense: May take actions to protect civilians from ongoing attacks.

- Justice: May take actions to fix an unfair distribution of pain or pleasure.

Players must follow two responsibilities:

- Harm Prevention: May not take actions that harm innocent people.

- Fidelity: May not take actions that violate promises or deceive people.

3.4.2. Virtues

Inspired by Aristotle's *Nicomachean Ethics* (Crisp, 2014), we selected three "mean" virtues that we felt were important to this scenario: deliberation, passion, and justice. An excess of deliberation is over-cautiousness while a deficiency is recklessness; an excess of passion is zeal while a deficiency is apathy; and an excess of justice is lenience while a deficiency is maleficence. Game masters and players are asked to rate the players' performance on these three scales.
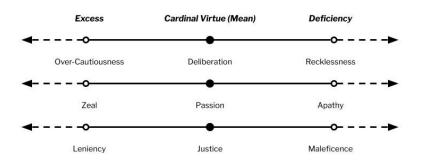

Image 3. Scales for the virtue ethics evaluation instrument.

3.4.3 Utilitarianism

For a utilitarian instrument, our plan was to anticipate common player actions and prepare a single score for how good or bad each action is. However, due to a lack of confidence in our predicted actions and an already-packed discussion agenda, we abandoned this instrument.

3.5. The Puzzles

The puzzles in an escape room challenge players, rewarding them with information that helps break the lock. Table 1 lists the five puzzles that deliver ethics content, from the third prototype of our game. We use puzzles to reward players for practicing the behaviors in the Markkula Center framework and to provide them with vocabulary and context they can use to make a choice between the authorization forms. Table 2 relates the puzzles and clues to opposing viewpoints on three questions that we felt captured the crux of the ethical debate around torture.

Name	Ethics Content	Puzzle Challenge
Shue/Historical Records	Is it ethical to torture someone who might be innocent? Shue proposes three categories of torture victims: ready collaborators, innocent bystanders, and dedicated enemies (Shue, 1978).	Find an excerpt from Shue's paper and use it to classify each victim from a set of five historical records into one of the three categories.
Bentham	How dangerous must a threat be to justify torture? Bentham offers four criteria for quantifying pain: intensity, duration, certainty, and nearness (Bentham, 1789).	Find four scales, one for each of the criteria. Listen to a risk assessment from a national security advisor and rate it on each scale.
Steinhoff	Steinhoff argues that torture may be permissible as self-defense out of necessity if the threat is imminent and proposes that the act of self-defense should be proportional to the threat, but does not have to be the mildest means (Steinhoff, 2015).	Match each of the four terms (necessity, imminence, proportionality, mildest means) to the example that best illustrates their definition.
Intervening Action	A principle which argues that we are responsible for our actions, even if someone else caused us to act that way, and we are not responsible for the consequences of actions taken by others because of us (Harris, 2008).	Read an email explaining the principle of intervening action. Then, unlock a list of three statements about national security and identify the ones that are supported by the principle.
Dirty Harry	Dirty Harry scenarios arise when: (1) a normally lawful person believes the only way for them to carry out (2) their legal mandate is to (3) take an action that is not morally permissible (Klockars, 1980).	Write a story about how the chosen authorization could lead to a Dirty Harry scenario. If the game master identifies all three requirements in the story, players unlock the custom authorization form (Form C).

Table 1. The five puzzles from the third prototype.

Is enhanced interrogation justified if...	Yes	No
...the threat may be fake?	Bentham	Shue/Historical Records
...it may not work?	Steinhoff	Pamphlet
...someone else performs it?	Intervening Action	Dirty Harry

Table 2. Puzzles mapped to moral questions in the ticking time bomb thought experiment.

Image 4 shows paths players can take to find the puzzles and

clues. The most notable clue is a pamphlet that provides background information on the fictional organization that the apprehended subject belongs to, Americans For Fairness (AFF). The group's beliefs include concern about increasing divides in the country, an imperative for children to defend their parents, and broad justifications for vague calls to respond to failures of government.
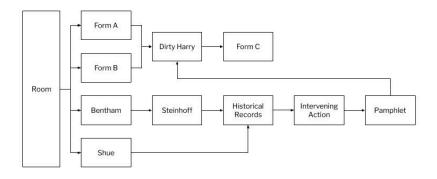

Image 4. Puzzle discovery paths in the third prototype. Form C is the custom form.

4. DEVELOPMENT

4.1. Puzzle Tuning

Brathwaite and Schreiber categorize puzzles as riddles, lateral thinking, spatial reasoning, pattern recognition, logic, exploration, and item use (2009). A weakness of our escape room is that all the puzzles are logic puzzles, where players receive

some information and must derive further information to solve the problem.

Image 5. Three players work together to solve the Historical Records clue.

To avoid overwhelming the players with logic puzzles, we applied several puzzle design principles from *The Art of Game Design: A Book of Lenses*. The Bentham puzzle illustrates puzzle principle #1: "make the goal easily understood" (Schell, 2008). Locks with numerical passcodes are so common in escape rooms that when players see Jeremy Bentham's four numerical scales, they recognize choosing the correct scale values will reveal the passcode.

Bentham's Criteria Scale

Email Dr. Carl Rush at _____ and use this form to rate his assessment of the threat.

Intensity Rating

How intense do they think the consequences will be?

Very tolerable	*Somewhat tolerable*	*Cannot be determined*	*Somewhat painful*	*Very painful*
1	2	3	4	5

Duration Rating

How long do they think the consequences will last?

Very short time	*Somewhat short time*	*Cannot be determined*	*Somewhat long time*	*Very long time*
1	2	3	4	5

Certainty Rating

How sure are they that threat will occur?

Very unlikely	*Somewhat unlikely*	*Cannot be determined*	*Somewhat likely*	*Very likely*
1	2	3	4	5

Nearness Rating

How close do they think the threat is occurring?

Very far	*Somewhat far*	*Cannot be determined*	*Somewhat close*	*Very close*
1	2	3	4	5

Image 6. Scales from the Bentham puzzle. Players find this document hidden in a cabinet. When they contact "Dr. Carl Rush," he gives his threat assessment. If the players correctly estimate the assessment according to Bentham's criteria, the numbers will form a four-digit code to unlock the Steinhoff puzzle.

Almost all escape rooms utilize puzzle principle #6: "parallelism

lets the player rest" (Schell, 2008). When a player gets frustrated or bored with one puzzle, they can simply transition to another one. Teams of players can work on multiple puzzles simultaneously and swap to get unstuck. Image 4 shows that our escape room offers multiple parallel paths.

The game master delivers on puzzle principle #8: "hints extend interest" (Schell, 2008). The game master can step in when players are stuck to offer them one of their three hints. In the Dirty Harry puzzle players must show their story to the game master for approval, who provides hints by poking holes in their story to show which of the three characteristics of a dirty Harry scenario it does not meet. This puzzle teaches the players to imagine negative consequences of their chosen form and fight confirmation bias.

4.2. Materials

Our prototypes were mixed-medium, including both physical and digital elements and puzzles.

For physical setting, we used a room in an office space and a room in a residential building. Many escape rooms have an immersive set design. Since our game takes place in a fictional government agency, we did not feel pressure to invest heavily in room decor. We instead spent that effort on puzzle design, totaling approximately 60 hours as a team, not including time spent in game sessions. Physical clues were either written on sticky notes or printed on letter-size paper.

We developed a web application where players could unlock digital clues. This allowed us to quickly iterate on the content or even fix mistakes in the middle of a gameplay session. Digital locks instead of physical locks allowed for quick configuration when updating puzzle answers. The application was developed and hosted on Glitch and the materials are available under an open-source license on GitHub, both free online platforms.

Enter a Clue Code

| S-HYR | Submit |

Document Viewer

Document look funny? Enter the password to decode it.

| | Decode |

42 32 75 106 122 116 114 115 112 99 100 109 39
82 104 100 118 114 103 116 39 111 105 33 80 110
119 102 121 114 114 104 104 116 108 112 117 32
86 118 122 112 104 100 123 115 13 11 69 32 90
102 39 97 117 102 39 109 108 116 122 105 113
104 39 108 100 99 108 108 118 33 109 111 117 33
123 104 104 116 108 32 43 83 74 44 35 74 73 44
35 69 76 41 47 33 106 97 113 33 122 111 112 102
118 110 104 33 106 104 104 100 114 32 122 105
112 99 107 33 118 110 104 116 39 97 117 102 39
119 107 106 106 104 66 33 80 102 35 122 118 117
35 100 118 117 113 117 39 117 115 33 111 111 122
33 116 97 113 122 39 101 123 98 116 112 111 102
122 32 119 105 108 114 104 33 104 114 104 33
118 102 35 117 111 101 35 117 111 114 104 102 39
116 124 113 108 115 35 112 109 32 118 118 122 112

Enter a Clue Code

| S-HYR | Submit |

Document Viewer

Document look funny? Enter the password to decode it.

| 7031 | Decode |

Historical Records of Interrogation Suspects

> We are missing labels for these (RC, IB, DE), can someone check which ones are which? If you count up how many examples there are of the three types of suspects, it should form the password for an email from Shannon about intervening action.

Subject (age 17) immediately agreed to cooperate before enhanced interrogation began, providing the names of individuals planning attacks targeted at a local concert. Further investigation

Image 7. Accessing a clue in the web application and unlocking it with a passcode.

4.3. Prototype Evolution

We iterated through three prototypes of the ethics escape room. The first version took one month to design. Over two weeks, we made two new versions and played six games in total.

The first prototype served as a proof of concept to validate the chosen scenario, demonstrate examples of ethical puzzles, and determine if the game would hold players' attention for an hour. This preliminary version included only the Bentham, Shue, and Historical Records clues and players could only choose between two forms. Both forms authorized enhanced interrogation, but the second form included blanks so that players could specify other constraints. Only one group played this version; the three players from this game joined the design team and served as

game masters for future games. Their familiarity with the game allowed us to run more sessions in a short time span.

The second prototype expanded the story of the suspected bomber, Anthony Haven. We kept the default enhanced interrogation form and juxtaposing a form that permits a staff psychologist to question Haven. The psychologist cannot lie to the subject, a condition inspired by the first group of players. We added the Pamphlet clue, as well as a clue with a real article from the American Psychological Association about motivations for terror and de-radicalization techniques. One team played this version. In the final five minutes, the players learn from the game master that Haven is not the bomber, his father is. Haven refuses to reveal information about the attack unless authorities guarantee that his father will not be tortured. When asked how they would proceed, players discussed Shue's ideas on how to treat accessories to an attack and referenced the Pamphlet clue to persuade Haven to share details.

The third prototype introduces more variety to the ethical perspectives, adding the Steinhoff, Intervening Action, and Dirty Harry clues, while revising the previous clues to account for previous players' misconceptions. The Dirty Harry clue allowed players to unlock a third, custom form. Previous players complained of too much time spent reading documents, so we removed the article about de-radicalization and changed the Bentham clue so that players call a phone number and listen to the report rather than reading it. We added an introduction script for the game masters and the two ethical evaluation instruments.

5. GAMEPLAY

We recruited 18 players through an online form with a teaser about the game. Our players were primarily undergraduate and graduate students studying technology and government. We are independent designers and chose not to find an institutional

review board for this game, instead we prefaced the signup form with a content warning about mentions of torture. We avoided any graphic descriptions or images of torture in the game. Players consented to their photos being used for this paper and their comments are presented anonymously.

5.1. Player Success and Struggles

Players generally did well with puzzle mechanics: they found tangible objects hidden in the rooms, properly navigated the digital interface, figured out what information was relevant for the puzzle they were trying to solve, and came to working solutions without needing clues from the game master. Players relied primarily on successes in comprehending the relevant ethical principles and correctly applying them to their specific contexts. In very few cases did they resort to solving puzzles via brute force. Additionally, players engaged in thoughtful discussions, most lasting the entire hour. No group tried to hand in a form early in order to leave; overall, players took our moral lock seriously, motivating their play.

When players struggled, it was most visible in their discussion. When players got stuck with puzzles it was usually because their discussion had moved away from the text that the relevant ethical principle was based on and towards personal opinions or morals. Players also noted this being the case when it came to choosing a form: some players had already developed moral instinct around the question of the acceptability of torture and found it difficult to separate these instincts from the information presented in the game. One player even noted, "In this scenario, I would never choose torture because I know it's not right." Game masters were there to issue reminders when the discussion was particularly unproductive, but players ultimately had to decide what kinds of discussions were going to help them choose an authorization form and how they were going to use the information presented to them. As one player said, "It was

frustrating when we realized that we could not solve the escape room by finding more clues, we had to write or say stuff instead."

5.2. Observations

Groups spent between 40-50 minutes completing puzzles and used the remaining time to discuss the authorization forms. Table 3 summarizes their final choices. There was variety in final form submissions, with no dominant form that everyone could agree on. Comparing the groups that played Prototype 3, we observed different outcomes even though all groups were presented with the same puzzles and information, confirming that we chose a sufficiently divisive scenario.

Group	Prototype	Chosen Authorization Form
1	1	Enhanced interrogation, custom constraint: interrogator may not lie
2	2	Psychological discussion
3	3	Custom form, constraints: interrogator may not lie or physically harm the subject
4	3	Enhanced interrogation
5	3	Psychological discussion
6	3	Enhanced interrogation

Table 3. Authorization forms chosen by each group.

Even if their conclusions differed, games shared common discussion elements. Players discussed the acceptability of physical harm and whether there were circumstances under which it could be justified. They also discussed psychological tactics, debating whether they would be sufficient to solicit information from the subject. Players questioned the effectiveness of interrogation, and whether it could solicit factually correct and actionable information from the subject. Players also contemplated the importance of the interrogator's identity, whether a different set of actions was acceptable if they

themselves did not need to carry out the actions, and what kind of precedent their decision would set for the future of their fictional society. Many of these ideas appear in the torture ethics literature, meaning that game masters could critically engage with players during post-game reflection.

Two groups rebuked the original two authorization forms, choosing to write their own custom forms. This demonstrated our goal of designing a moral lock that could admit multiple different solutions. Both custom groups were suspicious of the psychological interrogation form because they worried the interrogator might lie to the subject, a tactic they felt was immoral. Group 3 also added the constraint that the interrogator may not physically harm the subject.

Iterative prototyping makes outcome comparisons across games difficult. In Prototype 1, both authorization forms involved enhanced interrogation, so even players who found this morally dubious had to set aside their qualms in order to escape. In contrast, players of Prototype 3 games had a more diverse set of options, including the ability to completely specify their own terms. Still, groups chose forms authorizing enhanced interrogation; we suspect that the inclusion of the Steinhoff clue made the enhanced interrogation form a more morally viable choice. Additionally, we saw a shared moral aversion to lying across games: four out of six groups selected authorization forms explicitly prohibiting lying to the subject.

5.3. Player Engagement

To determine how players engaged with ethics escape room, we situate our game with LeBlanc's taxonomy, which proposes eight modes of "fun" (Hunicke et al., 2004):

1. Sensation: Game as sense-pleasure
2. Fantasy: Game as make-believe
3. Narrative: Game as drama

4. Challenge: Game as obstacle course
5. Fellowship: Game as social framework
6. Discovery: Game as uncharted territory
7. Expression: Game as self-discovery
8. Submission: Game as pastime

Traditional escape rooms invest heavily in fantasy, narrative, challenge, and discovery. We examined feedback from our players to place them on the taxonomy.

Positive reviews suggest that players were engaged through discovery and narrative:

- "It was lots of fun. I liked cracking the codes, the readings were interesting, and the clues turned out to be good reading comprehension checks."
- "The scenario was good. There was a sense of urgency and it was like we were forming the story as we found the clues."

As game designers, we feel pleased with the challenge level of the game. Every team finished all the puzzles without using up all three hints. We observed the players debate the puzzles and celebrate when their solution unlocked the next clue.

Criticisms focused on discovery overload, struggles with expression, and lack of submission:

- "Decisions didn't feel immediate. There were lots of papers and examples of philosophies, but there were no consequences. It was all to get to the next document or letter, which doesn't feel real."
- "Even though we learned new ethics topics, when the discussion started, we were still inputting our own morals and values into our opinions and justifications."

We view the unclear ending as the biggest detractor to player

engagement because players who enjoy challenge want a clear signal of victory. In the post-game discussion, players wanted to know if they had picked the "correct" authorization form, but each authorization form has positive and negative moral consequences. The ethical evaluations in the post-game discussion partially satisfied players' craving for feedback, but also made some players defensive. If we can limit the negative reactions, the evaluations could offer another avenue for enjoyment through expression.

One way to accomplish this could be to redesign the ethical instruments in a socially-shareable format, akin to how players share a photo after completing a traditional escape room. The results could highlight which of the three branches best represents each player's actions in the game, like an elaborate personality quiz. This avoids a deficit-based view of ethics and could encourage players to share and compare their results.

6. DISCUSSION

"I know about ethics, I watch The Good Place!," remarked one player during gameplay. The Good Place is an American comedy TV series about characters in the afterlife who must come to terms with the morality of their lives on Earth. The show's success demonstrates that ethics can succeed in popular culture and inspired us to consider innovative mediums to make ethics more accessible and tangible.

Image 8. Spoiler-free scenes from season two of The Good Place. Chidi, who was a moral philosophy professor on Earth, uses toys to teach the characters about the trolley problem. Michael, an afterlife employee, raises the stakes by simulating a realistic trolley and imperiled workers. Photo screenshots taken from: https://www.nbc.com/the-good-place.

Our game provides a proof-of-concept for how escape rooms can be an engaging vehicle for exploring ethics outside of traditional settings. As designers, we are proud of creating enjoyable puzzles from content that is normally considered frustrating and implementing highly-mutable escape room materials that could be refined over multiple prototypes. To have 18 players explore the scenario and participate in six thought-provoking post-game discussions is the best reward we could hope for.

When COVID-19 hit the United States in Spring 2020, physical escape rooms became impractical and people began looking for opportunities for remote-friendly social activities. As a result, we decided to use the idea of a moral lock to design a completely virtual sequel to this game.

In August 2020, we launched *Panopticon*, a virtual escape room rooted in questions about the ethics of surveillance technology. The design process for this game relied heavily on our experiences designing and playing the original, with a focus on

creating a more divisive ethical dilemma, diversifying our puzzles beyond reading documents, and designing an ending that gave more closure to players and spurred better discussion on this timely topic. Additionally, we observed that even though we released the materials online, writing instructions for others to recreate our physical escape room was challenging. The entirely virtual format of *Panopticon* makes adoption much easier.

As of November 6, 2020, almost 350 people have played the game. This includes three subgroups of players worth noting: ethics educators, young technologists working in government, and corporate teams (mostly in the technology sector). In these contexts, *Panopticon* provided an opportunity for these players to explore ethical quandaries they might experience over the course of their studies or careers. For ethics educators, *Panopticon* served as a classroom activity to introduce technology ethics into computer science classrooms. For the young technologists and corporate teams, *Panopticon* was used as a community building activity that opened dialogue at the intersection of technology and society in a low-stakes environment.

Escape rooms do not have to include a literal escape; players will engage and enjoy even with a moral lock. We hope that escape rooms will make ethical questions more accessible and enjoyable to gamers, professors, students, and other philosophy enthusiasts.

WORKS CITED

A Framework for Ethical Decision Making. (2009). Markkula Centre for Applied Ethics.

Adams, V., Burger, S., Crawford, K., & Setter, R. (2018). Can you escape? Creating an escape room to facilitate active learning. *Journal for Nurses in Professional Development*, *34*(2), E1–E5.

Chen, W. (2016, September 22). Top 11 Puzzle Ideas for Escape

Rooms. *Escape Room Tips*. https://escaperoomtips.com/design/escape-room-puzzle-ideas

Crisp, R. (2014). *Aristotle: Nicomachean Ethics*. Cambridge University Press.

Dancy, J. (2004). *Ethics without principles*. Oxford University Press on Demand.

Dancy, J. (2017). Moral Particularism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2017). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2017/entries/moral-particularism/

Escape Room Themes. (n.d.). Retrieved November 1, 2020, from https://escaperoomplayer.com/Escape_Room_Themes.html

Eukel, H. N., Frenzel, J. E., & Cernusca, D. (2017). Educational gaming for pharmacy students–design and evaluation of a diabetes-themed escape room. *American Journal of Pharmaceutical Education*, *81*(7).

Hare, R. M. (1963). *Freedom and reason* (Vol. 92). Oxford Paperbacks.

Hunicke, R., LeBlanc, M., & Zubek, R. (2004). MDA: A formal approach to game design and game research. *Proceedings of the AAAI Workshop on Challenges in Game AI*, *4*, 1722.

Nicholson, S. (2018). Creating engaging escape rooms for the classroom. *Childhood Education*, *94*(1), 44–49.

Schell, J. (2008). *The art of game design: A book of lenses*.

We Added New Words to the Dictionary in September 2019. (2019). Merriam-Webster.Com. Retrieved October 27, 2020, from https://www.merriam-webster.com/dictionary/escape room