Escape Room Experience for Team Building Through Gamification Using Deep Reinforcement Learning

Georgios Liapis, Aristotelis Lazaridis, Ioannis Vlahavas
Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece
gliapisa@gapps.auth.gr
arislaza@csd.auth.gr
vlahavas@csd.auth.gr

Abstract: Gamification, which is considered to be an efficient practice for learning through play, can be significantly expanded by Artificial Intelligence methods, and particularly Machine Learning. Nowadays, different industries employ a variety of applications based on gamification to create coherent and effective teams, e.g., by assigning roles based on the knowledge, understanding, and relationships between members. In this paper, we explore an online Escape Room experience that incorporates a variety of Raven-inspired intelligence tests and team-members communication, combined with Machine Learning methods. More particularly, we implemented state-of-the-art Deep Reinforcement Learning (Deep RL) agents, which are used for emulating human-like behaviour to navigate and interact with the 3D rooms, as well as to solve the tests. The RL agents simulate behavioural elements based on OCEAN personality traits model, such as openness, conscientiousness, and neuroticism, while also generating a big number of gameplay data. Analysis shows that their particular behavioural patterns have a significant effect on their performance, stability and time required to solve tasks. These findings allowed us to produce new performance metrics for a generic escape room model, which can categorize human play styles according to the OCEAN Five personality trait model. This approach effectively analyses the teams' behaviour concerning both individual and overall performance.

Keywords:
Gamification, serious game, escape room, team building, machine learning, deep reinforcement learning

1. Introduction

Advances in technology have helped games evolve in many ways in order to make them even more enjoyable, as well as improve the players' skills, critical thinking, and even their health and overall well-being. This new area that has been around for years is called gamification, and it is the application of game characteristics to change behaviours in non-game situations (Robson et al 2015). The main concepts of gamification have been used by many instructors and teachers. Things like rewards given for homework, challenges, and leader boards are the main elements of gamified systems (Dicheva et al. 2014). There have been several papers, as presented in the work of Dicheva et al. (2014) about gamification designs that have been used in education, gymnastics, and even in marketing.

One area that uses such methods is real-life Escape Room (ER) games (Pan, 2017). They require players to work together, hence they have become popular in companies as a team-building strategy. It is a widespread practice for companies to rent escape rooms and then have their employees work together to solve the room's mysteries. This activity lets the team members bond, and also gives the company a chance to assess each member's contribution.

There have been several attempts with real-life escape rooms to generate data to measure their effectiveness with respect to team bonding. The main skills that can be tested and measured are communication, teamwork, and leadership, and the only way that can be achieved is by conducting a survey after the room has been completed. There have been many problems with such experiments, as analysed in the work of Fotaris et al. (2019) about the evaluation of the room, the size, and the available resources.

An effective way to solve such issues is the collection, the analysis and procession of the data for each individual player. In order to translate all this information regarding real-life ERs into virtual ERs, a profile for each player must be created to keep track of different statistics (e.g., what the player said or solved, how many mistakes were made, etc.), and then create a summary for each player. Therefore, a detailed statistical report for a player can be created, which also includes a record regarding all the previous escape room games he participated in and understand more about his development. It is obvious that to generate accurate profiles, a large number of player data are needed.

In this paper, we present an online Escape Room game that implements intelligence and personality trait tests, alongside state-of-the-art Deep Reinforcement Agents that can emulate template behaviours and profiles by solving our game multiple times according to a variety of scenarios and behaviour characteristics based on the OCEAN five personality model. The ultimate goal is to compare and analyse the relationship between the results of a human player gameplay data with the agent's template behaviour data, and extract insights regarding his intelligence score and personality traits. The training results from the agents trying to emulate the OCEAN five characteristics are promising and depict the difference of the gameplay styles.

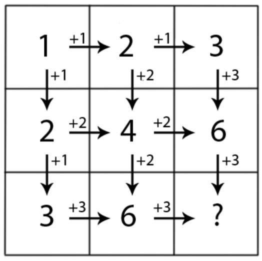


Figure 1: Raven IQ test example, the numbers increase by 1, 2, or 3, in each row/column, with the solution being the number 9.

2. Related Work

Extended research has been conducted about the importance of Escape Rooms as means for better team bonding and as a tool for gamification in a variety of scientific areas. Education is one of the most important and common areas for applying such techniques, as being demonstrated by Clarke et al., (2017). The authors created a real-life pilot room so that the players could develop their soft skills by solving puzzles as well as introduce this game as an educational tool. The room had a disarm-the-bomb theme, in which the players had 15 minutes to find clues and solve puzzles and riddles to stop a bomb from exploding. After the game ended, each team was asked about their opinion on the educational value of the room through a survey. The main difference with our work is that we created a simulated ER environment (i.e. video game) instead of building a real-life one, which allows us to compute and report several assessment metrics, consequently being more inclusive regarding each player's play style and overall performance. Furthermore, we implemented Deep Reinforcement Learning agents for simulating various behaviours within the room, which eliminated the need for a large number of players.

There have also been online escape rooms that have been used as a part of education. In games like the one created by Vergne et al. (2020), other means were used to create the environment, and specifically online Google Forms (Google 2017) and Zoom video conference (Zoom Video Communications 2012) applications. The game was used as a remote method using gamification and the players came across problems that they had to solve to progress each section (presented as rooms). The theme was a factory, and each room had a different kind of problems in various topics of the subject. The team communicated through the video and had to solve the room within a time limit of 20 minutes. Although this online game had a more direct approach in communication, there was not the ability to keep track of specific statistics (in contrast to Mind Escape), and even though the environment was accompanied by multimedia, there was no 3D exploration and interaction. Moreover, the implementation of Deep Reinforcement Learning agents generated a large amount of different gameplay data and a variety of results, so that a more complete evaluation of the game could be created.

3. Background

In the context of this work, a serious 3D Escape Room game was designed and developed using the Unity game engine (Unity Technologies, 2005), which consists of intelligence tests and puzzles instead of traditionally used riddles. These tests are based on the Raven Intelligence tests (Raven, 1936) and are scattered throughout the room and the player's goal is not only to discover but also to solve them.

Using intelligence tests is a widespread practice for evaluating one's IQ score. One of the most popular intelligence tests is the Raven intelligence test, which consists of a grid of 9 shapes that follow a pattern. Figure 1 depicts an arithmetic Raven IQ test, in which each row has an add operation pattern, and the number to the solution to this IQ test is 9.
Moreover, our game generates some metrics so that the results can be related to the OCEAN five Personality traits model (Jang et al, 1996), which is one of the most established and recognized approaches to describe and measure individual differences in personality (Power et al, 2015). The acronym OCEAN stands for *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness*, and *Neuroticism*.

The personality of a person generally consists of his temperamental, emotional, and mental traits. However, considerable controversy exists in personality research about the exact number of personality traits. There is a lot of research (Verywell Mind, 2020), e.g. by D.W. Fiske (1949), Norman (1967), Smith (1967), McCrae and Cost (1987), and later expanded by J.M. Digman (1990) and Goldberg (1993), that analyse these traits and consist of the basis in which the OCEAN five personality model was built on.

4. Methodology

In this section, we present the gameplay mechanics of Mind Escape, we define the metrics that are used for measuring player performance and analyse how the Deep Reinforcement Agents were implemented.

4.1 Escape Room Game

Mind Escape consists of several rooms that were designed with a space layout and decorated with furniture and assets, depending on the theme (e.g. as an office). The rooms are also grouped by their difficulty in five categories, ranging from easy to hard depending on the complexity of each room, the arrangement of the furniture, and the difficulty of the tests.

Inside the rooms, the player has to discover some test panels (usually four). By finding and interacting with each panel, a new screen appears which hosts the test that the player has to solve in order to continue his progress throughout the room. For a player to finish a room, all four puzzles have to be found and solved. To do so, the player has to search and find the first three visible puzzle panels and solve the respective test. In some difficulty modes, there are two visible panels and one hidden panel the player has to uncover by finding a button inside the room or entering some hidden numbers in a keypad that is located in different places inside the room. After the three puzzles have been solved, the fourth and last one (and a bit more difficult than the other) will be revealed for the player to solve and finish that room.

Since the aim of this project is the development of a serious game, some metrics were defined, collected, and analysed. One of the metrics is a walk heat-map that holds information about the player's movement inside the room and his overall game style, while the others are related to his performance regarding the intelligence test. Additionally, some indicators about the OCEAN five personality traits are included as well. More specifically, the player's final IQ score is related to the time spent on each test, the number of errors, and the use or not of a hint button. Regarding the OCEAN traits, some metrics were implemented based on the bibliography, related to the way the player explored the room, his commitment, and effectiveness in finding specific targets as well as his communication with the rest of the team.

4.2 Deep Reinforcement learning Methodology

In order to eliminate the need for gameplay data from human players, we implemented Deep Reinforcement Learning agents that acted based on particular behavioural templates, generating the necessary data so that the behavioural templates could be generated. These agents use the Proximal Policy Optimization (PPO) algorithm (Schulman et al 2017) to discover the best action they can choose, at each step.

The second primary interest besides the way the agent learned to understand, solve, and analyse the intelligence test patterns, is the analysis of the gameplay style based on characteristics of the OCEAN model, in terms of the movement of an agent in a room environment.
Table 1 presents a decomposition of real-life player behaviours into simpler ones, that can or cannot be replicated by an agent. It must be noted that not all OCEAN five personality traits can be tested with one agent or with a single-player game (e.g. extraversion, agreeableness) since there is a need for interaction between a group of agents and/or players so that the necessary information can be extracted to generate metrics. For this

reason, the tests of the agent we implemented approximately measure openness, conscientiousness as well as neuroticism behaviours.

Table 1: OCEAN five and movement ways

Category	Agent Movement
Openness	Moves a lot in space, exploring it in as many ways as possible.
Conscientiousness	Moves quite methodically and purposefully on each possible element before moving on to the next.
Extraversion	Needs to socialize with other potential agents in the field.
Agreeableness	Moves through space according to the orders and rules of his leader.
Neuroticism	Moves spasmodically and hesitantly through space, making undecided choices.

Based on these behavioural concepts, we simulated the OCEAN five personality traits to agent movements, as was mentioned before. In Table 2, one can find the approach we used to achieve this.

Table 2: OCEAN five and agent behaviour

Category	Agent Behaviour
Openness	Moves a lot in space, exploring it in as many ways as possible -> simple default agent.
Conscientiousness	Moves quite methodically and purposefully on each possible element before moving on to the next -> penalty on each step.
Neuroticism	Moves spasmodically and hesitantly through space, making undecided choices -> high epsilon parameter or random move (2 variations)

The agent during the training process uses the PPO from the ML agents package in Unity (Unity Technologies, 2017). To emulate the neuroticism, a training parameter of the agent related to the stability of the policy updates was tuned appropriately. More specifically, the epsilon hyperparameter was modified, which is the acceptable threshold of divergence between two consequent policies, and affects the speed of the training process, subsequently making the agent's behaviour more unstable and maladjusted. So, by increasing it, the agent is able to act in different manners, and therefore the agent's movement simulates neuroticism.

After implementing initially, a simple working model of the game for human players, we proceeded into creating a simplified environment for the agent that implements goals, rewards, and penalties, and applied Deep Reinforcement Agents available in the Unity engine. This new, simplified environment was used to help the agent learn to navigate in the 3D space and find the goals (puzzle boards) inside the 3D room.

For this task, we followed a specific methodology for the training of the agent. We created variations of the agent's behaviour depending on the category of the OCEAN five model we wanted to simulate, by properly modifying its training hyperparameters. For the agent in the 3D room, we used the PPO algorithm.

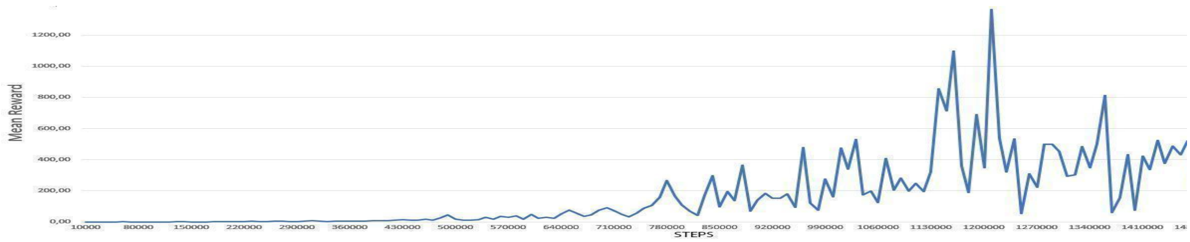
Regarding the training of the agent in solving the puzzles, the agent would analyse the 8 numbers of the raven test and the 6 possible answers to find the correct one. By choosing a wrong answer he would be penalized, in order to become more efficient. Two different experimentation techniques were implemented. For the sake of variety, and since the Unity ML-Agents package includes only one Deep RL algorithm (PPO), we used the A2C algorithm from the OpenAI Gym platform (OpenAI, 2016).

5. Experimental Results

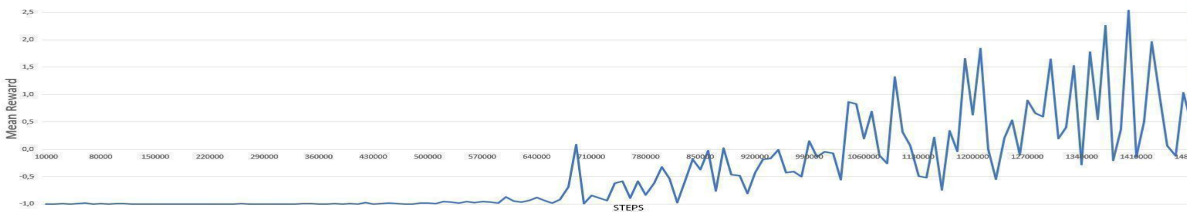
In the following sections, the results of the Deep Reinforcement Learning agents are presented. We demonstrate and analyse the results of the agents that learned to navigate inside the Escape Room, and how their performance depends on the OCEAN five trait they were emulating. Then we present the performance of the agents that were trained to solve the intelligence tests.

5.1 Agent Results - Escape Room Navigating

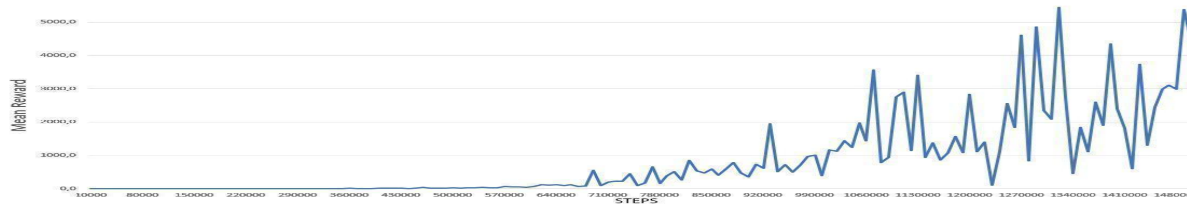
The training process consisted of 1500 thousand timesteps and the final results (i.e. rewards at each timestep) were obtained by averaging the results over 3 different seeds. Figure 2 depicts the results of the training process of the agents. As it can be noted from Figure 2a, the agent started to learn after the 500 thousand timesteps, where it marked his highest score till that point, and onward of that the rewards are increasing, reaching their highest score at around 1000 thousand steps. The results from Figure 2b show that the agent made a lot of mistakes till the 750 thousand steps, where it scored the first positive reward, yet the agent needed around 300 thousand more steps till it had more stable and positive results. Lastly, from Figure 2c, it can be seen that the agent learned a bit later than the first agent (around 700 thousand steps) yet it scores a bigger high reward at around the same step as the first one. Although the last agent was not so stable as the first one, it had better overall rewards than the second.



(a)



(b)



(c)

Figure 2: Performance of Deep RL agent emulating Openness (a), Conscientiousness (b), and Neuroticism (c). In the X-axis we can observe the timesteps of the agent while in the Y-axis the mean reward.

Furthermore, Table 3 presents a summary of the results generated by the Deep RL agents. In general, the agent with openness was stable and somehow efficient. Regarding the agent with conscientiousness, he learned not to make unnecessary movements to reach his goal, but it is reasonable that it took some time, especially if we observe that the first consecutive positive rewards came after the first 1000 thousand steps. Finally, the agents that simulated neuroticism are very efficient, though not so stable, something we expected to have as a result.

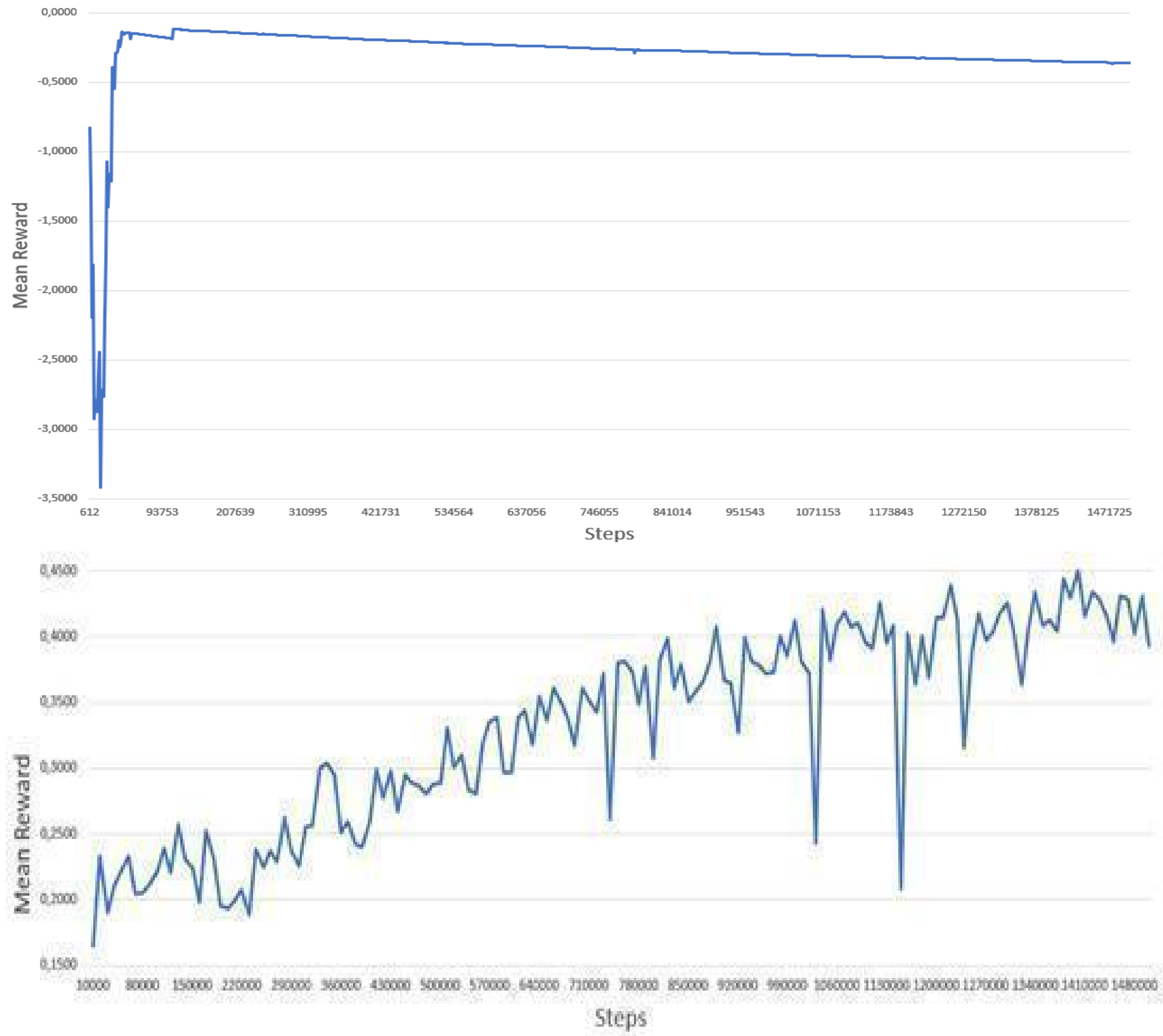
Table 3: Collective Results of agents emulating OCEAN five characteristics.

Trait	Result analysis

	First peak (step)	High/Low	Disturbances (biggest difference)	Negative rewards
Openness	500 thousand	1300/0	Around 900 points (medium difference)	No
Conscientiousness	750 thousand	2600/-1	Around 2500 points (big difference)	Yes
Neuroticism	690 thousand	5500/0	Around 4000 points (very big difference)	No

5.2 Agent Results - Intelligence Tests Solving

(a)



(b)

Figure 3: Performance of the agent-solving intelligence tests in Gym environment (a), Unity environment (b). The X-axis presents the timesteps of the agent while the Y-axis the mean reward.

As it can be seen from Figure 3a, the gym environment agent, starting by making a lot of mistakes, but very quickly (in less than 100 thousand steps), the agent learned to find the correct answer, but not with the first attempt. It should be highlighted that as the number of steps progressed, even though he learned quickly and found the correct answer, with not so many changes, he struggled to find a better way to do this, so there is a slight decrease in the efficiency.

On the other hand, by observing Figure 3b, it can be noted that the Deep RL agent, from the Unity ML package, starts by finding the correct answer at the last (sixth) try since it gets penalized with -0.2 on each wrong answer and gets rewarded 1 point on the correct answer. As the training process progressed, the agent learned to solve the puzzles with fewer tries, though its performance was so stable at times.

Table 4: Collective results of agents emulating OCEAN five characteristics.

Environment	Results analysis		
	First peak (step)	High / Low	Mean errors
Unity-ML agents	100 thousand	-0.1/-3.4	1
Gym	850 thousand (4th try to solve)	0.4/0.2	3

It is impractical to compare the results from the two experiments, as it can be observed from Table 4, that the Deep RL agent from the Unity Package needs around 850 thousand steps to learn to find the correct answer in the 4th try, while the gym environment agent needs 100 thousand steps to make the smallest number of mistakes possible. But we can also assume that given a lot more steps in the agent training, much better results could be extracted, although the correct results can lead to an early but positive indication for the efficiency of the agents.

6. Conclusion and Future work

In this paper, we presented Mind Escape, a 3D Escape Room that combines intelligence test solving with OCEAN five personality traits characteristics. The nature of the game allows us to generate metrics for measuring the game style of each player, not only regarding the way he/she navigates inside the room but also how he/she interacts with it and with the puzzles. The use of Deep Reinforcement Learning agents allowed us to generate gameplay data and profiles by emulating characteristics of personality traits, based on the analysis of the OCEAN five model. This allows us to create enough data to be able to understand the different human playstyles depending on the personality model.

Furthermore, after having analysed all the previous results, from the agents, we showed that agents not only can navigate and locate tasks in complex 3D rooms but also understand complex concepts and patterns in the Raven-inspired intelligence tests. There are also indications that an agent can emulate a person's behaviour, with respect to the OCEAN five personality trait category, by acting out with different aspects and actions, differentiating their playstyle inside the 3D Escape Room. We conclude that all the agents can learn to navigate efficiently the 3D Escape Room and also find the hidden targets inside the room. In more detail, the agent that stimulates openness and explores his environment learns rather quickly and has high enough scores, so we can say that a human player that has such a game style is performing well. A player who tries to work more methodically, like the agent simulating conscientiousness, will be late to be effective and may slow down the room progression due to instability of his performance. Finally, neuroticism generates very unstable results, though it can prove effective in the long run. But when referred to human players, even if these results are indicative since cooperation and communication are an important part of the game, they can prove useful for individual player profiling.

Future extensions of our implementation will include more agents with more complex methods of emulating behaviours that will generate a complete behaviour template. Additionally, a system that will create relations between the agents can be developed so that multiplayer behaviour and communication templates can be built. Furthermore, an online game system can be implemented, so that the metrics and the profiling will be based also on interaction and communication with other players.

7. References

Clarke, S. et al. (2017) 'escapED: A Framework for Creating Educational Escape Rooms and Interactive Games For Higher/Further Education.', International Journal of Serious Games, 4(3), p. 14.

Dicheva, D., Dichev, C., Agre, G., Angelova, G., 2015. Gamification in Education: A Systematic Mapping Study. Journal of Educational Technology & Society 18, 75–88.

Fotaris, Panagiotis, and Theodoros Mastoras. "Escape rooms for learning: A systematic review." Proceedings of the European Conference on Games Based Learning. 2019.

Jang, K.L., Livesley, W.J., Vemon, P.A., 1996. Heritability of the Big Five Personality Dimensions and Their Facets: A Twin Study. J Personality 64, 577–592. https://doi.org/10.1111/j.1467-6494.1996.tb00522.x

OpenAI (no date) *Gym: A toolkit for developing and comparing reinforcement learning algorithms*. Available at: https://gym.openai.com (Accessed: 15 April 2021).

Rui Pan, Henry Lo, and Carman Neustaedter. 2017. Collaboration, Awareness, and Communication in Real-Life Escape Rooms. In Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17). Association for Computing Machinery, New York, NY, USA, 1353–1364. DOI:https://doi.org/10.1145/3064663.3064767

Raven, J., 2000. The Raven's Progressive Matrices: Change and Stability over Culture and Time. Cognitive Psychology 41, 1–48. https://doi.org/10.1006/cogp.1999.0735

Schulman, J. *et al.* (2017) 'Proximal Policy Optimization Algorithms', *arXiv:1707.06347 [cs]*. Available at: http://arxiv.org/abs/1707.06347 (Accessed: 18 April 2021).

Robson, Karen & Plangger, Kirk & Kietzmann, Jan & McCarthy, Ian & Pitt, Leyland. (2015). Is it all a game? Understanding the principles of gamification. Business Horizons. 10.1016/j.bushor.2015.03.006.

Technologies, U., n.d. Unity Real-Time Development Platform | 3D, 2D VR & AR Engine [WWW Document]. URL https://unity.com/ (accessed 2.14.21).

https://www.verywellmind.com/the-big-five-personality-dimensions-2795422 (accessed 2.14.21).

Vergne, M. J., Smith, J. D. and Bowen, R. S. (2020) 'Escape the (Remote) Classroom: An Online Escape Room for Remote Learning', *Journal of Chemical Education*, 97(9), pp. 2845–2848. doi: 10.1021/acs.jchemed.0c00449.

Verywell Mind What Are the Big 5 Personality Traits? [WWW Document],