Does a players perception of AI complexity in games affect their experience during game play?

James Collins

Abstract—**AI in games is already an under represented area of computing, but this is only exasperated further when the player's experience and immersion is also taken into account. With so few papers currently focusing on this; this study aims to fill a gap in the literature about how players experience is effected by the complexity of AI in games. It will explore several definitions of immersion, and what affects a players level of immersion along with novel methods of creating interesting AI for games. It will then go on to detail a proposed method to study this effect.**

I. INTRODUCTION

AI for games is an area within the games industry that is garnering increased interest in both Game Design venues and AI Research Venues [1]. However, there is still a gap in the literature regarding how AI can effect players immersion within games. In order to understand how AI can affect the experience a player can have in a game I will need to look into several key areas. My first line of inquiry will be to define what is meant by 'Immersion' and whether it is possible to bring all immersive experiences under the broad term 'Immersion' or whether a more individual approach is needed. Generally speaking it has been found that there are several stages of immersion: engagement, engrossment and total immersion [2]. A player can progress through each stage and break down/remove several barriers before they can enter the next stage of immersion. Some barriers can be as simple as putting time into the game, while other can be influence by the game, such as realism [2].

Another interesting line of inquiry is whether the well known effect "The Placebo Effect" is present in AI for games too. That is to say, the player can be tricked into believing that certain mechanics are present within a game and thus immerse themselves further than they would without the addition of that mechanic. It is known that the impression of a mechanic that is actually not present can improve the user's experience [3]. It is also know that the placebo effect can help improve a players performance in game [4]. So the researcher should be mindful of this effect and research into it further in order to avoid this effect occurring during my test. It has the potential to significantly skew the results of the study to such a degree that it could render them useless. To further this significant care must be placed in making this experiment a 'double-blind' experiment in order to reduce the chance of any double guessing from the participants.

It is also known that AI for games can also affect a players choices in ways that they may not even perceive. For example, Ferstl et al found that NPC facial expressions can be used to affect a players decisions and moral choices [5]. An interesting point would be to see whether players are affected proportionately depending on their level of immersion. This is also something the researcher must be mindful about as I would need to ensure that any effects do not last longer than the game. It is a point of dispute as to whether games can cause effects that 'leak' over into real world decisions and actions however, so this may not need to be researched into. Due to the presence of and necessity for team work based AI systems this should be researched into heavily and many different methods for AI decision making and communication could be used. More complex methods for creating the AI could be researched into such as neural networks and the methods of how to train them, however, this is most likely outside the scope of this study so shall not be looked at in detail.

II. LITERATURE REVIEW

A. What is Immersion?

Immersion is a very large point of debate within the academic world. Many previous works have gone on to 'define' immersion but have not talked about what motivated that definition. It is almost impossible to define what immersion is as it is a very personal feeling, with many of the definitions being flawed in one way or another [6]. What one persons version of immersion could be very different to someone else's. Therefore during this literature review is going to go into detail about several different definitions of what influences immersion in order to decide which one best fits my study. The reasoning behind adopting a pre-defined model of immersion is that in order to create my own definition for immersion with any backing I would have to do an entire paper solely on immersion which, due to time constraints, I cannot. Therefore I must use a model of immersion that has been created and tested by other academics. While beginning to look at different definitions of immersion, it becomes clear that the that theories can roughly be split into two categories, Perceptual and Psychological as stated by *Carr et al* [7]. Researchers that view immersion as a perceptual phenomenon refer to immersion as the degree to which a technology or experience monopolizes the senses of a user. In contrast, researchers that view immersion as a psychological phenomenon emphasize cognitive rather than sensory features of the game, referring to immersion as involving the players mental absorption in the game [8]. Many researchers are starting to use different emerging platforms, such as VR [9] to see if they can induce a higher level of immersion, however, for this study I will just be using a game that can be played on a regular PC, so as such the research will focus on this as a platform.

Levels of Immersion

One theory of immersion is that it is experience in three levelled tiers. This is a 'Grounded Theory' [10], which means that it has been created through the rigorous testing of empirical data. This effectively means that this theory has been created by going through lots of empirical data such as test results to ensure it is correct. However, this statement is disputable as immersion is not an experience which can be defined for everyone thus this theory can not be correct for everyone.

The theory is that immersion is split into 3 distinct tiers: Engagement, Engrossment and Total Immersion. A player effectively moves up a tier when several 'Barriers to immersion' are broken [2]. A player will then move back down the tiers if barriers are remade.

1) Engagement: Engagement is the first tier of immersion. The first barrier in this tier is access. That is to say that a player must actually like the genre of game that they are playing. *Douglas et al* suggest that this can be done by adhering to a common schema [11] or in other words a set of rules set for the structure of specific genres. These will make the player feel comfortable in the game as they subconsciously know how the game's narrative will roughly go. For example, if the game is first person shooter (FPS) then they must like FPS games to break this barrier. The second main barrier is investment. That is to say the player must invest time into the game. "The amount of time, effort and attention required from the gamer increases for more immersive experiences" [2]. However, this does not necessarily mean that a large amount of time must be spent playing the game, as there seems to be an optimal exposure time to reach a immersed state [12].

2) Engrossment: The middle tier of immersion, engrossment is reached once the barriers during engagement have been broken. As stated by *Brown et al*: "The barrier to engrossment is game construction. This is when game features combine in such a way that the gamers emotions are directly affected by the game." [2]. This means that during engrossment the game itself is required to break barriers to immersion. Whereas previously the only barriers we're related to the player, the actual design of the game now plays a major part in how immersed the player is. Some game features that form this quality were visuals, interesting tasks, and plot. It has been shown that higher quality game visuals can lead to a higher level of immersion [13]. Players could also tell when a game was well constructed and could see when designers had put effort into construction. This added to their sense of respect for the game. It is this sense of respect that breaks this barrier to immersion. At this level of immersion players tend to be much more emotionally invested in the game [2]. This investment makes the player want to keep playing and can lead to them feeling emotionally drained when they stop playing. The game becomes the most important part of the players' attention and their emotions can be directly affected by the game. it is this state that can lead to total immersion.

3) Total Immersion: "Total immersion is presence" [2]. During total immersion players report feeling cut off from reality or detached to an extent where the only thing that mattered was the game. The main barrier at this stage is that presence is fleeting; it's a very hard thing to maintain in a player. However, *Cheng et al* dispute this, saying that once achieved, immersion will actually help 'cover up' any failures with the game [14]. The barriers to presence are empathy and atmosphere [2]. Empathy is an extension of attachment and atmosphere is created by the game construction. Brown et al found that, out of the games they tested with, " All but one game mentioned as totally immersive was a first person perspective game. Also role-playing games were mentioned, where the gamers assumed a character."[2]. The three main features that combine to create atmosphere are: graphics, plot and sound. A combination of all three at a very high level can create the atmosphere required for total immersion. Interestingly though *Cheng et al* found that once a sense of total immersion had been achieved, if any of the previously stated features were to drop out or be changed beyond reason, it would have little to no effect on the players state of immersion [14].

Embodiment

Another method for classifying immersion is embodiment. Many things can contribute to embodiment.

4) Graphical Fidelity: One common thought is that immersion is directly influenced by graphical quality and fidelity otherwise known as realism. Realism was found by several researchers to have a definite effect on immersion and engrossment of the player [6], [15], [13]. However, it has also been shown that this is not the case for all [14].

5) Audiovisual synchronization: The fact that both the video and sound react to each other, and to the players actions makes the world feel more vivid and compelling [15]. This factor might cause sensory immersion as defined by Ermi and Mayr [16].

6) Skill Level: It has also been noted that a player's skill level within the game can influence their level of immersion. That is to say, the higher the player's skill level, the higher their potential for immersion [13].

B. The Placebo Effect

The placebo effect is a common effect in many fields, from medicine, to sport. It is therefore possible that this effect could occur within games too, making it necessary to research whether this is the case and whether I should avoid it. Before beginning this review, a definition of the placebo effect is required. In medicine an example of the placebo effect would be: that the mere presence of a 'cure' (even if it is an inert, non-operational pill) heals people [17], [18]. Thus a translation of this to games would be telling a player that a certain mechanic exists within the game when it does not having an effect on the experience of the player[3]. It is known that the anticipation of the suggested reaction is said to lead to the generation of that reaction [19]. It has been confirmed by *Denisova et al* that when a player is falsely told that they are playing a game with adaptive AI they have a better experience in the game compared to those that are not falsely told [3]. This poses an interesting problem for my experiment as the group who are playing against a 'dumber' AI must not know that the experiment involves complex AI. This could lead to them

having a better experience in game due to the placebo effect. *Duarte et al* performed an interesting experiment using in-game power-ups to create the placebo effect, whereby they told the players these power-ups gave them bonus time to complete the stage, when in reality the power-ups did nothing. They found that players achieved a noticeably higher score when influenced by the placebo effect [4], [20].

C. Adaptive AI in Games

The main factor of games that I will be testing here is the AI. I am looking to see if there is a correlation between AI complexity and player experience, thus a delve into the current works on AI in games and more specifically adaptive AI is necessary. AI in games has been playing an increasingly important part in games in recent years but research into existing works is seemingly quite lacking [21]. When stating adaptive AI I mean an AI that can learn from it's surroundings but also, most importantly, can communicate between each other to overcome goals, or to create more interesting situations for the player [22]. This effect has been proven to be especially effective in the field of interactive story telling [23], [24] where it is used to create emergent and unexpected story lines, much like the AI director that is used in the game *Left4Dead 2* [25]. However, one issue this doesn't answer that is necessary for this study is whether AI could lie to each other or the player. In answer to this, *De Rosis et al* found that given the right implementation an AI is not only capable of lying to each other for personal reasons, but are also able to try and catch any deception [26]. The latter statement is of most interest as adding a mechanism for characters within the groups of AI agents in our game to lie to each but also to try and detect and remember the betrayals could massively contribute to the affect the AI has on the players decisions. *Ryan et al* created a interesting ontological design for an adaptive AI with a large focus on communication between AI. Their design was an ontology of linked mental models for a town environment where each NPC character would communicate and learn about each other by communicating. They would have opinions about things, and lie when they believe that it would benefit them. An interesting addition is that of forgetting information. After a period of time based on a memory stat, the characters would then also begin to forget information they have learned. This is an interesting design as it could fit quite well in my game as it would fit the needs of information propagation. They also used a novel method of action selection; utility based action planning. This is where an agent is viewed as a collection of competence modules. Action selection is modelled as an emergent property of an activation/inhibition dynamics among these modules [27]. Yannakakis performed an interesting experiment whereby he got university students to test a game that featured different variations of AI in order to find the best way of generating Interesting Interactive Opponents by the use of Evolutionary Computation [28]. He made heavy use of evolutionary neural networks [29], as they can learn the best ways to act in order to challenge players. However, I feel like this is outside the scope of my project due to time limitations.

III. RESEARCH QUESTION & HYPOTHESES

A. Research Question

My research question is as follows:
Does a players perception of AI complexity in games affect their experience during game play?
With this question the researcher intends to find out whether there is a correlation between the complexity of the AI opponents and a player's experience during game play. The question requires two different kinds of AI for participants to compete against, one simple, one more complex.

1) Simple AI: The simple AI will consist of 4 different AI agents that can navigate towards their goal, and then back towards the spawn location. They will not be able to communicate between each other, and as such they won't be able to work together to overcome any attempts made by the player to beat them.

2) Complex/Adaptive AI: The Complex AI will perform the same overall tasks as the simple AI. However, they will be able to communicate between each other to overcome the actions of the player. The idea behind this is that they will create a more interesting experience for the player.

B. Hypotheses

I have three hypotheses for this experiment:
a) Their will be a positive correlation between AI complexity and player experience.
That is to say the more complex a the AI is the better the player experience will be.
b) Their will be no correlation between AI complexity and player experience.
That is to say increasing the complexity of the AI will have no effect on the player experience.
c) There will be a negative correlation between AI complexity and player experience.
That is to say that as the AI complexity increases the user experience will decrease.

It is expected these hypotheses to be true for different demographics of players. I would expect more experienced gamers to have a better experience as the AI complexity increases as it would stimulate them more, creating more game play that inspires constant attention, and presenting more of a threat to them which is known to increase immersion [15] thus following hypothesis A.
However, for the players with less experience in gaming it is expected that they would either support hypothesis B or C. This is because if the player's skill level is not equal or higher to the AI then they will find the game too hard to be immersed [13].

IV. METHOD

The aim of this experiment is to measure the difference in experience between players competing against simple AI and complex AI. This will be done by giving the participants one of two versions of a game where the only difference will be the complexity of the AI. They will then complete a questionnaire about their experience during the game play.

A. Experiment Design

The experiment will take the form of an A/B test of two randomized groups of participants. This aims to remove or reduce data contamination whereby participants will act differently or have a different experience if they believe a complex AI to be present when it is not. This is also a form of placebo effect [3] that should be avoided throughout the experiment. Group A will play through the game with the simple AI, while group B will play through the game with the complex AI. Both groups will then be asked to complete a questionnaire about their experience during the game. The result of these questionnaires will be used to measure player experience.

B. The Game

The game will be a top down wave defence game, set in a mansion where the player is tasked with stopping a group of characters (AI Agents) from entering the house, retrieving items and then escaping from the house. To do this the player will be able to place a variety of traps down in order to scare/control the flow of the characters through the house or kill the AI to completely remove the threat. The simple AI will behave in very predictable ways, whereby they'll just travel to their goals, without looking for traps or trying to avoid them. They will then leave the house by the fastest route. Whereas the complex AI will be constantly trying to avoid any traps they've seen, and they will communicate the location of traps to other AI agents in close proximity, they will try and find the 'path of least resistance' to their goal. That is to say that they will try and find the path that involves the least number of traps, or in other words, the safest path to their goal, rather than just the fastest path. Each complex agent will behave differently based on their personality.

C. AI Agent Personalities

Each AI agent will have a different personality and will be able to perform different actions based on certain stats and traits. These stats will be things like strength, intelligence and perception not too dissimilar in functionality to games such as *For The King* [30] or *RimWorld* [31]. The traits will be tags that a character can have that will either change how they behave, or let them perform certain actions unique to that tag. The idea behind this is that this will add more depth to the AI and will create more game play that inspires constant attention, and presenting more of a threat to player, thus immersing them more [15]. A few examples of the stats and traits are as follows:

1) Stats: Strength: An increased strength stat will allow the character to break obstacles that are blocking them. However it will also reduce the chance the character will have of spotting traps. A character with a high strength stat is less likely to want to be in the same area as a character with a high intelligence stat.

2) Stats: Perception: An increased perception stat will allow the character to spot more traps even ones that are significantly harder to spot.

3) Stats: Intelligence: An increased intelligence stat will give the character a chance to disarm a spotted trap. The higher the intelligence stat, the easier it is to disarm traps. This will open up avenues for the AI to actively work against the players plans.

4) Traits: Lone Wolf: The lone wolf trait will make the character want to stay away from any other group member, effectively making them go their own way around the house. This will split the group up, giving the player more things to think about at any one time.

5) Traits: Fearless: This trait makes the character take half fear damage from any fear based traps. In effect this makes the character fit the role of a tank as they become significantly harder to kill. They also make any characters in the same room as them take reduced fear damage. This means the player will have to prioritize separating them from the group.

As you can see, the inclusion of these stats and traits will modify the game significantly, and give the player a lot more to think about and manage at any given time. This aims to add to the aforementioned game play that inspires constant attention that presents more of a threat to player.

D. Collecting Data

All the data collected will be through a questionnaire. This questionnaire has been made using previous questionnaires made for study into immersion as references such as that developed by *Witmer et al* [32] and *Yannakakis* [28]. These both involve a large amount of quantitative data that is used to judge the players experience during game play. It is worth noting that if I were to expend the study to include multiple platforms (multiple test games) of study I would have to consider the use of Krippendorffs alpha coefficient [33], while analyzing the data. This would allow me to find the general consensus between both of the games. Although this could be a possible avenue of further study, it is unlikely this will be necessary for this study.

E. Participants and Selection

The main method of recruitment will be through word of mouth. The main demographic will be university students aged 19 - 22. The researcher aims to keep my experiment as generalized as possible in terms of gaming skill, rather than narrowing down on one specific skill level. To do this they will be recruiting from the wider student pool at Falmouth University. However, the researcher will have to be careful as Falmouth University runs a large game development department, so as such I will have to take extra care that the participants do not solely come from there, as they will all high a much higher level of skill then students who do not regularly play games. To ensure that the researcher is aware of any imbalances in relation to this and to limit any imbalances that may occur, participants will be asked to record their course of study. This will not be linked to any specific questionnaire results, but will be used to further study any results.

Due to a power analysis using G*Power [34] it has been

determined that the total sample size of the experiment should be 70 participants, with a split of 35 in group A and 35 in group B. This will leave the experiment with a total power of 0.9523628. A non-centrality parameter of 3.3466401 was calculated, DF of 68 and t value of 1.6675723 were also calculated. When studying the results of the study I will use R [35].

F. Random Participant Assignment and Participant Blinding

During the experiment the participants will be randomly assigned to either group A or B. This will dictate whether the participant competes against the simple or complex AI. This randomization inhibits the participants ability to guess which AI they are facing before encountering them. This will reduce the likelihood of the placebo effect [3] occurring. To further this, the researcher will not know which version of the game the users will be playing either thus making this a 'double-blind' experiment. This means the participants won't be able to guess which version they are playing from any interaction with the researcher.

It is important to note that as the randomization is being handled by a computer it will not be truly random, rather it will be pseudo random. This does not pose a problem, however, as the participant will not be able to guess which version they are playing regardless of whether the randomization is true random or not. These measures are very important in order to stop or, at the very least, limit the likelihood a participant will try to double guess the system.

G. Variables

The independent variable during my experiment will be the 2 variations of the game the participants could play. One being simple the other complex. The dependant variable will be the user experience which should be influenced directly by the independent variable. This should allow me to study any correlations AI complexity and Immersion/Player Experience may have.

H. Issues

One major issue in the study will be the case of if participants try to double guess the version of the game that they are playing. The precautions taken to avoid this have been stated above but, to summarize, the main precaution is the implementation of the experiment in a 'double-blind' manner, whereby neither the participant nor the researcher know which version of the game the participant will be playing reduces the chances of this happening massively.

Another problem or question that the researcher has had to address is whether it would be useful to collect data during game play such as number of enemies killed or number of traps placed in order to get a better idea of how immersed a player is. However, it is thought that this is not going to be of a significant benefit over the questionnaire. This is because even though players can seem immersed on a video, they may not actually feel immersed themselves, it would be better to get their own feelings from them, rather than infer from a video.

V. DISCUSSION AND TOPICS FOR FURTHER STUDY

Obviously one of the most challenging points about this study is finding a suitable definition for immersion, as stated in the literature review. While the researcher is of the opinion that immersion is such a personal experience that it is nearly impossible to define, a definition must be used to standardize the experience as much as possible for the sake of the analysis. A combination of the current literature gives the impression that immersion is actually more like a huge web of intertwined factors and influences that are too numerous to measure at once. With many works such as that of *Emily Brown & Paul Cairns* [2], *Daniel Wilcox-Netepczuk* [6] and *Cheryl Campanella Bracken and Paul Skalski* [13] trying to pull together but a few of them at a time. The fact that so much literature is present but divergent from each other begs the question, is any of it wrong? Or is immersion perhaps such a broad topic that it can't be studied on such a small scale. If the topic of immersion is disputed so frequently, could it be that immersion as we know it now could be split into several different topics or effects that are related to each other? This could be especially true when the two main viewpoints on immersion are taken into account, that of whether immersion is a Perceptual or Psychological effect [7]. With both 'sides' attributing immersion to two very different phenomenons it becomes evident that perhaps immersion could be at least two separate but linked effects. Sadly this thought is far outside the scope of this study, but whether immersion as a term could be broken down into several different effects would be an interesting avenue for further study.

Another interesting point of development would be the use of a neural network for each AI agent in our game. This could lead to some quite unique behaviours. One way that a neural network could be trained is through the use of supervised and unsupervised learning methods to create emergence of cooperative multiagent spatial coordination as shown by *Yannakakis et al* [36].They also find that learning by rewarding the behavior of agent groups constitutes a more efficient and computationally preferred generic approach than supervised learning approaches in such complex multiagent worlds. The opinion of the researcher is that the inclusion of neural networks would have a benefit on a player's experience by increasing their immersion. This is because it would create interesting game play and present more of a challenge to a player [15] as the behaviours exhibited could be different from agent to agent and from play through to play through, meaning the game would continue to be stimulating even after the first couple of games. This idea would effectively extend from the current study as it ups the complexity of the AI even further. It would be interesting to use this idea to test whether there is an optimal level of AI complexity in relation to levels of immersion.

REFERENCES

[1] M. Mateas, "Expressive ai: Games and artificial intelligence." in *DiGRA Conference*, 2003.

[2] E. Brown and P. Cairns, "A grounded investigation of game immersion," in *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 2004, pp. 1297–1300.

[3] A. Denisova and P. Cairns, "The placebo effect in digital games: Phantom perception of adaptive artificial intelligence," in *Proceedings of the 2015 annual symposium on computer-human interaction in play*. ACM, 2015, pp. 23–33.

[4] L. Duarte and L. Carriço, "Blue pill or red pill?: placebo effect and the outcome on physiological & player performance metrics," in *Proceedings of the 4th International Conference on Fun and Games*. ACM, 2012, pp. 93–96.

[5] Y. Ferstl, E. Kokkinara, and R. Mcdonnell, "Facial features of non-player creatures can influence moral decisions in video games," *ACM Transactions on Applied Perception (TAP)*, vol. 15, no. 1, p. 4, 2017.

[6] D. Wilcox-Netepczuk, "Immersion and realism in video games-the confused moniker of video game engrossment," in *Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational & Serious Games (CGAMES), 2013 18th International Conference on*. IEEE, 2013, pp. 92–95.

[7] D. Carr, D. Buckingham, A. Burn, and G. Schott, *Computer games: Text, narrative and play*. Polity, 2006.

[8] C. Jennett, A. L. Cox, and P. Cairns, "Investigating computer game immersion and the component real world dissociation," in *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2009, pp. 3407–3412.

[9] M. Kyriakou, X. Pan, and Y. Chrysanthou, "Interaction with virtual agentscomparison of the participants' experience between an ivr and a semi-ivr system," in *Virtual Reality (VR), 2015 IEEE*. IEEE, 2015, pp. 217–218.

[10] K. Charmaz and L. L. Belgrave, "Grounded theory," *The Blackwell encyclopedia of sociology*, 2007.

[11] Y. Douglas and A. Hargadon, "The pleasure principle: immersion, engagement, flow," in *Proceedings of the eleventh ACM on Hypertext and hypermedia*. ACM, 2000, pp. 153–160.

[12] C. Zhang, S. Zadtootaghaj, A. S. Hoel, and A. Perkis, "How long is long enough to induce immersion?"

[13] C. C. Bracken and P. Skalski, "Presence and video games: The impact of image quality and skill level," in *Proceedings of the ninth annual international workshop on presence*. Cleveland State University Cleveland, OH, 2006, pp. 28–29.

[14] K. Cheng and P. A. Cairns, "Behaviour, realism and immersion in games," in *CHI'05 extended abstracts on Human factors in computing systems*. ACM, 2005, pp. 1272–1275.

[15] A. S. Bastos, R. F. Gomes, C. C. dos Santos, and J. G. R. Maia, "Assessing the experience of immersion in electronic games," in *Virtual and Augmented Reality (SVR), 2017 19th Symposium on*. IEEE, 2017, pp. 146–154.

[16] L. Ermi and F. Mäyrä, "Fundamental components of the gameplay experience: Analysing immersion," *Worlds in play: International perspectives on digital games research*, vol. 37, no. 2, pp. 37–53, 2005.

[17] D. E. Moerman, *Meaning, Medicine, and the" placebo Effect"*. Cambridge University Press Cambridge, 2002, vol. 28.

[18] D. Walker-Batson, S. Curtis, R. Natarajan, J. Ford, N. Dronkers, E. Salmeron, J. Lai, and D. H. Unwin, "A double-blind, placebo-controlled study of the use of amphetamine in the treatment of aphasia," *Stroke*, vol. 32, no. 9, pp. 2093–2098, 2001.

[19] A. L. Geers, P. E. Weiland, K. Kosbab, S. J. Landry, and S. G. Helfer, "Goal activation, expectations, and the placebo effect." *Journal of personality and social psychology*, vol. 89, no. 2, p. 143, 2005.

[20] L. Duarte and L. Carriço, "The cake can be a lie: placebos as persuasive videogame elements," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2013, pp. 1113–1118.

[21] J. O. Ryan, A. Summerville, M. Mateas, and N. Wardrip-Fruin, "Toward characters who observe, tell, misremember, and lie," *Proc. Experimental AI in Games*, vol. 2, 2015.

[22] G. N. Yannakakis and J. Hallam, "Interactive opponents generate interesting games," in *Proceedings of the International Conference on Computer Games: Artificial Intelligence, Design and Education*, 2004, pp. 240–247.

[23] H. Ten Brinke, J. Linssen, and M. Theune, "Hide and sneak: Story generation with characters that perceive and assume." in *AIIDE*, 2014.

[24] D. B. Carvalho, C. Pozzer, E. Clua, and E. Passos, "A perception simulation architecture for plot generation of emergent storytelling," in *Proceedings of the International Conference on Computer Games, Multimedia and Allied Technology*. Citeseer, 2012, pp. 6–11.

[25] "Left4dead2," Valve, 2009.

[26] F. De Rosis, V. Carofiglio, G. Grassano, and C. Castelfranchi, "Can computers deliberately deceive? a simulation tool and its application to turing's imitation game," *Computational Intelligence*, vol. 19, no. 3, pp. 235–263, 2003.

[27] P. Maes, "How to do the right thing," *Connection Science*, vol. 1, no. 3, pp. 291–323, 1989.

[28] G. N. Yannakakis, "Ai in computer games: generating interesting interactive opponents by the use of evolutionary computation," 2005.

[29] X. Yao, "Evolutionary artificial neural networks," *Encyclopedia of computer science and technology*, vol. 33, pp. 137–170, 1995.

[30] "For the king," IronOak Games, 2018.

[31] "Rimworld," Ludeon Studios, 2018.

[32] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[33] K. Krippendorff, "Computing krippendorff's alpha-reliability," 2011.

[34] A. G. . B. A. Faul F. Erdfelder, E. Lang, "G*power 3: A flexible statistical power analysis program for the social, behavioural, and biomedical sciences." in *Behavior Research Methods*, ser. Behavior Research Methods 39, pp. 175–191.

[35] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: https://www.R-project.org

[36] G. N. Yannakakis, J. Levine, and J. Hallam, "Emerging cooperation with minimal effort: Rewarding over mimicking," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 3, pp. 382–396, 2007.

APPENDIX A
FIRST APPENDIX

Appendices are optional. Delete or comment out this part if you do not need them.