

When the Game Shows the Player – A Study about Games for Leadership Assessment

Anastasia Thambwe
anastasia.thambwe.ser@gmail.com

Thesis MSc. Media Technology, Leiden University – July 2022
Supervision: Peter van der Putten & Marcello A. Gómez-Maureira



ABSTRACT

In recent years, games have shown a potential for personality assessment, becoming a new way to develop or assess specific traits in addition to regular assessment methods, such as questionnaires and tests for example.

This research focuses on a game which is designed to investigate leadership in a group context. This paper discusses leadership in society, the potential of serious games for leadership assessment and leadership behaviours in games.

A framework of leadership traits and behaviours is defined, and through an iterative and qualitative design process inspired by design science, the game is designed, prototyped and tested with participants. In total, 6 prototypes have been run where the behaviours of each player have been observed based on the framework of traits. Additional assessment tools such as self evaluation questionnaires are also implemented in some prototypes. Observations show how the game allows an external observer to assess leadership potential for each team. Results also show that some traits, such as conflict or stress management, seem to be more complex to distinguish among the players, whereas traits such as communication seem to be the most shown traits of all. It also shows how the group dynamic and players behaviours can be impacted by the team composition rather than by the game design itself.

Finally, the limitations and the potential of games for leadership assessment are discussed, as well as directions for future research or what questions it raises when it comes to how games can be used for leadership assessment.

1 INTRODUCTION

What is leadership to you? Take a few moments to think about it, about what the concept might mean to you. Maybe you have a definition in mind, or maybe a list of criteria and characteristics that a leader should have. Or you might even have someone specific in mind who fits the description of a leader for you, it might be someone known in human history or someone from your surroundings.

Leadership can take many forms and can be shown by a multitude of people. Several styles of leadership are defined in literature, not per se meaning that one is better than the other one, nonetheless, it can mean that one is more appropriate than another one according to the situation. For example, authoritarian leadership focusing more on getting the obedience from the followers in contrast to transformational leadership where encouragement and guidance of the followers is more a priority. Over the years, tools have been created to develop and assess such skills. In this research, the attention is brought upon the assessment of leadership, and more precisely, the potential of games for such task.

A few leadership assessment methods already exist and are used on a regular basis in a recruitment context, or other sectors and situations such as skill evaluations. Each of those tools have their ways of functioning or their framework of analysis to determine the type or efficiency of leadership of the person being put under the assessment. A few of them do mainly focus on self-evaluation or external opinions [11, 15, 19]. Here the attention is brought onto games since they focus on directly and actively analysing the behaviour of the person in-situation, which allows to directly observe the skills being applied [19, 21].

For this purpose, the following research question has been established: “How can games be used for leadership assessment in a group context?”. Overall, the research is focusing on the possibilities that games can offer regarding leadership assessment, as well as focusing on the assessment of leadership potential in the context of a specific team composition.

To answer this question, a multiplayer game is designed and tested with participants. The goal with this is to create an environment with the adequate circumstances to encourage the display of leadership behaviours within the teams of players testing the game. A framework of traits and behaviours was established, representing the leadership style observed in this research. An important part is that the behaviours are also analysed by an external observer, external meaning not participating in the game and in this case, the main researcher, to determine who in the team is showing the most potential for leadership compared to the other players. The game was iteratively designed, tested and reworked based on the observations. This was done following a methodology inspired by design science, and by Cole et al. (2005) [7] synthesized model. In total, 6 prototypes have been run with groups of 3 or 4 participants, and for each group, the player with the most leadership potential within their team was assessed. Some prototypes also made the use of additional external assessment methods, for example, through self-evaluation questionnaires.

Results show that for each group, the potential for best leadership was established for one participant within each team. It was also observed that the questionnaires seemed to show bias from the participants, therefore causing inconsistencies in their answers. Another observation is the fact that no player showed an adequate demonstration of conflict management, which was the only trait from the framework that could not be observed in any of the prototype runs. On the contrary, traits such as communication and problem finding/solving were some of the most common traits observed throughout the empirical phase of the research.

This research brings a new perspective to traits assessment using games as an alternative or additional tool, along with focusing on leadership in other ways than just training. The aim of the research is not to design the perfect tool or assessment method supposed to replace more conventional ones, but it is to design and explore the possibilities of a different tool, to bring a different perspective on leadership assessment.

Furthermore, this research covers on design science and iterative design processes as scientific and qualitative methods. As well as designing a framework of traits with the purpose of creating an artefact based on it. In addition to that, it also focuses on the actual assessment process of leadership in the groups that participated in the prototype iterations.

This research paper covers the scientific background of this research (section 2), as well as the methodology chosen (section 3). Next, the results and observations of each game prototype iteration are described (section 4). Finally, the insights gathered and interpretation done from the results and the limitations of the research are discussed (section 5), before concluding the paper (section 6).

2 BACKGROUND

This section covers what inspired this project and dives into the scientific background and literature around leadership in society and the potential of games for serious and assessment purposes.

2.1 Personal Motivation

The idea behind this thesis originated from a simple gaming afternoon with some friends. The four of us decided to attempt an escape game experience with a straightforward principle. We were brought into a locked room, from which we had to escape under 60 minutes.

It is only after losing 15 minutes of our precious time, unsuccessfully trying to complete a puzzle, that we realised how our current strategy was not going to allow us to escape on time. We had already used 1/4 of our time and had done close to no progress at all at this point. That is when one of the team members decided to set specific tasks for the team and encourage everyone to work more efficiently.

That player took the initiative to take the lead. Because of that, our team dynamic completely changed, and we were actually able to make up for the lost time. This allowed us to solve all the puzzles of the room and to escape from it before the end of the timer.

This simple gaming afternoon with friends ignited a spark of interest when it comes to the use of games and simulation for leadership emergence. I started to wonder, are games often used for leadership emergence and assessment? Is there a potential for it to become more of an assessment method of choice in diverse contexts? What do games specifically tailored for leadership-related purposes have in common and focus on?

Moreover, I had always shown interest for interpersonal skills and traits. I see leadership as this idea of people that are able to lead their followers in their ideology towards fulfilling a common goal. And those questions made me want to explore the concept more and to see how it could be paired with games for more purposes than simply completing a game.

This is what I wanted to explore through an iterative design process, on my quest to design a new tool to see the potential games have for leadership assessment.

2.2 Leadership in Society

This subsection covers leadership in literature. This takes into account how leadership in society and organizations is usually defined in studies and research, and how different leadership styles can influence its definition and characteristics.

Leadership is one of the most covered subjects in literature., with researches, studies, books and articles relating it to many disciplines and different fields. [9, 20]. The importance and need for leadership depends on the situation, however a few fields and context require the appropriate leadership for a proper completion of the tasks, for instance in health and medical care [19].

First, it should be clarified that managers are not per se leaders, and leaders are not per se managers. Those two roles are not mutually exclusive, but they also don't always apply to the same person. To differentiate both, Ackoff (1999) specifies that a manager focuses on getting tasks done by the use of means and by directing

others, while leaders focus on encouraging and guiding others to accomplish goals [1].

Depending on the literature, leadership can have more than one definition. A leader can be considered the one able to guide a group, represent its value and lead them towards their goals [14, 17].

However, Mazzarella et al. (1989) mention that leadership is not an innate trait, and that defining it through traits might be more tedious than expected as the context has a notable impact on whether they will take on the leader role or not, according to "situationists" [17].

Leadership has been defined by different leadership styles, which are defined by observable behaviours. More recent studies also relate the leader to the followers [5].

It does not necessarily mean that one style is better than the others, but it can mean that one is more appropriate than another one according to the situation. To illustrate how one style might drastically vary from another, a comparison can be made between authoritarian leadership and transformational leadership. Authoritarian leadership is usually characterised by the need for unquestioned obedience from and strong authority over the followers, while transformational leadership is usually characterized by the idea of inspiring and motivating the followers to achieve and develop themselves more [4].

Several studies have been conducted about the different styles of leadership, but also about the differences between good and bad leadership [17]. Indeed, it might be popular belief to think a leader is indeed always proactive and good in what they do, but some leadership styles might not be appropriate for a certain organization or situation, therefore making its impact weak.

Another aspect researched is leadership in a team context and how it impacts the team agitation and performance [8].

At this point, leadership assessment has become a very important process and has different methods ranging from comprehensive tests, written measures, interviews, observation, critical-incident techniques and more, with methods such as structural interviews being one of the most commons [11, 15]. It can be determined through solitary activities, missions, or assessment methods [19].

Larry Lashway (1998) dives in his article about measuring leadership and its effects, as well as the different measurement tools and their limitations and how each measurement tool focuses on a definition of leadership or the traits looked for that can vary from one organization or individual to another. He also argues that leadership assessment should not be left to one individual, and any evaluator should be properly trained and carefully prepare the process [15].

Overall, scientific literature shows how leadership can be defined through different styles and how this can greatly impact the behaviour and goal of a leader when it comes to their followers. Literature showed as well a focus on leadership styles being sometimes dependent on the situation, and on the importance that leadership assessment is gaining throughout the years.

2.3 The Potential of Serious Games

This subsection covers the potential of serious games and simulations for serious purposes. The literature covering assessment of traits, behaviours and skills through games is also discussed.

For a long time, games have been considered primarily as a tool for recreative purposes and not much for more serious tasks, however, in recent times this mindset has shifted [12]. Now that games are also being used for other purposes than just entertainment, a whole new genre of games destined to be used for serious purposes emerged, most commonly known as 'serious games' [21]. Those purposes can range from training, skill development to educational purposes and more [6, 21].

Games and simulations have the potential of gathering the conditions to create an environment that is very similar to the type of situations that might be encountered in real life [6], but also situations that could not happen in the real world [11]. According to specific design choices and game setting, an environment that will encourage certain behaviours can be created and used for a variety of diverse goals [6].

The potential of serious games for the military or healthcare has been researched because of the possibility to put the player into a situation without having consequences on the real world, since it remains contained in the simulated environment [16]. Although the game can be very close to reality in terms of immersion, it stays a fictitious context where any failure won't have an actual impact outside [19, 21]. This simulated environment allows testing their abilities and reflect on the players' behaviour in a safe environment where failure of the tasks would not have any impact on the real world [16].

It could be said that games propose unnecessary goals and challenges, since those have no impact or causal link with the real world. However, even though those are self-contained within the game, players can act very seriously towards the accomplishment of the tasks at hand. Serious games have a purpose that goes beyond the ones defined by the environment of the game itself. They can be used to influence the player's mindset and push the aims of the game experience to more than what is defined in the game [18].

In addition to development or teaching with serious games, assessment through the means of games has also been recently explored. Studies focused on game-based learning mentions there are three types of assessment that can be implemented into games: Game scoring, external assessment and internal assessment. Game scoring being related to the in-game achievement of the players, external assessment being not directly part of the game environment, so for example interviews and questionnaires. And internal assessment being part of the game experience and being the collection of information about the players' in-game behaviours [12].

Globally, even though games are frequently affiliated with entertainment, the whole genre of 'serious games' is gaining more attention throughout the years. Allowing to create a situation similar, or even different in some way, to the real world to be able to explore or realize specific tasks that might be easier than in the real world. With goals and challenges contained within its simulated environment, this allows to experiment without taking any risk related to the real world, therefore being a strength for training, learning or assessment purposes.

3 METHODOLOGY

The following section covers the research method chosen to answer the research question in more details. An iterative design method

is adapted to match the goal of the research and create the main artefact, being the game [10]. The framework of traits is described, as well as how it is used. The implementation of self-evaluation as external game assessment and experts' feedback through interviews are also discussed.

As previously said, this research aims to answer the following question: "How can games be used for leadership assessment in a group context". For this purpose, a game is prototyped, tested, analysed and from that analysis are changes and improvements determined to explore the potential games have for leadership assessment. Groups of 3 or 4 participants play the game, and the observations made from each prototype influences the next ones. All the observations are then compared to see what insight can be taken from the results (section 4).

With this research, the aim is to discover more insight about applied games for leadership assessment and to add to the existing scientific knowledge around games and leadership.

The leadership style observed in the context of this research is inspired by inspirational leadership, a type of transformation leadership [2]. Bass (1985) describes it a leadership focused on communicating their vision with the group, motivating and showing confidence to the team. The framework of traits is also based on the observations of games often used for leadership development, to determine what traits were mainly focused on within those. For that purpose, a table of different games paired with studies of leadership or used for leadership assessment was created (appendix C). Some of the traits assimilated to inspiration leadership are used to design a framework of 8 specific traits, and each of those traits is defined by clear behaviours to ensure a consistent observation throughout the prototypes. That framework of traits serves as a clear definition of what type of leadership is observed in the context of this research.

The game is designed in a way that encourages all players to participate and get together in collaboration to complete it. The gameplay experience aims to create similar behaviours from the players in real-life situations where leadership is required to successfully achieve the tasks at hand [21].

3.1 Iterative Game Design

This section goes more in detail about what iterative game design is and how the iterative game design part of this research was planned and dealt with. It also dives deeper into how design science, more specifically the version of Cole et al. (2005), was used and adapted to fit this project [7].

Iterative design can be described as a process where design decisions are made along the development based on how the artefact or game performs, most commonly by external testers. This allows to regularly adapt the game design in order to arrive to the desired play experience, since game designers and creators can't foresee how the final experience will be for the players [22, 23].

Philip Tan (2010) talks about iterative game design by stating that it provides continuous feedback through the process and iterations. Tan also mentions the difficulty behind that method, such as the need for a multitude of testers being a necessity, and therefore needing rigorous planning to ensure that [22].

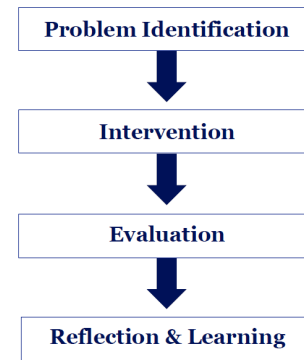


Figure 1: Cole et al. Synthesized Model of Design Science

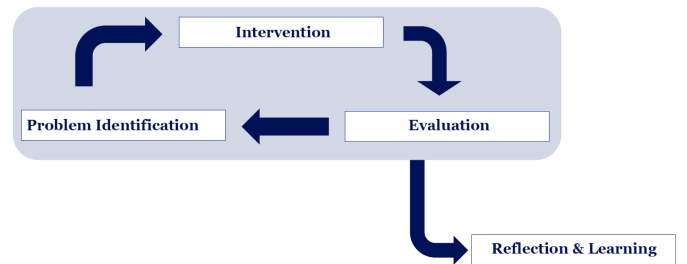


Figure 2: Iterative Design Methodology - Inspired by Cole et al.

Here, the methodology was inspired by the synthesized research approach based on design science by Cole et al. (2005) [7], which is suitable for the creation of such artefact (see figure 1). This methodology was chosen to explore the "how" aspect of the research question by performing changes throughout the iterations and seeing how this impacts the performance of the game for leadership assessment. Additionally, this allows to test different settings, as well as different assessment implementations for the game and to see how this could give more insight.

It is defined by the following steps: Problem identification, intervention, evaluation and finally, reflection and learning. Those steps are reiterated for each prototype, with the evaluation of each prototype defining the problems of the next iteration.

The problem identification step is understanding and defining what needs to be focused on, in this case it is to be able to assess leadership potential through the game. The intervention step represents the designing and creation of the artefact with the goal to solve the problem, therefore creating a game fulfilling the main goal. The evaluation step covers the game and observations that were made through the game, this is also a representation of the results of a prototype. Finally, the reflection and learning phase is where the observations and results are interpreted to contribute to the theoretical framework of the research.

The slight difference here is that the three phases of the "problem identification", "intervention" and "evaluation" are repeated

iteratively through each prototype before detailing the "reflection and learning", mainly through the discussion (see section 5 and figure 2).

Six prototypes were run in total, to see how the gameplay and design could be improved or changed, and to observe how it impacted the potential for each leadership trait. Some prototypes used the same game design to see if the observations made were related to the group and therefore the profile of the participants and the composition of the team, or if those observations were indeed more related to the actual game design.

3.2 Framework of Traits

As previously mentioned, leadership can take many forms and different styles. In this research, the leadership style is inspired by inspirational leadership, where meaning is given to the needs and actions of the team by the leader [3, 20]. Behaviours of the inspirational leader described by Bass in 1988 are being able to explain why to take certain actions, being persuasive, setting objectives for the team, being able to get their ideas across, motivate followers to exceed their own personal expectations, as well as contributing to the overall enthusiasm of the team [3]. The leadership style observed in this research can, additionally, be defined as the leader whose focus is brought on achieving the goals set for the team through proper communication to facilitate the completion of the tasks [13].

A framework of traits has been defined for the study to ensure the observation would be consistent. Each trait is paired with a clear description of behaviours that are considered as a display of that trait. Some aspects related to inspirational leadership or not taken into account in this research because it would be much more complicated to objectively observe in general or in the context of the game specifically (e.g. creativity).

Any player showcasing the behaviours from the framework would score one point for that trait, each time they do, allowing a more thorough tracking of leadership based on their interactions during the game and to see what traits should be more triggered in the next iteration.

The participant with the most potential for future leadership is determined by the amount of relevant behaviours they demonstrated, with diversity being more important than the repetition of the same behaviours. Indeed, if participant A showcased "communication" 9 times and not much of the other traits, but participant B showcased 8 different traits at least once, participant B is considered to have the most potential.

The framework is composed of the traits and behaviours listed in table 1.

3.3 User Evaluation as an External Assessment Method

User evaluation questionnaires were implemented as an external game assessment method for two of the prototypes, to explore that possibility. The results were analysed compared to the observations made to see what insights could be taken from them. The user evaluations were done through individual questionnaires (see appendix A and appendix B), where the participants had to fill in about their behaviour and the other players' behaviour during the

Confidence	The player displays confidence in their choices, themselves, the team or someone else. If during the game they seem to encourage a certain solution or seem to be sure about the choices and solutions presented to or by them. They show initiative.
Communication	The player keeps or strives to keep communication channels open. That means they try to get the team members to communicate with them and with each other.
Delegation	The player delegates and or makes decisions for the team or team members specifically.
Stress management	The player is able to manage the stress and tensions, resulting from the environment, within the team to still be cohesive in their actions and interactions with others.
Conflict management	The player makes an effort to address issues resulting from conflicts within the team members. If there is a problem, they will mention it or try to resolve it. For example, if tensions arise between 2 members, that player will point it out and try to handle it.
Problem finding & solving	The player proposes problem-solving solutions or actively aims to find another way to solve a problem and to deal with an obstacle.
Reward accomplishment	The player rewards achievement through verbal praise, encouragement or a physical action, for example, a friendly tap on the shoulder, to congratulate another player's or the team's effort.
Objective feedback	The player makes the effort to provide objective and constructive feedback to the other team members or the team in general.

Table 1: Framework of Traits and Behaviours

game. The questionnaires' questions were based on the framework of traits (see section 1). Those questionnaires were to be filled alone, after the game.

3.4 Expert Interviews for Feedback

The evaluation of the methodology was also based on leadership and game design experts opinion on the theories used as literature background, as well as the process, framework and methodology of the research, to determine if those were suitable choices for the research question.

The experts were presented with material, concept ideas and data for and from the game directly, while being explained the goal of the research and research question.

4 PROTOTYPES

To describe the observations of each prototype, the participants are referred to with letters (A, B, C, D). Each prototype run ended on a short group interview with questions related to the overall game experience, leadership in general and leadership assessment preferences.

The main differences between the prototypes are described in table 2.

For all prototypes, the players were gathered in the same room and could freely move around, interact with and speak to each other.

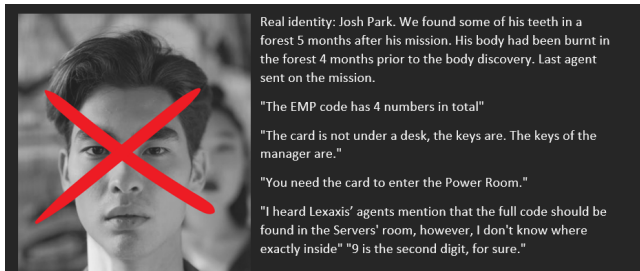


Figure 3: Game Information - Third Agent



Figure 4: Game Information - Lexaxis Staff

In the subsections "observations" of each prototype, the assessment of potential leadership is described and the main behaviours of each participant are chronologically listed.

Also, the description of the "communication" trait slightly differs from the other ones, as it is textually described in the paper for only the most relevant actions, however, communication is implied with every interaction aiming to share information in some way or another between the participants.

The narrative of the game is that the players are an emergency team of spies that need to help agent DEXTER fulfil an important mission. For this mission, DEXTER requires certain items and information that will help them sabotage the dangerous evil corporation that is being infiltrated. The players are provided with confidential files summarizing all information that the spy agency was able to collect from previous spies that died on the mission (see figure 3 and 4).

They are required to analyse the data at hand, since they are warned that the previous spies failed because of contradictory and flawed information. Therefore, the players need to distinguish what is actually useful from what is erroneous while guiding, through a digital game, agent DEXTER and gaining more information through the interactions with DEXTER.

	Setting differences
Prototype 1 (3 players)	Interaction with a computer by typing. Music played with audio time reminders. Information available on a tablet.
Prototype 2 (4 players)	Interaction with a game on a computer by clicking on options. Music played with audio time reminders. Different information (textual and vocal) available on 4 different devices. Audible and visible timer.
Prototype 3 (4 players)	Interaction with a game on a computer by clicking on options and need to enter answers by typing when required. Music played with audio time reminders. Different information (textual and vocal) available on 4 different devices.
Prototype 4 (3 players)	Only vocal information is available on devices.
Prototype 5 (4 players)	Same setting as for prototype 4.
Prototype 6 (3 players)	Use of a projector to allow all players to see the screen at all time. Change of timing and situation to increase stress levels. Important information being punctually played audibly, thus, the players have to take notes of it

Table 2: Main differences between the prototypes

4.1 Prototype 1: Testing the Design and Gameplay

Three players participated in this prototype. For this first instance of the game, the focus was to determine in what direction the gameplay and game design should go. The players could interact with agent DEXTER by typing and reading instant messages. They had access to all confidential information on a digital tablet.

The team did not manage to complete the game in time. From the observation of that prototype, we can deduct that participant B was the one showing the least potential for leadership in that context and participant A seemed to be the one with the most potential. We can also observe that participant C demonstrates more leadership traits toward the end of the game.

4.1.1 Prototype design. The material used were a laptop with the digital communication software installed, a tablet used to give access to the documents and a speaker to play music and sounds during the game. Audio messages reminding how much time the players have left were played on a regular basis.

The players, being in the same room, and being able to freely move around and communicate with each other, are presented with a computer where they need to type to communicate with fictitious characters that are part of the game. When the game begins, a suspenseful music is started to symbolize the beginning of the experience. On the computer, the players are in contact with characters that present themselves as a spy agency. That agency explains more about the mission that they need to complete under 35 minutes.

They are then put in communication with agent DEXTER and have to ask questions or give orders to the spy, through typing

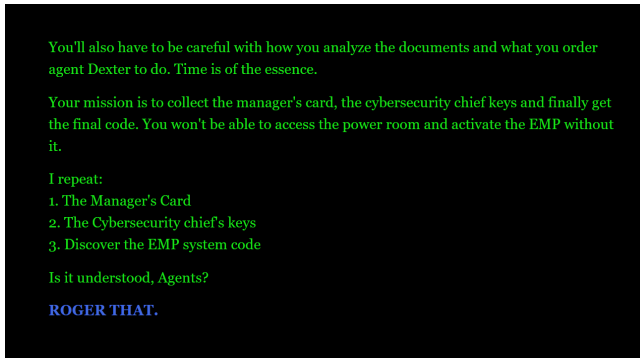


Figure 5: Digital Game - Summary of the Mission

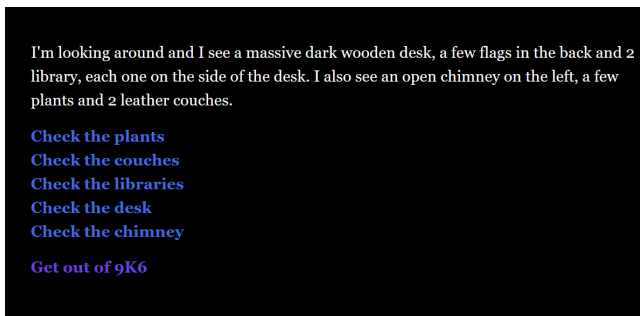


Figure 6: Digital Game - Interactions

on the laptop, to gain knowledge about the situation and make progress in the mission. The participants are also provided with blank sheets of paper and pens for plausible note-taking.

4.1.2 Observations. When it comes to the overall group dynamic, what is observed with this prototype is that for the majority of the game, the players are calmly exploring the elements they have and are trying to read through all the information that is being sent to them by the fictional characters. There were very few interactions between the players in the first part of the game, since they were more focused on observing. It also did not seem as if they felt immersed in the urgency of the narrative until the last 5 to 10 minutes. Since all the information was on a tablet and the communication happened by typing on the computer, one player was assigned the role to read out loud all the conversations and type with the fictional characters.

Globally, the participants did show the emergence of some traits from the framework but as it was previously mentioned this prototype was more focused on the overall design of the game to see how the game design should be improved. The main and most dominant traits that were observed during this prototype were communication, delegation and constructive feedback.

In more details, participant A is the first one to move around, they do that to be able to read the information better (problem finding/solving). All participants also read the information on the tablet silently (problem finding/solving). Participant B starts interacting with the other players, to ask about the resources they have (communication). Participant A mentions that a map might be part of

	Participant B	Participant A	Participant C
Confidence	Participant C	Yes, but no name	N/A
Communication	Participant B	Team	Participant B
Delegation	Participant B	Participant C	N/A
Stress management	Participant A	Team	Participant A
Conflict management	N/A	N/A	Participant B
Objective feedback	Participant A	N/A	Team
Problem finding and solving	Participant C	N/A	Participant C
Reward accomplishment	Participant B	N/A	N/A

Table 3: Results of Prototype 1's Questionnaire

the resources, the team is able to find the map in the documents and use it to understand the situation better (problem finding/solving, communication). Participant A brainstorms information and summarizes what the team needs to focus on (problem finding/solving, communication, delegation, confidence). Participant A and B delegate participant C with asking the fictional character information (delegation). All participants decide to analyse the map together (communication, problem finding/solving). Participant C reads the information from the game and proposes an idea that might help the team go forward (problem finding/solving, confidence). Participant C takes the initiative to interact with the game (confidence). Participant C answers and confirms a question confidently (confidence). Participant C proposes a solution to a problem (problem finding/solving, confidence), the rest of the team agrees.

4.1.3 User evaluation questionnaire. At the end of the experiment, the players had to fill in a form and to state for each trait of the framework of traits, if they believed someone showed those traits, and if yes, who they thought did. The results of those forms are listed in table 3.

4.1.4 Interview. This section lists the main answers and information given by the participants during the interview.

The participants believed the game helped encourage leadership behaviours to a certain extent. It was entertaining and collaborative, and it encourages communication. However, it was also sometimes confusing, and the participants did not feel confident enough to take action sometimes. Such an activity could also be too fun, and thus distract from the leadership need.

They did feel pressure and stress because of the last alert signifying they had about 5-10 minutes left, as they did not realize how much time passed before that.

They felt detached from the characters and consequences.

They think there is potential for games and leadership assessment, specifically through having a situation where different roles and tasks are required, similar to a professional environment.

4.1.5 Conclusion. Globally, this prototype made it hard to efficiently observe the behaviours looked after because of the setting and lack of interaction or dynamism coming from that game. The players did not feel immersed in the narrative and felt detached from the consequences, which is the opposite of what is looked for with the prototype.

It could also be observed that the participants did not understand right away that the tablet could be swiped to show more information, thus reducing the amount of interaction they could have with each other and not allowing them to fully play the game in the beginning. With this first prototype, the dynamic looked for in the gameplay was not attained yet, thus there were in general less behaviours to track compared to with the following prototypes and that is why another method was chosen to display the information.

From the interview, we can take as insight that there is a potential for the game to encourage leadership, as it requires collaboration, cohesion and strategy. However, the current prototype does not immerse the players enough to really accentuate that need for leader behaviours.

The next iteration will focus on having a better game flow, a better structured narrative and better visual and audio cues to create a more tense environment.

4.2 Prototype 2: Testing the Design and Gameplay

Four players participated in this prototype. This time, the digital game was ready to be used, so the interaction with agent DEXTER happened through an interactive textual game instead of an instant messaging software. The players each had one device, and each of the devices had access to a different set of information. Nonetheless, the players were allowed to move around and to show their devices to others. There was also the implementation of a visual and audible timer that was visible during the whole game. The team managed to complete the game in time.

From the observation from the game, participant A and B showcased the most potential for leadership through their behaviours, with participant A being the one with more diverse traits' demonstration, therefore, being the one with most leadership potential.

The participants filled in a questionnaire at the end of the game. They had to rank every player from best to worst, themselves included, for each trait of the framework and for an additional "good leader" trait. The questionnaire shows that some players put themselves as the worst player on numerous traits, even though this does not align with the external observations. Participant A was chosen as first for "good leader" by all participants, but was not always the participant other players would put the most often as first for the other traits.

4.2.1 Prototype design. For the second prototype, the overall design and material used for the game was improved. Foremost, instead of having the players interact with the computer by typing, they were able to interact with a digital game that was run on a computer. They were only required to interact by clicking and selecting options that were given by the game. Therefore, simplifying the exchanges between the fictional character and the team, and leaving more room to let the players interact with each other rather than with the game.

Each participant was using an additional computer. Each device had access to a digital platform that would show a set of information, and each set of information was different. This was described in the narrative of the game as a more secured way to display confidential information. Although each received one exclusive set of information, they were not required to be the only one to look at it.

Therefore, the team had to decide their strategy when it came to analysing the resources efficiently.

This decision was made to encourage the need for interaction, as each participant had valuable information for the mission. If one player decided to not share their documents, the completion of the game would hardly be realizable. More content was also added to the game documents in general, to allow the players to have more to analyse and look through. Consequently, encouraging the need for a proper strategy and leadership. This prototype also introduces vocal messages that are part of the information available, where before only text and photos were included in the documents.

The game also had a visual and audible timer, that showed minutes and seconds while making a ticking sound at every second passed. The audio reminder for the time left were kept for this prototype too.

A survey and a leadership test were completed after the game.

4.2.2 Observations. Directly early in the process, the team decides to move all the computers so that everyone is able to see everything. This is an idea that is proposed by participant A (delegation, confidence, problem finding/solving). Participant A is the first one to interact with the game (confidence). Later in the game, participant B asks the team what they have for information (communication, problem finding/solving). Participant C is the first one to react to this by reading their information out loud (communication, confidence). Participant A is the first one to move around the room to be able to interact more efficiently with the game and resources (problem finding/solving, confidence). Participant B often takes the time to propose to the other players ideas and solutions (problem finding/solving, confidence). Participant D decides to draw on the map, attempting to understand where DEXTER is better (problem finding/solving). Participant B is the second one to move around and goes to help participant D with the map (problem finding/solving, communication). Participant A and C start interacting more (communication). Participant A gives information from the game and C analyses their resources (communication, problem finding/solving).

Stress builds up because of the clock ticking, the players acknowledge it but bring their focus back to the game (stress management).

Participant B rewards the accomplishment of the team when finding an item required in the game by telling the team they did a good job (rewarding accomplishment). Participant C proposes to check all the options of the game given at some point (problem finding/solving, delegation). Participant B proposes to check one option specifically, and they managed to find another element part of the game (problem finding/solving, delegation). Participant A mentions toward the end of the game that the team still needs an important element before going to the end. (feedback, communication)

4.2.3 User evaluation questionnaire. The participants had to fill in a survey where every question had to be answered, and each question covered the traits of the framework (see subsection 3.2). The participants were required to rank every member of the team from best to worst for each trait, themselves included. It was also specified that rating someone as best or worst does not per se signify that the person is good or bad regarding the trait, but that they seemed to perform better or worse than the other players in the context of that game, at that time. This survey was adapted in this manner to avoid having the participants not choose someone

Participant A

	First to last	Coming most as 1st
<i>Good leader</i>	A, B, D, C	A (4), B (2)
<i>Confidence</i>	A, B, C, D	Best leader
<i>Communication</i>	B, C, D, A	A
<i>Delegation</i>	A, B, D, C	
<i>Stress management</i>	A, D, C, B	Coming most as 4th
<i>Conflict management</i>	D, A, B, C	C (4)
<i>Problem finding/solving</i>	B, A, D, C	Worst leader
<i>Reward accomplishment</i>	C, B, D, A	C

Participant C

	First to last	Coming most as 1st
<i>Good leader</i>	A, B, D, C	A (3), B (3)
<i>Confidence</i>	B, D, A, C	Best leader
<i>Communication</i>	A, B, D, C	A
<i>Delegation</i>	B, A, D, C	
<i>Stress management</i>	B, A, D, C	Coming most as 4th
<i>Conflict management</i>	A, D, B, C	C (6)
<i>Problem finding/solving</i>	C, A, B, D	Worst leader
<i>Reward accomplishment</i>	C, D, A, B	C

Participant B

	First to last	Coming most as 1st
<i>Good leader</i>	A, D, C, B	D (5), A (3)
<i>Confidence</i>	D, A, C, B	Best leader
<i>Communication</i>	D, A, C, B	A
<i>Delegation</i>	D, A, C, B	
<i>Stress management</i>	D, A, C, B	Coming most as 4th
<i>Conflict management</i>	A, D, C, B	B (8, all)
<i>Problem finding/solving</i>	A, D, C, B	Worst leader
<i>Reward accomplishment</i>	D, A, C, B	B

Participant D

	First to last	Coming most as 1st
<i>Good leader</i>	A, B, C, D	B (3), C (3)
<i>Confidence</i>	C, D, A, B	Best leader
<i>Communication</i>	B, A, C, D	A
<i>Delegation</i>	A, C, B, D	
<i>Stress management</i>	C, B, A, D	Coming most as 4th
<i>Conflict management</i>	C, A, B, D	D (6)
<i>Problem finding/solving</i>	B, A, C, D	Worst leader
<i>Reward accomplishment</i>	B, C, D, A	D

Table 4: Results of Prototype 2's Questionnaire

specific, as they did with the first questionnaire (appendix A) by selecting the 'team' instead of a player.

Globally, participant A has been rated first 13 times, second 13 times, third 4 times and last 3 times. Participant B has been rated first 8 times, second 9 times, third 4 times and last 11 times. Participant C has been rated first 6 times, second 3 times, third 13 times and last 10 times. Participant D has been rated first 6 times, second 7 times, third 11 times and last 8 times. With participant A being voted the most as first and with participant C being voted the most as last. Other observations show that participant A rated

themselves last for 2 traits, participant C rated themselves last for 6 of the traits? Participant B rated themselves last for all behaviours, when they were part of the highest rated participants overall, and participant D rated themselves last for 6 traits.

4.2.4 Leadership survey results. Regarding the leadership test results, participant C came in first position with a score of 100 out of 100. In second position, participant B with a score of 63. In third position, participant A with a score of 53. And finally in fourth and last position, participant D with a score of 37.

Consequently, the results given by the leadership test vary substantially from the external observations of the game or the survey results. However, it should be taken into account that the leadership test and the prototype of this research are not testing for the exact same behaviours or situations. As it was mentioned in the background of this research (see section 2), leadership behaviour can differ according to the context and the situation. Someone might be a great leader in one situation and a simple follower in another one. The leadership test also focuses on questions related to habits of the person taking the test and does not actively take the group context into account. After this prototype run, the focus was brought on the game design And how this impacts the leadership behaviour rather than comparing the game to other leadership assessment tools (e.g. existing leadership surveys).

4.2.5 Interview. This section lists the main answers and information given by the participants during the interview.

The participants stated that the experience was entertaining and that they could communicate well.

They also stated the information was sometimes too contradictory, and they were not sure what to believe or not believe at some point during the game.

They believe the game encourages leadership behaviours since it creates a need to take a role leader with all the tasks that need to be dealt with in parallel.

They believe that most players unintentionally took a role during the game.

They stated that usual leadership methods can be rough estimations for leadership but also have bias, since it relies on how you perceive yourself for some of them.

They do believe that there is a potential for games and leadership assessment. One participant states that it is probably already done with professional retreats.

Globally, the participants would probably prefer to go through an experience that is between a regular leadership test and the game. They would feel more stress with the game, so some participants stated they would prefer the test since it is easier to cheat.

4.2.6 Conclusion. With this prototype, there is a drastic improvement with how immersive the experience was. More interactions and behaviours from the framework could be observed and therefore the establishment of leadership potential could be done in a more thorough manner. The players also felt immersed much earlier in the gameplay than with the first prototype. The survey showed some potential as the best leader chosen by the players correlated with the external observations done, however, that result did not match the player rated highest for most traits within each survey,

which could indicate that determining the best leader could have been done based on other requirements.

The leadership test showed very different results from the external observations and the survey, which could show that both the prototype and leadership test focus on different aspects. Therefore, it does not seem relevant to compare both tools.

The interview indicate that the game is now entertaining but still encouraging leadership behaviours, with all the tasks required to be completed in time. Players also notice that they sometimes naturally take on roles without the need for delegation. They also state the bias that other leadership methods might have, but also how it is easier to cheat with regular tests, for example. They explain they might not prefer to go through such a game activity, in a hiring context, because of those reasons.

4.3 Prototype 3: Testing the behaviours

Four players participated in that game. The self evaluating questionnaires were removed, as the previous results tended to show that too much bias was influencing the results. The team managed to complete the game.

Participant A was the one to show the most potential for leadership during this prototype, as they showed the most amount of behaviours from the framework, as well as the most diverse traits' demonstrations throughout the game.

4.3.1 Prototype design. For prototype 3, the setting was very similar to the one of prototype 2 to see how this game design does with a different group of participants. The position of the elements during the game was slightly different, but basically the same.

4.3.2 Observations. Participant C takes the initiative to read out loud the game for the other players (confidence). Participant A is the first one to ask that someone interacts with the game (delegation). Participant D takes on the role of interacting with the game for the team, since they were closer to the computer (confidence). Participant A is the first one to communicate what type of information they have, and encourages others to do the same (communication, problem finding/solving, confidence).

Participant A asks questions to the team (communication), they also ask the team to find specific information (delegation, problem finding/solving). Participant D decides to keep track of the timer for time management (stress management, problem finding/solving). Participant A and D propose a solution regarding one of the puzzles of the game (problem finding/solving, communication). Participant C encourages starting drawing a map as this could help work more efficiently (Problem finding/solving, feedback, delegation). Participant A tries to interact with participant B to get them more involved in sharing information (communication, confidence, delegation). Participant D proposes to do an action to try to find more clues (delegation, problem finding/solving). Participant C and participant D propose to do a specific action (problem finding/solving), participant A is the one to confirm to do that action (confidence, delegation). Participant B and C express success and enthusiasm when one part of the puzzle is solved (rewarding accomplishment). Participant A takes initiative and decides to delegate the team what to do from there onwards, by either proposing to take an action or by making the decision for the team when a question is asked

(confidence, delegation, feedback). Participant C reminds the team of important information they need to take into account (Problem finding/solving, feedback).

4.3.3 Interview. This section lists the main answers and information given by the participants during the interview.

They stated that everyone was active and if some people were less active in another context, maybe that would have encouraged more leadership behaviours.

The participants believe players showed signs of leadership during the game.

They felt stress, especially towards the end, and about 20-25 minutes before the end of the game.

They felt immersed. The music and clicking timer helped with that.

They mentioned that having the information printed could be more interesting for interactions, as now they had to go back and forth between all sources of information, which could have slowed down their interactions.

In a hiring context, they think the game would provide better engagement from the participant. It is a more practical way to assess leadership compared to written tests, for example.

In a non-hiring context, they think this game would work for leadership assessment, since it's collaborative and needs dynamic leadership.

4.3.4 Conclusion. Observations show that, on one hand, some players decided to interact less and take on the position of a follower, such as participant B. On the other hand, participant A showed several times signs of confidence, initiative and delegation by naturally taking the lead during the game. Participant C also showcases throughout the game behaviours from the framework, however, those actions were more related to problem finding and solving. With this prototype, one instance of stress management, could be observed with time management.

The interview shows that the participants believe the game encourages leadership behaviours, and that players did showcase signs of leadership. They would prefer to play this game in a hiring context and think it has potential in a non hiring context.

4.4 Prototype 4: Testing the Behaviours with Printed Resources

Three players participated in this prototype. The main difference with the setting is that the game documents were printed instead of displayed on digital devices (see figure 7). The team did not complete the game.

Observations indicate that participant A showed the most potential for leadership.

4.4.1 Prototype design. With this prototype, instead of having the information available on one or several electronic devices, most of the resources have been printed and gathered on a table in the middle of the room (see figure 7). Only a few audio recordings were available on electronic devices, and they were accessible by everyone. They also had access to a blank sheet of paper and pens for note-taking, similarly as for the previous prototypes. The main computer and timer were placed on another table than the one

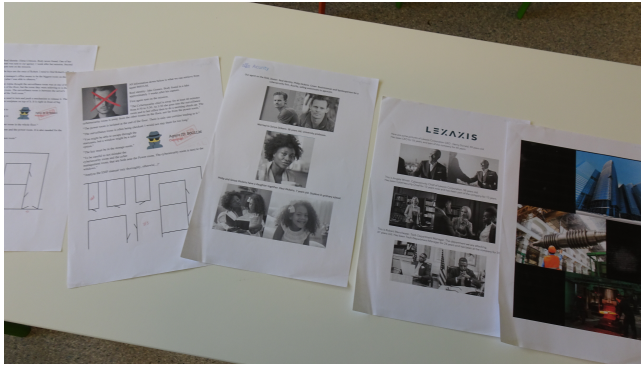


Figure 7: Game Information on the Table

with the printed document, nonetheless, participants could move everything around.

4.4.2 Observations. Participant A takes the initiative to do the first action and goes to the game to read the text out loud (confidence). Participant C encourages the behaviour (rewarding). Participant A summarizes the important information for the other player and encourages the team to take notes (confidence, communication, delegation). Participant A and C work in collaboration to read out loud and take notes together (communication, problem finding/solving). Participant A and C go together to the other devices to listen and analyse the audio recordings (communication, problem finding/solving). Participant B starts analysing and taking notes on the printed material to highlight important clues (confidence, problem finding/solving). Participant B explains their initiative to the rest of the team (communication). Participant A proposes to the team to interact with the game (delegation). Participant B communicates that it's a good idea. (Feedback) Participant A reminds the team of the time left and encourages them to take action (stress management, delegation, feedback). Participant B proposes to look around to help better understand the map (delegation, problem finding/solving). Participant B also reminds the importance of certain actions that could cause a game over (feedback, communication). Towards the end of the game, participant A and B tend to feel the stress and lose patience, however they globally ask for participant C's input before taking action at that time (stress management, communication). Participant C gives a proposition (problem finding/solving).

4.4.3 Interview. This section lists the main answers and information given by the participants during the interview.

The participants do believe the game can encourage leadership behaviours. Especially with the need for collaboration and a good team dynamic and cohesion to complete the game.

They believe everyone showed signs of leadership. Speaking up is an example. They do not know if there was a specific leader, as it seemed more equal.

They felt stress during the game, especially 20 minutes before the end until the end. They were also much more efficient from there on, according to them.

They state that the only conflict was debating over the documents and information.

They do believe that the experience is immersive, through the setup and environment.

In a hiring context, globally, they would prefer to play such a game, but it would also make them more nervous.

In a non-hiring context, they also would prefer the game, as it is more similar to real life. One of the participants mentions that, however, with the game people would tend to react differently than in real life since they would most probably want to be seen as a leader. It might be interesting to see how the game does when people aren't aware leadership is assessed.

4.4.4 Conclusion. Observations show that participant A seems to be the one showing the behaviours of the framework the most, with participant B probably being the second one to show leadership potential. Overall, this team was a lot more cautious with the game since they did not want to put agent DEXTER in danger with poor decisions, so there were fewer actions within the game than for other teams. However, it still allowed the players to debate with each other, and therefore to have substantial content for external observation.

During the interview, the participant stated that they believe the game encourages leadership and that everyone showed signs of leadership. They would be open to play such a game in a hiring and non-hiring context to determine in a group who has the most potential for leadership, but mention they might be nervous.

4.5 Prototype 5: Re-testing the Setting of Prototype 4

Four players participated in this prototype. The setting was globally the same as for prototype 4, to see if a different group and a group of 4 players, instead of 3, could change the overall dynamic. The team dynamic was indeed different, with more behaviours being showcased by this team. They managed to complete the game.

Observations show that participant D is the one showcasing the most potential for leadership for this prototype, with participant B being the second one to also show potential for leadership within this group.

4.5.1 Prototype design. For prototype 5, the main focus is seeing if the observations of prototype 4 were mainly related to the setting and game as well as information they had access to, or if it was more related to the participants that were selected for the actual experiment. Globally, corrections were made to the game and documents were needed, but the setting was the same.

4.5.2 Observations. Participant C is the first one to propose a solution when it comes to the strategy for the game (problem finding/solving, confidence). Participant A reminds the team of the extra material that they should probably analyse and asks the team if they should do that now (problem finding/solving, communication, feedback). Participant D takes the initiative to delegate specific tasks to members of the team and to guide the group in the beginning of the game, therefore increasing efficiency (delegation, confidence, communication, problem finding/solving). The team decides to do a specific action, but participant D questions the choice of the team and asks for further explanations (feedback, communication). Participant B engages the team to analyse the written information and takes pictures of some parts of the game to ease

problem-solving (problem finding/solving, communication, delegation). Participant C celebrates solving one part of the game as a team (rewarding accomplishment). Participant B insists on checking for specific actions before going further (delegation, feedback). Participant D encourages the team to keep on going how they are currently (reward accomplishment). Participant D mentions that the team went the wrong way and should go back (Feedback, delegation, confidence, problem finding/solving). Participant B asks that someone goes to a specific place to check an information (delegation, problem finding/solving). Participant D asks the team to check if everything has been gathered or if something is missing (delegation, communication, problem finding/solving). Participant C reminds the team important information related to the game, and gives one of the solutions (problem finding/solving, communication, feedback). Participant B is enthusiastic in the proposition given by the other players, they do the effort to confirm questions that are being asked and delegate actions (feedback, communication, delegation).

4.5.3 Interview. This section lists the main answers and information given by the participants during the interview.

The overall experience was fun and enjoyable. The participants felt immersed, one of them states that they felt more immersed with that game than with usual similar games they have tried. The ticking and vocal announcement really helped with the engagement.

They believe the game encourages leadership behaviours. There was a need to complete tasks and to collaborate since there were so many details, a strategy was required.

The participants believe people showed signs of leadership during the game.

They felt stress during the last 5 minutes, mostly, but good stress and adrenaline.

They stated that there was healthy conflict over ideas, but no interpersonal conflict.

In a hiring process, the participants globally stated they would prefer a conventional test or survey since it would be less stressful. According to them, leadership tests would not realistically show how they react. They do state that it also depends on the content of the assessment method. Also, for the game, it would depend on the conditions, as they might really not prefer the game if they're being observed by a committee, for example.

In a non-hiring process, some participants stated they would like to use such a game to determine in a group who might be fitter to be the leader. One of the participant states that they would prefer discussing it since it might be more efficient than playing a game. Also, the game might assess one type of leadership that is not required for the project the group needs to work on.

4.5.4 Conclusion. Globally, observations show that participant C and A take on the role of the follower more within this team, whereas participant D and B seem to be the ones to give the most feedback and take the most initiative. Participant B shows more leadership behaviours towards the end, however, participant A showed it the most consistently throughout the game.

The interview shows that again the experience was immersive and encourages leadership behaviours, according to the players. They state they would prefer a conventional leadership assessment method in a hiring context, because of how stressful the game might

be. However, this also depends on the conditions of the game. In a non-hiring process, some players state they would prefer to go through such a game, another participant mentions they would prefer simply discussing it.

4.6 Prototype 6: Testing the behaviours

Three participants played this version of the game. For this prototype, the game was projected onto a wall to allow all players to see the game at all time. Also, some of the information from the game documents were removed and played audibly throughout the game. The team did not complete the game.

Observations show that participant A was the one to show the most potential for leadership. Participant B did showcase some behaviours of the traits, but far less than participant A.

4.6.1 Prototype design. Having the game projected made so that there was no need for someone to read all the options and information out loud. This decision was made to see if removing that predisposition for the role of the person relaying the information from the game could make a difference in the leadership dynamic of the team.

The printed resources were shown on the table where the computer with the main game were showcased, such as in the previous two prototypes. The rest of the information was vocally transmitted through a speaker throughout the game. The goal was to observe the team behaviour when the participants have to be attentive at all time, in case they would receive a vocal message that they might not be able to listen to later on. The goal was to encourage the need for a strategy to make sure they would retain the information, for example, by taking note of them.

Another difference is that the timer could be reduced or increased because of spontaneous events happening in the game narrative to create additional tension due to unforeseen changes.

Other than that, the setting and game design was similar to the previous one.

4.6.2 Observations. All the participants decide to gather around to analyse the documents and participant B is the first one organising the files while communicating the information they see (communication, confidence). All the participants tended to write information and read the resources on their own without sharing much. This is also the only group that didn't read out loud the game. Different A and B discuss the resources (communication). Participant C proposes an idea to the team (communication, problem finding/solving, confidence), this participant also delegate the others with specific tasks related to it (delegation, confidence). Participant B communicates information and summarizes it, they also actively ask the others for complimentary information (communication, confidence, problem finding/solving). Participant A takes on the role of writing down the audio information during the game. Participant A proposes options to solve a puzzle (problem-solving/finding). Participant A hesitantly proposes an action to take at some point (communication, problem-solving/finding), they also enunciate information that could help solve a problem. Globally participant A and B try exchanging information the most while participant C is more reserved, not talking much, and seems to be more focused

on the documents. Participant B proposes an action (communication, problem-solving/finding), participant A mentions that they lack information (communication, feedback), participant C encourages the action (confidence, feedback). Participant A proposes that the team writes down information from the game (problem finding/solving, communication, delegation). Participant A delegates the other player to execute several interactions to do with the game from there on until the end of the game (delegation, communication). Participant C proposes an idea (communication, problem finding/solving), but participant A rectifies them on that idea (feedback). Participant C delegates someone to check a specific part of the game and try a few actions (delegation, communication). Participant A mention that some of those actions might not work and explains why (feedback, communication). Stress can be felt a lot more with this group since they are stuck towards one of the ending puzzles and are reminded of the time pressure. Small conflicts arise, but no conflict management behaviour is observed. Participant C and A propose solutions (problem finding/solving). After a while, participant C proposes to take control of the computer to try to solve the puzzle (confidence, problem finding/solving). Participant A finally proposes an alternative solution, which actually is the right one, but the timer runs out.

4.6.3 Interview. This section lists the main answers and information given by the participants during the interview.

The participants appreciated the game, which reminded them of an escape game.

They believe the game encouraged some leadership behaviours to a certain extent. One person mentioned that it might have done that in the beginning, but it was not continuous.

They believe that players showed signs of leadership during the game.

They think there is a potential for games and leadership assessment.

One player mentioned that they think there was some friction or conflict at some point.

They felt the stress during the game and felt immersed.

In a hiring context, one participant mentioned they went through a hiring process where they used games as filters, and they don't know if they prefer that since it over encourages behaviours people would not show in real life, because they are actively observed. Another participant states that they do not like some other leadership assessment methods, such as personality surveys, because it is easier to lie or not assess skills objectively with those. The last participant mentions that maybe having a series of game instead of just one could help see more accurate and natural behaviours from the participants.

In a non-hiring context, the participants think this type of game could be useful to assess leadership, especially since this would be a more comfortable setting where they know the other participants better.

4.6.4 Conclusion. Team 6 showed fewer interactions in general compared to other teams, therefore fewer behaviours were observed. The addition of the projector seemed to diminish the amount of interactions, considering the players could simply observe the screen to understand the state of the game. The players ended up being

stuck very close to the end of the game, where a puzzle was misunderstood. Even though they made the fastest progress during the majority of the game, they ended up being stuck at that point for over 15 minutes, which discouraged them and decreased the leadership behaviours.

Some insights from the interview are that the participants believe the game might have encouraged leadership, however not in a substantial and continuous manner. They mention how being actively observed playing the game can over encourage certain behaviour and not portray the natural demeanour of a participant.

5 DISCUSSION

In this section, the underlying meaning of the research is explored, and the results are discussed further. The main limitations and future prospects are covered as well.

What can be observed is that this format of the survey showed limitations, as it was not expected that the participants would occasionally put in the "team" as an entry, instead of an actual player. What is also observed is that the players that are focused on the activity might not be as observant of certain behaviour compare to someone that is analysing from an exterior perspective, also would need to be taken into account is the bias of each participant's relationship with other participants during the activity and the fact that they actually have to, in a way, thank their fellow players which had an impact on the overall result. This shows that maybe it is more interesting to force the participants to actually put someone's name as the best one, maybe even to rank the other players. Also, this type of form does not take into account the amount of time a certain behaviour has been shown by the player examined, which is something that is relevant to the observations of each prototype.

Furthermore, the survey of prototype 2 also showed inconsistencies between who they voted as best leader and what participant came as first the most often for each participant. This is most likely due to a different definition of what leadership is to them, which could differentiate from the one defined for this project.

Another observation is that some behaviours related to the framework chosen for the study could not be observed that much during the experiments. Conflict and stress management. A possibility is that the timeframe of the game was too short to really create tensions that would encourage the need for conflict management. Even though some groups showed signs of stress, there was no significant demonstration or attempt of real stress management. It could be explained by the fact that the teams only need to cooperate for the time of the game, therefore conflict or stress management was less relevant during the prototypes of this research, compared to, for example, group activities or projects that last days, weeks or months.

On the other hand, communication was the one trait that was consistently shown by all players of every prototype, some players did communicate and interact less than others, but there was no case of someone actually isolating themselves from the team.

When it comes to the assessment of the behaviour and analysis of them, there is a limit to what can be reported since the testing environment is so open with endless possibilities of interactions and strategies. Indeed, with such experiment, there is the possibility to analyse the demeanour of the participants on a very deep level or

a more surface level. With a clear definition of traits and behaviours observed, significant and consistent observation was able to be done throughout the prototypes (subsection 3.2).

One thing that can be said about how observable some of these behaviours are is that some are definitely easier to identify than others. To illustrate, behaviours such as rewarding accomplishment or feedback are more clearly defined than, for example, confidence. Some of those behaviours are also more subject to subjectivity, however the framework of behaviours does help with that.

Related to that, even though a clear framework was used for the analysis of behaviours, it is acknowledged that there is a limit to how objective the observation can be, since it also depends on the perspective and the person observing.

Besides, some prototype runs also showcased how the dynamic of the team is not only influenced by the setting and design of the game, but can also highly depend on the composition of the team. This was observed between prototype 3 and prototype 4, where much fewer behaviours were observed in prototype 3 compared to the next iteration, but it did not seem to be because of the setting change since prototype 4 was basically the same.

When it comes to the interviews, we notice throughout all the prototypes that, globally, all participants believe there is a potential for leadership assessment through games. Also, that players, in some cases all players, showed signs of leadership, and that most participants did believe the game encourages leadership through the need for strategy and all the tasks that need to be completed simultaneously to achieve the game.

The interviews also show how participants can find it interesting to go through such activity in a hiring process, as a candidate, and in a non-hiring process, to determine the best leader in a team of co-workers, for example. They also mention how the game might not be preferred because of how nervous they might be while actively being observed. Nevertheless, some participants state how it might be easier to cheat with other leadership tests, therefore, showing an additional potential with games where, even though you could act differently than in real life, it might be harder to actually cheat the results.

Finally, seeing how assessment and training could be paired for such a project would also be an interesting take, since games and simulations are already used for training purposes. The game developed in the context of this research could also be adapted to fulfil an assessment purposes, as well as a training purpose with the right design decisions.

5.1 Limitations and Future Prospects

Ideally, the experiment would be carried on with more participants and more iterations, each iteration exploring different game design changes to see what other insights can be gathered regarding the potential of games for leadership assessment. For example, testing with more than 4 players could be interesting to see how this impacts the group dynamic and to observe if it could help trigger some of the traits more, such as conflict and stress management.

As it was mentioned in the background section of this paper (section 2), the definition of leadership is something that can differ according to the style. This research defines, and focuses on, one

specific type of leadership through the framework of traits (section 3.2), with that being the definition of the leadership observed. Therefore, future research could cover other styles of leadership as well, with the focus being brought on other traits and behaviours, which, in turn, could impact the overall game design.

One other aspect to take into account is the setting in which the prototypes were run, that, to a certain extent, most probably had an impact on how the participants acted. Ideally, that could be improved to make the game experience itself more immersive, for example, by not having the observer be in the same room or by directly adding more immersive elements to the environment to match the game narrative and story better. Another interesting take, for future research, would be to see how this game could be adapted in an online context. The 6 prototypes that were tested were all run in a physical space, with each participant being in the same room and being able to freely communicate and interact with each other. That offered the possibility to observe their behaviours, not only through vocal communication, but also through body language. An online multiplayer form of the game would definitely change the interactions and players' demeanour to a certain extent, however, the game design and observation framework could be adapted to match the new forms of communication that the online setting might offer. Maybe even offering new observations that would not be made without the online setting.

Additionally, as it was also mentioned during the interview of prototype 4, almost all participants knew that the experiment was related in a way or another to leadership assessment, which could have created a bias in their behaviour. Having the players not be aware, beforehand, and during the game, of what this research is observing or focused on could also allow gathering interesting insights in terms of results.

6 CONCLUSION

To recapitulate, this research allowed to answer the research question of "how can games be used for leadership assessment in a group context".

By iteratively designing, testing, observing and analysing the prototypes of the game, the results positively supported the hypothesis that an external observer, in that case, the main researcher, is able to identify potential for future leadership within each team, for that specific group context.

The observation was done through defining a clear framework of traits, where each behaviour of the players defined in the framework is tracked throughout the game. For some prototypes, self-evaluating questionnaires were designed and used to see the potential of pairing that external assessment tool to the game itself. However, this method appeared limited within this research, as we can observe tendencies of bias or a lack of consistency in the participants' answers.

Globally, the game was able to encourage the demonstration of all traits of the framework, except for conflict management, which was not significant enough to be considered as an adequate demonstration of that behaviour. Other than that, traits such as communication and problem finding/solving were shown by players the most, and traits such as stress management were shown the least throughout the prototypes. Besides, some prototypes with

identical settings showed how the composition of players in the group also has a significant correlation with the dynamic of the team and the behaviours of the players.

In conclusion, this research investigates the possibilities for games as an assessment method, specifically for leadership. Games already show a potential for training and assessment in general, however, more could be covered about it in scientific literature. This research could open a door on using an iterative design methodology for the creation of such serious game purposes, along with offering a knowledge base for this specific research question that could be extended or worked further on in the future. Such researches could explore more of the limitations and other opportunities of games for leadership assessment.

7 ACKNOWLEDGMENTS

I would like to thank and acknowledge the following people that helped and guided me at some point or during the whole process of this thesis. First and foremost, my supervisors, Peter van der Putten and Marcello Gómez Maureira. I would like to thank them for their supervision, guidance and precious advices every step of the way, which helped me shape and conduct this research by giving me recommendations and all the information needed to make this initial idea a reality.

Secondly, I would like to thank Meggy Pepelanova and Mathijs van Meerkerk that were able to share their expertise and experience working on games used for serious purposes, as well as by giving me feedback on my own research. I also would like to thank Sara Burks, from Adaptis, for taking the time to meet with me and give me advices and feedback regarding the overall game, as well as for the methodology, theories, and framework used to complete this research. Mark Noort, teacher from the Leading and Managing People course at Leiden University, who shared valuable resources in the context of my thesis. Also, I would like to thank both my critics, Edith Järv and Caitlin Schäffers, for taking the time to read this paper, attend my presentation in order to review my work and for giving me feedback.

Finally, I'm thanking all the participants that took part in the prototype runs.

REFERENCES

- [1] Russell L. Ackoff. 1999. Transformational leadership. *Strategy & Leadership* 27, 1 (Jan. 1999), 20–25. <https://doi.org/10.1108/eb054626>
- [2] Bernard M Bass. 1985. *Leadership and Performance Beyond Expectations*. Free Press, New York, NY.
- [3] Bernard M. Bass. 1988. The Inspirational Processes of Leadership. *Journal of Management Development* 7 (1988), 21–31.
- [4] Bernard M. Bass and Bruce J. Avolio. 1990. Developing Transformational Leadership: 1992 and Beyond. *Journal of European Industrial Training* 14, 5 (may 1990). <https://doi.org/10.1108/03090599010135122>
- [5] Lars Borgmann, Jens Rowold, and Kai C. Bormann. 2016. Integrating leadership research: a meta-analytical test of Yukl's meta-categories of leadership. *Personnel Review* 45 (2016), 1340–1366.
- [6] Robyn P Cant and Simon J Cooper. 2010. Simulation-based learning in nurse education: systematic review. *Journal of advanced nursing* 66, 1 (2010), 3–15.
- [7] Robert Cole, Sandeep Purao, Matti Rossi, and Maung Sein. 2005. Being proactive: where action research meets design research. *ICIS 2005 proceedings* (01 2005), 27.
- [8] David V Day, Peter Gronn, and Eduardo Salas. 2004. Leadership capacity in teams. *The leadership quarterly* 15, 6 (2004), 857–880.
- [9] Sara de Freitas and Helen Routledge. 2013. Designing leadership and soft skills in educational games: The e-leadership and soft skills educational games design model (ELESS). *British Journal of Educational Technology* 44, 6 (2013), 951–968.
- [10] Aline Dresch, Daniel Pacheco Lacerda, and José Antônio Valle Antunes. 2015. Design science research. In *Design science research*. Springer, 67–102.
- [11] Mariluz Guenaga, Andoni Eguiluz, Alex Rayón, Asier Núñez, and Elena Quevedo. 2014. A serious game to develop and assess teamwork competency. In *2014 International Symposium on Computers in Education (SIIE)*. IEEE, 183–188.
- [12] Dirk Ifenthaler, Deniz Eseryel, and Xun Ge. 2012. Assessment for game-based learning. In *Assessment in game-based learning*. Springer, 1–8.
- [13] Timothy A Judge, Ronald F Piccolo, and Remus Ilies. 2004. The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of applied psychology* 89, 1 (2004), 36.
- [14] Muhammad Saqib Khan, Irfanullah Khan, Qamar Afaq Qureshi, Hafiz Muhammad Ismail, Hamid Rauf, Abdul Latif, and Muhammad Tahir. 2015. The styles of leadership: A critical review. *Public Policy and Administration Research* 5, 3 (2015), 87–92.
- [15] Larry Lashway. 1998. Measuring Leadership. *Research Roundup* 14, 2 (1998).
- [16] Mauricio Capobianco Lopes, Francisco AP Fialho, Cristiano JCA Cunha, and Sofia Inês Niveiros. 2013. Business games for leadership development: A systematic review. *Simulation & Gaming* 44, 4 (2013), 523–543.
- [17] Jo Ann Mazzarella and Thomas Grundy. 1989. Portrait of a Leader. (1989).
- [18] Konstantin Mitgutsch and Narda Alvarado. 2012. Purposeful by design? A serious game design assessment framework. In *Proceedings of the International Conference on the foundations of digital games*. 121–128.
- [19] Elizabeth D Rosenman, Jonathan S Ilgen, Jamie R Shandro, Amy L Harper, and Rosemarie Fernandez. 2015. A systematic review of tools used to assess team leadership in health care action teams. *Academic Medicine* 90, 10 (2015), 1408–1422.
- [20] Andrés Salas-Vallina, Cristina Simone, and Rafael Fernández-Guerrero. 2020. The human side of leadership: Inspirational leadership effects on follower characteristics and happiness at work (HAW). *Journal of Business Research* 107 (2020), 162–171.
- [21] Tarja Susi, Mikael Johannesson, and Per Backlund. 2007. Serious games: An overview. (2007), 1–13.
- [22] Philip Tan. 2010. Iterative game design in education. *International Journal of Arts and Technology* 3, 1 (2010), 118–123.
- [23] Katie Salen Tekinbas and Eric Zimmerman. 2003. *Rules of play: Game design fundamentals*. MIT press.

Appendices

A APPENDIX A

Group 1 - Prototype 1

Codename: _____

Leadership Experiment – Prototype 1

Post Experiment Survey

GROUP: 1

PROTOTYPE: 1

AGENT CODENAME: _____

For each of the following behaviour, indicate if you think someone in your team showcased such behaviour, and if yes, specify who you think was showing the behaviour the most by writing their **codename** in the adequate space.

I believe someone displayed confidence in their choices, themselves, the team, or a team member.	<input type="radio"/> Yes <input type="radio"/> No
According to you, who was showing that behaviour the most? Leave that box empty if you believe no one did.	
I believe someone made effort on keeping communication open in the team. They tried to encourage the team to communicate and tried to communicate with all team members.	<input type="radio"/> Yes <input type="radio"/> No
According to you, who was showing that behaviour the most? Leave that box empty if you believe no one did.	
I believe someone took the initiative to delegate and or make decisions for the team or team members specifically.	<input type="radio"/> Yes <input type="radio"/> No
According to you, who was showing that behaviour the most? Leave that box empty if you believe no one did.	
I believe someone was able to manage the stress or tensions in the team, and still be cohesive in their actions or interactions with others.	<input type="radio"/> Yes <input type="radio"/> No
According to you, who was showing that behaviour the most? Leave that box empty if you believe no one did.	
I believe someone made the effort to address issues. if there was a problem, they did mention it or tried to resolve the issue. (E.g.: Some tensions arise between 2 members; they will point it out and try to handle it)	<input type="radio"/> Yes <input type="radio"/> No
According to you, who was showing that behaviour the most? Leave that box empty if you believe no one did.	
I believe someone made the effort to provide objective feedback to me, someone else or the team in general, during the game.	<input type="radio"/> Yes <input type="radio"/> No

Group 1 - Prototype 1

Codename: _____

According to you, who was showing that behaviour the most? Leave that box empty if you believe no one did.	
I believe someone was finding a new way to solve the problem or deal with an obstacle during the game.	<input type="radio"/> Yes <input type="radio"/> No
According to you, who was showing that behaviour the most? Leave that box empty if you believe no one did.	
I believe someone was rewarding the accomplishment of a team member or the team in general, through encouragement and verbal praise (E.g., "Well done!")	<input type="radio"/> Yes <input type="radio"/> No
According to you, who was showing that behaviour the most? Leave that box empty if you believe no one did.	

Enter your codename

Rankings

A few behaviours will be described. You will have to rank your team members from best to worst, including yourself. If you are hesitating, simply pick a choice that seems to make the most sense according to the behaviour. Each ranking is composed of the following choices: 1st (Being the best place), 2nd place (being the second-best place), 3rd place (being the second-worst place), 4th place (Being the worst place). Ranking someone in 4th place does not mean per se that they are bad at showcasing a speciHc behaviour, but rather that the other members showed that behaviour more. Same thing for ranking someone as 1st place, it does not mean per se that they are good at a speciHc behaviour, but rather that the other members showed that behaviour less.

Rank from Best (1) to worst (4) the players that were good leaders during this game.

	Rhea	Cronos	Gaia	Crius
1st	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2nd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3rd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4th	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Rank from Best (1) to worst (4) the players that displayed confidence in their choices, themselves, the team or a team member.

	Rhea	Cronos	Gaia	Crius
1st	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2nd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3rd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4th	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Rank from Best (1) to worst (4) the players that made effort on keeping communication open in the team. They tried to encourage the team to communicate and tried to communicate with all team members.

	Rhea	Cronos	Gaia	Crius
1st	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2nd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3rd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4th	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Rank from Best (1) to worst (4) the players that took initiative to delegate and / or make decisions for the team or for team members.

	Rhea	Cronos	Gaia	Crius
1st	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2nd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3rd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4th	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Rank from Best (1) to worst (4) the players that were able to manage stress or tensions in the team, and still be cohesive in their actions or interactions with others.

	Rhea	Cronos	Gaia	Crius
1st	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2nd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3rd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4th	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Rank from Best (1) to worst (4) the players that made effort to address issues. If there was a problem, they did mention it or tried to resolve the issue. (E.g. Tensions arise between 2 members, they will point it out and try to handle it)

	Rhea	Cronos	Gaia	Crius
1st	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2nd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3rd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4th	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

B APPENDIX B

Anastasia Thesis: Prototype 2 Post-Experiment Survey

Rank from Best (1) to worst (4) the players that were finding a new way to solve the problem and deal with an obstacle during the game.

	Rhea	Cronos	Gaia	Crius
1st	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2nd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3rd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4th	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Rank from Best (1) to worst (4) the players that were rewarding the accomplishment of a team member or the team in general, through encouragement and/or verbal praise. (E.g. "Well done!")

	Rhea	Cronos	Gaia	Crius
1st	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2nd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3rd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4th	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

C APPENDIX C

Aspects + Games	Inspire / Influence / Guide	Cooperation	Group communication	Empathy	Team-building / cohesion	Problem finding	Problem solving	Creativity	Adaptability	Strategy	Confidence	Accountability	Risk taking	Focus	Organisation &/or Quality	Stress management	Meta-Gains
Escape games	x	x	x		x	x	x	x	x	x			x	x	x	x	x
MOBA games		x	x		x		x		x	x				x	x	x	
FPG team games		x	x		x		x	x	x	x				x	x	x	
Pass the Hoop	x				x?		x	x?									
Manoeuvre the minifield							x										
Partner stand up		x			x		x										
Tallest tower					x			x		x			x	x	x		
Werewolf	x	x	x	x	x			x	x	x	x	x	x	x	x	x?	x
Resistance (Board game)	x	x	x	x	x			x	x	x	x	x	x	x	x		x
Magic Carpet		x	x		x		x	x	x	x			x				
Who ya gonna call		x															
Ducks in a row		x	x														
Move the Egg		x	x		x		x	x	x	x			x	x	x		
Lead the blindfolded		x	x				x			x						x	
Brainstorming for change	x		x		x		x	x?									
Buckets and balls		x	x		x		x	x		x					x	x	
Team jigsaw		x					x										
Team guesswork		x	x			x		x	x					x?	x		x?
Fort Boyard End Game	x	x	x		x			x	x	x				x	x	x	x
Keep talking and nobody explodes		x	x				x?	x									
Team sports (Captain)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		
MMORPG Guilds									x	x	x				x		
Management	x	x	x	x	x				x	x							x
MMORPG Raids and Dungeons					x				x					x	x	x	
It takes two		x	x		x	x	x	x	x	x							
Phantomophobia	x	x	x		x	x	x			x	x		x			x	
Ludicrous Pandemic (Board Game)		x	x		x		x			x							
Tabletop RPG	x	x	x	x	x	x?	x	x	x	x	x				x	x	x
Minecraft (Team Survival)					x					x			x	x	x		
Don't Starve Together		x	x		x		x	x	x	x			x	x	x	x	x
Team PVP Games		x			x		x	x	x	x				x	x	x	
Fort Boyard Guided games	x		x	x		x	x	x	x		x			x			