Matching the effects of immersion with game mechanics in a cooperative game

Rúben de Carvalho Nogueira Vines

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

Supervisor: Prof. Carlos António Roque Martinho

Examination Committee

Chairperson: Prof. Luís Manuel Antunes Veiga
Supervisor: Prof. Carlos António Roque Martinho
Member of the Committee: Prof. Hugo Miguel Aleixo Albuquerque Nicolau

May 2019
Acknowledgments

Firstly and deservedly, a huge thankful note to Prof. Carlos Martinho: for supporting me and this project; for believing in Dark Things About, and pushing every possibility in turning it into reality; and for the eventual insomnia I may have caused - I apologize. The availability and encouragement in every step of the way were absolutely unconditional, this truly would not have been possible without his support, thank you so very much.

To my mother and father, both loving role models whom I owe everything. My sister, who makes life feel so uncomplicated and worth living. To my grandma, a star I have the privilege of admiring everyday. Thank you all for making me who I am.

I also thank the people who can bring joy for nothing in return. The people who made life in college bearable, and the ones that confirm that good friends are for any occasion: Diogo, Gonçalo, Pedro and André. To the guys at Dark Things About, not forgetting the ones who came and go. I’m truly excited to see what the future brings us. Wish you all the best.

I usually say that there is no good time for a bad thing to happen. Luckily with you I also found out that anytime is perfect for the right person to come. You make me feel so loved Beatriz, and I love you just as much.
Abstract

In current times we grant questionnaires as the standardized way of measuring several experience-driven feelings. The same applies for the likings of immersion. However, taking a questionnaire is usually a tedious and leisurely task that test subjects usually show aversion or discontent in performing. Therefore, in this work, we explored if a complementary or alternative tool could be created to measure immersion.

We hypothesized that the actions a player takes when playing a game can be inline with their reported immersive feeling. In case this revealed to be true, we could now have an in-game implementation that would have measuring capabilities comparable with a questionnaire, while also providing several other benefits. For instance, we could now withdraw the dull task of answering a questionnaire; give game developers detailed information on how their game mechanics were affecting the immersive experience; and lastly, open up run-time video game adaption possibilities.

A test was conducted to analyze the correlation between some pre-established game mechanics and actions - metrics captured by our implemented model - i.e. with immersion values reported via questionnaire. Pairs of player (n = 20) were asked to play the first level of the cooperative game Dark Things About. The most promising correlation value with immersion (W=-0.352, p-value=0.128) happened with the in-game distance covered, using a Spearman’s rho correlation test. This result is inconclusive when trying to establish a connection with immersion, but the used methodology suggests a new manner of analyzing gameplay data, opening a guideline in possible future approaches.

Keywords

Immersion; Game development; Game mechanics; In-game metrics; Cooperation
Resumo

Nos tempos que correm damos por garantido o uso de questionários como a maneira padrão para medir sentimentos que ocorrem aquando de experiências. Fazemo-lo também para o senso de imersão. No entanto, preencher esse tipo de questionários é normalmente uma tarefa desinteressante ou fastidiosa, vista com desinteresse pelos sujeitos de teste. Desta forma criámos um estudo focado em encontrar uma ferramenta complementar ou alternativa para medir imersão.

Tomámos como hipótese que as acções efetuadas por um jogador quando joga um jogo podem estar alinhadas à sensação de imersão que estes relatam. No caso de tal ser corroborado, poderíamos então criar uma implementação dentro de um jogo de modo a ter capacidades de medição comparáveis às de um questionário, para além outras vantagens adicionais. Por exemplo, poderíamos remover a enfadonha tarefa de resposta a um questionário; oferecer a desenvolvedores de jogos informações mais detalhadas acerca de como as mecânicas dos seus jogos influenciam a experiência de imersão; e por fim, abrir portas a possibilidades de adaptação do jogo em tempo-real.

Conduzimos um teste para analisar a correlação entre algumas mecânicas e acções do jogo previamente estabelecidas (métricas) com os valores de imersão capturados via questionário. Pares de jogadores (n = 20) foram solicitados para jogar o primeiro nível do jogo cooperativo Dark Things About. O valor de correlação com imersão mais promissor (W=-0.352, p-value=0.128) aconteceu com a distância percorrida dentro do jogo, usando um teste de correlação do ró de Spearman. O resultado mostra-se inconclusivo ao tentar obter uma ligação com imersão, mas a metodologia usada sugere uma nova maneira de analisar dados durante uma sessão de jogo, possibilitando futuras abordagens.

Palavras Chave

Imersão; Desenvolvimento de jogos; Mecânicas de jogo; Métricas de jogo; Cooperação
# Contents

1 Introduction ............................................. 2
   1.1 Motivation ........................................ 3
   1.2 Problem .......................................... 4
   1.3 Hypothesis ........................................ 5
   1.4 Contributions ..................................... 5
   1.5 Document Outline .................................. 6

2 Related Work ........................................... 8
   2.1 Social Play & Social Presence ..................... 9
   2.2 Immersion, Flow & Presence ....................... 9
   2.3 Experience Evaluation ............................. 12
   2.4 Discussion ....................................... 14

3 Case study ............................................. 16
   3.1 Dark Things About ................................ 17
      3.1.1 Game characteristics .......................... 17
      3.1.2 Reasons for DTA ............................. 17
      3.1.3 Game Mechanics ............................. 18
   3.2 Game Scenario ..................................... 20
   3.3 Game Analytics ................................... 23
      3.3.1 Data Gathering .............................. 23
      3.3.2 Data Treatment .............................. 25

4 Evaluation ............................................ 27
   4.1 Preliminary Tests .................................. 29
      4.1.1 Social Presence Variation ................... 29
         4.1.1.A Sample .................................. 29
         4.1.1.B Measurement ............................. 31
         4.1.1.C Procedure ............................... 32
         4.1.1.D Results & Discussion ................... 32
List of Figures

3.1 Medic and Priest side-by-side ....................................................... 18
3.2 Players interchanging lantern with camera ........................................ 18
3.3 Lantern player facing enemy .......................................................... 19
3.4 Camera player perspective mechanic ................................................. 19
3.5 Player collecting a fuel flask ........................................................... 20
3.6 Village sky overview with interaction labels ....................................... 21
3.7 House indoor plant .......................................................................... 22
3.8 Activity diagram for the run-time data gathering process .......................... 24
3.9 Logging of an action ........................................................................ 25

4.1 Lisboa Games Week test setting ......................................................... 30
4.2 Game related questions - Part 1 ......................................................... 30
4.3 Game related questions - Part 2 ......................................................... 31
4.4 Conditions of preliminary test .......................................................... 35
4.5 Game and level design changes .......................................................... 36
4.6 Game related questions - Final Test/ Part 1 .......................................... 37
4.7 Game related questions - Final Test/ Part 2 .......................................... 38
4.8 Dispersion line between Immersion and Distance Covered ....................... 40
4.9 House blueprint and scale ............................................................... 41
4.10 Forest overview and scale ............................................................... 42
4.11 Dispersion of data on metric-metric correlation ..................................... 44
List of Tables

4.1 Social Presence reported by test subjects ........................................... 32
4.2 Spearman's Rho correlation between in-game metrics ............................. 43
Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTA</td>
<td>Dark Things About</td>
</tr>
<tr>
<td>IEQ</td>
<td>Immersion Experience Questionnaire</td>
</tr>
<tr>
<td>SPGQ</td>
<td>Social Presence in Gaming Questionnaire</td>
</tr>
</tbody>
</table>
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Motivation</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Problem</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Hypothesis</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Contributions</td>
<td>5</td>
</tr>
<tr>
<td>1.5 Document Outline</td>
<td>6</td>
</tr>
</tbody>
</table>
1.1 Motivation

Since the appearance of video games there has been the need, as creators, to express ourselves and appeal to audiences in this type of media. Just like it is has been done long before in other non-computerized ways. When evidencing the existing amounts of different kinds of games, we are also evidencing the appeal of each individual, as players, in searching and fulfilling their own needs. Therefore, video games differ enough to suit several personalities or motivations (socializers, explorers, etc.), levels of skill or dexterity (point-and-click and controller-driven games), technological requirements towards realism (2D, 3D, VR), genre preferences and many other personal characteristics.

With the computerized representation of games there was the arising of new manners of perceiving and feeling towards this media. For instances, in the turn of the millennium the concept of immersion came into play. Douglas and Hargadon - in The Pleasure Principle: Immersion, Engagement, Flow - introduced this affective experience into the gaming world. They highlight the interactivity of the interface and the “ability to take situated action and to savor the results [...]”. In their eyes “interactive games fulfill their promise as immersive when they offer us [...] interactions enabling us to enjoy virtual experiences either unattractively risky or denied to us in everyday life” [1].

Immersion became an important keyword both in the scientific community and the marketing side of the gaming industry. Its meaning and effect is not the same between authors and has evolved in the last two decades. However, even though unclear, it is seen as fundamental part of the experience by either developers or players.

In this project immersion shall be interpreted as the cognitive experience of a player being absorbed within the game, unaware of their surroundings and incapable of perceiving the passing of time [2]. Being such an intense point of player experience, there are several properties within a game that can be altered to create a different experience, therefore and perchance, towards an overall different immersive one. There has been an extensive amount of research that explores this latter point: affecting (and enhancing) the player’s experience. Importance of the HUD [3], intuitiveness and familiarity with controllers [4–6], impact of sound and music [7, 8] or the social setting while playing [9, 10] are just some of the many paths of study player experience has been studied upon.

One common misconception when talking about immersion is that a game with such characteristics is trying to replace a person’s reality. It is, however, making a good job in complementing it. Frank Lantz, in GDC 2005\(^1\), rejects the idealism and expectations that game development should keep “evolving towards an infinitely detailed and utterly seamless simulation [...] which is indistinguishable from direct experience”. This high-mindedness is known as the “Immersive Fallacy” [11]. Game’s should not try to feel in a gap between reality and its representation, otherwise and exemplifying, players would feel

\(^1\)https://youtu.be/6JzNt1bSk_U - Frank Lantz’s GDC 2005 presentation
compelled to run away from dangerous [in-game] situations instead of feeling excitement in facing them\textsuperscript{2}.

This line of thought became really noticeable with the arrival of VR. The argument arose that feeling immersed was then so much simpler just by the fact of wearing the respective ocular set and being surrounded by another world. But without engaging gameplay, fun mechanics, interesting story line or so many other vital characteristics that can hook a player to a game, the feeling of immersion may never be experienced.

One way to have a better understanding of player and player-base preferences, is knowing what players do and like, so we can strive for a better experience or overall product. This knowledge is supported by metrics that are recovered from the players as individuals or as a whole - i.e. player analytics. In the industry, we can see this being taken advantage either financially, development-wise or gameplay-wise. For instance, an extremely popular tool that game developers take advantage of are analytic systems for mobile game monetization. Several metrics regarding games’ player-activity and financial measures can be drawn: e.g. ‘daily active users’, ‘cost per installation’ or ‘non-paying to paying player conversion rates’.

Metrics can also be captured in order to give developers a tool to assess which content in their own games can provide useful information either to make better gameplay or marketing decisions. Google’s digital video game distributor - Google Play - does not provide a pre-integrated platform for this purpose, and suggests instead a more customized implementation of the retrieval system\textsuperscript{3}. And so, appropriate metrics for a given game (e.g. player actions and game events) can be drawn and stored in order to make that assessment.

Even though most developers would argue their intent is to create the most immersive experience possible, their pursue is probably empirical and unaligned with their intentions. So what holds a developer from purposely pursuing a more immersive game? Maybe the ambiguity of the term; the standardized static way of measuring it; or the lack of tools to use player analytics to adapt it into a more immersive one. This project means to address these potential reasons.

\subsection*{1.2 Problem}

Even though questionnaires may take some time to be administered, they are seen as the most standard and validated way of measuring effects such as immersion. Their usage comes usually after user-testing sessions which may be inconvenient for two main reasons: it extends the session’s duration, and it is performed after the highest levels of engagement between the player and the game are over. These processes are usually tedious and dull, making this very important tool for developers unappealing to

\textsuperscript{2}https://youtu.be/Xnxs1mrMV6o - Video with views on immersion and the immersion fallacy

\textsuperscript{3}https://cloud.google.com/solutions/mobile/mobile-gaming-analysis-telemetry - Google’s analytics implementation methodology
participants.

With the increase in popularity of real-time video game adaptation, it is more and more common to retrieve players’ in-game inputs in order to dynamically change, place or generate content to better suit their needs. For that reason, a dynamic method of measuring effects like immersion could potentially be equated and used in such methodology. Adapting a player’s immersion to create the best possible experience. However, since the current method is purely static it obviously does not take into account for this potential. Therefore, and perceiving questionnaires as rather impractical and outdated tool, can we consider another way of measuring immersion?

1.3 Hypothesis

Since immersion is a feeling that is present while playing a game, it is reasonable to think that the way we feel towards the game we are playing can be caused by the way we play the said game. Either by being more or less effective, more or less goal-aligned, having more or less skill, and many other performing qualities and metrics of a given player when playing a specific game.

Metrics are quantifiable actions that occur recurrently throughout the gameplay. For this particular project and game, some of these actions can be: successful and unsuccessful interactions, number of deaths, distance covered or distance between both players. With a cooperative two-player game at hand named Dark Things About we hypothesize that:

- The effects of immersion can be measured via in-game metrics captured by a given player’s actions;
- Some of these metrics can correlate, and admittedly cause, the immersion values obtained by a standard questionnaire (particularly Immersion Experience Questionnaire (IEQ)).

1.4 Contributions

As many other works in this field of study tend to prove: there is a high variance in the terminology of immersion. From the previously mentioned pursue for a hyper-realistic simulation (immersive fallacy) to perceiving immersion to be aligned with engagement and skill (inline with the concept of flow - to be discussed in a later chapter), comes the first contribution from this project: an overview of the literature regarding the keywords of immersion, flow and presence. Narrowing down these large amounts of concepts into stricter definitions that fits both this and future projects.

After doing so, we want to address the static nature of questionnaires. Firstly, with their state of the art, and then, capturing their purpose and usage methodologies. This way we can endorse our
process in order to obtain a measuring tool that goes inline with the standards of other validated questionnaires.

However, doing so is dependent on the main issue of our project: trying to attain an answer to the existence of correlation between immersion and in-game metrics. Understandably, these metrics are not directly assignable to other games, and only make sense in the context of the one here used. Therefore this project will try to argue in favor of analyzing and perceiving what makes a piece of data a valuable metric, so that we possibly provide an overall methodology adaptable to another game’s mechanics.

Potentially, matching immersion with in-game player actions can have a very impactful outcomes in game development. It can provide developers enough tools to: reduce user-testing session times; have informative measures in what can be more immersive in a given game; or ultimately, open the possibility of real-time immersion capture for video game adaptation - i.e. turn an experience more immersive when a player is failing to feel accordingly. However, in case our project does not achieve convincing positive results towards the correlation of metrics and immersion, we will still attempt to provide considerations regarding the obtained metric’s information, giving designers the possibility of understanding the impact and importance of in-game actions to the game experience quality in a more general fashion.

Lastly, working in Dark Things About alongside with this project, we can take this testbed game further in its development state, implement player analytics recording systems and, hopefully, use the acquired conclusions to create the most immersive experience possible for DTA.

1.5 Document Outline

For the sake of marketing, many keywords are tossed away lightly (even if completely misused) to capture interest. Being immersed differs from feeling presence or enjoying a game. Extensive research has been done in order to clarify and create clear bounds between these terms; practical experimentations have tried to prove the feasibility of achieving these states of mind and engagement; some keywords have been adapted from others areas of study into video games; questionnaires have been made and validated to measure whether or not these conditions have been accomplished within a game session.

The second chapter of this document (i.e. Related Work) reports some of the investigation, theory, experimentation and questionnaires for the fundamental concepts of this study, in particular: immersion, flow, presence, social play and social presence. The inclusion of the latest two keywords are based of the cooperative nature of the test bed game.

Case Study, the third chapter, illustrates the approach to test our hypothesis. Firstly, with a detailed explanation of the cooperative game in development: Dark Things About, its background, gameplay
characteristics, game mechanics and reasoning behind this choice. Followed by the assembled in-game scenario that was presented to our participants in the testing phases of our project. To finalize this chapter we present the data gathering implementations and reasoning behind them. From the in-game input retrieval, to the external data parsing and our view in what constitutes a metric.

Afterwards, on the Evaluation chapter, we minutely display the practicality of our experiments. Going from the reasons to proceed with some preliminary trials; description of the events before, during and after every test, and respective considerations; lastly, the obtained test results and mild reflections on them. The final conclusions have a chapter of their own, concluding this way this document’s outline.
# 2 Related Work

## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Social Play &amp; Social Presence</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Immersion, Flow &amp; Presence</td>
<td>9</td>
</tr>
<tr>
<td>2.3 Experience Evaluation</td>
<td>12</td>
</tr>
<tr>
<td>2.4 Discussion</td>
<td>14</td>
</tr>
</tbody>
</table>
2.1 Social Play & Social Presence

The expression social play is currently used as a definition and not as a matter of study in the gaming scientific community. However, in the theme of child behaviour and social activities this is not the case. For instances, Frost et al. - based on Parten's work [12] - break this concept into six categories: unoccupied behaviour, onlooker behaviour, solitary play, parallel play, associative play, cooperative play [13].

Yet, and as previously stated, in the gaming community and when usually investigating test subjects of a higher age range this is taken much more lightly. In the research referenced below, social play is seen as playing in the company of someone. This type of play can happen by playing a game with or against someone, surrounded by a crowd or any other type of way that can simulate this interaction (e.g. via internet). To see if this simulated way of interaction produces a comparable effect to the real-life interaction, the levels of social presence are measured.

Biocca and Charms created a questionnaire - Networked Minds Measure of Social Presence [14] - for this exact purpose: understanding one's sense of social presence when interacting with others, in this case, in a mediated way. This and other questionnaires are going to be further analyzed later in this chapter. However, it is important to note the that all these serve as a basis for interaction evaluation in multiple scientific researches (for instances, Leite in 2013 [15]).

In Guide to the Networked Minds Social Presence Inventory - also by the authors of the last mentioned questionnaire - social presence is seen as:

"the degree to which users of a medium feel that mediated others are spatially co-present, psychologically accessible, and behaviorally interactive." [16]

Additionally, Cairns et al. define recognizes social presence to happen "when there are actions that have social meaning [...] or social actors who can affect your mood and your actions as a response to the social situation" [10].

The previous quote is referenced from Who but not where: The effect of social play on immersion in digital games - Cairn et al.'s endeavor on understanding the influence of social presence on immersion (as a result of social play). This research was influenced by Gajadhar et al.'s earlier effort on understanding the same [9]. Results showed that higher social presence (by either playing online or in co-location) meant higher immersion. Not only so, but restating Ravaja's [17] claims, playing with friends (as opposed to strangers) resolves in an overall better game experience.

2.2 Immersion, Flow & Presence

Gordon Calleja¹ states that regarding virtual environment habitation (being “in there”) people tend to define it from a spectrum of definitions that can go from “the deeply engaging to the hopelessly addictive”.

tive, or any shades inbetween”. Amongst that spectrum, *immersion* and *presence* usually come along, sometimes separately, many times interchangeably. For disambiguation “the game’s world either being games research, game’s industry or general gaming population [...] uses immersion as the preferred term”.

*In-Game: from immersion to incorporation* [18] is Calleja’s contribution dedicated to the exploration of all headings of many ambiguous terms used for engagement by the gaming and scientific community. For instances, a strong argument against Slater’s view on presence and immersion being necessary to each other was held:

“Suppose you shut your eyes and try out someone’s quadraphonic sound system which is playing some music. “Wow!” you say, “that’s just like being in the theatre where the orchestra is playing.” That statement is a sign of presence. You then go on to say, “But the music is really uninteresting and after a few moments my mind started to drift and I lost interest.” That second statement is nothing to do with presence. You would not conclude, because the music is uninteresting, that you did not have the illusion of being in the theatre listening to the orchestra. The first statement is about form. The second statement is about content. Presence is about form, the extent to which the unification of simulated sensory data and perceptual processing produces a coherent “place” that you are “in” and in which there may be the potential for you to act. The second statement is about content. A [virtual environment] system can be highly presence inducing, and yet have a really uninteresting, uninvolving content (just like many aspects of real life!).” [19]

This look on presence as something that comes from high-fidelity mediation is pointed out as a flaw by Calleja. An example is given where someone who has never heard an orchestra or been to a theatre before wouldn’t feel present in the ambient created by the sound system because there is nothing familiar that the said person can be transported to. Therefore in his eyes high-fidelity systems are not the cause of presence but representing an ambient with fidelity may be a big step in feeling present. This way, the feeling of *immersion comes from a continuous stream of stimuli* and can’t be turned into an “on and off switch” as Slater seems to state.

Even though Calleja’s work is important and extensive and can provide an insight of the inconsistencies existent in this types of studies, it is far from becoming a staple in the scientific community as a converging line of work. From the early studies until now there has been discrepant approaches from a broad amount of work researchers, which means that many studies are not usually in-line with each other concerning terminology. Therefore, a look in retrospect will be made and then a selection of the terms that seem to suit better our needs will be discussed.

**Flow** is probably the least misinterpreted term out of the ones associated with enjoyment. Mainly because Csikszentmihalyi’s early endeavours [20,21] gave foundation to many other researches. Here, described as the process of optimal experience, flow, is the “**state in which individuals are so involved**
in an activity that nothing else seems to matter” [21]. In The evolving self: a psychology for the third millennium, Csikszentmihalyi describes what he thinks are the eight dimensions of the flow experience:

1. **Clear distinctive objectives** or goals and immediate feedback: one knows instantly how one is doing;
2. High opportunities for acting matched by one’s perceived ability to act (i.e. personal **skills match given challenge**);
3. **Actions and awareness merge**: one-pointedness of mind;
4. **Absolute concentration** on task at hand; irrelevant stimuli, worries and concerns are temporarily blocked or suspended;
5. **Sense of potential control**;
6. **Loss of consciousness** and transcendence of ego;
7. **Altered sense of time**, usually seems to pass faster;
8. Experience becomes autotelic: **action conducted completely from inside** (but only when several of the previous conditions are present).

These dimensions are be reported by race motorcyclists, surgeons, basketball players, chess masters and others [22]. Being transversal to so many areas, flow was naturally carried into the gaming world. This gaming experience was named by Sweetser et al. as **GameFlow** [23].

Brown & Cairns, in a user study, proposed immersion to be three-layered state where a person needs to break consecutively barriers of immersion to achieve higher layers or levels. Firstly comes **engagement** which is “the lowest level of involvement [...] and must occur before any other level. To lower the barriers to enter this level, the gamer needs to invest time, effort and attention”. Then it comes **engrossment**, where a player needs to “become further involved [...] and become engrossed”. Specifically, when a player feels engrossed its “emotions are directly affected by the game”. Lastly, and more controversially, comes **total immersion** which the authors describe as being the same as presence. In this state “participants described being cut off from reality and detachment to such an extent that the game was all that mattered” [24].

In the research’s discussion section several points on immersion can be noted. Based on the users experiences immersion is a “shared concept but [...] not a static experience”, also “immersion was not necessary [...] for enjoyment” however “no one described an experience of immersion that they did not enjoy”.

In-line with other authors - Radford’s work on entering a game through its controls [25], Winograd’s on accessment for engagement [26], and Norman’s on software usability [27] - Brown & Cairns work
states that “enjoyment through immersion is not possible if there are usability and control problems”, therefore, “there needs to be an invisibility of controls for total immersion to take place” [24]. More importantly, his work makes the bridge between being highly immersed and Csikszentmihalyi’s concepts of flow.

This bridge between immersion and flow is also explored by Jennett et al.. In a study that hypothesizes several ways of measuring immersion (to be discussed later), Jennett et al. characterizes immersion as lack of awareness of time, loss of awareness of the real world, involvement and a sense of being in the task environment [2].

Cairns et al. takes this latter approach and extends immersion to “a cognitive experience wherein thoughts of the player are wholly absorbed in the action within the game[...]”. Once again, they make a correspondence between immersion and flow. This time, in their understanding “flow corresponds to the extreme end of immersion where a person is so immersed [...] they enter a flow state” [10].

Lastly, presence needs to be addressed. Lombard and Ditton separated the concepts of spatial and social presence as a way of trying to break down this complex and controversial term. Presence is a way to address the feeling people have when interacting with media. Being present in the medium means having the illusion of non-mediation when experiencing mediated environments (i.e. losing the perception of media interaction when that is happening in the actuality) [28].

Spatial presence, denominated by Lombard et al., and further analyzed by Wirth [29] is “the sense of being physically located within a virtual environment”. The same meaning goes for games, this is noted when the game “gives them ability to navigate through the world” having “an effect on the game world” and using game objects “the way real objects would [be used]” [10]. Social presence, as previously explained, is about the social meanings of actions.

Finally, Floridi suggests a slightly different view on presence. Instead of not feeling mediated when interacting with media, presence should be having the ability to achieve an observable or exploratory effect when interacting with media [30].

## 2.3 Experience Evaluation

This chapter serves the purpose of understanding the intents, aims and potentials of what the literature uses as measuring tools for the gaming and social experiences we will strive to achieve.

Biocca and Harms created Networked Minds Measure of Social Presence not as a questionnaire designed for games but instead as a more general way of assessing the user’s sense of social presence [14]. Social presence in this questionnaire was conceptualized as a combination of six different sub-dimensions:
1. **Co-presence**: degree in which the observer does not believe is alone or secluded;

2. **Attentional allocation**: amount of attention user allocates and receives from an entity;

3. **Perceived message understanding**: ability to understand a message from an entity and perceive the level of understanding of that entity;

4. **Perceived affective understanding**: ability to understand the emotional state of an entity and perceive the understanding of that entity;

5. **Perceived affective interdependence**: extent to which the user’s emotional and attitudinal state affects and is affected by the emotional and attitudinal states of an entity;

6. **Perceived behavioural interdependence**: extent to which user’s behaviour affects and is affected by the behaviour of an entity.

**Social Presence in Gaming Questionnaire** (SPGQ) also regards social presence, but this time with several applications in experiments with video games. This questionnaire designed by Kort et al. consists of 17 items scored on a Likert scale (0 to 4) [31]. Those items are unevenly dispersed in three sub-scales:

1. **Psychological Involvement - Empathy** (6 items): sense of feeling to be in the same enjoyable situation as the other players;

2. **Psychological Involvement - Negative feelings** (5 items): extent to which players were competitive or aggressive towards other players;

3. **Behavioural Engagement** (6 items): degree of feeling that the other players were influencing the game.

**Game Experience Questionnaire** (GEQ) the questionnaire used by Gajadhar et al. [9] and refuted by Cairns et al. [10], is a very extensive questionnaire that serves the purpose of measuring several aspects of the game experience. The questionnaire is divided in three modules that serve to address the gaming experience as a whole, each using a 5-point Likert scale [32]. The said modules are:

1. **Core module**: assesses the game experience as scores in immersion, flow, competence, positive and negative affect, tension and challenge;

2. **Social presence module**: investigates psychological and behavioural involvement of the player with other social entities; assesses scores in empathy, negative feelings and behavioural involvement (should only be administered when one of this entities is involved in the game);
3. **Post-game module**: perceives how players feel after they stopped playing; assesses scores in positive experience, negative experience, tiredness and returning to reality.

Finally, the **Immersive Experience Questionnaire** (IEQ), designed by Jennett et al., is a 31 question-based tool which utilizes a 5-point Likert scale to score immersion in a range from 31 to 155 points [2]. Six particular items in this questionnaire are valued by the symmetrical answered value (subtracting '6' to the value). It is composed by five different components with an uneven amount of items:

1. **Cognitive Involvement** (9 items): experience of focusing on the game;
2. **Emotional Involvement** (6 items): strength of feelings experienced while playing;
3. **Real world dissociation** (7 items): sense of losing awareness of the world around and increasing awareness of the game;
4. **Challenge** (4 items): experience of being challenged by the game;
5. **Control** (5 items): extent to which player felt in control whilst playing.

Conventionally, either handing out a questionnaire or performing an interview is the way to measure a player’s experience. However, sometimes other methods need to come into play. For instances, the way Jennett et al. validated their study was with additional sets of methodologies. Those methodologies were essentially three uncommon scenarios performed by test subjects and assessed by the authors. In the first scenario a user would have to conclude a control task (resolving a tangram), followed by a randomly assigned immersive task (play a video game) or non-immersive task (perform a button clicking activity), and then, repeat the control task once again. This showed that subjects that went through the immersive task would take longer to conclude the last control chore because their minds would still be “stuck” in the virtual world. Second and third conditions were related with eye-tracking and changing the pace of a game, respectively [2].

To conclude, what is important in this process is understanding that there are validated and established ways of measuring a player’s experience, however, this process has to be sometimes unconventional, informal, exploratory or simply iterative.

### 2.4 Discussion

Cairns et al.’s work on immersion is seen by many as a mark in the literature. Thus, this research is going to take into account these authors views, terminology usage, approaches and further details henceforth. More specifically, using or adapting several points into this project: the same game experience measuring questionnaires (IEQ and SPGQ); ‘co-location’ scenario; and accounting for the nature of social connection (i.e. playing with friends).
Regarding terminology, and using the same reasoning of the prior paragraph, **immersion** shall be seen as the lack of awareness of time, loss of awareness of real world and a sense of being in the task environment (Jennett 2008 [2]); **presence** as a way to address the feeling people have when interacting with media (Lombard and Ditton 1997 [28]); **social presence** happens “when there are actions that have social meaning[...] or social actors who can affect your mood and your actions as a response to the social situation”; and finally, a bridge between two terms where **flow** is seen as “the extreme end of immersion where a person is so immersed [...] they enter a flow state.”
3

Case study

Contents

3.1 Dark Things About .................................................. 17
3.2 Game Scenario ....................................................... 20
3.3 Game Analytics ....................................................... 23
This chapter details the practical challenges of this project. From the testbed game mechanics, game engine supporting it, reasoning for its usage and more overall characteristics. Complementing the game overview, we have the designated scenario prepared for players to test our hypothesis. Containing design choices and illustrations for a better understanding of the scene.

Lastly, the implemented metrics. Which information is recovered in-game and what we consider feasible to constitute a metric. Then, the treatment process from recovery to their exporting [to a statistical analysis program].

3.1 Dark Things About

3.1.1 Game characteristics

Dark Things About (DTA) is a survival horror two-player local cooperation game. Its development started in the beginning of 2017 in the Information Systems and Computer Engineering Master’s course with the partnership of Laboratório de Jogos do Instituto Superior Técnico, and has continued in parallel as a side project ever since.

Unreal Engine 4 is the supporting game engine for this game, aided by an audio workstation named Wwise and 3D-modeling tools such as Blender and 3ds Max. All this stated programs are highly compatible with the game engine, hence their choices. Apart from compatibility, this engine seems to be keeping up in the vanguard of engine technology comparing to its free-to-use competitors. Opening up the possibility of importing external plugins and tools with no great issue. The engine uses C++ as a programming language and an alternative visual scripting method called Blueprints.

3.1.2 Reasons for DTA

Using DTA as this project’s basis for scenario testing, offers a few advantages:

1. no need to adapt to another engine, study tools to mod an existing game, or start a testbed game from scratch;

2. time and effort put into creating a better game experience (already happening as a side project’s objective), will also result in a better game experience for the future test subjects of this game;

3. results from this project will be conveniently used in the game (either gameplay or knowledge wise), hopefully creating a better project and game overall.
3.1.3 Game Mechanics

The core mechanic of the game is a compromise that happens by one of the players being attached to the camera (game’s view perspective) while the other player holds a light source. The main idea is to cooperate in such a way that the camera player can keep the light player at focus, while the light player enlightens the path of his partner (represented in Figure 3.1).

![Figure 3.1: Medic and Priest side-by-side](image)

(a) Medic holding the camera on the left. Priest with the lantern on the right
(b) Priest moving freely from the Medic’s perspective

Even though it may seem that these mechanics (camera and light) are assigned to a given player, both are interchangeable at any time (Figure 3.2). Meaning that by being close to one another and willingly clicking a given button at the same time will make the camera player hold the light, and vice-versa.

![Figure 3.2: Players interchanging lantern with camera](image)
Apart from lighting the world, the lantern one of the players holds can also slow down or stun enemies (Figure 3.3). In the case of holding the camera - character in focus camera setting (terminology by Seif El-Nasr et al. [33]) - a player can also see through its character perspective. Which means if a character holding a camera is a medic then medicine related objects will be highlighted in its view; while if it is a priest he will see a more divine version of the world (Figure 3.4).

(a) Enemies affected by light are slowed down  
(b) Enemies hit by a flash of light run away

Figure 3.3: Lantern player facing enemy

(a) Medic sees blood in this particular lever  
(b) Priest is guided by a ‘holy’ light

Figure 3.4: Camera player perspective mechanic
In figures 3.4(b) we see a power up performed by the light player. This is a mechanic that can only be activated by the person holding the light at any given time, if enough fuel flasks are available to the pair. These flasks are consumable items that exist throughout the world that can be picked up by any of the two players (Figure 3.5). Powering up the light uses a flask, increasing the light’s intensity in addition to scaring off enemies in the path.

![Figure 3.5: Player collecting a fuel flask](image)

Even though it was previously mentioned that this is meant to be a horror game, studies were conducted only during the first level (tutorial). In this level, most game mechanics are introduced and a few puzzles are to be cooperatively beaten - aligning much more with an adventure-type game. Most of the existing interactions happen in the same slow paced fashion as previously shown (picking up a flask). The players can interact with several objects in the world (levers, letters, doors, etc.) by standing close enough to them and pressing the assigned button.

Lastly, players can die in this game by either letting their health bar reach a given minimum (by being attacked several times by an enemy); or if they hit or get hit by certain objects (e.g. falling in a river or a trap; get crushed by a spike wall).

In conclusion, in this game both players share importance, capabilities and essential need of cooperation to progress, and should not feel impaired on in disadvantage by their partner’s current role.

### 3.2 Game Scenario

The statistics studies website Statista recorded the adventure video game genre as the 5th most popular in terms of sales\(^1\), while the horror game type genre doesn’t make the top 10 list. Naturally, it is seen more as a niche than the first one. In the presented test scenario, other than slightly darkened or foggy scene no other components should intimidate or throw off players with less likability for this types of games.

Even though, the lack of these characteristics make this level of the game more aligned with a cooperative adventure, the game mechanics are still considerably uncommon comparatively to many other games. Therefore, the players are presented with the tutorial level, so that they can potentially learn and exercise the mechanics while also exploring the scenery.

Prior gameplay testing has proven that the time to beat the level is around the 20 minute mark, which usually is enough to become engaging without becoming exhausting. The two following images (Figures 3.6 and 3.7) show an overview of the map played:

![Figure 3.6: Village sky overview with interaction labels](image_url)

**Figure 3.6:** Village sky overview with interaction labels
In the figure 3.6, on the reachable player ground (delimited with a green shading) we can see several possible interactions a player can go through while completing the level. When starting the game both players spawn next to each other on a village (route indicator 1) and are prompted, via a cutscene, with the immediate task of reaching a house (route indicator 2).

Tutorial points 1, 3 and 4 are connected with the previously mentioned fuel flasks. The first, teaching the players how to collect them; the second demonstrating their use - increasing the intensity of the light; and the last giving a purpose for its use - scaring an enemy that is present in danger zone 1. Tutorial point 2 is a small challenge that helps the players settle in the unusual setting of only one of the players controlling the game’s perspective.

Clickable items 1, 2 and 3 are objects that can be interacted while nearby, serving the purpose of guiding the player to the pretended destination (route indicator 2). Clickable items 4 and 5 (figure 3.7) can be interacted in the same fashion, but in this case they are levers that integrate two puzzles that are to be cooperatively beaten. In here, danger zone 5 concludes the possible places a player can die in this level - in this case, by pressing an incorrect lever.

Most consumable items represented around the map are mostly fuel flasks that can be collected and used. The main exception is consumable item 1 (figure 3.7) that is a key to the destination door (route indicator 3). Interacting with the given door while holding the key concludes the level successfully.
Lastly, the remaining type of interaction existent in this world is the possibility of opening a letter. Doing so happens with the action of one of the players, but can only end when both accept to close the letter.

3.3 Game Analytics

Scrutinizing the game mechanics and scenario had the intent of giving a better understanding about the types and amount of information that is generated and can be captured for each individual player. The following chapter demonstrates how that capture was perceived and achieved - both the run-time data retrieval and post-gameplay data treatment.

3.3.1 Data Gathering

A 3D open-ended world played by two players at the same time can have a near infinite set of actions during gameplay, making each play session unique from one another. Since our hypothesis means to match in-game actions with the values of player immersion, we can’t be sure at first of which actions can be responsible for that. Therefore, we needed to implement a data retrieval system that could gather all the necessary information to create our quantitative measures, while also opening up the possibility for creating future ones with the data extracted. For this reason we divided the capture and transformation processes of the data in two separated steps.

Data gathering process was implemented within the game’s code using Unreal Engine 4 blueprint scripting. The main goal was to capture every action of both players during a session, making it so it would be theoretically possible to reconstruct the entire session after it was over. Following the world’s description, given in the latter sub-chapter, we can divide the possible outcomes of data at a given time in three key types: discrete and without gameplay value (e.g. player pressed a button on a controller); discrete and with gameplay value (e.g. player pressed a button and interacted with a game object - changing the world’s state); and lastly, continuous (e.g. the players’ positions at all times). The following activity diagram (Figure 3.8) envisions the flow of this process:
Figure 3.8: Activity diagram for the run-time data gathering process

From the diagram above we can perceive the ongoing activity that is executed after one of the players presses the start button on the main menu of the game. When doing so, the game fetches the current file number from a given local directory, stores it and immediately creates a file that will save both player positions at every possible timestamp. This file’s naming convention is “Distance_FILEID”, where “FILEID” is an integer correspondent to the amounts of games already existent in that directory. The reasons behind separating this continuous data gather from the remaining has to do with the large amounts of information comparatively generated. On the discrete side every variable is updated in run-time and dispatched to a file when the session ends. As for the continuous (i.e. player positions), the respective variable is printed to its file in accordance with a timer.

“Metric_Report_FILEID” is the file’s naming convention for every other discrete piece of information. Every time an action occurs either in a player controller or in the game world, a correspondent log is added to a preliminary metric value. These values are grouped among homologous actions, where every line for a given group means a new action of the same kind.

An example is given in Figure 3.9, where “Player 2” pulls a lever. When that action happens it is included with other existing ones that are from the same category (“lever interactions”). Along with that action other four key points are captured: when was it performed (“timestamp”), by whom was it performed (“pulled by”), and both players positions at that moment (“p1 xyz” and “p2 xyz”). Retrieving these additional details opens the possibility of creating metrics derived from the primary act.
The data capturing process resolves when the last action on the level is accomplished (i.e. opening up a door). Doing so will result in the two log text files above mentioned (and exemplified in Appendix B). Lastly, a preventive way of printing the files without reaching the end of the level was also implemented with a keyboard input.

### 3.3.2 Data Treatment

Upon finishing a gameplay session two text files would be created containing thousands of line entries. But as previously briefly described, these were not more than preliminary metrics. Most do not constitute what we find to be worthy of representing the gameplay experience on their own.

We conjecture this line of thought on the literature, and argue that since immersion is achieved in the presence of several factors that only make seen for an overall experience, the metrics that compose the experience should also describe it as a whole. We tried to transpose to metrics the actions that could potential either bring cohesion or disruption to a player’s satisfaction. Doing so was based on three factors: whether the measure is based on an important gameplay mechanic; if an action happens usually either constantly or very frequently while playing; and finally, if an action is something players are familiar with upfront - by either being perceptibly intuitive or commonly seen in other games. With this in mind and from the existing data, we created the following quantitative in-game metrics:

- **Deaths**: Number of times a given player died [and respawned];
- **Successful interactions**: Number of times a player pressed the correct input near an object that required that same input;
- **Failed interactions**: Amount of non-fruitful or unresponsive input actions - i.e. when a player pressed a button and a successful interaction didn’t occur;

- **Time as Camera** (ms): Time spent while holding the game’s perspective (one of the main mechanics in the game);

- **Time away from focus** (ms): Time spent while holding the lantern and not appearing on the viewport of the game;

- **Distance covered** ([in-game] meters): Amount of ground that was stepped on;

- **Time stopped** (ms): Time spent while not changing position;

- **Average distance to other** (m): Average distance between players during the session.

The task at hand was now to parse the existing information from the generated text files to usable metrics. For that, a *Python* script was created using *Thonny*², a Python’s IDE. When executed, it went through every existing log with the predefined names, and recovered information from both files for each pair. Dealing with each category of information separately, sequenced by the order coming in from the metrics file.

Apart from capturing (and counting) the amounts of times an event occurred, it also dealt with converting the formats of timestamps and player positions into more convenient units. Some measures were directly obtained by counting the amounts of occurrences (e.g. deaths); others made use of the stated format conversion to get the pretended sums (e.g. distance covered and time stopped); and a few were derived from other simpler metrics (e.g. time as camera was comprised from knowing which player started with the lantern, and the timestamps where players swapped the camera with the light).

The exporting process of these values to a statistical program only made sense, however, after obtaining the immersion values for each session. That process was achieved via a questionnaire after the conducted test scenarios - described in the next chapter. After doing so, these immersion values would be saved in another text file in the same directory as the others. Finally, a cvs file would be created containing the above mentioned metrics in addition to the immersion value. Having a line entry for each individual participant. This way, importing the values into *IBM’s SPSS Statistics 25*³ was a simple procedure.

---

²[https://thonny.org/](https://thonny.org/) - Free Python IDE
Contents

4.1 Preliminary Tests ....................................................... 29
4.2 Final Evaluation ....................................................... 37
4.3 Discussion ............................................................. 44
4.1 Preliminary Tests

This chapter is composed by the two preliminary tests that were carried out before proceeding with our finalized test. We present their intents, preparation and execution of the experiences, and a brief discussion and conclusions taken from them.

On a first hand, we wanted to address the problem raised by this project and understand if we can take steps further in creating tools to support or either replace the current standardized ways of measuring the experiences felt during gameplay. Immersion was aligned with our intents for this project, but at this stage we conjectured if other experiences alike could also be studied. So we conducted the first preliminary test focused on the values of Social Presence while playing DTA.

Then, and in order to validate our process and better execute the last evaluation phase of the project, we conducted a pilot test. This experiment was carried in parallel with the development of the data retrieval and transformation processes.

4.1.1 Social Presence Variation

When considering our hypothesis on whether a player's feeling towards a game could correlate with in-game player metrics, there were several of those feelings that could be put to the test. Immersion seemed like an important feeling to base our study for two main reasons. Firstly, it's a very marketable word, however, its definition varies greatly. So, producing research in this field of study in order to create a more consistent interpretation of the term is beneficial. Secondly, immersion is a very personal experience varying much without external factors on a given experience.

The latter aspect was seen as broad enough to be used as case study to measure a personal experience felt towards a medium in digital gaming. We pondered however if there were other more sociology-driven aspects that could be brought to be also part of the study. The one we felt was more relevant for our case was Social Presence. Since the testbed game has a purely cooperative component associated with it, we wanted to understand if that component revealed itself through questionnaires after a gameplay session. So, a preliminary test for this matter was carried out.

4.1.1.A Sample

During Lisboa Games Week - a games’ convention that happens every year in Portugal - the game "Dark Things About" was present as a finalist for the PlayStation Talent Awards. This event occurred between the 15th and 18th of November 2018 in a major exhibition room in Lisbon. Since there would be a large affluence of people with high interest in games, a test was arranged to better understand the levels of Social Presence when playing a session of ours (as depicted in Figure 4.1)
During the four days of the event, 22 participants agreed to participate in the experiment and 21 valid answers were taken from the test (1 user didn’t properly submit his form). Out of these 21, 14 were male and the remaining female. Their ages were between 16 and 32 years-old (M = 20.48, SD = 4.43).

Regarding their professional relation in the video game industry (Figure 4.2(a)) 9 people answered with none, and other 9 said they were students in that area. Only 3 users were professionally related with video games. Inquired for the frequency of which each person played video games: 14 said that they scheduled their time in order to play, while the rest only played occasionally and when given the opportunity (Figure 4.2(b)).
We tried to get a better understanding of the type of game the players were familiar with, so we questioned for multiplayer games (Figure 4.3(a)): we had 1 person that played rarely; another said that every game he played was multiplayer; 9 users reported that the majority of games they played were multiplayer, and the remaining 10 said only some of the games they played were of that type. Concerning the 3rd person adventure genre (a broad definition in which Dark Things About fits). Results showed that 2 players don’t play this kind of games; 9 had played at least one of these games in the past while the remaining sample said it was one of their most preferred genres (Figure 4.3(b)). Eight people had played a Dark Things About demo before that day.

(a) Frequency of playing multiplayer games  
(b) Familiarity with 3d adventure genre

Figure 4.3: Game related questions - Part 2

4.1.1.B Measurement

A two-section questionnaire was created in order to inquire the participants (Appendix A). Firstly, a section that was meant to characterize the sample’s demographics. This section was reused for all the other evaluation questionnaires of this project, having only minor corrections and adjustments until the finalized version. And then, a Social Presence related section used to determine the involvement of player with their gaming partner.

Social Presence in Gaming Questionnaire (SPGQ) was used as the the core for this measurement (A). This questionnaire is based on the GEQ’s Social Presence Module having, however, a more extensive analysis in the subject by adding 4 additional items to the original questionnaire. This survey is composed by three sub-scales: psychological involvement - negative feelings, psychological involvement - empathy and behavioural engagement. All the questions had their order randomly assigned and no association to which sub-scale they belonged to.
4.1.1.C Procedure

Since this preliminary test was done in an uncommon set of conditions (i.e. large gaming fair with thousands of people attending at a given time) some adjustments to the expected procedure had to be done. Firstly, as described by the literature, the experience is overall more positive when players play with humans that they have a relationship with [17]. Therefore, we only invited pairs of users that were already familiar with each other. This practice was used in every other gameplay session procedure that occurred in this project.

Furthermore, it is common in this kind of exhibition to see people eager in experimenting dozens of games and technologies in a short span of time. Thus making many players leave the playing booth without completing or even exploring the prepared gaming build for that given event. To make up for such situations, a pair would only be invited to take part in our study after reaching the final end of the demo. Note that this would be independent of the enjoyment and difficulty shown for the duration of the playthrough.

Lastly, in the introductory section of the questionnaire a box to fill in with a team name was added. This had two main reasons: it was meant to result in a enjoyable moment between the pair as they agreed on something to name the duo, at the same time as it would create an easier way to backtrack information and understand which players formed a pair (without breaking anonymity).

Summing up, the experience resulted in the following procedure: if a pair of players (already known to each other) willingly played the first level of Dark Things About until the end, we would invite them to take part in our study. In case they agreed, we would administer the questionnaire detailed above created using Google Forms. All users accessed it using their smartphones via a shortened url created for the purpose of easing said access. In the end, there would be a small summary on some reasoning for the study, usually complemented by feedback given by the players regarding the overall game experience.

4.1.1.D Results & Discussion

Upon concluding this preliminary test it was time to treat the gathered data. Since the delivered questionnaire had randomly assigned question orders it was necessary to reorder the questions to conveniently calculate the subject’s answer values for the respective SPGQ’s sub-scales.

Awarded in a 5-point Likert scale, each question can be valued between 0 and 4 - respectively from lowest to highest agreement for a given item. Calculating the mean and standard deviation of every subject’s answers for each SPGQ sub-scale, we got the results given in Table 4.1.

Table 4.1: Social Presence reported by test subjects

<table>
<thead>
<tr>
<th>PI – Empathy</th>
<th>PI – Negative feelings</th>
<th>Behavioural Engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.05</td>
<td>2.87</td>
</tr>
<tr>
<td>Std Dev</td>
<td>0.37</td>
<td>0.21</td>
</tr>
</tbody>
</table>
Psychological Involvement - Empathy - i.e. feeling to be in a situation as enjoyable as the other player - came up as the highest difference to minimum or maximum obtainable values with also highest standard deviation among all three sections. This section appears to be the one that can vary the most depending on the alignment of the player's preferences and the existing game mechanics. The empirical argument to support the latter statement has to be with the fact that in this particular game each player is playing with very different game mechanics at each given time (one player is carrying a light and the other holds the game's camera).

Psychological Involvement - Negative Feelings - i.e. competitiveness and aggressiveness shown towards others - produced a convincingly low result. Feelings of competitiveness are almost mitigated by the purely cooperative nature of the game. There's no real competition with an individual when we are dependent on said person to successfully persevere. Aggressiveness, though, can be a result of frustration surfacing from misalignment between players styles of play, levels of skill or aptness to communicate. Howsoever, those feelings barely arose.

Ultimately, Behavioural Engagement - impression on whether the other player influences the game - also came across as highly affected in the measuring of Social Presence levels for this test. Evidently, when the game mechanics promote such bilateral dependency on each member of a pair to progress through the game, one can assess that when one of the individuals is not fulfilling their part they are directly influencing the other player's experience.

To conclude, even in such undesirable test conditions the results on all three sections of the questionnaire came near either the obtainable maximums or minimums, without an ample variation. Thus, suggesting that our hypothesis of finding a correlation using values with such characteristics would theoretically be very difficult. So, we discarded the study of social presence in addition to immersion, and decided to narrow our focus solely to the latter.

4.1.2 Model and Process Validation

After confining our study to deal with immersion it was time to implement the mechanisms to retrieve the player's generated data. Since that mechanism would recover every player input, in-game action and position it naturally created large amounts of data for each play session. Via a pilot test, we sought validation both on the data gathering and its treatment. Besides this, we also prepared this test to better establish the non-automated parts of the experiment - the experiment setting and the delivered questionnaire.

4.1.2.A Sample

This preliminary test was conducted with the participation of 4 MSc and PhD students (2 of each) in the field of study of 'Games', from IST's Information Systems and Computer Engineering course. One 23
year-old female and three 25 year-old males formed the two pairs of participants. Every participating member had great interest and knowledge in playing games, and only two of them had already played a DTA demo before.

4.1.2.B Measurement

By this time we were trying to fully reproduce an early version of what would later be the finalized experiment. That meant also capturing the player’s immersion levels by questionnaire. We administered an immersion-focused questionnaire consisting of 31 items, named Immersion Experience Questionnaire (IEQ present in Appendix A). The items were randomly ordered composing the second section of the delivered form. The first one was the reused demographic’s section from the latter experiment.

IEQ’s items are awarded in a 5-point Likert Scale from 1 to 5 resulting in a possible outcome between 31 and 155 points. The higher the value, the more we can consider a player felt immersed in a game. Almost every value can be directly summed to the final outcome, except for 6 particular items that are worth their symmetrical value (i.e. ‘6 - item value’). The results from these questionnaires are briefly presented on the respective sub-chapter below.

4.1.2.C Procedure

Two pairs of volunteers accepted the invitation for participating in a preliminary test to assist a Master thesis experiment (each pair at a time). They were informed that the test would have an approximate duration of under 30 minutes, and it would mainly consist in playing the first level of the game Dark Things About. Seemingly, they would be playing in order to give overall feedback of the game experience.

Upon being accompanied to a quiet and empty workspace with tables and chairs, each subject would be invited to sit and choose one of two remotes available (Figure 4.4 (a)). A laptop would have DTA already launched in the main menu, awaiting for the game session to be started. Players would willingly start playing when they felt comfortable to do so (Figure 4.4 (b)).

Players’ in-game actions were being monitored without their knowledge while being saved in logs. Sporadically, players would ask for assistance that would be given in the manner of very short and clear spoken interventions.

After reaching the final part of the level, the game would automatically save the generated log and close. Resending the players back to the main menu. By that time, a brief conversation would be initiated informing the end was successfully reached and asking permission to fill a game experience form. Once more, a shortened url was created to facilitate the access of the users. The questionnaire was once again created using Google Forms. Finally, when submitted, the players would be informed on some more details of the experiment they performed. Not only so, but also encouraged to give feedback about the experiment and the game itself, and lastly, properly thanked for their participation.
4.1.2.D Results & Discussion

In addition to the obtained immersion values (which were inconsiderable), the reasoning for this test was to structure the guideline of the experiment, detect eventual bugs in the data retrieval implementations, and also, to prepare the script to treat the retrieved information.

About the guideline of the experiment, we concluded that the best way to approach potential testers would be by asking for help and feedback on the game experience of the game itself. This way, we could advertise the experience as a playthrough, and not as test. Meaning that we could provide details on the game genre (for instance), without compromising telling about the details of the study.

After the first preliminary test, we dropped the query about a team name and replaced it with another differentiation factor - in this case, the color of the controller the players used. Apart from this, some minor misspellings and formatting problems were corrected in the questionnaire. As for the implementation side of the experience, we detected some flaws in two metrics. Having to do with incorrect information upon its retrieval - i.e. wrong or incomplete data in some lines of the generated logs. Some slight adjustments to the level design were also reviewed and rectified:
(a) House with non-perceptible entrance path  
(b) House with more evidenced entrance  
(c) Walls blocked intended entrance  
(d) Stairs resolved blocked path issue  
(e) Priest was using a female model  
(f) Priest model now equal to Medic model

Figure 4.5: Game and level design changes

The reasoning behind changes (b) and (d) are due to confusion when players tried to get to the pretended destination (house). Alteration (e) was firstly to avoid bias in preference towards one of the two characters, and also to match the existing male voice-actor that currently voices the said character. Implementing these changes concluded the arranged game scenario for play testing our demo.
4.2 Final Evaluation

4.2.1 Sample

After having a functioning implementation of the player’s in-game metrics gathering and treatment processes, it was possible to address the main question between a player’s performance and their feeling of immersion. The carried out experiment counted with twenty (20) participants - 5 females and 15 males - that played the first level of Dark Things About. This sample was composed of 10 pairs of Instituto Superior Técnico’s students (4 female/male pairs, and 6 only male pairs), registering an average of 22.3 years (SD = 0.19). The already mentioned demographic’s section of the questionnaire was once again used, characterizing the subject’s background and desirability to play video games.

On the topic of the testers’ professional relation with games (Figure 4.6 (a)) 11 users responded with having no relation, 7 having a purely academic relation and 2 being researchers in this field of study. When asked about how often they played video games, 1 person admitted to never play video games, 9 showed that they played when the opportunity presented to themselves, and the remaining 50% of the population revealed to purposely reserve time on their schedule to play video games (Figure 4.6 (b)).

![Professional Relation with Games](image1)

![Frequency of playing games](image2)

**Figure 4.6:** Game related questions - Final Test/ Part 1

Regarding the question on the frequency of playing multiplayer games (Figure 4.7 (a)): 6 reported that they have played only a few games while 12 play mostly this type of game; the remaining 2 played exclusively multiplayer games. As for the interest in the 3rd person adventure genre (Figure 4.7 (b)): 45% of the sample was familiar with it and had played at least one game of this genre, 5 people said it was their favorite and other 5 said it wasn’t part of their preferences; 1 person restated not to be interested in video games. Lastly, 15 out of the 20 players had never played any demo of the testbed game.
Analyzing further the one individual that stated not to have interest in games we considered it not to be an outlier. The reasons for this choice are based on: the reported values seeming inline with rest of the sample; apparently not affecting that person’s pair values; the regularity in which the testing scenario occurred; and finally, that person reporting to have actually played multiplayer games.

### 4.2.2 Measurement

In order to try to test our hypothesis and correlate the in-game player metrics and their own immersion levels, information was gathered by two means: in-game data - captured by the game while the players went through the test level; and immersion level - collected via questionnaire presented after the gameplay session.

The items used to analyze this latter value formed the second section of the delivered overall gameplay experience questionnaire. The one used (and adapted) from the preview preliminary experiment. The mentioned items of IEQ are scored in a 5-point Likert scale from 1 to 5: where 1 usually the highest disagreement depending on the formulation of the question; opposed to 5 which means the symmetric value in this said spectrum. The resulting sum of the items varies between 31 and 155. A lesser number means a lesser degree of immersion, and a higher number means a higher degree.

The gathered in-game data is preserved as a single file for both persons on that test. Processing the retrieved information is the coming step in transforming the large amounts of data into usable metrics for each single player. The criteria we used to select which information should be transformed to a metric was: relevancy in terms of gameplay mechanics, recurrent use throughout a session, and finally, commonness (of a condition) between this and other games. As described in the previous chapter, the following metrics were considered: Deaths, Successful interactions, Failed interactions, Time as Camera, Time away from focus, Distance covered, Time stopped and Average distance to other.

All the latter metrics were the result of transforming the accrued information using a post-gameplay data treatment Python script. Correlating metrics to immersion levels, and every other statistical chal-
4.2.3 Procedure

IST Alameda’s campus site was where every pair of volunteers participated in the final experience for this project. When accepting the invitation to join, the pair would be slightly briefed of the game’s mechanics and expected session time. Just like in the pilot test, they did not know the real intents of the study.

While playing the game, minor spoken interventions could be made in case of game design discussion or to resolve eventual in-game setbacks. When the pair reached the end of the demo, a log containing information for that session would be saved automatically. The pair would then be thanked for playing the game, and requested to fill the experience questionnaire (prepared from the last preliminary test).

After accessing a provided link to the Google Forms’ survey, and completing the required task, the player’s information would be stored in a Google Spreadsheet for later analysis. As for the subject’s participation, it would end there. Shortly after a summary on the real purpose of the study and a piece of candy would be offered as an appreciation for their collaboration.

Following the user tests, every file recovered from the gameplay sessions would be parsed and agglomerated turning its raw information into the previously described metrics. This process was achieved running a Python script for all existing files. Then, a single csv file would be generated containing the metrics’ information for every individual.

This process could actually only be fully achieved after calculating the immersion results for the participants before running the script, saving these values in a particular file. Doing so, would then merge them to the output file. In the end, this csv file was imported to SPSS 25 where the upcoming explained data analyses were performed.

4.2.4 Results #1 - Immersion to metrics correlation

Tackling our hypothesis was the main priority at this stage. We needed to understand if there were any indications on whether a player’s feeling of immersion can be identified by quantitative actions performed while playing. Or if it is solely a subjective feeling that always dependent on the subject’s state of mind, preferences and personality, completely unaligned with said measures.

Having every metric value and immersion level discriminated for each individual was the first step in this challenge. To understand whether or not gameplay and immersion correlate, we needed to perform statistical operations to get said correlation values. We did them using SPSS 25.

First we needed to find if the data values were normally-distributed. So, a Shapiro-Wilk test was performed to study the normal nature of our group’s metrics. The values shown by this group (n =
20), claimed the metrics “Deaths”, “Time Stopped” and “Average distance to other” as not following a normal distribution. Since their $p$-values were all below the level of significance of 0.05 for W-values of 0.574, 0.871 and 0.898 respectively. All other metrics had significance levels above the aforementioned value ($p$-value$<0.05$), having simultaneously greater W-values (nearing 1). Therefore, all the remaining metrics couldn’t be labeled as non-normal. After discarding these three metrics, we proceeded with a Pearson’s correlation test. However, no significant correlation was found.

Even though, the parametric correlation test didn’t confirm results to support our hypothesis, the non-parametric test was still a valid option. Firstly, because normality tests are only indicative of which values are not normally-distributed and no other conclusions can be taken. Meaning that some values might not be indicated as not-normal even though they might be. Then, because the sample size may be a hinder in assuring the normalcy of the data.

Consequently, a non-parametric correlation test was also performed: Spearman’s rho. This time, it appeared to exist a slight indication of correlation. Spearman’s rho ranked the correlation between “Immersion” and “Distance Covered” on a value of -0.352 for a $p$-value = 0.128 ($n = 20$). This value couldn’t be considered a high correlation, but it was a first step in considering new other approaches. The achieved significance value may once again be connected to the diminished sample size, therefore, not granting a value under the desired $p$-value. Because of that, we settled from this moment on with a non-parametric approach.

![Figure 4.8: Dispersion line between Immersion and Distance Covered](image)

### 4.2.5 Results #2 - Gameplay-based metric division

The criteria that came into play when subdividing the players’ data into metrics was meant to make a rundown of the objective gameplay experience as a whole - i.e. capture the most significant and continuing
actions that would occur throughout the entire gameplay for the great majority of all gamethroughs.

However, when analyzing the previous results, we felt the need to refine the most promising metric that obtained correlation with immersion: Distance covered, by further exploring the moments of gameplay where this action took place. For this particular game and level, the main action occurs in two very different settings. And on a superficial overview, it makes sense to affect considerably the values of the mentioned metric, since in this level, the players play in two different spatial scenarios: a forest and a house.

This forest (Figure 4.9), is a spread out field that, while fairly linear, it is a place that can incite exploration or cause players to simply wander around - for its large dimensions and various distractions along the path. The house (Connor's Mannor - Figure 4.10), however, doesn’t seem to share the same characteristics. It is divided in puzzle-rooms that have no other real objectives than to solve its main quest, appearing consecutively one after another. It has way more constrict width limitations (as understood by the already mentioned figures) and a less overall ground to cover. In-engine values estimate the covering ground of the forest to be around 80m$^2$, while only 11m$^2$ for the house’s - over 7 times smaller.

![Figure 4.9: House blueprint and scale](image)
So, we created a separation on the metrics’ values before and after they enter the house. Hypothetically, if one of the incited gameplay-types (covering large amounts vs. shorter amounts of ground) really stood out when performing this division, it could potentially interfere with the metrics’ previously calculated outputs.

To achieve this, the data treatment Python script was adapted to fit the needs. Now, it would perform a split in the metrics calculation before and the after the timestamp of entering the house, and then, new metrics would be created accordingly. After that, the new version of the cvs file would be exported.

Right upfront the supposition confirmed to be contrary to the conducted line of thought. The 20 players presented an average 172.5m (SD = 20.3) of covered ground, and even though the forest had a wider area the players walked an average of 81.8m (SD = 32.9), while the more constricting house area was walked on for 90.7m (SD = 30.5).

Non-parametric correlation tests proceeded but no matching values were found. Distance Covered not only did not assert its correlation values, it significantly dismissed it. Spearman’s Rho evidenced no correlation between Immersion and Distance covered, both for the forest and house settings. The resulting rho’s were -0.028 and 0.069 for the significance values of 0.907 and 0.823, respectively.

### 4.2.6 Results #3 - Additional analyses

After dealing with the main question raised in this project, the recovered data still had potential to be further explored. For instances, another approach to the issue could be by understanding patterns
between users who felt more immersed opposed to the ones who didn’t.

A non-parametric test for independent samples was then conducted - Mann-Whitney’s test. With a grouping variable, this test allows to understand the heterogeneity of the data between groups split with a different characteristic (defined by the given variable). In this case, we used the median value score for immersion: 118 (included). Results showed that there weren’t significant differences in the groups of players reporting higher and lower immersion values.

On another approach, we explored the correlation not with immersion, but within the detailed metrics themselves. When reevaluating the already performed non-parametric test (Spearman’s Rho), several correlations were revealed. The more relevant ones are demonstrated in Table 4.2 (and the dispersion of data in Figure 4.11):

<table>
<thead>
<tr>
<th>Table 4.2: Spearman’s Rho correlation between in-game metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Failed Interactions</strong></td>
</tr>
<tr>
<td>Successful Interactions</td>
</tr>
<tr>
<td>Time as Camera</td>
</tr>
<tr>
<td>Average distance to other</td>
</tr>
<tr>
<td>Distance covered</td>
</tr>
</tbody>
</table>

Even though these correlations did not come from the intended approach for this project, they still deserve some considerations for future ones. Some of these considerations will be present in the discussion chapter below.
4.3 Discussion

When comprehending the alignment of immersion with gameplay mechanics we were trying to canvass the possibility of this feeling being more than purely subjective. We were thriving for a more theoretical iteration on the immersive experience that is commonly accepted in today’s industrial and academical video game studies. Doing so could potentially open doors to run-time video adaptation possibilities.
For that, immersion needs to surface as being reproduced (or correlated) by what a player objectively does while playing, therefore, adjusting these actions could ultimately also tune the experience.

Results obtained for this correlation were not conclusive but still incited future solutions to explore. In the first approach, when trying to understand if the mentioned correlation existed, the metric of “Distance covered” seemed to slightly align with our assumptions. By scoring -0.352 in Spearman's Rho test, player’s appeared to be less immersed the more ground they covered. There are potentially several reasons for this to happen: the game may have had a lot of unnecessary, long or confusing paths - giving a sense of dissatisfaction to the ones who ended up walking the most.

Splitting the information into two zones (“house” and “forest”) did not show to be enough to identify the issue. This line of thought ended up contradicted by the values of ground covered by the players in and outside the house - showing larger values for indoors (house: M = 90.7m, SD = 30.5; forest: M = 81.8; SD = 32.9). This may be because of the existing puzzles inside, where players may have travelled much in order to assess which were the correct levers to press to beat that challenge.

Immersion continues to been seen as an experience that is fairly more recognizable by its effects on a person using a medium, than by the ability of using that medium. However, there are still reflections to take from this study. Both in terms of results and approach. The intended contribution from this study was to provide game developers a strict and theoretical approach to their methodology of producing more engaging and immersive games. Or at least explore the feasibility on achieving that approach. The most impactful results of this study that could potentially serve the game development contribution as a tool is the correlations of metrics between themselves.

Looking at some of the strongest correlations inbetween metrics (latter presented in Table 4.2) we can try to understand their meanings. Some will seem very obvious: correlation of “Successful interactions” with “Time stopped”, while others can be considered counter-intuitive: “Successful interactions” with “Failed Interactions”, respectively. Scrutinizing and understanding the reasons behind them can become a viable way to tune the experience of the players.

For instances, it is very understandable that the metric of making fruitful actions (“Successful interactions”) has a parallel with the time players spends stationary (“Time stopped”). Since every time an interaction of this sort occurs, the in-game character displays an animation where it locks the player briefly to a given position. In the optimal case these metrics would also align with immersion. If so, we could adjust the experience by either inciting the player to either perform more interactions of this sort (if the correlation was positive); or otherwise, by locking the player’s position for lesser span of time.

“Successful interactions” to “Failed Interactions” correlation seems to match positively. Meaning that the more times a player performs a successful interaction, he also performs more incorrect ones. This result may seem abnormal or at the least counter-intuitive. A possible reason for this to happen is due to what we consider "successful" and "failed". Successful is when a player presses a given button, near
enough an interactable object that requires that input - ultimately interacting with it. Failed means not
doing so. It may be that players that strive more to interact with every object in the game may not be
receiving a clear enough feedback, therefore, pressing the same button over and over until achieving the
desirable outcome - increasing the overall number of failed interactions. These are mere conjectures.
They serve but to show the direction in which this project would go in case there was alignment with
them and immersion.
5 Conclusion
5.1 Conclusions

In the opening chapter of this project we raised the following issue: *if questionnaires are an impractical and outdated tool, can we consider another way of measuring immersion?* Ultimately, our approach cannot provide a concrete answer, but several points can be taken to further explore this work’s basis.

Exploring our main hypotheses of finding a correlation between immersion and game metrics values was inconclusive. We believe that our process of metric selection was substantiated, and that not every aspect of the game can be responsible for affecting or even inducing the state of immersion.

The obtained most valuable converging point between the relation of immersion with gameplay was the Spearman’s rho correlation ($W=-0.352$, *p-value*=0.128) between said immersion and distance travelled by players. The value was not strong enough to proceed with more conclusive answers on the influence of this metric in immersion. However, as a developer, making efforts in understanding the real impact of this value in the experience, may well be a enough in producing a better one. All in all, this project raised the degree of knowledge of the relation with DTA and the cooperative aspects of players who explore this game. And here, a parallelism with other developers and games can be drawn.

5.1.1 Future Work

Nevertheless, the real potential from this project may be achieved if a correlation with immersion can be strongly asserted. Therefore we ponder some possible variations to this project’s approach that may see different results from what were here obtained:

- **Single-player game.** We were supported by the literature’s positive endorsement on playing games in the presence of someone. However, considering a single-player game might reduce the disruption of a player’s data from outside factors (in this case another player);

- **Account for player preference.** Some players might enjoy playing the game as they want, and for as long as they want. Having a predefined locked goal to achieve (e.g. we asked people to play until reaching the end of the level) may be unaligned with what some people expect from the experience. Some might find it exhausting while others would have liked to proceed further. Giving the player the freedom to play as much as they want (in a game where doing so would be acceptable) might be enough to obviate this situation. In the case of reaching further in exploratory work, an effort could be made in understanding if personal preferences have in importance in identifying positive aspects of the game.

- **Larger user sample.** Large amounts of test subjects are usually crucial in statistical-based experiments. We inferred that our obtained data followed non-parametric distributions for this same
reason. Some eventual approaches (below presented) might even compel in having much wider sample. However, finding test subjects willing to participate in such experiences can prove to be quite a challenge, specially when using local-multiplayer games.

- **Simplified game mechanics.** Using a game with less complexity in mechanics or stricter gameplay branching possibilities can help narrow down the overall actions that can happen during a session; and so, giving more importance and simplifying the definition of what we can consider crucial (and therefore a metric) for a given game.

A conceptual error we might have incurred was with the linear-type correlation we were hoping for some metrics. For instances, when we studied the relation of the distance between the two players and immersion, we were conjecturing that a higher average distance value would result in a lower immersive one. What we didn’t account for was that this linearity may not exist. There may be a perfect value which may neither be too apart or too adjoining.

Creating a model to study such relations would need a more fitting approach. A Machine Learning model could be equated in relating larges amount of gathered data with reported immersion values. Doing so could increase the chances of finding these eventual patterns, but would be dependent on having a greater sample size.

Lastly, if the correlation between immersion and game metrics proved to be existent, achievable and strong, real tools could emerge from this knowledge. As an example, if a given metric proved to be aligned with immersion, a game developer could now take that into account in order to adjust a mechanic; and doing so, would consequently adjust a game’s perceived immersion. Another instance for a possible usage of this correlation would be with run-time adaptation. Where a game could monitor a player’s performance, extrapolating and adjusting (if necessary) that player’s immersion.
Bibliography


Appendix - Questionnaires
A.1 Demography section

This was the first section of the handed out forms that characterized our samples. The red marked question was present on the first preliminary test, and then replaced by the green one. Which was then maintained from that moment on. After the demographic's section, one of the two following sets of questions (present in sub-chapters A.2 and A.3) composed the rest of the overall questionnaire.
Age *

A sua resposta

Sex *

- Female
- Male
- Would rather not say

What is your professional relation with video games? *

- I am a researcher in a field related to video games
- My profession is related to video games
- I am a student in a course related to video games
- None

How often do you play video games?

- I make some time in my schedule to play video games
- I play video games occasionally when the opportunity presents itself
- I don't play video games
How often do you play multiplayer games? *

- I only play multiplayer games
- Most games I play are multiplayer
- Only a few of the games I play are multiplayer
- I generally don’t play multiplayer games

Are you familiar with the 3rd person adventure genre? (e.g. Uncharted, Brothers: A Tale of Two Sons, A Way Out, etc.)

- This genre is one of my favorites, and I played several games of this genre
- I am familiar with the genre and played at least one game of the genre
- I play video games but not of this genre
- I don’t play video games

Have you played a “Dark Things About” demo before? *

- Yes
- No

Please rate your agreement with the following statement

I am comfortable with using a game controller *

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>totally disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>totally agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A.2 Immersion Experience Questionnaire

Immersion Experience Questionnaire (IEQ) used to measure immersion values. This version of the questionnaire does not represent visually what was handed out to the participants as this is the original version present in the literature and not the online adaptation that was created for the purpose.

To what extent did the game hold your attention?
Not at all 1 2 3 4 5 A lot

To what extent did you feel you were focused on the game?
Not at all 1 2 3 4 5 A lot

How much effort did you put into playing the game?
Very little 1 2 3 4 5 A lot

Did you feel that you were trying you best?
Not at all 1 2 3 4 5 Very much so

To what extent did you lose track of time?
Not at all 1 2 3 4 5 A lot

To what extent did you feel consciously aware of being in the real world whilst playing?
Not at all 1 2 3 4 5 Very much so

To what extent did you forget about your everyday concerns?
Not at all 1 2 3 4 5 A lot

To what extent were you aware of yourself in your surroundings?
Not at all 1 2 3 4 5 Very aware

To what extent did you notice events taking place around you?
Not at all 1 2 3 4 5 A lot

Did you feel the urge at any point to stop playing and see what was happening around you?
Not at all 1 2 3 4 5 Very much so

To what extent did you feel that you were interacting with the game environment?
Not at all 1 2 3 4 5 Very much so

To what extent did you feel as though you were separated from your real-world environment?
Not at all 1 2 3 4 5 Very much so

To what extent did you feel that the game was something you were experiencing, rather than something you were just doing?
Not at all 1 2 3 4 5 Very much so

To what extent was your sense of being in the game environment stronger than your sense of being in the real world?
Not at all 1 2 3 4 5 Very much so

At any point did you find yourself become so involved that you were unaware you were even using controls?
Not at all 1 2 3 4 5 Very much so
To what extent did you feel as though you were moving through the game according to your own will?
Not at all 1 2 3 4 5 Very much so

To what extent did you find the game challenging?
Not at all 1 2 3 4 5 Very difficult

Were there any times during the game in which you just wanted to give up?
Not at all 1 2 3 4 5 A lot

To what extent did you feel motivated while playing?
Not at all 1 2 3 4 5 A lot

To what extent did you feel the game easy?
Not at all 1 2 3 4 5 Very much so

To what extent did you feel like you were making progress towards the end of the game?
Not at all 1 2 3 4 5 A lot

How well do you think you performed in the game?
Very Poor 1 2 3 4 5 Very well

To what extent did you feel emotionally attached to the game?
Not at all 1 2 3 4 5 Very much so

To what extent were you interested in seeing how the game's events would progress?
Not at all 1 2 3 4 5 A lot

How much did you want to "win" the game?
Not at all 1 2 3 4 5 Very much so

Were you in suspense about whether or not you would win or lose the game?
Not at all 1 2 3 4 5 Very much so

At any point did you find yourself become so involved that you wanted to speak to the game directly?
Not at all 1 2 3 4 5 Very much so

To what extent did you enjoy the graphics and the imagery?
Not at all 1 2 3 4 5 A lot

How much would you say you enjoyed playing the game?
Not at all 1 2 3 4 5 A lot

When interrupted, were you disappointed that the game was over?
Not at all 1 2 3 4 5 Very much so

Would you like to play the game again?
Definitely not 1 2 3 4 5 Definitely yes
A.3 Social Presence in Gaming Questionnaire

Social Presence in Gaming Questionnaire (SPGQ) used to obtain social presence felt during a testing session. This version of the questionnaire does not represent visually what was handed out to the participants as this is the original version present in the literature and not the online adaptation that was created for the purpose.

Rate the following statements from 0 (completely disagree) to 4 (completely agree):
(note: the following categories did not appear explicitly in the delivered questionnaire)

Psychological Involvement – Empathy:

- When the others were happy, I was happy.
- When I was happy, the others were happy.
- I empathized with the other(s).
- I felt connected to the other(s).
- I admired the other(s).
- I found it enjoyable to be with the other(s).
- I sympathized with the other(s).

Psychological Involvement – Negative Feelings:

- I tended to ignore the other.
- The other tended to ignore me.
- I felt revengeful.
- I felt schadenfreude (malicious delight).
- I felt jealous of the other.
- I envied the other.

Behavioural Engagement:

- My actions depended on the other’s actions.
- The other’s actions were dependent on my actions.
- What the others did affected what I did.
- What I did affected what the other did.
- The other paid close attention to me.
- I paid close attention to the other.
- My intentions were clear to the other.
- The other’s intentions were clear to me.
Appendix - Data retrieval files
B.1 Distance file

Truncation of an existing file created for logging distances within a gameplay session.
B.2 Metrics file

Simplified version of an existing file created for logging metrics within a gameplay session.