# EVALUATING THE USER ENGAGEMENT AND THE TECHNOLOGY ACCEPTANCE OF AN AUGMENTED REALITY PERVASIVE GAME FOR URBAN AWARENESS

FEDERICO FABIANO, AUGUST 2018, UNIVERSITY OF TWENTE

SUPERVISORS: RANDY KLAASSEN (UT), MARIËT THEUNE (UT), PALOMA DÍAZ PÉREZ (UC3M)

*Dedicated to my parents, for their unconditional support and their silent presence.*

# ACKNOWLEDGEMENTS

I would like to thank every person that supported and helped me to complete this thesis. Firstly, I would like to express my gratitude to Paloma Diaz who welcomed me in the DEI Lab and gave me crucial observations and suggestions during the meetings. Also, I want to thank all the colleagues that worked with me in the DEI Lab, who made me always feel at home.

I am thankful for the precious comments received from Randy Klaassen and Mariet Theune, especially while writing this thesis.

Besides those who helped me carrying out this project, I'd love to thank the people who made it possible to achieve this. My parents and Matilde, for being next to me with endless patience. All the friends met in Enschede: Mauro, Nicola, Giacomo, Teresa, Pietro, and Ozgur. The Italian friends for having supported me virtually: Edoardo, Matteo, Gianvito, and Vincenzo.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# 1. INTRODUCTION

In the last decades, the technology is advancing fast-paced and over time it's becoming more compact and ubiquitous. Thanks to some useful functionalities such as the GPS and the wireless connection, which helped the portability of the smartphones, the mobility and pervasiveness of the technology have become a central topic in the design of new digital products. In parallel to this change, the technology evolved becoming an integral component of the environment where we are living; now, more than ever, we are residing in a deeply connected urban environment, where the virtuality is well integrated into the real world. The evolution from a physically-limited technology into a pervasive scenario, had also affected the concept of cities and their design turning most of them into Smart entities, that is more intelligent, interconnected, and instrumented (Harrison et al., 2010). However, this progress towards intelligent cities assisted another evolution concerning the citizens, in fact, the city dwellers gained a key role by becoming nodes of a wide network and being capable to have a direct impact in the city-making (Ampatzidou et al, 2014).

This progress has had an impact also on digital games and their design. In fact, the idea of games passed from being static and tied to the console to a mobile and pervasive experience, where the game is not played only digitally on a console but it's experienced in the physical environment and the game information is well-integrated in the physical space. These types of games are called pervasive. They are used not only for the mere enjoyment of the player, rather for meaningful purposes such as civic engagement, informal learning or environmental awareness (Neuenhaus et al., 2017; Humphreys et al., 2011). One of the main issues concerning pervasive games is the union of the virtual environment of the game and the physical one where the games actually occur. In order to tackle this issue, different solutions have been thought out, but the most effective is the adoption of Augmented Reality (AR). In fact, AR allows to place the virtual information of the games consistently with the physical space (Wang et al., 2013).

The research presented in this thesis took place in the Interactive Systems laboratory of the Universidad Carlos III de Madrid[1] and contributed to Project PACE (TIN2016-77690-R), which investigates citizen engagement in different forms. One of the goals of the project is to analyze how to engage citizens in meaningful experiences using affordable pervasive

[1] dei.inf.uc3m.es

technologies. In particular, an AR pervasive game has been designed for raising awareness about the urban environment and the history of a building of the university (i.e. the Sabatini building) that otherwise goes unnoticed by most of the people that dwell in that space. Hence, the goal of this game is promoting the urban environmental awareness through an engaging informal learning experience.

This research has the purpose to evaluate to what extent the pervasive game engages the users (i.e. citizens) and to assess the level of acceptability of the chosen technology, namely an AR pervasive game running on a smartphone. Besides these two evaluations, the research aims at contributing to the definition of user engagement for pervasive games in the context of playable cities. According to the study of the literature, it is hypothesized that two diverse factors (i.e. Urban Environmental Awareness and Perceived Walked Distance) might be considered as measures of the user engagement (Howe et al., 2016; Bursztyn et al., 2017).

## 1.1. Research Questions

Prior to this thesis, an exploration of the topics of the research has been carried out in order to know more how to deal with the subjects involved and to define the Research Questions of the thesis (Fabiano, 2018). The research questions (RQ) and the correspondent goals divided according to the topic are described as follows:

| | | |
|---|---|---|
| User Engagement | Goal | Measuring the engagement of the mobile application and have a benchmark about how engaging it is for users to be compared eventually with other metrics. |
| | RQ 1 | To what extent is the mobile application engaging for the city dweller? |
| Technology Acceptance | Goal | Assessing the acceptability of the technology chosen for the playful experience and check if the UE and TA are connected. |
| | RQ 2 | a. To what extent is the mobile application accepted technologically?<br>b. To what extent are the Technology acceptance and User Engagement related? |

| | | |
|---|---|---|
| Perceived Walked Distance | Goal | Try to find other indicators of the User Engagement in the context of pervasive games. |
| | RQ 3 | In the context of pervasive games, to what measure can the Perceived walked distance be considered as an indicator of the User Engagement? |
| Urban Environmental Awareness | Goal | Measuring how the familiarity with the surrounding environment can influence the game experience and to what extent this factor of UEA is a measure of the user engagement. |
| | RQ 4 | a. To what extent does the familiarity with the environment influence the engagement while playing the game?<br>b. To what extent does the familiarity with the environment influence the acceptability of the mobile application? |

## 1.2. Overview of this Research

The design of the pervasive game and a Beta version existed before the beginning of this research. The Beta has been improved with some little adjustments in order to reach a first working version (v1) of the pervasive game. In parallel, the topics of the research were investigated in order to establish valid Research Questions and their correspondent goals according to the examined literature.

Once the goals have been set, the studies necessary to answer the Research Questions have been planned. In order to reach an adequate evaluation of the user engagement and technology acceptance and investigate properly the potential determinants of user engagement, a usability study was required to evaluate the pervasive game itself and to collect user requirements for the second version (v2). In fact, the usability study allowed us to gather the user requirements and to convert them into features for improving the game. The second version (v2) of the game has been used in the evaluation of the user engagement. Although the results of this evaluation answered all the Research Questions, some results were uncertain. This led to the last phase of the research: the planning of the future work, that is a proposal for the improvement of the game (v3) and a suggestion on how to proceed further with the research according to previous findings.

The thesis is structured as follows. Chapter 2 describes the theory on which the research is based. Chapter 3 explains the PACE project objectives and presents the AR pervasive game for urban awareness used in the research and its technology. Chapter 4 and 5 report the two studies that have been carried out, namely the usability testing and the evaluation of the user engagement. As conclusion of Chapter 4 the new features of the version 2 are described. Chapter 6 zooms out from the experiments and discusses the results of the research in the context of playable cities. The Research Questions are answered in Chapter 7. Finally, the proposals for the future progress for both the research and the game itself are presented in Chapter 8.

# 2. BACKGROUND

This chapter describes the theory on which the research is based. Section 2.1 explains the concept of smart cities to set the context of the research. Section 2.2 addresses one subcategory of the smart cities, namely playable cities and gives some examples of applications. Particular attention is given to the subtopic of pervasive games, which is the foundation of the game for urban awareness (Section 2.3). Subsequently, in Section 2.4, user engagement and its determinants are described as well as what are the available tools for the evaluation of the engagement of a system. Finally, in Section 2.5 the topic of technology acceptance is explained starting with the creation of the definition of technology acceptance over years until arriving to the choice of the evaluation method for the assessment, namely UTAUT2. This chapter is a condensed version of (Fabiano, 2018).

## 2.1. Smart Cities

The concept of smart cities has been described in different ways and in diverse contexts. In fact, the literature doesn't give a unique and straightforward definition, but multiple ones depending on the investigation and the approach used.

In order to explain the versatility of the topic and the different interpretations, I will give an example. Let's assume that a smartphone application exists for connecting citizens of the same neighborhood in order to make them communicate better locally. This can be considered a smart city mobile application as well as a very complex system for improving how to manage emergencies in real time scraping data from social networks (Díaz et al., 2017). As shown in this example, the topic of smart cities is versatile to different views and contexts (Chourabi et al. 2012). In general, the gap in the literature of not having a unique definition of Smart cities has been closed with the definition of the quality "smart" when considering cities, which can be related to the images of sustainability and liveability (Chourabi et al. 2012). These are the two qualities that a physical space should have for being considered as Smart. In addition to this, a city should use the technology to take advantage from it for creating smarter cities, that is more intelligent, interconnected and instrumented (Harrison et al., 2010).

## 2.2. Playable Cities

### 2.2.1. Definition

In order to explain what playable cities are, it is important to define what are the two different approaches used in the context of smart cities. First, a top-down approach in which the technology is developed and controlled centrally by municipalities or corporations. For this approach, usually big investments for crafting the solutions are necessary. The second is a bottom-up approach, which doesn't necessarily involve a big investment in terms of money and resources. In this second scenario, citizens are involved in the system both as nodes of the network and they are directly implicated in the design process in a participatory way.

An example taken from (Fabiano, 2018) can make this distinction on the diverse approaches clearer: "*let's assume that we need to detect in real time the traffic jams in the city. A top-down solution might be to place sensors all over the city to monitor the traffic in real time. On the contrary, if most of the drivers possess a smartphone with a mobile app with a community-based GPS (such as Waze), this can be considered a bottom-up solution for the same issue*"(Fabiano, 2018, p.6).

After defining what a bottom-up approach is in the context of smart cities we can define the concept of playable cities, which is a sub-group category of smart cities that follows a bottom-up approach and has a vision of the city as a place where *"hospitality and openness are key, enabling residents and visitors to reconfigure and rewrite city services, places and stories"*(Nijholt, 2017, p.11). In fact, playable cities shift the focus from a data-driven system of the smart cities with a top-down approach to a people-centered (i.e. citizens-centered) solutions with the aim to turn cities into something gameful, hackable, playable and playful (Nijholt, 2017). Therefore, the playable cities want to pay attention to the role of the citizens, trying to change the concept from smart cities to smart citizens (Schouten et al., 2017). In fact, the involvement of the citizens is crucial while designing playable city applications in order to engage them in the experiences. Hence, a playable city presupposes that city dwellers are also able to develop their own application to hack the city in their home, in the public spaces or in the urban environment. One of the proposed technological methods for producing such playability in a physical space is the adoption of actuators and sensors well integrated ubiquitously in the physical environment. In this scenario, the physical spaces in the cities are not only passive places anymore, but they change into being a dynamic and interactive mean for living smartly, which means that the citizens are engaged in the design of communities or the processes in the city. In conclusion, the physical space in the context of playable cities is the product of social practices and designed directly by people.

### 2.2.2. Examples

After having defined what are playable cities and what is their purpose, we discuss two examples of past projects, considered matching optimally with the vision of playable cities.

The first example is called the Shadowing project[2] and it is an example of a playable city application consisting in embedding physically sensors and actuator in sidewalks or squares of the city. The Shadowing project lies in recording the shadows of all the people passing underneath a spotlight and reproduce the shadow, that has been saved beforehand, whenever the following pedestrian is walking through the same spot (Figure 1). Their design concept wanted to make the city alive by making it possible to memorize what was happening in a fixed spot of the city, in a way this system enables the city to remember things and events.

The second example concerns a research project called Hackable City[3] which aims at exploring new modes of collaborative city-making in order to make the city more democratic and connected (Ampatzidou et al, 2014). In this project, citizens, institutions and researchers are involved in order to co-create new technologies for empowering the quality of liveability in the city. One of the Hackable City case studies took place in the north of Amsterdam, in the area of Buiksloterham. In this occasion, game mechanics have been applied to the brainstorming process to gather ideas and insights on how to improve the area. Three sub-scenarios have been created according to a different purpose:

1. **Buiksloterham Matrix**. It is a tabletop-role game that helps users to conceptualize the urban environment;
2. **The Neighborhood**. It enhances the brainstorming and the creation of ideas using the storytelling and a collaborative map-drawing of the public spaces;
3. **The Water Must Flow.** It is a serious game; the players have as a goal to manage the public resources in their neighborhood.

[2] www.playablecity.com/projects/shadowing/
[3] www.hackablecities.com/

In conclusion, these are just two case studies of playable cities applications to show the versatility and the diversity of the designed solutions. On the one hand, a creative installation for its mere enjoyment by creating surprise in the pedestrians. On the other hand, the introduction of game mechanics in the city-making process.

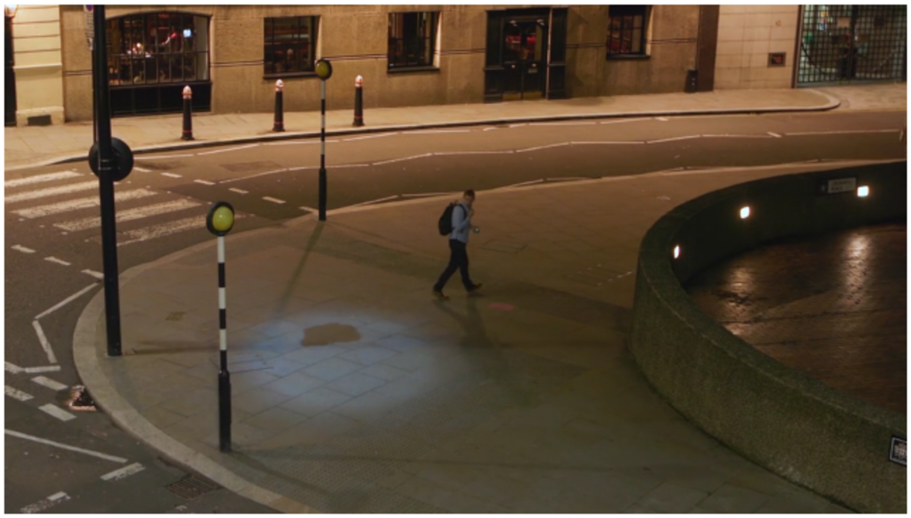


*Figure 1 - Screenshot from the Shadowing project movie[4].*

## 2.3. Pervasive Games

### 2.3.1. Definition

In the literature there are two main ideas of the concept of pervasive games. Generally, it is described as the extension of the game from being limited to virtual environment, to the extension to the physical space as well. Therefore, a game to be pervasive needs to break the spatial boundaries of the regular games, that is, the experience should be designed to take place everywhere and not only being limited to the console. Pervasive games should bring the game experience to the real world (Benford et al., 2005), which can be everywhere around the player. The two notions of pervasive games differentiate on the inclusion of the technology in the definition. On the one hand, pervasive games are described as "*games that have one or more salient features that expand the contractual magic circle of play spatially, temporally, or socially*" (Montola et al., 2009, p.12). This interpretation does not include the technology as a key element of the experience. On the other hand, the second vision of pervasive games is mainly focused on the connection of the physical space and the virtual one, in fact, it does

[4] vimeo.com/236887546

include digital technology as a prerequisite and as interface for the game, while the physical space is the game board.

In general, if we consider digital pervasive games, they lie on the intersection of two diverse worlds: the virtual and the physical one. Their goal is to place digital elements contextualized in the physical space, sometimes using Augmented Reality (AR), creating interactions involving both the worlds. This leads to the creation of a hybrid reality of the playground where the pervasive games are played. It has been demonstrated that Augmented Reality can be efficiently used to create a digital layer for contextualizing information in the physical environment (Wang et al., 2013).

Pervasive games have been used in diverse contexts with different goals. In the past years, they were not created only for the enjoyment of playing, rather they were designed for a more serious purpose. These types of pervasive games are called Serious pervasive games, that is experiences that are both entertaining for the players and, at the same time, having a meaningful impact (e.g. crowdsource information, solve problems, promoting awareness on the urban spaces, etc.).

### 2.3.2. Examples

In the last decade, pervasive games case studies have been carried out and, in the literature, there are plenty of examples on how to develop pervasive applications extendable to the real world. In this section, two examples of the most known pervasive games are explained.

The first is called Geocaching[5], it is a treasure hunt-like game which takes place all around the physical world. The user needs to seek small boxes containing low-value objects and the logbook with the help of the mobile application containing a GPS map (Figure 2). The exact position of the boxes and the user are displayed on the map. However, even though the position of the boxes is known, they are well hidden in the physical environment (e.g. under a rock or hung on a tree). Whenever the users find a box, they are supposed to write their name and the date in the logbook and, subsequently, update their online profile in the mobile app. Nowadays, the game is played by more than 3 millions of people around the globe[6].

---

[5] www.geocaching.com/play

[6] newsroom.geocaching.com/

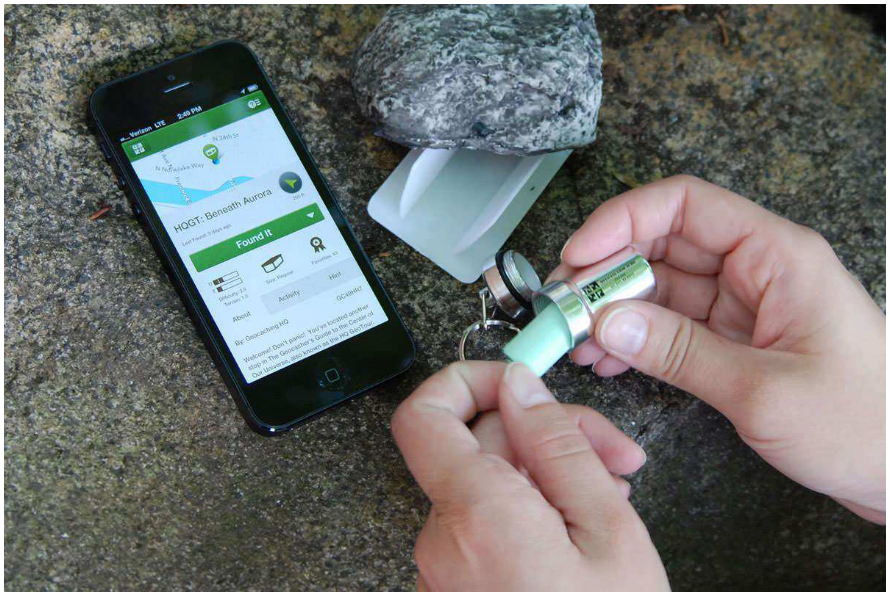



*Figure 2 - The mobile app of Geocaching and the box with the logbook.*

The second example is Pokémon Go[7]. The game consists in the transposition of the Pokémon game in the real world. In the original game, the main character needs to walk around different cities, finds Pokémon and defeat enemies. The same game mechanics happen in the physical environment, where the players need to walk around in the real cities finding Pokémon and catching them with the help of an augmented reality interface.

This pervasive game went viral after its release in the digital stores in 2016. In fact, it has been downloaded 752 millions of times in almost one year[8]. Besides its virality, Pokémon Go started diverse discussions in the field of augmented reality pervasive games. In general, they can be categorized into three different areas: health, safety, and humorous situations.

1. **Health.** In order to be played, the game forced players to walk around the city. This has raised a discussion on the role of pervasive games for motivating people to walk more daily. In fact, the number of obese people grows year by year (Caballero, 2007), therefore this discussion generated interest in researchers. Several investigations demonstrated that using Pokémon Go increases significantly the number of daily steps taken by the users (Althoff et al., 2016; Howe et al., 2016).

[7] www.pokemongo.com/
[8] venturebeat.com/2017/06/30/pokemon-go-passes-1-2-billion-in-revenue-and-752-million-downloads/

2. **Safety**. The virality of the game was accompanied by unfortunate events that happened to people while playing the game. There were a few cases of players encountering safety problems, such as being assaulted and their phone got stolen while playing the game (Khomami, 2016) or teenagers being a target for thieves that used the game to know in which spot there were more people playing the game (Yuhas, 2016). The safety in the context of pervasive games is crucial for the final user experience. Therefore, the playground must be safe.
3. **Humorous situations**. While merging the digital and the physical world, several weird and accidental situations can occur while playing (Andujar et al., 2017). According to Andujar et al., humorous situations can happen in three different scenarios: incongruity, relief, and superiority/disparagement. Incongruity happens when the player finds a discrepancy between two objects in the same "frame". If we consider an augmented reality feature of some pervasive games, this situation can occur when the digital elements are merged into the real world and turn out to be incongruent. Relief is putting a lot of effort and struggling a lot in order to achieve a goal, and, to suppress these negative feelings, a pleasant sensation is experienced. Superiority/disparagement consists in having fun by the misadventures of others. In this occasion, the "winner" is laughing at the "loser", that is the target of the humorous situation.


*Figure 3 - Augmented reality interface of the Pokémon Go app.*

Serious games can be also pervasive. The REXplorer game (Ballagas et al., 2007) aimed at promoting knowledge on the history of buildings all over the city to players. The knowledge is hidden in the scenario of the game, and, according to the position of the players, different information is displayed on the smartphone. REXplorer is an example of how digital information has been contextualized in the physical environment.

Pervasive games have been used for crowdsourcing geo-referenced data about the urban accessibility (Prandi et al. 2016). In Geo-zombie, players are asked to walk in the city and take pictures of inaccessible places which have been transformed into zombies in the game. The users can kill the virtual zombies taking a picture of the inaccessible spot (Figure 4). In the game some gamification functions (such as points, leaderboards, etc.) are available in order to keep users motivated.


*Figure 4 - User interface of Geo-zombie (Prandi et al. 2016).*

Pervasive games have not only been applied to the city context, they have been adopted also in the workplace (Tolias et al. 2015). In a pervasive game called IdleWars, players (i.e. workers) need to walk around the office and scan inactive computers. This had the purpose of reducing the time of the inactive computers, thus reducing the energy consumption of the company. The more scanned inactive computers, the more points the user is collecting.

## 2.4. User Engagement

### 2.4.1. Definition

The user engagement is a quality of the user experience that has been described in different ways and in different contexts. In general, it can be summarized as "*the explanation of how and why applications attract people to use them*" (Sutcliffe, 2010, p.3) and *"the emotional, cognitive and behavioural connection that exists, at any point in time and possibly over time, between a user and a resource*" (Attfield, 2011, p.3). In general, the concept of being engaged is very similar to the idea of Flow described by (Csikszentmihalyi, 1990). Csikszentmihalyi explained the concept of Flow in the field of positive psychology as "*the state in which people are so involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at great cost, for the sheer sake of doing it*" (Csikszentmihalyi, 1990, p.4). In order to achieve this Flow state, the person should be involved in an activity that is well-balanced between being challenging and being doable with the skills of the user. In this case, it is more likely that the user will enter in the state of Flow.

Another perspective of the user engagement consists of considering engagement as a process with different stages rather than just a unique stage. This process has been called Process of Engagement (PoE) (O'Brien et al. 2008). The PoE (depicted in Figure 5) has four main stages: Point of Engagement, Period of Engagement, Disengagement and Re-Engagement. Besides of these, another state, the Non-engagement, was included in the schema to represent users who aren't engaged, that is the users do not enter in the process. When users are implicated in an action and it starts to be engaging, the Point of Engagement is reached. The Period of Engagement follows the Point of Engagement and it ends when the user starts to be disengaged, in the schema this is called the Disengagement stage. Furthermore, internal and external factors can influence the process of engagement ending it or affecting the measure of the user engagement. This process usually occurs as a loop process. This means that whenever the engagement is interrupted, thus the user is disengaged, the process can occur again in the near or far future.

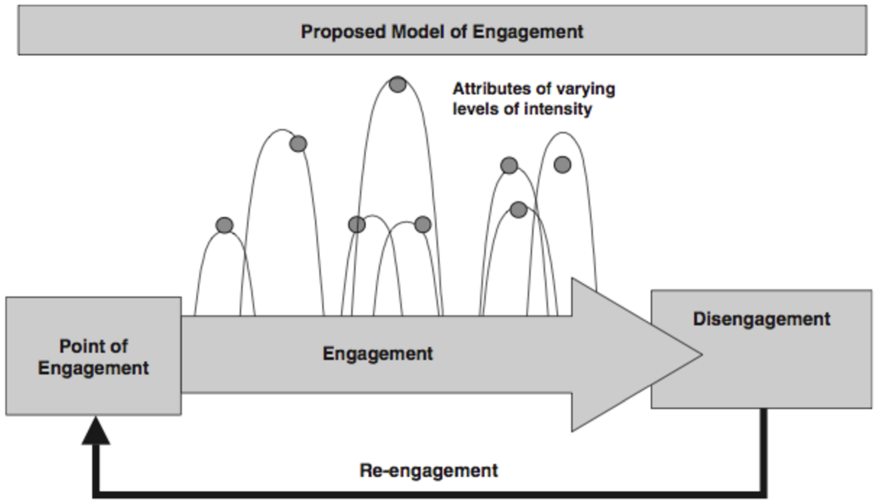



*Figure 5 - Model of the Process of Engagement (O'Brien et al., 2008).*

The user engagement has been addressed as a scheme as well, where several attributes are having an influence on how engaging a system is. (O'Brien, 2008) created a list of attributes of the user engagement that are impacting on the level of user engagement. The attributes were developed by merging the findings of four different theories related to the topic. These theories include the Flow theory (Csikszentmihalyi, 1990), Aesthetic theory related to design principles crafted by (Beardsley, 1982), Play theory (Rieber, 1996) and Information interaction (Schneiderman, 1997). Conclusively, O'Brien summarized the attributes that were matching in the four theories, namely Aesthetics, Positive Affect, Focused Attention, Endurability, Feedback, Motivation, Novelty, Perceived Time and Perceived Usability. They are described in Table 1 which has been created in the preliminary research on the state of the art (Fabiano, 2018).

*Table 1 - Attributes of user engagement and their definition.*

| Attribute | Description | Papers |
|---|---|---|
| Aesthetics | The visual appeal of the user interface. | O'Brien, 2008 |
| Positive Affect | The emotional response to the system. | O'Brien & Toms, 2008 |
| Focused Attention | In order to be engaged, the user needs to be focused on the accomplishment of the task. | Webster & Ho,1997; O'Brien,2008 |

| | | |
|---|---|---|
| Endurability | The users who remember experiences as enjoyable are more likely to repeat them. | Read, MacFarlane, & Casey, 2002; O'Brien, 2008 |
| Feedback | Response from the system whenever a task is accomplished. It is useful to demonstrate to users the progress towards the goal. | O'Brien & Toms, 2008 |
| Motivation | Elements that create desire and willingness to perform a task. | O'Brien & Toms, 2008 |
| Novelty | Characteristics of the system which results to the users unexpected, surprising, new and unfamiliar in a positive meaning. | Webster & Ho, 1997; O'Brien,2008 |
| Perceived Time | The user perception of the flow of time while performing a task. | O'Brien & Toms, 2008 |
| Perceived Usability | It's a negative effect controlled by the perception of the required effort in using the technology and the control over it. | O'Brien & Toms, 2010 |

### 2.4.2. Measurements

The level of how much a user is engaged is a functional indicator and it can be evaluated in several ways. In the preliminary work (Fabiano, 2018) all the methods have been described, and they can be divided into four main categories described as follows:

1. **User reported measures**. They are made with self-report methodologies, where the users are asked specific questions and they need to evaluate or express their opinion on a system or a service. Users are directly involved in this process and the data is fully subjective. The most used methods are questionnaires, interviews, and observational protocols.
2. **Interaction engagement.** It is the evaluation of the user engagement through the log data, thus they are an objective measurement because the users are analyzed in their familiar environment without their direct involvement in the measurements. In order to evaluate to what extent a person is engaged while using a product, it's necessary to establish the KPIs (Key Performance Indicators) depending on the goal of the system. For instance, a metric that can express an indicator of the level

of engagement is the user retention over a fixed number of days, which is how much a person is using the system over a timeslot, or the conversion rate, which is the percentage of users that are completing a prefixed action.

3. **Body related measurements.** These measurements evaluate the cognitive engagement, which is the involuntary response of the body to a specific circumstance. These methods, such as eye-tracking (Navalpakkam et al., 2012), electrodermal activity analysis (Bardzell et al., 2008), and brain activity analysis (Fairclough et al.,2013), usually involve expensive and intrusive instruments for the evaluation.
4. **Combination of approaches**. They consist of using multiple methods combined in order to have more reliability and validity with combined results. This has been used for tackling the bias of using only one measurement (Kobayashi & Boase, 2012). For example, a mixed-method approach can include self-report data as well as the measurement through logs data, both combined.

### 2.4.3. User Reported Measure, the User Engagement Scale

For the purpose of evaluating a pervasive AR game, Fabiano (2018) found that the most appropriate method is to use various self-reported measurements. This type of measurement can be applied in different experimental settings and the gathered data allows to perform a qualitative and quantitative analysis. Moreover, several questionnaires have been tested and are available to the public. This section explains the tool that has been chosen in the preliminary work for the evaluation of the user engagement while playing the pervasive game.

As mentioned in the Chapter 2.4.1, one of the self-reported method for the evaluation of the engagement is the completion of a questionnaire after using a system or product in order to provide impressions or evaluation. For assessing the extent of the engagement of the user and the evaluation of its attributes, the User Engagement Scale (UES) has been created (O'Brien & Toms, 2010) and subsequently empirically validated (O'Brien & Toms, 2013). The scale evaluates six different attributes of the engagement separately, namely Focused Attention, Perceived Usability, Aesthetics, Endurability, Novelty and Felt Involvement. From these six factors, a 31-items questionnaire has been developed and statistically proved as a reliable instrument for the assessment of user engagement (O'Brien & Toms, 2013). Subsequently, the generalizability of the UES questionnaire has been investigated in diverse fields of study (O'Brien & Toms, 2013), the results of the same scale in different areas has been compared and the correlations in the results matched. The findings of the study showed the versatility of some of the factors, namely Perceived Usability, Focused Attention, and Aesthetic Appeal. Differently, changing the settings in the investigations and their context had an impact on the cohesion of the remaining attributes: Novelty, Felt Involvement, and Endurability. For this very reason, O'Brien investigated towards the possibility of creating another scale, the User Engagement Scale-Short Form (UES-SF), a more flexible solution if compared with UES (O'Brien et al., 2018). The UES-SF includes only four factors of the engagement (Focused Attention, Perceived Usability, Aesthetic, Reward), instead of the original six. It has 12 items, consisting in statements, which have as anchors "strongly disagree" and "strongly agree". Thus, they are assessed through a 5-points Likert scale (the list of all the UES-SF items is available in Appendix C).

## 2.5. Technology Acceptance: UTAUT2

The investigations made on the topic of Technology Acceptance have a considerable distant past. Over the years, the theories have been changed, merged and improved. However, the genesis of this research subject can be found in the Information Integration Theory (IIT) (Anderson, 1981), which theorized how information from multiple sources is processed by humans. Anderson proposed a mathematical model (represented in Figure 6) which combines different stimuli in order to predict how humans are making judgements, that is taking decisions.

The theoretical model established by Anderson inspired many works among which the Theory of Reasoned Action (TRA) (Ajzen & Fishbein, 1980; Fishbein, 1963, 1967, 1980; Fishbein & Ajzen, 1975) and Theory of Planned Behavior (TPB) (Ajzen, 1991), which are the fundamental basis for understanding the currently used theoretical frameworks.

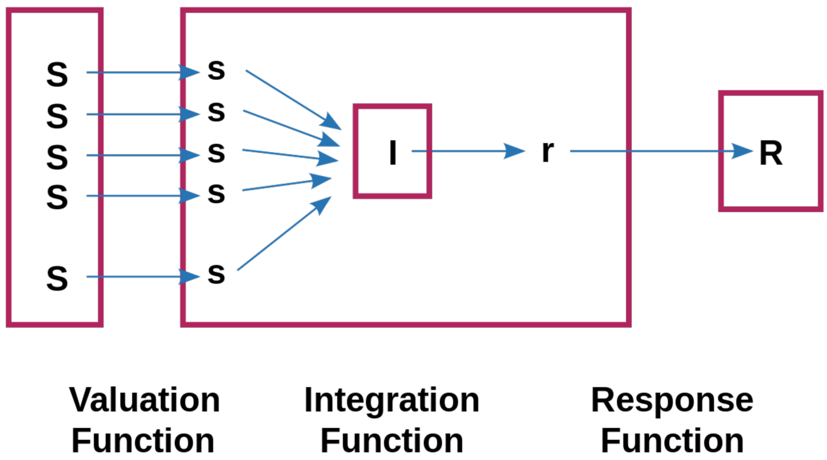



*Figure 6 - Representation of the mathematical model of IIT (Anderson, 1981).*

A graphical overview of the theoretical frameworks concerning technology acceptance (Figure 7) has been created in the preliminary research (Fabiano, 2018). This had the purpose to show entirely the behavioral theories or models that have been theorized, validated and subsequently applied in diverse contexts. Moreover, it shows the full picture and the contributions of each model. The overview shows how the most recent models have been created. In particular, the most known is the Technology Acceptance Model (TAM) (Davis et al., 1989) from which have been derived four different extensions, namely TAM2 (Venkatesh & Davis, 2000), TAM3 (Venkatesh & Bala, 2008), UTAUT (Venkatesh, 2003) and UTAUT2 (Venkatesh, 2012). The UTAUT2 has been chosen in the preliminary research as the most complete measurement for evaluating the Technology Acceptance of a system, as well as the most accurate for the purpose of evaluating to what extent the examined pervasive game is accepted by the users.

The UTAUT2 model is an extension of the UTAUT, which was created with the purpose of detecting what determinants are contributing towards the final behavior of the users and what elements are moderating their influence. UTAUT was created by Venkatesh et al. (2003) through a longitudinal study where eight different models were compared and merged into one. The data gathering, which was carried out using a questionnaire, made it possible to draw as a conclusion of the study the UTAUT theoretical framework. It shows what are the determinants of the Behavioral intention of the individual which influences directly the Use Behavior. Moreover, in the research by Venkatesh et al, the authors also defined the moderators that are influencing the relationship of the determinants on the Behavioral Intention. The moderators can contribute both positively and negatively on the influence of a construct to the behavioral intention. Starting from this model, the UTAUT2 has been built improving the UTAUT framework by adding three more determinants and deleting the moderator "Voluntariness of Use", removed in order to make the model suitable for the voluntary behaviors.

*Table 2 - Description of the determinants of UTAUT2 (Venkatesh et al., 2012).*

| **Name of the construct** | **Description** | **Influenced construct** | **It is moderated by** |
|---|---|---|---|
| Performance Expectancy | To what extent the person believes that the use of the system is helpful in the task. | Behavioral Intention | Gender, Age |
| Effort Expectancy | To what extent the system is perceived as easy to use. | Behavioral Intention | Gender, Age, Experience |
| Social Influence | To what extent the individual feels that for the people close to him/her it is important to use the system. | Behavioral Intention | Gender, Age, Experience |
| Facilitating Conditions | To what extent the person believes that an infrastructure is supporting the system use. | Use Behavior | Age, Experience |
| Hedonic Motivation | To what measure the user is having fun or pleasure while using the system. | Behavioral Intention | Gender, Age, Experience |
| Habit | It happens when the users tend to perform behaviors and actions automatically. | Behavioral Intention, Use Behavior | Gender, Age, Experience |
| Price Value | The monetary cost of such technology according to the perceived value of the product. | Behavioral Intention | Age, Gender |

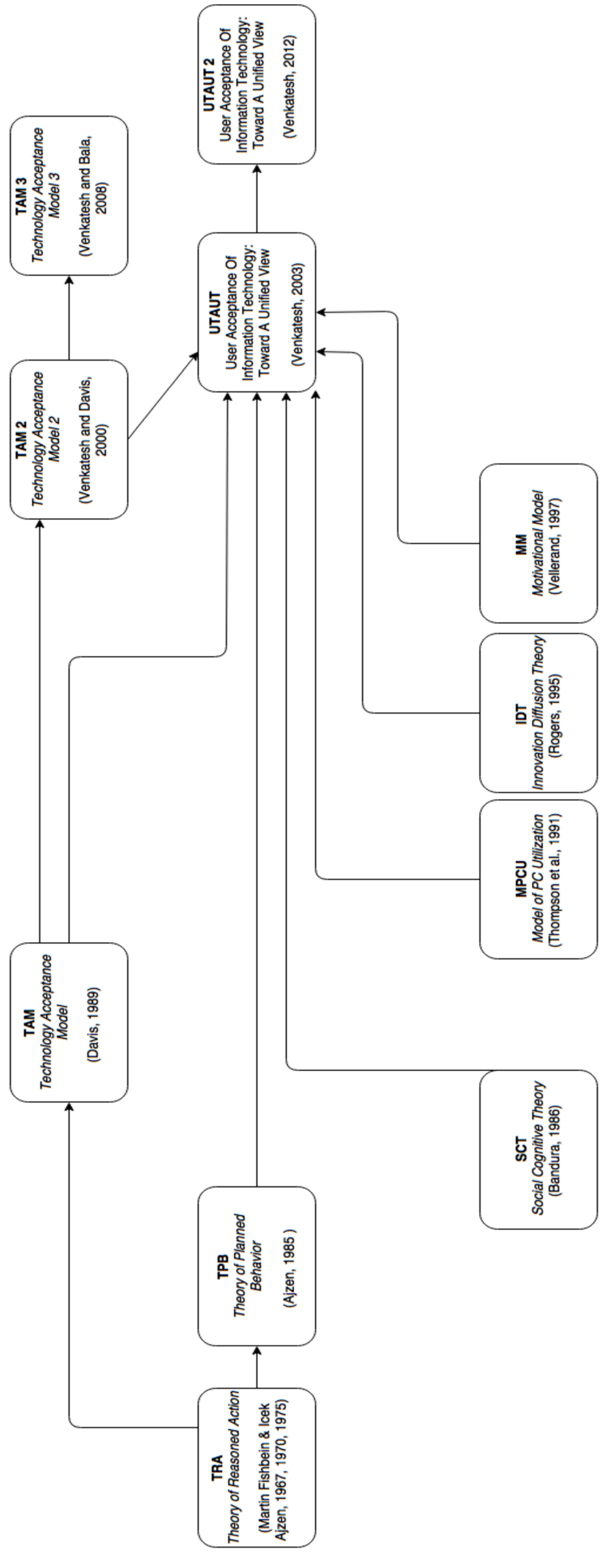



*Figure 7 - Overview of all the theories contributing to the technology acceptance models.*

The determinants of the UTAUT2 model are described in Table 2 and the theorized model is presented in Figure 8.

Practically, UTAUT2 consists of a questionnaire with 28 items evaluated through a 5-point Likert scale (Appendix D). Moreover, it has been already proved to be reliable (Venkatesh, 2012). The items consist of statements, which have as anchors “strongly disagree” and “strongly agree”, concerning the eight constructs of UTAUT2, namely Performance expectancy, Effort expectancy, Social influence, Facilitating conditions, Hedonic motivation, Price value, Habit, Behavioral intention. Every section of the questionnaire is dedicated to a construct of the model and includes either three or four items.

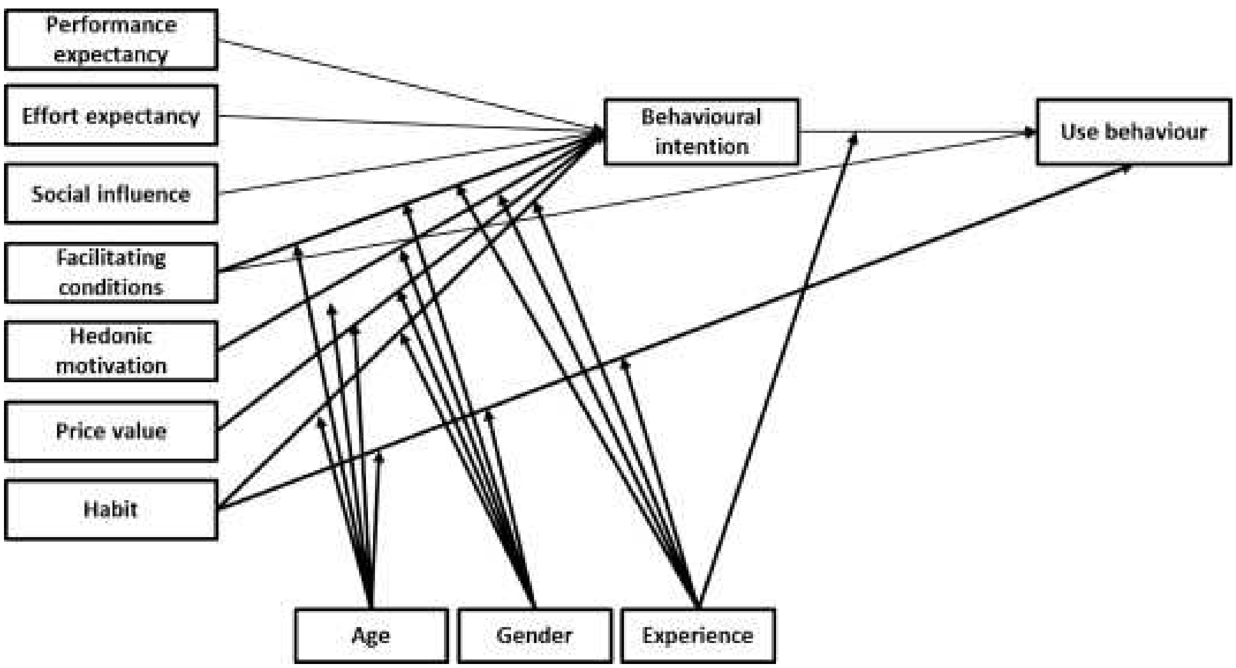



*Figure 8 - UTAUT2 model (Venkatesh et al., 2012).*

## 2.6. Conclusion

To summarize, Chapter 2 explained the theory on which the research is based. The definition of what can be considered a Smart city is defined as well as the two different approaches for designing smart cities applications, namely top-down and bottom-up. After that, the notion of playable cities according to (Nijholt, 2017) and some case studies have been presented. The two different interpretations of pervasive games have been explained, one including technology as a prerequisite and the second general and applicable to all the types of games. Particular attention has been paid to the Pokémon Go case study and the discussions that started with its huge popularity.

The last two subsections focused on explaining the two independent variables that are evaluated in the research, namely user engagement and technology acceptance. For these two topics, have been presented the available methods for their assessment and the rationale behind the choice of adopting them in the research. In particular, for the user engagement is used the UES-SF questionnaire and for evaluating the technology acceptance, the UTAUT2.

# 3. CONTEXT OF THE RESEARCH

This chapter explains the context of the research in terms of the general project to which this thesis contributed. First, the objectives and the vision of the project are presented. After that, an overview of the pervasive game in all of its aspects is explained. In particular, the historical scenario of the game, the user interface, the game mechanics, the technical aspects and the user-game interactions are addressed.

## 3.1. The PACE Project

PACE (Pervasive and Affordable technologies for Civic Engagement) is a project[9] funded by the Spanish Ministry of Economy and Competitivity (TIN2016-77690-R) that has as its main objective to create interactive systems that eventually engage citizens in local activities that might imply coproduction of knowledge and informal learning. In general, the project is investigating what are the best means to engage them in the urban context with ubiquitous and augmented technologies.

The PACE project has two different planned case studies. The first focuses on how to integrate citizens' information about early warnings into an efficient process (Díaz et al, 2017). The second one is situated into the smart and playable cities context (explained in Section 2.1 and 2.2) and aims at exploring the use of pervasive games for informal learning about urban environments (Sánchez-Francisco et al, 2018). This research lies in the second scenario. In general, the PACE project target is limited to affordable technologies from a cognitive, usability, and socio-economic point of view. In addition to this, the chosen technology needs to be perceived as acceptable by the final user.

## 3.2. A Pervasive Game for Urban Environmental Awareness

Environmental Education is a notion created by (Stapp, 1969). It highlights how citizens that are aware and educated about the environment around them are also more motivated in living in an optimal way in the environment itself. In addition to this, it has been proved that GPS-based games can improve the motivation of people to learn about the surroundings (Bursztyn et al., 2017) and it has been investigated how pervasive games can be used for urban

[9] dei.inf.uc3m.es/dei_web/dei_web/index.php?page=projects&id=100

exploration (Montola et al., 2017). As explained in Section 2.3.2., pervasive games are meant to extend the physical world with digital elements. It has been proved in (Kasapakis et al., 2013), that pervasive games can actually create a connection with the physical environment through digital interactions. Moreover, it was found that in order to place digital information in the physical environment (Liao & Humphreys, 2015) as well as form social practices in the urban life (Bursztyn et al., 2017), Augmented Reality (AR) is a valid mean. According to these investigations, in the PACE Project a pervasive game in the form of a smartphone application has been developed in the context of smart and playable cities. The pervasive game is described in (Sánchez-Francisco et al., 2018). It has the purpose of promoting awareness about the urban environment through a pervasive playful experience, that aims to develop knowledge about the history of the university. In particular, about one of the buildings of the university campus (Sabatini building) that is an old military headquarter built by the architect of King Carlos III, Francesco Sabatini.

### 3.2.1. Scenario of the Game

In order to explain the history of the building and show what was its role in the past, the pervasive game presents the scenario of the Spanish Independence war (Uprising of 2nd of May) that occurred in 1808. In 1808, the Sabatini building was occupied by the French Hussar Regiment and the civilians from Leganes (which is the city where the university campus is located), decided to assault, set the building free from the French occupation and helped the rebellion. The scenario is crucial in order to explain the game and the questions asked during the game. In fact, the users play the role of the citizens of Leganes that are helping to assault the building. In order to help in the incursion, the player needs to collect several pieces of the cannon all around the university, mount the cannon and then shoot it towards the Sabatini building to help the insurrection.

### 3.2.2. Game Mechanics and User Interface (v1)

At the beginning of the game, a short introduction is presented to the players in order to explain the historical scenario, the instruction on how to play and the goal of the game. Right after, users need to start walking towards the first cannon piece which is visible from the map of the mobile application, but not in the physical environment (see Figure 9A). The cannon pieces are scattered outdoor all around the Sabatini building. There is no defined order to collect them, the players decide on their own how to explore the surroundings.

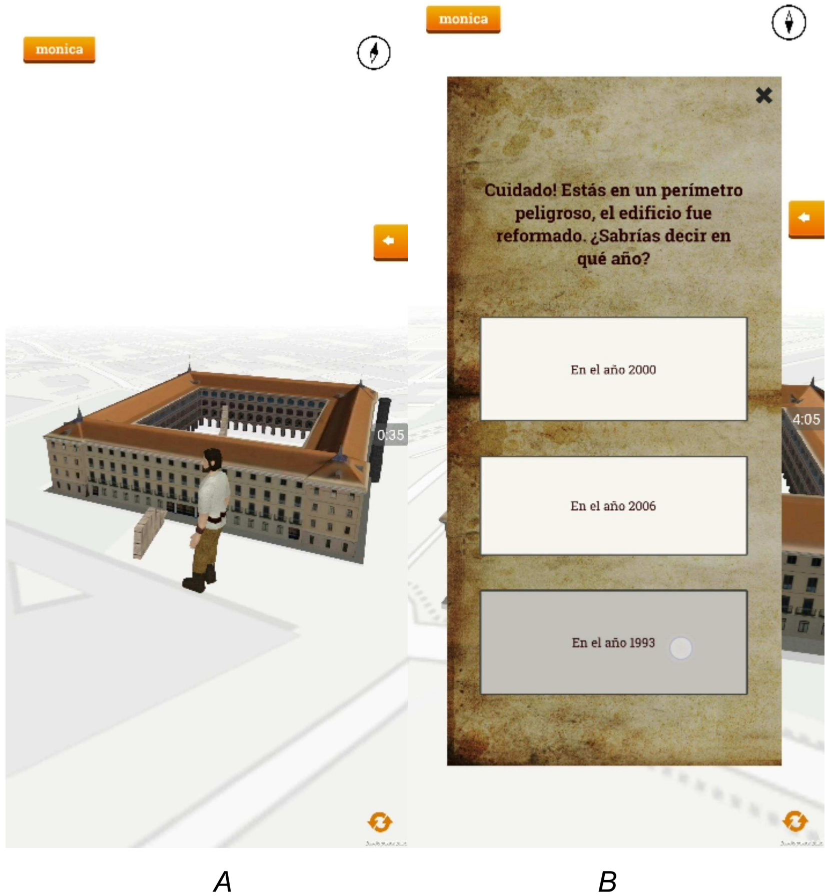
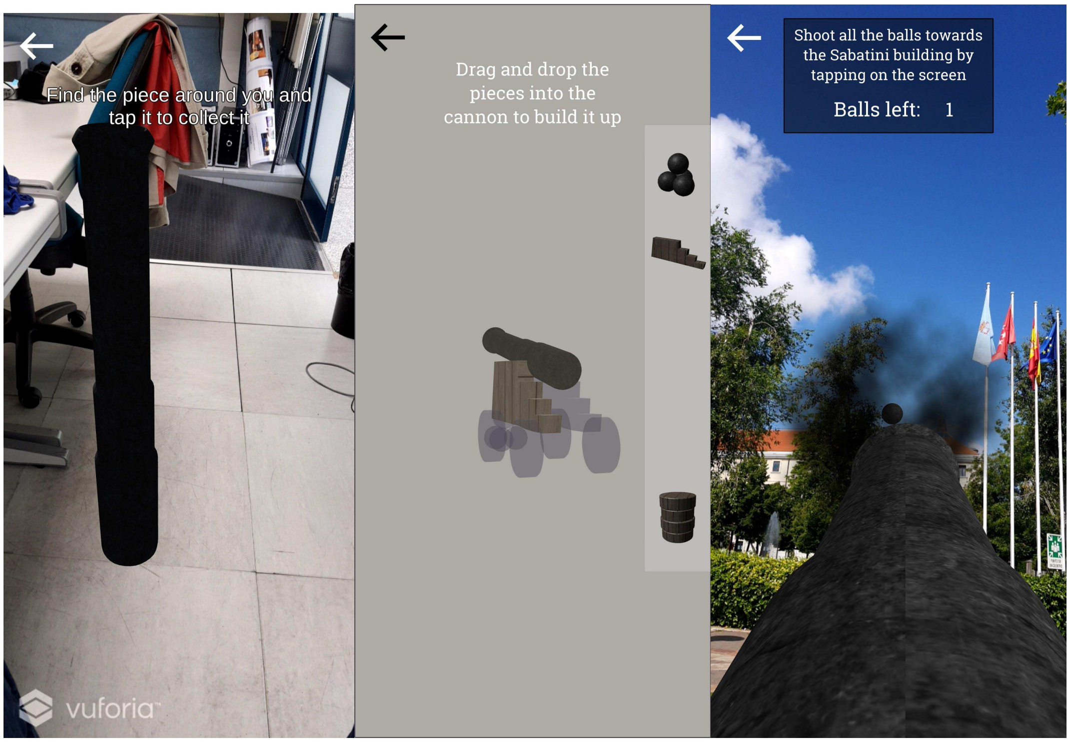



*Figure 9 - Screenshot of the user interface of the game (v1).*

Once players are close enough to the area of one piece, they can tap the virtual object on the screen in order to collect it. A question concerning either the historical scenario or a historical fact about the building itself will appear and the player needs to answer (Figure 9 B). After choosing the answer, a short text that explains the question and the correspondent historical event will appear to promote knowledge in the user.

In this phase, the players have not collected the object yet, they need to gather it with the Augmented Reality view that opens in the mobile screen, where the user sees the virtual cannon piece in the real environment (see Figure 10 A). The virtual object is arranged as a layer in the capture of the camera and it's placed randomly around the players who are asked to detect it and tap on it to finally collect it. The pieces are stored in the inventory and they are accessible through the orange button placed on the right side of the screen (Figure 9A). This process needs to be repeated for five times, that is the total number of the cannon pieces.




A B C

*Figure 10 - Screenshot of the Augmented Reality interface (v1).*

When all the pieces are collected, users need to walk to the shooting point, which is a default spot in the garden of the campus with a clear view of the Sabatini building. Once there, the players need to compose the cannon (Figure 10 B) and eventually shoot to the building with the AR view (Figure 10 C). When the player has shot all the three cannonballs, the mission is accomplished, and the player has helped the insurrection. An explanatory video has been created in order to show the process and the interactions of the game[10].

In the game design, particular attention has been paid to the aesthetic of the game. In fact, (O'Brien & Toms, 2010) defined the Aesthetic as a factor that influences the final user engagement while using a system. Hence, the Aesthetic of the user interface was strongly inspired by Pokémon Go, which is considered one of the most successful examples of pervasive game, thus with a high level of engagement (Zach & Tussyadiah, 2017).

[10] https://drive.google.com/drive/folders/1q3n6MghDqKL8VDn5ZNUtoF5J5KXsRcE4

### 3.2.3. Technical Specifications

The game has been developed with Unity version 2017.3.1F1, it runs on a Samsung Galaxy S8 owned by the laboratory running an Android version 8.0.0. The game requires a fast and stable internet connection because of the constant update of the GPS coordinates, which determine the position of the avatar in the map according to the movements of the user. Due to the impossibility of having a sim card exclusively for the experiments and the instability of the university Wi-Fi, the phone was connected with the tethering to another smartphone. In order to display the AR screen for collecting the object, the game uses a Vuforia[11] , which is a plug-in for Unity which allows to include the augmented reality feature in the camera as well as customize it.

The game stores the data of the game in the database as JSON files. In particular, it saves the username and the GPS coordinates (i.e. the walked path) with the correspondent time and date. In addition to this, all the answers to the historical questions are stored. All the stored data in the database is totally anonymous and not attributable to users.

The game is available in two different languages, namely Spanish and English. In the first screen of the application it can be selected according to the user's preferences. The game is not available in the Play Store; it is downloaded only on the smartphone of the laboratory for research purposes. In particular, the game will be used to explore which features make it engaging and under which circumstances.

## 3.3. Conclusion

To summarize, Chapter 3 gives a preliminary explanation of the state of the research with a focus on the version (v1) of the pervasive game which will be evaluated in the study explained in Chapter 4. In particular, the game design and the rationale behind the game and its development have been addressed. The player-game interactions have been explained as well as the technological settings of the game for urban awareness.

[11] www.vuforia.com/

# 4. USABILITY TESTING

This chapter addresses the usability study that has been carried out during the research (Fabiano et al., 2018). In general, the research aims to explore to what extent the pervasive game was engaging the users and contribute to the creation of a theoretical framework for the evaluation of the engagement in pervasive AR games. However, in order to do that, a prior assessment of the usability is necessary because the game has never been used by the users.

The usability evaluation that has been carried out had a three objectives. Firstly, the experiment was aimed at assessing the usability of the system as a factor that influences the user engagement as stated by (O'Brien & Toms, 2009). The authors confirmed that the system's perceived usability is an indicator that eventually helps to determine the engagement, that is if the technology is perceived as requiring effort in the usage, this will have a negative effect on the level of engagement. Therefore, this usability evaluation addresses not only whether the game is comprehensible and usable without any explanations, but also aims to discover new technical problems and solve them before the evaluation of the user engagement. Bugs and issues in the usability would play the role of external factors and would ruin the whole experience, thus negatively influencing the engagement of the user.

Second, this test was meant to check the safety of the game. Since the pervasive game is taking place in a physical environment and the player was asked to move around while immersed in a different activity, it is important to check whether the mobile application is safe for the users. For instance, the Pokémon Go case study (Zach et al., 2017) highlighted some drawbacks concerning the security of players (e.g. the game was used to locate people in isolated places and then steal their phones (Yuhas, 2016)). The safety is considered both in terms of accessibility of the spaces of the playground (e.g. presence of architectural barriers or problems in the navigation) and in terms of potential dangers during the experience.

Finally, the last goal of this evaluation consisted in gathering feedback and impressions on the game from users in order to improve the whole experience or eventually develop new features.

In conclusion, we can summarize the purpose of this usability study with 3 sub-research questions:

1. **sub-RQ1.** To what extent is the system perceived as usable by the users?
2. **sub-RQ2.** What are the features of the game that can be improved?

3. **sub-RQ3.** Is it safe to play the pervasive game in the chosen playground?

The next sections describe the methodology, results, and findings of the evaluation. The chapter ends describing the second version of the pervasive game (v2) that included improvements to deal with the flaws and issues detected during this evaluation.

## 4.1. Methodology

### 4.1.1. Procedure

Before starting with the experiments, a pilot experiment with an expert user (i.e. a professor of Human Computer Interaction) was carried out to check if the app was working smoothly, the data gathered during the study were saved correctly and, if the procedure of the experiment was well planned. After checking that everything worked fine and changing minor things, such as misspellings in the text and some output texts that were missing, the usability testing started.

Participants were equipped with the Samsung Galaxy S8 of the laboratory with the game installed. Subsequently, they were asked to complete one round of the game without any extra help in order to evaluate objectively to what extent the app is easy to use, usable and accessible. Afterwards, participants completed the questionnaire in the laboratory and they were asked if they wanted to add some questions or remarks about the experience or the game itself.

### 4.1.2. Participants

A total of 5 participants were recruited through convenience sampling inside the computer science department, which means that all the people were chosen by a non-probabilistic sampling, that is because they were a convenient source of data (Lavrakas, 2008). A sample with a limited size of 5 users has been chosen according to (Nielsen et al., 1990), which showed that this amount is necessary in order to find a reasonable amount of usability problems. Moreover, all the recruited participants can be considered as expert users with a high familiarity with technology and knowledge of usability guidelines, as a matter of fact, they were all Master (n = 2) and PhDs students (n = 3) in computer science, specializing in Human-Computer Interaction. This allows us to find a bigger amount of usability issues instead of carrying out the experiment with more not expert.

### 4.1.3. Materials

For this experiment a mixed-methods approach has been chosen, which consists of two diverse mechanisms to collect information, namely a questionnaire and a qualitative analysis of the users' behavior.

Firstly, the concurrent think-aloud protocol (Charters, 2003) has been used for collecting qualitative data about the user behavior, where participants are asked to say whatever comes to their mind while they are playing the game and interacting with the mobile application. This protocol helped to collect as many comments as possible, not only about the experience as a whole but also on the details of the user interface and to see whether the explanations within the app were clear. During the whole test, the participants' voice and the screen have been recorded using the "AZ Screen Recorder"[12], which saves easily the recorded audio and the screen into a video format and, moreover, highlights the taps and gestures of the users in the video recordings.

The questionnaire used includes the System Usability Scale (SUS) questionnaire (Brooke, 1996) and 8 extra questions for collecting feedback on the game experience (see Appendix A). The SUS questionnaire has the purpose to provide a quick but reliable measurement of the perceived usability of the mobile app in the form of a 100-points score (Bangor et al., 2008).

Besides of the SUS questionnaire, eight extra questions were added in order to investigate the user interface and the features of the application (see Appendix A). The questions requested an answer in the form of a short text, so that participants could express themselves freely. Specifically, they had three diverse purposes: an evaluation of the user interface, investigate on adding extra features in the game and to assess the clarity of the game.

For the user interface it was asked whether the sizes of the text and buttons were big enough to interact with them. For the features, users could express if they wanted to add or remove some feature of the game. In order to check the clarity of the game, they were asked if the text descriptions were understandable and in what situation they felt confused in the game.

[12] play.google.com/store/apps/details?id=com.hecorat.screenrecorder.free&hl=en

### 4.1.4. Data Processing

All the experiments were carried out in Spanish to make participants feel comfortable while describing their feelings on what was happening. Therefore, all the cited comments made by the users in this thesis have been translated into English. Subsequently, all the video recordings have been watched and analyzed separately and for every user all what was said by the participants was transcribed. In fact, during the whole experiment, the sentences concerned either how to improve the experience or the discovering of bugs within the game were considered as potential insights. Moreover, I noted all the phrases spoken by participants that were a symptom of confusion or uncertainty on the use of the application that can eventually lead to usability problems. For instance, sentences like *"What's happening now? Should I click on it?"* (User 1) or *"So? Now what I need to do?"* (User 3) have been noted down. For every noted sentence, the moment of the interaction or the category of the interaction has been added in the transcripts. For instance, User 5 stated *"in all the questions, when the background is dark there isn't a good contrast for the reading"* which belongs to the category "text". Moreover, in the transcripts have been added and categorized the eight extra questions of the second part of the questionnaire as well as all the notes taken during the game.

As far as the SUS questionnaire is concerned, scores have been transformed into a 100-score following the method by (Bangor et al., 2008).

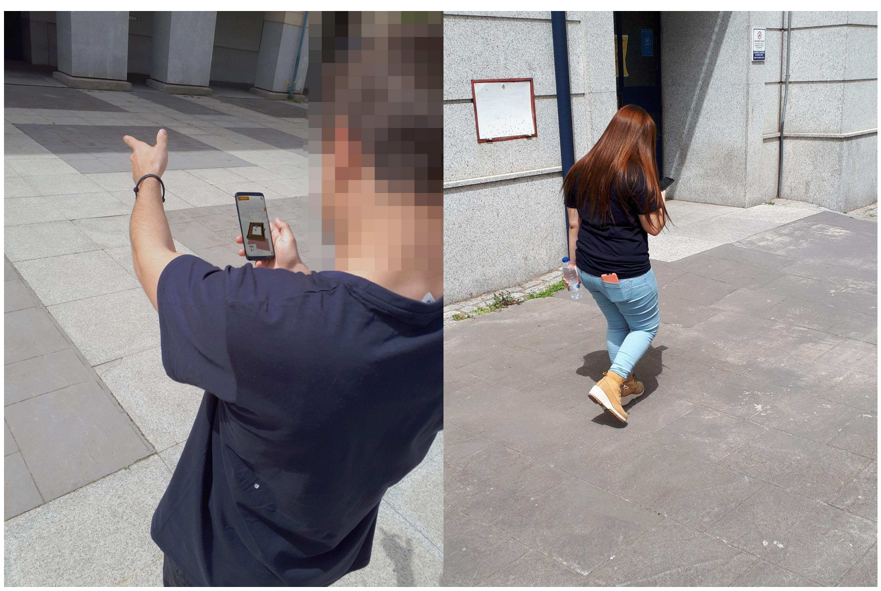


*Figure 11 - Participants playing the game and orienting in the university campus.*

## 4.2. Results

### 4.2.1. System Usability Scale (SUS) Questionnaire

In Figure 12 are represented the scores of the SUS Questionnaires filled in by participants after trying the game. All the scores of the five users are equal or higher than 82.5 (AVG = 88,5/100; SD = 6.275), thus, the perceived usability of the game can be considered as acceptable. In fact, as depicted in Figure 13, the minimum threshold in order to consider a technology usable has been fixed at 70/100 by (Bangor et al., 2008). Moreover, according to the same scale (Figure 13), the perceived usability of the game can be considered as "Excellent".

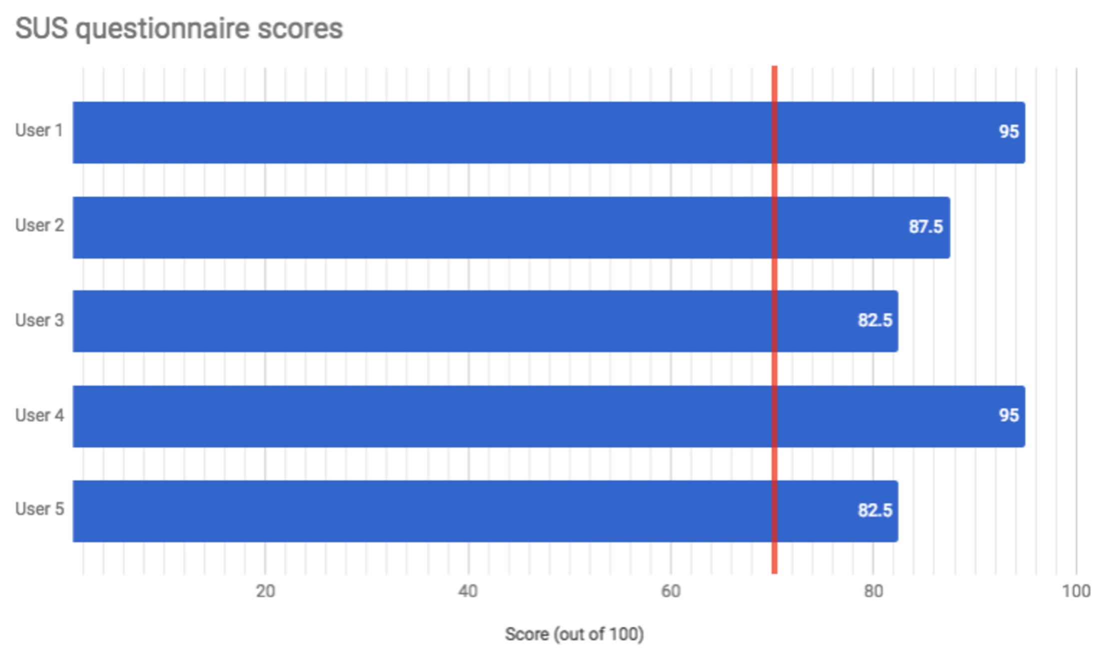



*Figure 12 - Scores of SUS questionnaire, red line for indicating the acceptance threshold (70/100).*

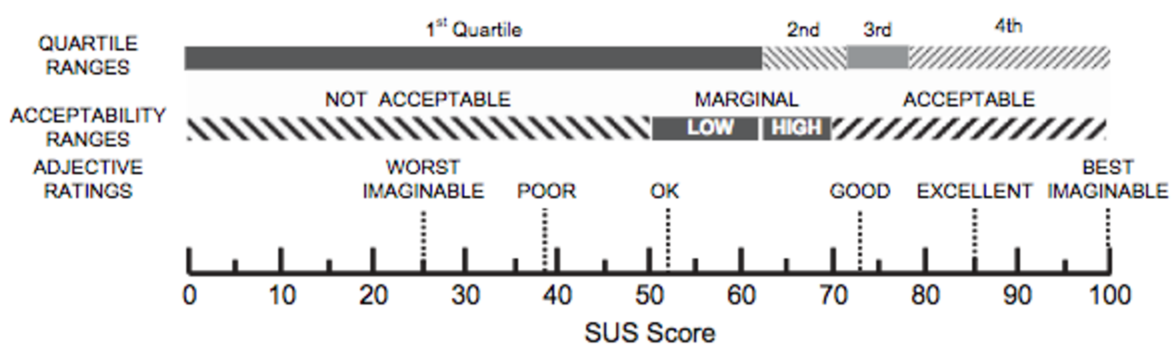



*Figure 13 - Scale of the scores of the SUS questionnaire, retrieved from (Bangor et al., 2008).*

### 4.2.2. Users' Behavior Analysis

The method used for the qualitative analysis led to summarizing the insights and the issues by five diverse categories:

1. **Orientation.** While walking around the university campus, three users (User 1, User 3, User 4) had difficulties in orienting themselves in the game and walk in the correct direction. In general, users found difficult matching the direction of the virtual character with the real position and orientation.
2. **Collection of pieces.** This step consists in tapping the virtual object displayed on the map. Some users (User 1, User 3, User 5) found it hard to distinguish the virtual piece from the model of the building or to figure out how they were supposed to collect it for the first time.
3. **Text.** User 1, User 2 and User 5 expressed frustration when reading the text. In some screens it wasn't big enough or hadn't enough contrast with the background.
4. **Instructions:** In general, users felt lost too many times while playing. Frequently, users were waiting like they were expecting an instruction from the game on what to do and they felt confused. This difficulty was due to the lack of guidance by the game itself.
5. **Position Update.** With all the users, there was a complication with the movement of the avatar in the game. Due to the slow connection and the inaccuracy of the GPS coordinates, updating the position in real time didn't work smoothly but it required some time while the participants were walking. Moreover, participants explicitly expressed their discontent about the issue, both during the test and in the following questionnaire. Nevertheless, this technical bug wasn't an obstacle in order to conclude the game. The problem occurred only in a specific position in the building, in fact, in that area of the university the connection is very fragile, and it affected the signal of the smartphone.

The usability study was meant to evaluate whether the game was safe, that is to check if the spaces were accessible and if the participants were exposed to dangers during the game experience.  The game didn't present any problem or danger in the outdoor environment during the whole duration of the study. The playground, where the pervasive game was played, consisted of a closed area where the access to vehicles is limited. In addition to this, the space was accessible without any architectonic barrier.

## 4.3. Discussion and Identification of User Requirements

The previous sections have presented the methodology used in order to evaluate the Augmented Reality pervasive game for urban awareness and its results. From these results can be reported how users responded to the usability of the system and to the overall experience. Even though the scores of the SUS questionnaire demonstrated a good perceived usability of the game, an in-depth qualitative analysis indicates that the system can be improved in several ways. In order to transform the comments stated by the users into practical user requirements, different clusters have been created in order to merge insights according to their category (the full list of user requirements is available in Appendix B). These items were turned into new features aiming to tackle already identified problems in the second version of the game. The user requirements, which have different priorities according to the number of users that mentioned the problem or the entity of the problem, are meant to fix technological problems and improve the former version of the mobile application, that is making it easier to user and more understandable. Moreover, it improves the overall experience adding new features and guidance in the game.

Table 3 describes the elements that have been changed from the first version and how the usability issues have been tackled.

*Table 3 - Overview of the issues and the correspondent solution developed in v2.*

| **Issues** | **Solutions** |
|---|---|
| Orientation | In order to help users finding the direction to follow in the game, a top-view little map has been added in the bottom left section of the user interface (see Figure 14). In the little map, the users are represented as dots and they can see where they are pointing the smartphone. This feature was already present in the Google Maps mobile application and it has been adopted because of its high popularity and its familiarity for |

| | |
|---|---|
| | the users. Moreover, it has been mentioned by User 4 *"I would like to have a view like Google Maps that tells you where you are pointing the phone"*. The map shows also the cannon pieces that the user still need to collect (see Figure 14) |
| Collection of Pieces | The pieces are not motionless anymore. In order to make them more distinguishable from the Sabatini building, in the second version of the game they are floating and rotating on themselves to differentiate the virtual 3D model of the piece from the model of the building. |
| Text | The size of the text has been increased and a higher contrast between the background and the text color has been adopted. |
| Instructions | Whenever an action is required from the user, an instructional text is now prompted in the interface to guide the users, especially in between the interactions. |
| Position Update | Due to the impossibility of solving this issue in a technical way, an in-game solution has been designed which consists in making the problematic area inaccessible in the scenario of the game. In fact, in the version 2, the map displays a red area where the users must not walk through because it is full of enemies, thus they needed to walk around it and avoid the area with low signal (see Figure 14). |

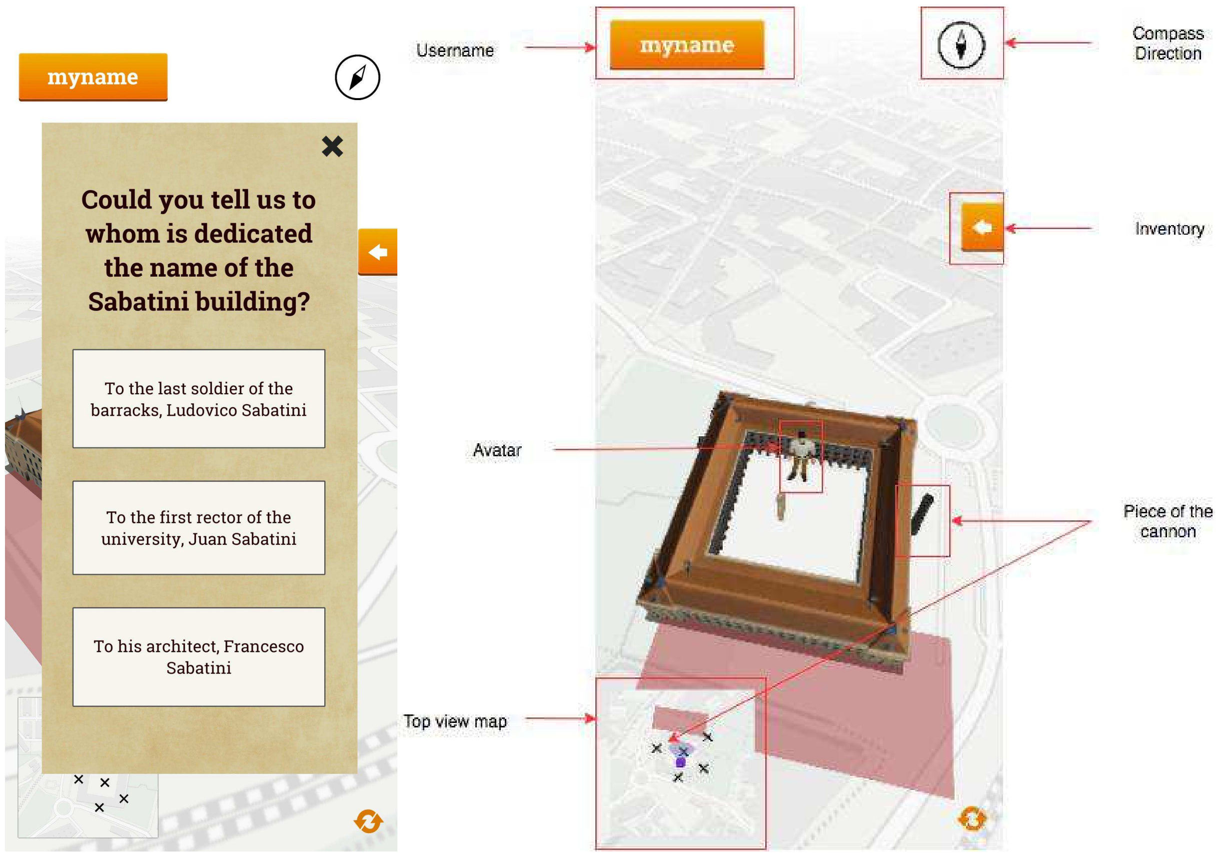



*Figure 14 -User Interface of the second version of the game.*

In conclusion, the sub-research questions of this usability study can be answered as follows:

1. **Sub-RQ1.** According to the SUS questionnaire results, the perceived usability has been evaluated as "excellent", with the average of the scores of 88,5/100 (SD = 6.275).
2. **Sub-RQ2.** Even though the game was perceived as usable, the game and its interactions can be improved in several ways. In order to answer this sub-research question, a list of user requirements for the following version of the game was created.
3. **Sub-RQ3.** While carrying out the experiment, no architectonic barrier or potential danger has been found during the game. Therefore, we can infer that the playground is safe for the players.

# 5. EVALUATION OF THE USER ENGAGEMENT

The usability of the pervasive game has been evaluated and it has been proved that is over the acceptable threshold, moreover, all the technological issues have been fixed in game v2. Thus, the next step is to evaluate to what extent the pervasive game v2 is engaging and if the technology used is accepted by the users.

In general, the experiment wants to contribute to the creation of a theoretical framework for the evaluation of the engagement of pervasive games in the context of playable cities. The experiment aims to investigate in three different directions. First, the evaluation of the User Engagement and the Technology Acceptance. Second, the assessment of the Perceived Walked Distance as a factor related to the User Engagement. Finally, the Urban Environmental Awareness in two different interpretations. On the one hand whether the familiarity with the urban environment impacts the final user engagement. On the other hand, if being more aware of the urban environment is related to a higher user engagement.

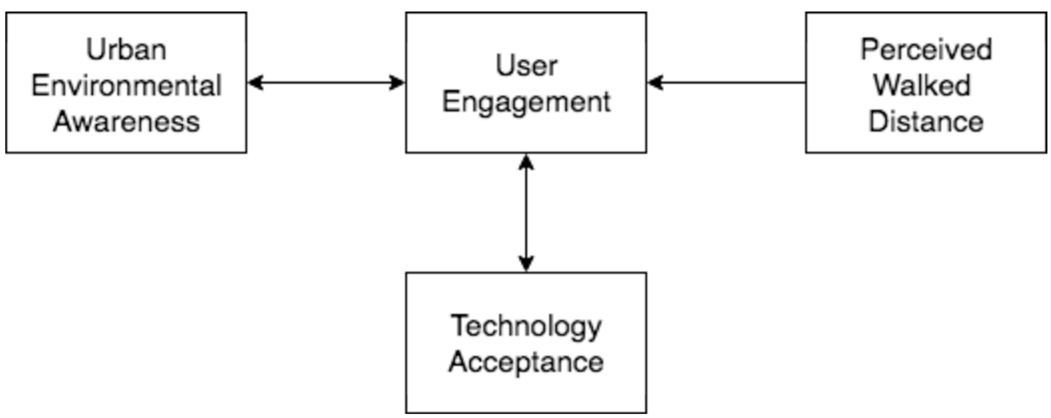



*Figure 15 - Overview of the experiment for the evaluation of the User Engagement.*

The goal of the game is to engage citizens into a playful and informal activity with the use of an augmented reality pervasive game. However, it's still unknown whether the pervasive game is eventually engaging the users and whether the technology chosen for engaging users is accepted by them. Therefore, the first focus of the experiment is to assess as independent variables the user engagement (RQ1) and the technology acceptance (RQ2 a). Besides these two independent appraisals, the investigation aims to establish whether there is a link between the user engagement and the technology acceptance (RQ2 b).

The two remaining aspects of the experiment are meant to explore if some indicators can be defined reflecting the level of engagement of the users. Finding new factors would help to create a better definition of the user engagement, especially in the context of pervasive games for playable cities. In particular, it's supposed that two different factors might reflect the user engagement, namely Perceived Walked Distance and Urban Environmental Awareness. In the definition of the user engagement, O'Brien et al. showed that one of the determinants of the engagement is the Perceived Time (O'Brien et al., 2008). This means that users that are engaged in an activity are also underestimating the flowing of time while accomplishing the task. Moreover, in the context of pervasive games, it has been proved that pervasive games can increase the daily walked distance when the user is engaged while playing (Howe et al., 2016). According to these two works, it's possible to infer that the perceived walked distance might be an indicator of the engagement of the users playing a pervasive game.

As far as the Urban Environmental Awareness is concerned, the theory has been already explained in Section 3.2, where the combination of knowing more about the surroundings (Stapp, 1969) and using pervasive games can improve the users' motivation (Bursztyn et al., 2017), and thus might keep them more engaged. In the context of pervasive games for the urban awareness, it is important to assess to what level the familiarity with the "playground" influences the engagement of the users. This factor has two facets that need to be considered. On the one hand, the investigation aims to know who the best target users for the pervasive game are. The definition of playable cities uses applications for all the city dwellers, that is both visitors and locals (Nijholt, 2017), where the differentiation between the two categories lies on the level of familiarity with the surrounding environment. On the other hand, the research has the purpose of exploring whether the urban environmental awareness, in terms of familiarity of the users with the urban environment, can be considered as an indicator of the user engagement in pervasive games.

In the following chapters are addressed the methodology used in this experiment, the results, and the findings.

## 5.1. Methodology

### 5.1.1. Study Design

In this study, two different types of user have been involved using the same condition. On the one hand, people who are not familiar with the university campus, that is that have never been there or just a few times. On the other hand, people that are considered familiar with the environment, namely people that are in the university campus on a regular basis. The recruitment followed a convenience sampling approach and the unique requirement for being recruited was to consider the familiarity with the environment in order to have a balanced number of participants for each group.

Differently, for the evaluation of the user engagement, the technology acceptance of the pervasive game, and for comparing the perceived walked distance and the engagement, all the participants have been considered as one single group, without differentiating the type of the user.

### 5.1.3. Participants

A total of 27 participants has been recruited for this experiment. 13 users were not familiar with the university campus - an international fair hosted in the university campus allowed to recruit 7 of them and the rest have been convinced to come on campus on purpose for the experiment. The remaining 14 users were familiar with the environment and they were all either working or studying in the university. The familiarity with the university campus has been tested before the beginning of the experiment and double checked with an item in the questionnaire. In fact, the question D4 (see Table 4) has the purpose of categorizing participants according to the number of times they have been on campus. This approach has been effective in order to detect one user that stated to be Not familiar but, in the questionnaire, selected the answer "I come here regularly (e.g. every week)", thus he/she has been categorized as Familiar.

Among the 27 participants, the gender and the experience with augmented reality apps were well balanced. According to the gender of the participants, 15 users were males (55.6%) and 12 females (44.4%). Concerning the experience, 15 users were beginners, that is that they have used this type of Augmented Reality games less than 6 times before the experiment, and 12 were expert users. However, if these two variables are analyzed by the two groups, the gender is not well balanced. Within the Familiar users, 10 are males and 4 females. Concerning the Not familiar, 5 users are males and the remaining 8 participants are females. The experience within the Familiar is perfectly balanced (7 experienced users and 7 beginners), instead, in the Not familiar group, 5 were experienced and 8 participants used Augmented Reality games less than 6 times. Most of the users fit in the age range of 18-24 years old (n = 16), 8 participants are in the range of 25-34 years old, 2 users in the 35-50 range and only one was under 18 years old, turning 18 shortly after the experiment.

### 5.1.2. Procedure

Once the participants agreed to take part in the experiment, they were asked to sign an informed consent form which states that their attendance was voluntary and the collected data was totally anonymous. Moreover, they were asked if they wanted to give the permission to take pictures during the experiment, which eventually would have been anonymized. Afterwards, users were provided with a Samsung Galaxy S8 already set-up for playing and they were asked to complete one full round of the game which took place in the Sabatini building, one of the university buildings. They all started in the same location of the building in order to prevent any bias on the first object to pick and for consistency in the measurement of the path length. While participants were playing the game, the Pedometer app[13] was running in background for calculating the exact walked distance, thus it was not visible to the player. Once the game was completed, the real walked distance was saved with a screenshot of the Pedometer app and, subsequently, the distance was erased for the following participant.

Right after the participant completed the game, he or she was asked to answer a single question printed on paper, that is the estimation of the Perceived Walked Distance (see Figure 18), and to fill in the remaining part of the questionnaire on a Samsung Galaxy Note 10.1 tablet. To save all the answers of the questionnaire Google Forms has been adopted.

The collected data is totally anonymous and not attributable to the participants. Moreover, no sensitive information has been asked to participants. If considering the introduction of the General Data Protection Regulation on the 25th of May 2018, all the data has been gathered before the implementation date.

[13] https://play.google.com/store/apps/details?id=com.tayu.tau.pedometer&hl

### 5.1.4. Material

The method chosen for the assessment of the experiment is a self-reported measurement, namely a questionnaire composed by several parts with diverse objectives. The sub-questionnaires described in this section are five, namely Demographic, User Engagement, Technology Acceptance, Urban Environmental Awareness, and Perceived Walked Distance.

#### *Demographic Questionnaire*

The first section of the questionnaire consists of five general demographic questions. These items are expected to collect some background information about the participants. The questions asked are covering the moderators shown in the UTAUT2 model (Figure 8), namely Age and Gender of the participants and their Experience with the type of technology used (Venkatesh et al. 2003). Moreover, in this section it was added a question intended to be a double check concerning the familiarity of the user with the environment (question D4 in Table 4). Table 4 shows the items included in this section of the final questionnaire. The item D3 collects information about the highest level of education of the participant.

*Table 4 - Questions of the sub-questionnaire about the demographic information.*

| ID | Items |
|---|---|
| D1 | What is your gender? |
| D2 | What is your age? |
| D3 | What is the highest level of education you have completed? |
| D4 | How many times have you been in this campus before today? |
| D5 | How many times have you used an Augmented Reality app before today? |

#### *User Engagement*

In order to assess to what extent a technology is engaging the user, diverse methodologies are available, tested and proved to be reliable measurements. In Section 2.4.2. three different approaches for evaluating user engagement have been described. The chosen method for this research consists in a user-reported measurement, namely the use of a questionnaire. The most common instrument for achieving this evaluation is the User Engagement Scale (UES). However, this study includes multiple factors with questionnaires and using multiple questionnaires leads to a considerably high number of items, thus, a big effort for the participants. For this reason, the UES-SF (O'Brien et al., 2018), which is a shorter and more flexible version of the UES, will be adopted in the final survey. This method has been proved to be shorter but still reliable (O'Brien et al., 2018).

***Technology Acceptance***

In Chapter 2.5 the topic of technology acceptance has been discussed and the reason why the model UTAUT2 was chosen. The proposed questionnaire includes items for 8 different factors of Technology Acceptance, namely Performance Expectancy, Effort Expectancy, Social Influence, Facilitating Conditions, Hedonistic Motivation, Price Value, Habit, and Behavioral Intention (Siow, 2016). However, not all of the factors are coherent with the purpose and the type of the mobile application used in the experiment, thus, they have been included in the final questionnaire only partially. This sub-questionnaire contains only four of the eight factors of the UTAUT2 questionnaire, namely Effort Expectancy (EE), Facilitating Conditions (FC), Hedonic Motivation (HM), and Behavioral Intention (BI). Table 5 summarizes all the changes applied to the UTAUT2 questionnaire (Siow, 2016). All the discarded questions are represented in red and the items with the strikethrough have been replaced with the subsequent item in the same cell. Performance Expectancy, Social Influence, Price Value and Habit were left out because they were not relevant for the purpose of the pervasive game. For example, Price Value was removed from the questionnaire because the game doesn't have a commercial value and a price. Besides these categories, also FC3 was left out from the questionnaire because the pervasive game doesn't communicate with any other technologies yet. BI1 was changed from "*I intend to continue using the game in the future*" to *"I would like to use the game in my life*", which expresses the behavioral intention not in using the game all the time, rather if the users see the application as valuable in some occasions of their life. BI2 was changed from "*I will always try to use the game in my daily life*." to "*I would like to continue to use the game frequently*". It has been rephrased in order to give less restrictions to the word “always”. In fact, the goal of project PACE is not to engage the players constantly, rather engage them in specific circumstances.

*Table 5 - Items of the sub-questionnaire for the assessment of the technology acceptance.*

| ID | Items |
|---|---|
| PE1 | I find the game useful in my daily life. |
| PE2 | Using the game increases my chances of achieving things that are important to me. |
| PE3 | Using the game helps me accomplish things more quickly. |
| PE4 | Using the game increases my productivity. |
| EE1 | Learning how to use the game is easy for me. |
| EE2 | My interaction with the game is clear and understandable. |
| EE3 | I find the game easy to use. |
| EE4 | It is easy for me to become skillful at using the game. |
| SI1 | People who are important to me think that I should use the game. |
| SI2 | People who influence my behavior think that I should use the game. |
| SI3 | People whose opinions that I value prefer that I use the game. |
| FC1 | I have the resources necessary to use the game. |
| FC2 | I have the knowledge necessary to use the game. |
| FC3 | The game is compatible with other technologies I use. |
| FC4 | I can get help from others when I have difficulties using the game. |

| | |
|---|---|
| HM1 | Using the game is fun. |
| HM2 | Using the game is enjoyable. |
| HM3 | Using the game is very entertaining. |
| PV1 | The game is reasonably priced. |
| PV2 | The game is a good value for the money. |
| PV3 | At the current price, the game provides a good value. |
| HT1 | The use of the game has become a habit for me. |
| HT2 | I am addicted to using the game. |
| HT3 | I must use the game. |
| HT4 | Using the game has become natural to me. |
| BI1 | ~~I intend to continue using the game in the future~~ I would like to use the game in my life. |
| BI2 | ~~I will always try to use the game in my daily life.~~ I would like to continue to use the game frequently. |

***Urban Environmental Awareness***

In order to explore to what level the urban environmental awareness has consequences on the engagement of the user, two different approaches for two different interpretations of urban environmental awareness are embraced.

First, the experiment has been carried out with two different groups of users and then the results of user engagement and technology acceptance have been compared. In the case of this experiment, the user engagement of the two sets will be compared and if there is a statistical significance we can thus infer that the urban environmental awareness is influencing the engagement depending on a specific category of user. The experiment is matching results of users that have never been in the Sabatini building with engagement scores of users that go often there, i.e. are familiar with the "playground" of the game. In this approach, the urban environmental awareness is interpreted as the users' prior knowledge of the environment, that is if the users were aware of the surrounding environment before playing the game.

The second approach consists of a 3-questions survey which is meant to assess the level of urban awareness of the university campus. This approach evaluates whether playing the pervasive game helps the perceived urban environmental awareness and if this factor can reflect the user engagement.

*Table 6 - Items of the sub-questionnaire about the environmental awareness.*

| ID | Items |
|---|---|
| UEA1 | After playing the game I feel more attached with this environment. |
| UEA2 | After playing the game I feel more connected with the university. |
| UEA3 | I would like to use Augmented Reality games to be more aware about the urban environment surround me. |

***Perceived Walked Distance***

As mentioned in the introduction of Chapter 5, we can infer that the perception of the walked distance while playing in the game might be an indicator of the user engagement. Thus, the purpose of this sub-questionnaire is to estimate the perceived walked distance as well as to evaluate the adopted methodology itself. In order to assess how users perceived the walked distance, in the game two different approaches are applied.

First, two questions (PWD1 and PWD2) are present in the questionnaire aiming to investigate what is the perception of the distance walked during the game. The questions explained in Table 7 are included in the final questionnaire. They are evaluated through a 5-point Likert scale (1=Strongly Disagree; 5=Strongly Agree) for consistency with the rest of the questionnaire. They are statements with a negative meaning compared to the hypothesis, so during the phase of evaluation, the score will be inverted.

*Table 7 - Items for the evaluation of the PWD.*

| ID | Items |
|---|---|
| PWD1 | During the game I felt that I walked a lot. |
| PWD2 | Throughout the game I needed to move around too much. |

These two questions are formulated in order to have the same meaning and to be the evidence for the perception of having walked intensely during the game and, further, to be compared with the second approach (PWD3) and evaluate its methodology.

This second approach (PWD3) consists in asking the user to draw a cross on a meter printed on paper right after finishing the experiment. This allows to prevent any bias through the distance between the ending point of the game and the laboratory, where the questionnaire is filled in. The methodology is intended to prove whether the estimations of the walked distance are linked in some way to the score of the UES-SF scores. This approach for the estimation of the distance has been adopted by (Lee, 1970), where users were asked to estimate the distance between two points in the city. In PWD3, a meter with a predefined range of distances in meters is given to the users. In order to know the exact values of the minimum and maximum distance, the shortest and the longest possible paths in order to complete the game have been calculated. The distance has been calculated with Google Maps, which was proved to be a really accurate instrument for the measurement of distances in a map. In fact, to prove this, one of the sides of the Sabatini building has been taken into account for the evaluation; Google Maps measured 90 meters and a mobile application for keeping track of the walked distance (Pedometer), measured 91 meters. Thus, we can state that the Google Maps measurement is reliable with a very little percentage of error ( about 1.11m of error every 100m walked).

After confirming that Google Maps is a reliable measurement, according to the disposal of the cannon pieces, the shortest and the longest paths have been calculated: the shortest measures roughly 460 meters and the longest is about 630 meters. Both are presented in Figure 16 with the Google Maps measurements.

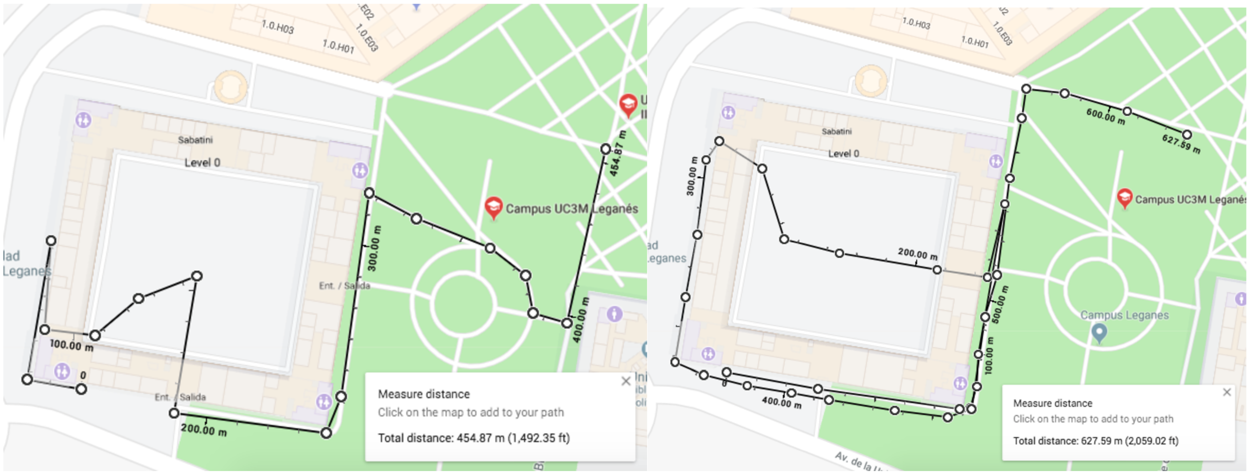



*Figure 16 - Shortest and longest paths in order to complete the game.*

Once the minimum and the maximum values are found, the meter with the real distances can be completed. The central values are the "correct" ones, thus, we leave the chance to overestimate and underestimate the distance leaving an extra distance in the meter. Figure 17 shows the procedure for the establishment of the distances in the range. Subsequently, the distances found have been rounded to multiples of 50 in order to prevent an estimation bias for the users, that is if users see a 672,50m in the meter, they might think that they are accurate values and try the guess. The final question with the final rounded range is represented in Figure 18.

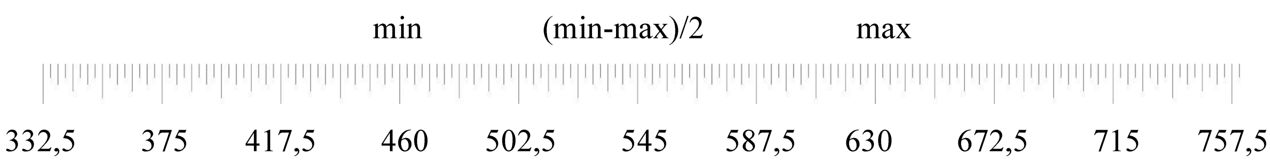



*Figure 17 - Creation of the meter and establishment of the range.*

While the users are playing, the backend of the game is collecting the position of the player in the form of GPS coordinate. A Python script has been created to calculate automatically the actual walked distance. However, the GPS coordinates aren't accurate enough: if compared with the Google Maps distances, the error is 20/30m every 100m walked. Thus, the gathering of the real walked distance will be carried out using Pedometer app running in the background in the Samsung s8 in the hands of the player, which is a more reliable source.

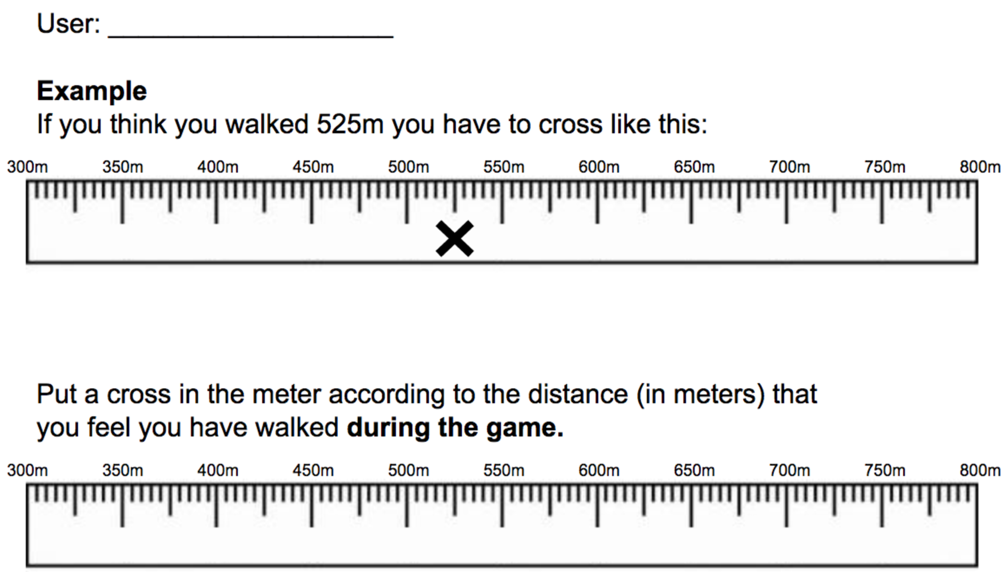



*Figure 18 - Final question for estimating the perceived walked distance.*

Once the real walked distance and the perceived one are stored, it's possible to calculate if the guess of the user is an underestimation, overestimation or if it's the exact value. To the real value has been subtracted the perceived walked distance, thus the result of this calculation indicates the overestimation if the value is negative, the underestimation if positive and the correct estimation if the value is equal to zero. The calculation has been made reversed on purpose according to the hypothesis, where the underestimation should be a positive indicator of the user engagement.

## 5.2. Results

This chapter addresses the results of the experiment and the procedure of the data analysis. First, the evaluation of the user engagement and technology acceptance as independent variables is explained. Subsequently, the results about the urban environmental awareness and the two approaches used in the data analysis. Third, the validation of the methodology of the perceived walked distance is considered as well as the correlation with the user engagement. Finally, the analysis on the results according to the demographics of the participants is described.

### 5.2.1. Evaluation User Engagement and Technology Acceptance

The User Engagement Scale - Short Form provides an assessment technique of the user engagement (O'Brien et al., 2018). The answers of the questionnaire, which are in a 5-point scale (1=strongly disagree, 5=strongly agree), are converted directly as the score of the single question, thus, for each category (i.e. Focused Attention, Perceived Usability, Aesthetic, Reward), every user receives a score out of 5. They are necessary for calculating the UE General (UE General in Table 8), which consists of calculating the score for every category of the user engagement and then averaging the scores of the five categories.

In general, the results of the user engagement are positive (see Appendix E for the full overview of the results). In fact, all the UE General are higher than 3, which is the score of the neutral answer. Even though the average of all the UE General scores is 3,836 (SD = 0,46), if the scores are analyzed by categories they are notably different. The highest scores average belongs to the category Perceived Usability (4,6/5; SD=0,52), moreover this category has overall consistent results. On the other hand, the results for the Focused Attention category are the lowest means with a higher Standard Deviation (SD = 0,87), meaning that the scores aren't very regular, and, 9 results out of 27 are negative (i.e. lower than 3). The remaining categories scored positively with very few negative cases, the Aesthetic average scores is 3,605 out of 5 (SD = 0,67) and the Reward category 3,975 out of 5 (SD = 0,58).

As far as the scores of the UTAUT2 are concerned, the same approach and scale as the user engagement is used for the calculation of the scores (see Table 9).

In general, the UTAUT2 questionnaire has acceptable results. The final scores of the survey (TA General in Table 9) are never negative (see Appendix F for the full overview of the results). The average of the final scores is 4,06 out of 5 (SD = 0,53) which is positive and highly acceptable. Nevertheless, the category Behavioral Intention scored only 3,59 out of 5, which on average is still above the neutral threshold but it has a Standard Deviation of 0,87 (see Table 9). The remaining categories, namely Effort Expectancy, Facilitating Conditions, Hedonistic Motivation, have been perceived positively.

*Table 8 - Scores of the UES-SF and total averages per category.*

| ID user | Focused Attention | Perceived Usability | Aesthetic | Reward | UE General |
|---|---|---|---|---|---|
| Average | 3,160 | 4,605 | 3,605 | 3,975 | 3,836 |
| St. Dev | 0,87 | 0,52 | 0,67 | 0,58 | 0,46 |

*Table 9 - Scores of the UTAUT2 and total averages per category.*

| ID user | Effort Expectancy | Facilitating Conditions | Hedonic Motivation | Behavioral Intention | TA General |
|---|---|---|---|---|---|
| Average | 4,36 | 4,33 | 3,95 | 3,59 | 4,06 |
| St. Dev | 0,53 | 0,51 | 0,73 | 0,87 | 0,53 |

The final scores concerning the user engagement and the technology acceptance are displayed in Figure 19 with a scatter plot, where the horizontal axis consists of the UE General and the vertical one the TA General. Using this representation allows to display the merged results with a graphical visualization. The top-right green square in the graph (see Figure 19) highlights the area where both of the general scores are positive. On the contrary, if a score is in the red area (bottom-left), it means that a user evaluated both of the independent variables below the neutral threshold. Therefore, the Figure 19 confirms again that all the final scores are positive.

The graph in Figure 19 shows also the trendline of the plot, which shows a strong positive correlation (Pearson Correlation Coefficient = 0,764) between the two variables.

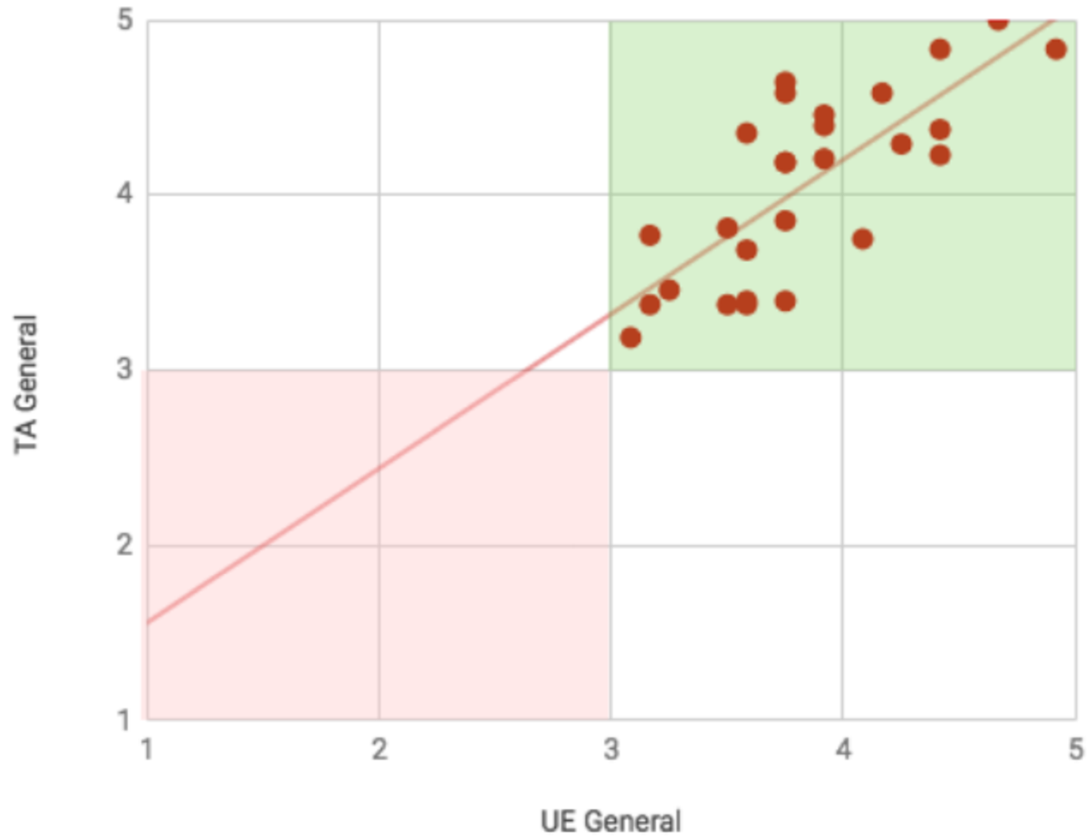



*Figure 19 - Plotted results of User Engagement and Technology Acceptance.*

**5.2.2. Urban Environmental Awareness**

In order to answer to the research questions about the environmental awareness (RQ4 a, RQ4 b), the data analysis followed two different methods. The first approach consists of comparing the two sets of final scores, namely UE General and TA General of the Familiar (F) and Not familiar (NF) samples to see if a statistical significance occurs. Results are shown in Tables 10 and 11, which also display the p-values obtained through the ANOVA test. The two values are above 0,05, thus are not statistically significant. The ANOVA analysis allows to analyze the difference among the means of two samples and it doesn't require to have the size of samples equal.

*Table 10 - Comparison of UE General scores and the p value from the ANOVA.*

| | UE General Familiar | UE General Not Familiar | |
|---|---|---|---|
| F01 | 3,583 | 4,083 | NF01 |
| F02 | 3,500 | 3,917 | NF02 |
| F03 | 3,750 | 4,167 | NF03 |
| F04 | 4,250 | 3,583 | NF04 |
| F05 | 3,167 | 3,750 | NF05 |
| F06 | 3,250 | 3,583 | NF06 |
| F07 | 4,417 | 3,917 | NF07 |
| F08 | 3,750 | 3,500 | NF08 |
| F09 | 4,417 | 3,167 | NF09 |
| F10 | 3,750 | 3,750 | NF10 |
| F11 | 3,583 | 4,417 | NF11 |
| F12 | 3,917 | 4,667 | NF12 |
| F13 | 3,750 | 4,917 | NF13 |
| F14 | 3,08 | | |
| | | | ANOVA |
| Average | 3,726 | 3,955 | P-value = 0,2 |

*Table 11 - Comparison of TA General scores and the p value from the ANOVA.*

| | TA General Familiar | TA General Not Familiar | |
|---|---|---|---|
| F1 | 3,688 | 3,750 | NF1 |
| F2 | 3,375 | 4,208 | NF2 |
| F3 | 3,854 | 4,583 | NF3 |
| F4 | 4,292 | 3,396 | NF4 |
| F5 | 3,771 | 4,188 | NF5 |
| F6 | 3,458 | 3,375 | NF6 |
| F7 | 4,229 | 4,458 | NF7 |
| F8 | 4,646 | 3,813 | NF8 |
| F9 | 4,833 | 3,375 | NF9 |

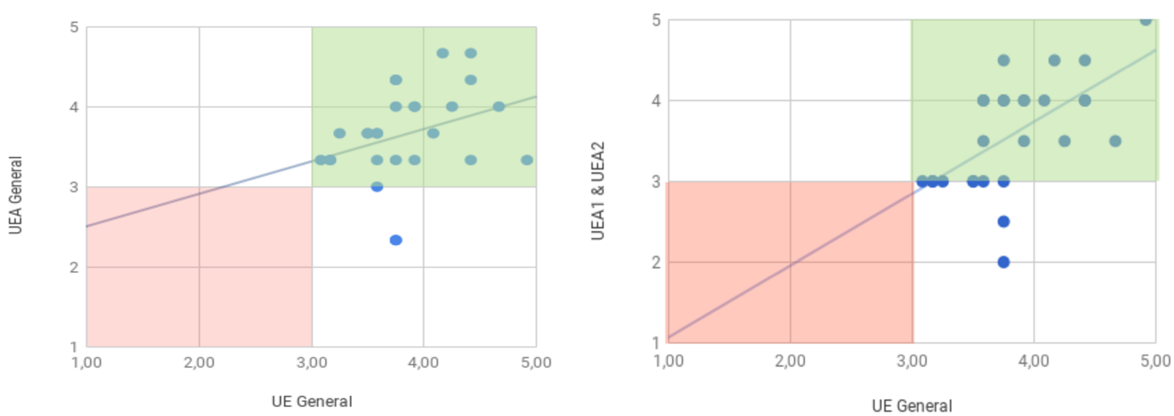
| | | | |
|---|---|---|---|
| F10 | 4,188 | 3,396 | NF10 |
| F11 | 4,354 | 4,375 | NF11 |
| F12 | 4,396 | 5,000 | NF12 |
| F13 | 4,583 | 4,833 | NF13 |
| F14 | 3,19 | | |
| | | | ANOVA |
| Average | 4,061 | 4,058 | P-value = 0,986 |

The second method wants to look for insights on the topic and the relationship between the two independent values. Three different steps are carried out. First, the reliability of the three items of the urban environmental awareness, namely UEA1, UEA2, UEA3 (see Table 6), added in the questionnaire has been evaluated through the calculation of the Cronbach's alpha. Considering all the three items of this category, the value of alpha ($\alpha$ = 0,58) is lower than the acceptability threshold, set at 0.7 by rule of thumb (Kline, 1999; Sekaran & Bougie, 2016). However, checking the correlation between the three questions, it can be noted that the correlation between UEA3 and the other two questions (UEA1 and UEA2) is pretty low. This means that the question UEA3 has been perceived with a different meaning by the users. In fact, it might be closer to a Behavioral Intention item instead of evaluating the environmental awareness and the connection with the environment. In view of the facts, if only UEA1 and UEA2 are considered for the calculation of the internal reliability, the value of alpha increases to 0,65. The calculation of Cronbach's Alpha is presented in Table 12.

*Table 12 - Analysis of the internal reliability of the methodology.*

| | **UEA1** | **UEA2** | **UEA3** |
|---|---|---|---|
| **UEA1** | 1 | - | - |
| **UEA2** | 0,482 | 1 | - |
| **UEA3** | 0,187 | 0,278 | 1 |
| | | | $\alpha$ = 0,58 |

Further, the evaluation of the environmental awareness has been evaluated according to the answers of the survey. If all the three questions are taken into account, the results (see Figure 20 a) are mostly positive. Nonetheless, two users scored are negatively and one neutrally. Subsequently, in order to see if there is a correlation between the user engagement and the environmental awareness, the Pearson Correlation Coefficient is calculated (r = 0,32). Thus, there is a weak positive correlation. Because of the increase of the reliability while considering only the questions UEA1 and UEA2, they also have been analyzed separately without UEA3. In general, if the two situations are compared, the scores of only UEA1 and UEA2 are inferior than if all the three items are considered. In fact, the average of the final scores of UEA1 and UEA2 is 3,59 out of 5, which is lower than the average of the three questions (3,65 out of 5). However, an interesting trend is found: the correlation between UEA1 and UEA2 and the user engagement rises (the Pearson Correlation Coefficient increased from 0,32 to 0,593) with the increase of the reliability (alpha value from 0,58 to 0,65). Furthermore, the coefficient passed from being weak to be moderate. The rise of the correlation requires further analysis on this set of questions. As happened with the correlation, if only the UEA1 and UEA2 scores are considered and compared between the two samples, Familiar and Not familiar, a notable difference can be found. Thus, the results about connection with the urban environment have been compared between the two groups. The mean for the Not familiar is 3,846 and for Familiar users 3,357. The ANOVA showed that the p-value while considering only the two questions scored 0,066, while examining all the three questions had 0,23. Both are not significant but proceeding further on UEA1 and UEA2 might lead to concrete results to prove that the connection for Not familiar is higher.




a) UEA General b) UEA1 and UEA2

*Figure 20 - Results of UEA and UE General plotted on a scale from 1 to 5.*

### 5.2.3. Perceived Walked Distance

The approach used for the evaluation of the Perceived Walked Distance (PWD) is in some terms similar to the second method used for the assessment of the urban environmental awareness. In fact, the procedure expects an evaluation of the internal reliability of the used methodology and a subsequent investigation on the correlation between the PWD results and the UE General.

The reliability has been calculated using the Cronbach's Alpha and it is presented in Table 13. The obtained alpha value is 0,597 which is lower than the acceptability threshold (0,7). However, the item PWD3, which consists in evaluating the distance, reduces the reliability of the method drastically. In fact, users during the experiment expressed the difficulty of estimate a distance without any reference. Therefore, considering only the two items of the questionnaire, namely PWD1 and PWD2, the reliability reaches almost the acceptance, with an alpha value of 0,684.

As far as the correlation with user engagement, the level of correlation between the UE General and the average of PWD1 and PWD2 is almost null (Pearson Correlation Coefficient = -0,046). In a further analysis, an interesting finding appears: if the perceived walked distances from Familiar users and Not familiar are compared, a difference appears. The scores for Familiar users are lower than the Not familiar ones. Although performing the ANOVA has a not significant output (p-value = 0,13), it might need some follow up investigations with more data. If we analyze the correlation between the average of PWD1 and PWD2 and the UE General divided by groups, an unexpected result appears. The scores of Familiar users have a weak positive correlation (Pearson Correlation Coefficient = 0,20), instead, the Not familiar a weak negative correlation (Pearson Correlation Coefficient = -0,31) between the two variables.

*Table 13 - Summary of correlations for calculating the Cronbach's Alpha.*

| | **PWD1** | **PWD2** | **PWD3** |
|---|---|---|---|
| **PWD1** | 1 | - | - |
| **PWD2** | 0,52 | 1 | - |
| **PWD3** | 0,11 | 0,37 | 1 |
| | | | $\alpha$ = 0,597 |

### 5.2.4. Demographics

Besides of the analysis on the four areas of interest of the research, an extra analysis of the scores according to the demographic data of the users has been carried out. In particular, this section focuses on the differences of the final scores of user engagement and technology acceptance according to the demographic data of the users, namely Gender (male or female), Age (under 18; 18-24; 25-34; 35-50 years old), Experience with the pervasive games with AR (very beginner, beginner, expert, very expert) and higher level of education (High school, Bachelor degree, Master degree, Doctorate degree).

An ANOVA single factor has been used to explore whether there were some significant differences in the scores if compared according to the demographic categories. The analysis has been carried out for each category and for both for TA General and UE General. Table 14 shows all the means for every category and subcategories and the correspondent p-value obtained from the ANOVA. All the values are not significant, however, the comparison of the UE General compared by age is close to the significant value. It is an interesting insight which can be elaborated more in future works because it might indicate the ideal age of the target users of the pervasive game.

*Table 14 - Means and correspondent p-values of the between group analysis of the demographic data.*

| | **Gender** | **Age** | **Experience** | **Education** |
|---|---|---|---|---|
| **User Engagement** | Male: 3,73<br>Female: 3,97 | under 18: 4,08<br>18-24: 3,67<br>25-34: 4,16<br>35-50: 3,75 | Very beginner: 3,90<br>Beginner: 3,75<br>Expert: 3,89<br>Very expert: 3,86 | High school: 3,61<br>BSc degree: 3,95<br>MSc degree: 3,92<br>PhD degree: 3,97 |
| | **p: 0,177** | **p: 0,09** | **p: 0,92** | **p: 0,374** |

| | | | | |
|---|---|---|---|---|
| **Technology Acceptance** | Male: 4,00<br>Female: 4,13 | under 18: 3,75<br>18-24: 3,99<br>25-34: 4,28<br>35-50: 3,88 | Very beginner: 3,98<br>Beginner: 4,08<br>Expert: 3,87<br>Very expert: 4,16 | High school: 3,87<br>BSc degree: 4,18<br>MSc degree: 4,15<br>PhD degree: 4,04 |
| | **p: 0,538** | **p: 0,56** | **p: 0,85** | **p: 0,625** |

## 5.3. Discussion

In Section 5.2 were presented the results of the study assessing the level of engagement and the technology acceptance of the users. In general, the scores of the questionnaire UES-SF were positive, which means that participants were engaged in the experience. In particular, the category with the highest score is Perceived Usability. This confirms what has been found in the usability study, and, having made improvements to the game usability helped to improve this score. Participants perceived as not requiring effort to be play the game. On the other hand, the lowest score belongs to the Focused Attention category. This is an indicator that the users were not deeply involved in the experience. However, this result matches with the type of application, in fact, a pervasive game that would totally involve the user might also lead to dangerous situations because of the lack of attention to the surroundings while playing the game. The category Aesthetic scored positively, meaning that users liked the visual appearance of the user interface (UI) in general. During the study, some users paid attention to the visual elements and expressed some confusion about specific sections of the UI, such as confusion about the name on the top-left part of the screen or the unreality of the size of the avatar (see Figure 14). The last category, Reward, showed a good result, which can be translated as the users felt that the experience was worthwhile for them.

As far as the technology acceptance is concerned, the technological instrument used, namely an augmented reality pervasive game, has been perceived as acceptable by the users. Participants felt that the system had a technological infrastructure ready to help them and they had fun during the experience. As a matter of fact, the category of Hedonistic Motivation received a score of 3,95 out of 5. On the other hand, the category of Behavioral Intention scored only 3,59 out of 5. The fact that it’s not a remarkable score means that most of the users are willing to use the app in the future but not very frequently (according to the results of BI1 and BI2). Nevertheless, this doesn't represent a problem. In fact, the main purpose of the mobile application is to promote environmental awareness while having a playful experience, which is a secondary goal.

The results of the questionnaire that investigated the urban environmental awareness did not bring any significant results, thus, the urban environmental awareness cannot be declared as an indicator of the user engagement so far. Nevertheless, further investigation showed that if we consider only the first two questions, they were almost a reliable source while considering all the three items it is not. Using the first two items (UEA1 and UEA2) demonstrates interesting insights on the correlation with the user engagement, which grew with the increase of the internal reliability of the scale. A further analysis was focused on examining the perception and the connection with the urban environment comparing the two groups, Familiar and Not familiar. The results for the Not familiar participants is higher if compared to the Familiar group. However, the analysis didn't show any statistical significance although the p-value ($p=0{,}066$) is close to the acceptability threshold, which will be interesting to investigate more in the future works. In order to establish whether the familiarity with the urban environment has an impact on the promotion of the urban awareness, the experiment would require a more reliable instrument for the evaluation and more participants. So far, the results suggested that Familiar users felt less connected with the surrounding after playing if compared with Not familiar users, probably due to their previous knowledge of the surrounding or because they are used to the environment.

The familiarity with the environment might also have an influence on the perception of the walked distance. In fact, according to the items PWD1 and PWD2, participants familiar with the environment had the perception that they walked more during the game if compared to Not familiar users. Concerning the methodology of asking to estimate the distance (PWD3), it does not seem to be a reliable measurement and no correlation with the user engagement has been found. This result suggests that even if the user underestimates the walked distance this is not an indicator of the level of the engagement but mere coincidence. In fact, the walked distance is a physical subjective estimation, thus, there are many variables to take into account that can influence the result. For instance, if a participant is tired and has played after a workday, the perception will be different compared to if the user plays at the beginning of the day. Another example consists of how often a user is exercising. For instance, a runner has a better perception of the distance because he/she is used to keeping track of it. In conclusion, the perceived walked distance is a topic that would require a complete research by itself and cannot be evaluated with only two items in the questionnaire or only asking to estimate the walked distance. However, this potential complete investigation is outside the scope of this thesis, where the perceived walked distance was only a mean to evaluate the user engagement.

To sum up, the pervasive game has been perceived as a fun experience and the users felt engaged. In addition to this, citizens felt more connected and aware of the urban environment after playing with the game. Once the system and its concept have been validated, it would be possible to extend the mobile game to other points of interests, as planned in the original design. Moreover, the technology used, namely a pervasive game with augmented reality features, has been evaluated as acceptable. Therefore, if the system will use the same technology and in the same context it won't need further investigation on the evaluation of the acceptability of the technology. Intead, if extra technology will be added, another evaluation will be necessary. In conclusion, it can be inferred that the technological mean used is accepted by the users.

## 5.4. Limitations

The experiment presented three main limitations which can be categorized as technical, methodological and procedural.

The technical limitation consisted of the strength of the connection of the mobile phone on which the game was running. In fact, using the tethering connection revealed to be impractical and required the researcher to be close to the participants all the time. The tethering connection was weak and it was limited to the 3G connection of the phone used as hotspot Wi-Fi. This fragile connection made it difficult to store accurately the GPS coordinates. In fact, they were not precise and this revealed to be inaccurate for the calculation of the walked distance using GPS coordinates (see Chapter 5.1). Every GPS coordinate had an error of several meters, and in phase of calculating the length of the total walked path, it reaches an inaccuracy of 30%. Moreover, it happened that the game wasn't smooth because of the fragile connection and this influenced the users who were confused or asked why the avatar was not moving or updating its position.

The second limitation lies on a methodological problem. In fact, as explained in the chapters of the results (see 5.2.2 and 5.2.3) the questions of the survey had an unacceptable reliability which needs to be fixed in the next experiment as well as more data to prove the reliability, especially for the UEA.

Finally, the last limitation consists in the difficulty of recruiting Not familiar participants that have never been on campus. The choice of using this between-group approach made the recruitment complicated. The campus is positioned 30 minutes far from the city center and users needed to come to the university just for the experiment which must take place on campus because of the game scenario. This limitation led to a limited number of recruited participants and a not well-balanced samples when considering the demographic categories. In fact, the gender in the two groups (Familiar and Not familiar) is unbalanced. If we consider the results of the user engagement and technology acceptance compared by the gender, there is not a remarkable difference between the results (both UE General and TA General). Instead, the imbalance of the age in the sample might have influenced the results of the study. In fact, the results of the user engagement (UE General) are considerably different between the categories and the subsample sizes are quite different.

# 6. DISCUSSION

## 6.1. The research in the context of Playable Cities

The contributions and limitations of the experiments have been discussed in the previous chapter. This chapter addresses the general results of this thesis, contextualized in the topic of playable cities.

In general, this research helped to describe a possible approach for the evaluation of the user engagement of pervasive AR games. If zoomed out from the experiments, it becomes clear why evaluating whether a system is engaging citizens or not in the context of playable cities is important. As a matter of fact, the concept of playable cities (Nijholt, 2017) delves into the shift from smart city to smart citizens, enabling them to have interactive experiences and to reshape the urban environment. Therefore, assessing whether the user (i.e. the citizen) is involved in the experience is a crucial aspect for an application in the context of playable cities.

Summarizing, the application developed was perceived positively by the users. Specifically, it was found as engaging, affordable and as a potential way of promoting urban environmental awareness. Besides this, the application presents some drawbacks and limitations, especially if it is analyzed broadly. The first one concerns the effort for extending the game to the entire city, which might require a remarkable amount of work, not optimal cost-wise and effort-wise. In fact, if we compare the design of this pervasive game with the one used for Pokémon Go, we can notice some differences; Pokémon Go was released as a worldwide game with a general approach. In fact, the game is not tailored to a specific physical space, but the mapping of the environment and the subsequent placement of Pokémon and Poképspots has been made automatically on a database of Ingress, a game by Niantic (Speed, 2016). Instead, if considering the pervasive game developed for this research, it is developed specifically for determined spots, which makes it difficult to expand it to the entire city. Thus, it should be limited only to a few points of interests, exclusively designed for enhancing those physical spaces.

Another limitation of the game consists of the lack of social elements and a link between the players of the game. The vision of pervasive games lies on engaging participants in experiences that are played spatially, temporally, or socially (Montola et al., 2017). The introduction of social game mechanics might better connect the players and turn the game into a game played collectively instead of a single-player approach.

During the experiments, several Humorous acts (HAs) occurred. These have been described by Nijholt as humorous interaction generated from the merging of the augmented reality world and the real one (Nijholt, 2014). Humorous acts have been described with three theories of humour, namely incongruity, relief, and superiority/disparagement (Andujar et al., 2017). In the case of these experiments several users, started pointing the cannon at people passing by the university campus, and shot at kids on the bike or old people walking in front of them. This situation generated humour in the participants because of the incongruity between the scenario of the game and the reality of the surroundings (see Figure 21). They were not focused on shooting the building and helping the rebels anymore, but to see the discrepant situation that they were facing.

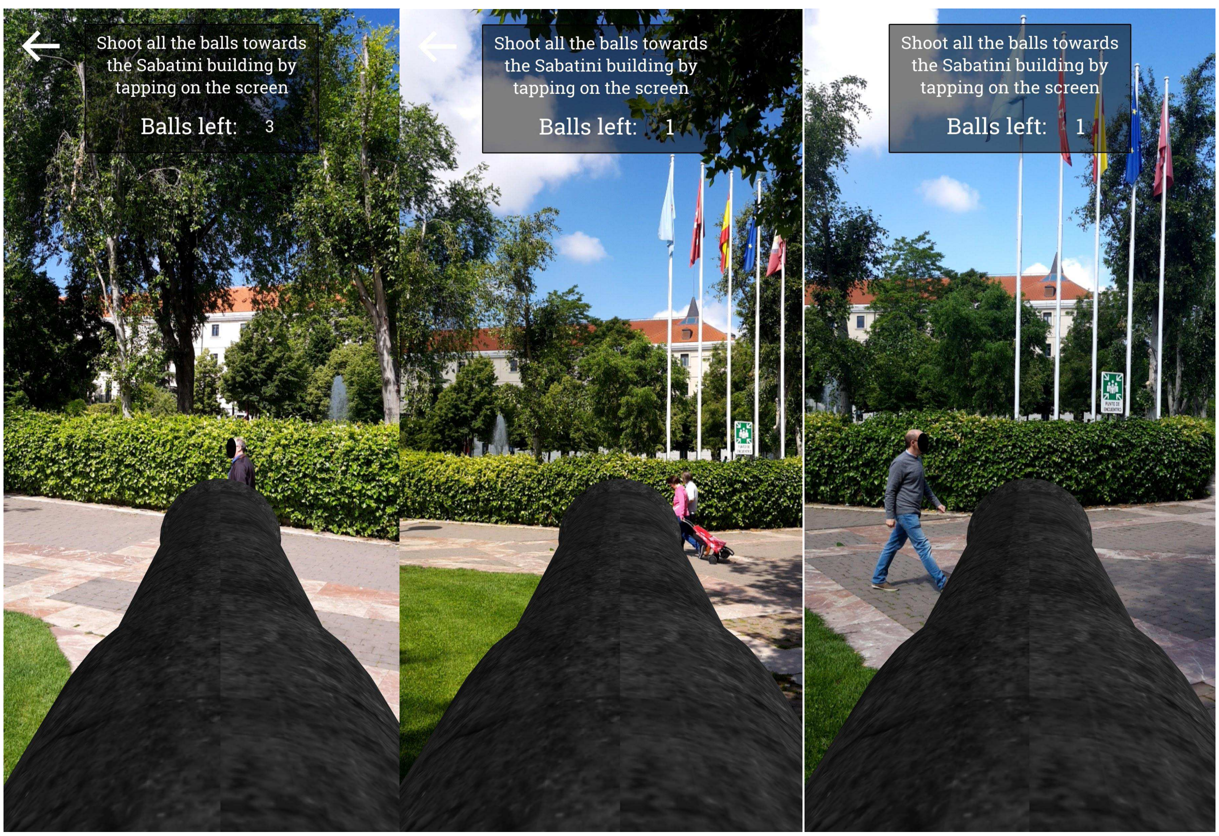



*Figure 21 - Humorous interactions occurred during the experiment.*

Another remarkable event that constantly occurred during the experiments that refers to playable cities, concerns the merging of the physical reality and the virtual one. The game presupposes the interaction to be digital committing to a physical environment. In fact, the game has puzzles that are describing physical elements (such as the statues in the building) or asking information about them (e.g. "how old is the Sabatini building?"). In addition to this, the virtual pieces of the cannon were positioned in the physical world, thus the virtual world and the physical one are sufficiently interwoven. This mix of the two realities disoriented the players mainly on two occasions. First, at the start of the game, when the players read the first question, they started searching for clues in the physical environment that may have had the answer to the in-game question. For instance, a lot of players, when the question concerning the age of the Sabatini building appeared, started searching clues in the sign at the entrance or in some signs on the building, which didn't have any information about it. Asking a question about the physical element took for granted that they were supposed to search all around and the treasure hunt approach has been triggered automatically. The second scenario that occurred frequently concerns the position of the pieces of the cannon in the physical world. When the users are close to the cannon they can tap the virtual object in the map in order to arrive at the question, right after they are supposed to collect the object with the augmented reality view. What happened when the augmented reality camera view opened, is that the players were looking only towards the direction that they were pointing already expecting that the virtual cannon piece was located in the real world as it was in the map. Instead, the object can appear randomly all around the players, regardless of their position on the map. If considering this pervasive AR game in the context of playable cities, it is clear that there is a restriction concerning the virtual-physical interaction. According to the vision of playable cities (Nijholt, 2017), a technological mean for producing the smartness and the playability in the city can be the involvement of actuators and sensors well integrated ubiquitously in the physical environment. Therefore, the combination of the lack of interaction with the physical environment and the vision of ubiquitous applications leads to the conclusion that the introduction of sensors for triggering some components in the game might be an appropriate direction to pursue.

## 6.2. The execution of the research

Conclusively, we can outline some considerations on the execution of the research. What turned out to be profitable is the choice of designing a prior usability study before evaluating the user engagement. This allowed collecting a considerable amount of feedback and technical issues that have been fixed before the second experiment. Moreover, this consideration is confirmed by the results of the second experiment, where the perceived usability scored the highest score in the UES-SF questionnaire.

Another execution aspect that turned out working very well is the design of the out-of-the-box solution for the connectivity problem due to a low signal of the smartphone in an area of the university campus. Designing an in-game resolution turned out working smoothly in the second experiment, where participants never asked why the area was a danger area, that is full of enemies in the game scenario. On the contrary, they were curious to know what would have happened if they have passed through it. A lot of users accepted the fact without any further questions, some users, when it was said that they couldn't walk in the red area, replied curiously: *"what happens if I walk in there then?"*.

Instead, one of the weak points of the research concerns the low reliability of the questions created. It would have been better to ask not only a few questions each category, but more items (at least two times to the prior number) in order to explore what the most reliable combination was and have more concrete results for the following experiment. However, the choice of keeping a low number of items has been made on purpose because of the higher effort required to complete a 30-items questionnaire.

# 7. CONCLUSION

The current research describes the adoption of an augmented, pervasive and playful system for improving the awareness of the surroundings, giving gamified notions about the surroundings to the players. This investigation exemplifies how the user engagement and the technology acceptance of an AR pervasive game can be assessed adopting user-reported measurements. In the research, two main phases were carried out. First, a usability study was useful to assess if the system was perceived as easy to use and safe. Moreover, this study allowed to create system requirements according to the comments and feedback of the users, which were analyzed thanks to the think aloud protocol, and, subsequently converted into game features for the second phase of the research. This second phase consisted of evaluating to what extent the pervasive system was engaging the users and the technology chosen as acceptable. In parallel, a theoretical framework for a more complete evaluation of the user engagement of pervasive games in the context of playable cities was hypothesized. Table 15 gives the conclusions of the research answering all the research questions.

*Table 15 - Conclusions of the research according to the research questions.*

| **Research Questions** | **Conclusion** |
|---|---|
| RQ1: To what extent the mobile application is engaging for the city dwellers? | The AR pervasive game has been perceived as positively engaging by the users that have participated in the study. If we consider that there aren't negative or neutral results of the UES-SF questionnaire, and the scores are in the medium-high range, we can conclude that, according to the user-reported measurements, the pervasive game engaged the users in a significant way. |
| RQ2 a: To what extent the mobile application is accepted technologically? | The technology used for empowering urban awareness, namely an AR pervasive game, was perceived as acceptable. The results from the chosen user-reported method show that there |

| | |
|---|---|
| | aren't negative or neutral scores. In addition to this, the majority of the scores are positioned in the high score-range (i.e. from 3,5 to 5). Conclusively, we can infer that using an AR pervasive game for enhancing the urban surrounding is accepted by the players. |
| RQ2 b: To what extent the Technology acceptance and User Engagement are related? | The results of the second experiment showed that a strong positive correlation between the two independent variables does exist (Pearson Correlation Coefficient = 0,764). |
| RQ3: In the context of pervasive games, to what measure the Perceived Walked Distance can be considered as an indicator of the User Engagement? | In the second experiment, two different methodologies have been adopted for assessing the Perceived Walked Distance. Estimating the walked distance in meters turned out to be a unreliable source. On the other hand, using regular questions appeared to be over the threshold of reliability, but it shows a null correlation when compared with the user engagement. Therefore, we can conclude saying that, for the state of this research, the Perceived Walked Distance is not correlated with user engagement and when taking it into account in the investigations it should be done in parallel with multiple factors, not only behavioral ones, rather physical. |
| RQ4 a: To what extent the familiarity with the environment influences the engagement while playing the game? | Using the between-groups approach in order to compare and evaluate the user engagement do not present any significant difference. Nevertheless, comparing the results from the |

| | |
|---|---|
| | UES-SF and the questions about UEA, indicates that there might be a correlation between the two variables.<br>Therefore, the familiarity with the urban environment do not have a significant impact on the user engagement, rather, how the user perceives the urban environment might influences the results. |
| RQ4 b: To what extent the familiarity with the environment influences the acceptability of the mobile application? | The between-groups method does not show any significant differences in terms of the acceptability of the technology when compared the scores of the two groups. |

In conclusion, the pervasive game was perceived as engaging, useful for empowering the urban environmental awareness and the technology adopted in the system was perceived as acceptable. However, due to a low reliability and correlation with the level of engagement, we cannot add the urban environmental awareness and the perceived walked distance as determinants of the user engagement in the hypothesized theoretical framework. These two topics should be addressed in two separate future researchers because they require more aspects to be thought-out.

Although the game has been perceived as engaging and the technology used acceptable, participants showed confusion in the merging of the virtual and the physical world or inexperience on how to approach the game, as discussed in Chapter 6.1. This leads to the necessity in the future to improve the game itself in diverse aspects.

# 8. FUTURE WORK

In the view of facts explained in the previous sections, several questions were raised during the research as well as diverse ideas for improving the pervasive game itself. Therefore, this section explains a possible follow-up of the research pursuing two different objectives. Section 8.1 shows an attainable and affordable proposal on how to improve the system according to the findings of this thesis. Considering the proposed new version of the mobile application (v3), a future experiment is designed for the evaluation of the new version of the system and a deeper investigation on the factors that might influence the final user engagement is explained in Section 8.2.

## 8.1. Improvement of the game

The game can be improved in several ways on the basis of the second experiment. The proposed changes derive from the discussions of the previous chapters and they can be summarized in 4 different categories.

1. **User interface.** Some elements of the graphical user interface of the game were confusing for the players. In particular, the username, which appears on the top-left of the display in the orange box (see Figure 14), confused the participants because it is very similar to the visual appearance of the buttons. In fact, participants frequently asked what was the purpose of that button. Changing its style, that is the background and the text color of the username will benefit the comprehension.
2. **Virtual-physical interaction.** As explained in Chapter 6, the merging of the virtual world with the physical one and the design of the game were not very well optimized. That is, the pervasiveness of the game is not well integrated into the physical environment. As a consequence, the participants appeared disoriented on how to approach the game and the required tasks. For this very reason, an adaptation of the positioning of the objects is necessary. The proposal for the new interaction consist on the introduction of iBeacons[14] (or another similar technology that allows the proximity communication) hidden in the physical

[14] en.wikipedia.org/wiki/IBeacon

environment where the cannon pieces are positioned. With the installation of the iBeacon in the physical position, two main problems can be solved. First, the users are not required to tap on the screen for collecting objects anymore. This would be changed into an automatic message appearing on the display saying to the player that she or he is close to a cannon piece. Second, when the question is displayed and the users need to respond, they can select the option *"search for the clue"* with which they can seek in the surroundings with the augmented reality view and a hint will appear hidden in the environment. For example, when the question asking which regiment the player is facing appears, the users can find as a clue in the environment a French flag or something related to France to show that the French regiment is the correct answer.

3. **Audio interaction.** It has been proved that using the audio and the background music has an influence on how users perceive the time and how they feel immersed in the game (Zhang & Fu, 2015). Therefore, an improvement of the game can consist of the inclusion of sounds when an action is accomplished and a background music during the game.
4. **Social interaction**. In the game v2, there was a complete lack of social interaction. This can be introduced with the introduction of some gamification elements for the peer-competition as well as for the reward of the player. It means that the motivation of all the players performing the same task can be improved with some gamification mechanics. However, the introduction of social elements is not coherent with the current state of the game which is limited to the university building and it's not available for the public. The current modification should be applied when more players can play the game in parallel and in the long term, not just in laboratory settings.

## 8.2. Proposal for the future experiment

The following experiment should continue the investigation according to the findings of the last study adopting the new version of the game proposed in Section 8.1. The pervasive game would have new features as well as a new technology. Nevertheless, it is unknown whether the new version of the game will improve the overall experience, that is the level of engagement of the user. It is possible to investigate this with as a benchmark (or control version) the results of the prior experiment, which evaluated the engagement of the current version. Therefore, an A/B test would define if the improved version (or variation version) actually helps to engage more the citizens. Hence, decide whether it's acceptable to apply the proposed changes. For this very reason, it is important that the level of engagement of participants is assessed with the same method used in the previous study, namely the UES-SF questionnaire.

Besides assessing the new version of the game, the future experiment should also analyze better some findings and explore new directions. As shown in Chapter 5.3, the area of the research that requires further investigations is the Urban Environmental Awareness in terms of the perception of the surroundings while playing the game. In order to have a more complete evaluation of this topic, more categories of the so-called Urban Experience (Gensler Research, 2016) need to be considered. Accordingly, besides the estimation of the awareness of the environment, the proposed questionnaire also includes items for the categories of the Urban Experience theoretical framework (Gensler Research, 2016), which describes a model for assessing the perception of the urban spaces. The considered factors that influence the perception of the urban spaces are Ambient Comfort (AC), Amenities, Built Form (BF), Mobility (M), and Safety (S). From these categories, Amenities has been excluded because this not relevant to the purpose of the pervasive game. The remaining categories are converted into multiple items and added to the proposed questionnaire, presented in Table 16.

In conclusion, this proposed future work investigates in two directions. On the one hand, it wants to evaluate whether the new game features and the introduction of the iBeacon technology improve the game experience in terms of user engagement. On the other hand, it wants to explore whether the perception of the environment, namely the urban experience, influences the user engagement of pervasive games. Moreover, the planned questionnaire includes a revisioned version of the UEA items from the survey of the previous study. The follow-up can be done in two separate studies, one for the evaluation of the proposed new version (v3) and the second in order to investigate the new determinants of the user engagement. According to the results of the future first study, the version which performs better will be used in the second study.

*Table 16 - Proposal of the section of the questionnaire about UEA.*

| ID | Item |
|---|---|
| UEA1 | After playing the game I felt more attached with this environment. |
| UEA2 | After playing the game I felt more connected with the university. |
| UEA3 | The game helped me be more aware of this surroundings. |
| UEA4 | The game improved my attention to this location. |
| S1 | During the game I felt safe. |
| S2 | During the game the surroundings were free of harm. |
| S3 | The space where I played was not dangerous. |
| M1 | I had easy access to the spot where the cannon pieces were. |
| M2 | Where I played there were a few of people passing by. |
| M3 | Reaching the location required in the game was not difficult. |
| AC1 | While playing, I felt like the temperature was too hot.<br>(or too cold, to be changed according to the season) |
| AC2 | The space where I played was quiet. |
| AC3 | During the game, the weather was optimal. |
| BF1 | There was enough space where move freely. |
| BF2 | There was enough open space where I played. |
| BF3 | The places where I used the game were accessible. |

# References

Althoff, T., White, R. W., & Horvitz, E. (2016). Influence of Pokémon Go on physical activity: study and implications. Journal of medical Internet research, 18(12).

Ampatzidou, C., Bouw, M., van de Klundert, F., de Lange, M., & Waal, B. G. M. (2014). The hackable city: a research manifesto and design toolkit. Knowledge Mile.

Anderson, N. H. (1981). Foundations of information integration theory. Boston: Academic Press. ISBN 0-12-058101-9.

Andujar, M., Nijholt, A., & Gilbert, J. E. (2017). Mobile Augmented Games in Playable Cities: Humorous Interaction with Pokémon Go. In International Conference on Distributed, Ambient, and Pervasive Interactions (pp. 575-586). Springer, Cham.

Ajzen, I., and Fishbein, M. (1980). Understanding attitudes and predicting social behavior. Englewood Cliffs, NJ: Prentice-Hall.

Ajzen, I. (1991). The Theory of Planned Behavior. Organizational Behavior and Human Decision Processes, 50, 179-211.

Attfield, S., Kazai, G., Lalmas, M., & Piwowarski, B. (2011, February). Towards a science of user engagement (position paper). In WSDM workshop on user modelling for Web applications (pp. 9-12).

Ballagas, R. A., Kratz, S. G., Borchers, J., Yu, E., Walz, S. P., Fuhr, C. O., Hovestadt, L., & Tann, M. (2007, April). REXplorer: a mobile, pervasive spell-casting game for tourists. In CHI'07 extended abstracts on Human factors in computing systems (pp. 1929-1934). ACM.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. Intl. Journal of Human-Computer Interaction, 24(6), 574-594.

Bardzell, J., Bardzell, S., Pace, T., & Karnell, J. (2008, April). Making user engagement visible: a multimodal strategy for interactive media experience research. In CHI'08 extended abstracts on Human factors in computing systems (pp. 3663-3668). ACM.

Beardsley, M. C. (1982). The aesthetic point of view: Selected essays, Monroe C. Beardsley. Ithaca: Cornell University Press.

Brooke, J. (1996). SUS - A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

Bursztyn, N., Walker, A., Shelton, B., & Pederson, J. (2017). Increasing undergraduate interest to learn geoscience with GPS-based augmented reality field trips on students' own smartphones. GSA Today, 27(5), 4-11.

Caballero, B. (2007). The global epidemic of obesity: an overview. Epidemiologic reviews, 29(1), 1-5.

Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. Brock Education Journal, 12(2).

Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., ... & Scholl, H. J. (2012, January). Understanding smart cities: An integrative framework. In System Science (HICSS), 2012 45th Hawaii International Conference on (pp. 2289-2297). IEEE.

Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. New York: HarperCollins

Davis, F. D.; Bagozzi, R. P.; Warshaw, P. R. (1989), "User acceptance of computer technology: A comparison of two theoretical models", Management Science, 35: 982–1003.

Díaz, P., Onorati, T., & Aedo, I. (2017, May). A digital knowledge ecosystem to increase participation in emergency warnings and alerts management. In International Conference on Green, Pervasive, and Cloud Computing (pp. 700-711). Springer, Cham

e Silva, A. D. S., & Sutko, D. M. (Eds.). (2009). Digital cityscapes: Merging digital and urban playspaces. Peter Lang.

Fabiano, F. (2018). User Engagement and Technology Acceptance in the context of Playable Cities. Research Topics report. University of Twente.

Fabiano, F., Sánchez-Francisco, M., Díaz, P., and Aedo, I. (2018). Evaluating a pervasive game for urban awareness. Poster session presented at: 20th International Conference on Human-Computer Interaction with Mobile Devices and Services. ACM.

Fairclough, S. H., Gilleade, K., Ewing, K. C., & Roberts, J. (2013). Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation. International Journal of Autonomous and Adaptive Communications Systems, 6(1), 63-79.

Fishbein, M. (1963). An investigation of the relationships between beliefs about an object and the attitude toward that object. Human Relations, 16, 233–239.

Fishbein, M. (1967). Attitude and the prediction of behavior. In M. Fishbein (Ed.), Readings in attitude theory and measurement (pp. 477–492). New York: Wiley.

Fishbein, M. (1980). Theory of reasoned action: Some applications and implications. In H. Howe and M. Page (Eds.), Nebraska Symposium on Motivation, 1979 (pp. 65–116). Lincoln: University of Nebraska Press.

Gensler Research. (2016). Measuring Urban Experience. Retrieved from www.gensler.com/research-insight/gensler-research-institute/measuring-urban-experience.

Harrison, C., & Donnelly, I. A. (2011, September). A theory of smart cities. In Proceedings of the 55th Annual Meeting of the ISSS-2011, Hull, UK (Vol. 55, No. 1).

Howe, K. B., Suharlim, C., Ueda, P., Howe, D., Kawachi, I., & Rimm, E. B. (2016). Gotta catch'em all! Pokémon GO and physical activity among young adults: difference in differences study. BMJ, 355, i6270.

Humphreys, L., & Liao, T. (2011). Mobile geotagging: Reexamining our interactions with urban space. Journal of Computer- Mediated Communication, 16(3), 407-423.

Kasapakis, V., Gavalas, D., & Bubaris, N. (2013). Pervasive games research: a design aspects-based state of the art report. In Proceedings of the 17th panhellenic conference on informatics (pp. 152-157). ACM.

Khomami N. Pokémon Go: London players robbed of phones at gunpoint. Guardian 2016 Jul 30. https://www.theguardian.com/uk-news/2016/jul/30/three-pokemon-go-players-robbed-of-phones-at-gunpoint-in-london

Kline, P. (1999) A Handbook of Psychological Testing, 2nd edn. London: Routledge.

Kobayashi, T., & Boase, J. (2012). No such effect? The implications of measurement error in self-report measures of mobile communication use. Communication Methods and Measures, 6(2), 126-143.

Lavrakas, P. J. (2008). Encyclopedia of survey research methods Thousand Oaks, CA: SAGE Publications Ltd doi: 10.4135/9781412963947

Lee, T. (1970). Perceived distance as a function of direction in the city. Environment and Behavior, 2(1), 40-51.

Liao, T., & Humphreys, L. (2015). Layar-ed places: Using mobile augmented reality to tactically reengage, reproduce, and reappropriate public space. New Media & Society, 17(9), 1418-1435.

Navalpakkam, V., Kumar, R., Li, L., & Sivakumar, D. (2012, July). Attention and selection in online choice tasks. In International Conference on User Modeling, Adaptation, and Personalization (pp. 200-211). Springer, Berlin, Heidelberg.

Nielsen, J., & Molich, R. (1990, March). Heuristic evaluation of user interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 249-256). ACM.

Nijholt, A. (2017). Towards playful and playable cities. In Playable Cities (pp. 1-20). Springer, Singapore.

Nijholt, A. (2014). Towards humor modelling and facilitation in smart environments. In: Ahram, T., Karwowski, W., Marek, T. (eds.) Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics (AHFE 2014). pp. 2992–3006. Krakow, Poland

Neuenhaus, M., Aly, M. (2017). A key HCI based-concept for Creating Empathy for Refugees. Workshop in Refugees & HCI Workshop: The Role of HCI in Responding to the Refugee Crisis.

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. Journal of the Association for Information Science and Technology, 59(6), 938-955.

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. Journal of the Association for Information Science and Technology, 61(1), 50-69.

O'Brien, H. L., & Toms, E. G. (2013). Examining the generalizability of the User Engagement Scale (UES) in exploratory search. Information Processing & Management, 49(5), 1092-1107.

O'Brien, H., Cairns, P., & Hall, M. (2018). A Practical Approach to Measuring User Engagement with the Refined User Engagement Scale (UES) and New UES Short Form. International Journal of Human-Computer Studies.

Prandi, C., Salomoni, P., Roccetti, M., Nisi, V., & Nunes, N. J. (2016, February). Walking with Geo-Zombie: A pervasive game to engage people in urban crowdsourcing. In Computing, Networking and Communications (ICNC), 2016 International Conference on (pp. 1-5). IEEE.

Read, J. C., MacFarlane, S. J., & Casey, C. (2002, August). Endurability, engagement and expectations: Measuring children's fun. In Interaction design and children (Vol. 2, pp. 1-23). Shaker Publishing Eindhoven.

Rieber, L.P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. Educational Technology Research and Development, 44(2), 45–58.

Sánchez-Francisco, M., Díaz, P., & Aedo, I. (2018, May). Improving urban environment awareness through pervasive AR games. In Proceedings of the 2018 International Conference on Advanced Visual Interfaces (p. 82). ACM.

Sekaran, U., & Bougie, R. (2016). Research methods for business: A skill building approach. John Wiley & Sons.

Schouten, B., Ferri, G., de Lange, M., & Millenaar, K. (2017). Games as Strong Concepts for City-Making. In Playable Cities (pp. 23-45). Springer, Singapore.

Schneiderman, B. (1997). Direct manipulation for comprehensible, predictable and controllable user interfaces. In Designing the user interface (3rd ed., pp. 33–39). Reading, MA: Addison-Wesley.

Siow, Benjamin. (2016). Validated UTAUT2 questionnaire for elderly. 10.13140/RG.2.2.34892.21124.

Speed, B. Where did Pokemon Go get its maps from?. CityMetric (2016) Jul 13. https://www.citymetric.com/fabric/where-did-pokemon-go-get-its-maps-2254

Stapp, W. B. (1969). The concept of environmental education. Environmental Education, 1(1), 30-31.

Sutcliffe, A. (2010). Designing for user engagement: Aesthetic and attractive user interfaces. In J. M. Carroll & E. M. Frymoyer (Eds.), Synthesis lectures on human-centered informatics, vol. 5. Morgan Claypool.

Tolias, E., Costanza, E., Rogers, A., Bedwell, B., & Banks, N. (2015, September). IdleWars: an evaluation of a pervasive game to promote sustainable behaviour in the workplace. In International Conference on Entertainment Computing (pp. 224-237). Springer, Cham.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science, 46(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425-478.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision sciences, 39(2), 273-315.

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS quarterly, 157-178.

Wang, X., Kim, M. J., Love, P. E., & Kang, S. C. (2013). Augmented Reality in built environment: Classification and implications for future research. Automation in Construction, 32, 1-13.

Webster, J., & Ho, H. (1997). Audience engagement in multimedia presentations. ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 28(2), 63-77.

Yuhas, A. Pokémon Go: armed robbers use mobile game to lure players into trap. Guardian 2016 Jul 11. theguardian.com/technology/2016/jul/10/pokemon-go-armed-robbers-dead-body

Zach, F. J., & Tussyadiah, I. P. (2017). To catch them all—the (un) intended consequences of Pokémon GO on mobility, consumption, and wellbeing. In Information and communication technologies in tourism 2017 (pp. 217-227). Springer, Cham.

Zhang, J., & Fu, X. (2015). The influence of background music of video games on immersion. Journal of Psychology & Psychotherapy, 5(4), 1.

# Appendices

## Appendix A - Questionnaire Usability Testing

| Code | Items |
|---|---|
| SUS1 | I think that I would like to use this system frequently. |
| SUS2 | I found the system unnecessarily complex. |
| SUS3 | I thought the system was easy to use. |
| SUS4 | I think that I would need the support of a technical person to be able to use this system. |
| SUS5 | I found the various functions in this system were well integrated. |
| SUS6 | I thought there was too much inconsistency in this system. |
| SUS7 | I would imagine that most people would learn to use this system very quickly. |
| SUS8 | I found the system very complicated to use. |
| SUS9 | I felt very confident using the system. |
| SUS10 | I needed to learn a lot of things before I could get going with this system. |
| EXTRA1 | Which was the hardest task of the game? |
| EXTRA2 | Have you ever felt confused about the game? If yes, when? |
| EXTRA3 | Do you think the text of the app was big enough? |
| EXTRA4 | Do you think the buttons in the app were big enough? |
| EXTRA5 | What do you think about the aesthetic of the app in general? What would you change? |
| EXTRA6 | If you would have the chance to add something to the game, what would you add? |
| EXTRA7 | If you would have the chance to remove something from the game, what would you remove? |

## Appendix B - User Requirements from the usability testing

Legend: High Priority, Mid Priority, Low Priority

| Topic | Insight | Subject |
|---|---|---|
| Text | Place titles in the introduction text: “Scenario” and “goal of the mission”. | S1 |
| | Increase the size of the text | S1, S2 |
| | Add in the instructions text that users need to walk towards object | S4 |
| | Change background while displaying the text. Not enough contrast (find a lighter background). | S4, S5 |
| Orientation | Add a little top-view map on the bottom left of the screen + direction of compass. Example | S3, S4 |
| Buttons | Increase a bit the size of the buttons. | S5 |
| Find Piece | Highlight the border of the pieces to make them more visible. Example | S1, S3, S5 |
| | Add instruction for the next task “Now you have to collect the next object”. | S1, S3 |
| Collect object | Add instruction the first time “click on the object to collect it”. | S1 |
| | Add progress when an object is collected. | S1, S4 |
| | The object cannon can be confused with the sabatini tower, put it horizontal. | S5 |
| Collect object with camera | Show the first time a text saying “Search the object turning all around you and tap the object to collect”. | S3, S4, S5 |
| Zoom | Add possibility to move the view. | S1, S4 |
| Compone cannon | Add arrows below the model to give the impression of possibility of moving it. Example | S2 |
| Shoot the cannon | Increase size of instruction text. | S1 |
| | Add text “touch the screen to shoot the cannon”. | S2, S4 |

| | | |
|---|---|---|
| | Change the “0/3” text into “You have 3 cannon balls to shoot”. | S4 |
| Avatar | Possibility to change size of the avatar. | S1 |
| | Possibility to customize avatar. | S4 |

## Appendix C - *User Engagement Scale, Short Form*

| Code | Items |
|---|---|
| FA-S.1 | I lost myself in this experience. |
| FA-S.2 | The time I spent using the game just slipped away. |
| FA-S.3 | I was absorbed in this experience. |
| PU-S.1 | I felt frustrated while using this game. |
| PU-S.2 | I found this game confusing to use. |
| PU-S.3 | Using this game was stressful. |
| AE-S.1 | This game was attractive. |
| AE-S.2 | This game was aesthetically appealing. |
| AE-S.3 | This game appealed to my senses. |
| RW-S.1 | Using the game was worthwhile. |
| RW-S.2 | My experience was rewarding. |
| RW-S.3 | I felt interested in this experience. |

## Appendix D - *UTAUT 2*

PE1 I find the game useful in my daily life.

PE2 Using the game increases my chances of achieving things that are important to me.

PE3 Using the game helps me accomplish things more quickly.

PE4 Using the game increases my productivity.

EE1 Learning how to use the game is easy for me.

EE2 My interaction with the game is clear and understandable.

EE3 I find the game easy to use.

EE4 It is easy for me to become skillful at using the game.

SI1 People who are important to me think that I should use the game.

SI2 People who influence my behavior think that I should use the game.

SI3 People whose opinions that I value prefer that I use the game.

FC1 I have the resources necessary to use the game.

FC2 I have the knowledge necessary to use the game.

FC3 The game is compatible with other technologies I use.

FC4 I can get help from others when I have difficulties using the game.

HM1 Using the game is fun.

HM2 Using the game is enjoyable.

HM3 Using the game is very entertaining.

PV1 The game is reasonably priced.

PV2 The game is a good value for the money.

PV3 At the current price, the game provides a good value.

HT1 The use of the game has become a habit for me.

HT2 I am addicted to using the game.

HT3 I must use the game.

HT4 Using the game has become natural to me.

BI1 I intend to continue using the game in the future.

BI2 I would like to continue to use the game frequently.

## Appendix E - *Scores of the UES-SF questionnaire*

| ID user | Focused Attention | Perceived Usability | Aesthetic | Reward | UE General |
|---|---|---|---|---|---|
| NF1 | 4,00 | 4,67 | 4,00 | 3,67 | **4,08** |
| NF2 | 2,67 | 4,67 | 4,00 | 4,33 | **3,92** |
| NF3 | 4,33 | 5,00 | 3,00 | 4,33 | **4,17** |
| NF4 | 4,00 | 3,00 | 4,00 | 3,33 | **3,58** |
| NF5 | 3,00 | 5,00 | 3,00 | 4,00 | **3,75** |
| F1 | 3,00 | 4,33 | 3,67 | 3,33 | **3,58** |
| NF6 | 2,67 | 5,00 | 3,33 | 3,33 | **3,58** |
| NF7 | 3,67 | 5,00 | 2,67 | 4,33 | **3,92** |
| F2 | 2,33 | 5,00 | 3,00 | 3,67 | **3,50** |
| NF8 | 3,00 | 4,67 | 2,67 | 3,67 | **3,50** |
| NF9 | 2,00 | 4,33 | 2,67 | 3,67 | **3,17** |
| NF10 | 2,33 | 4,33 | 4,33 | 4,00 | **3,75** |
| F3 | 3,67 | 5,00 | 3,00 | 3,33 | **3,75** |
| F4 | 3,33 | 5,00 | 4,67 | 4,00 | **4,25** |
| F5 | 2,00 | 4,00 | 3,33 | 3,33 | **3,17** |
| F6 | 1,67 | 4,33 | 3,00 | 4,00 | **3,25** |
| F7 | 4,00 | 5,00 | 4,33 | 4,33 | **4,42** |
| NF11 | 4,00 | 5,00 | 3,67 | 5,00 | **4,42** |
| F8 | 3,33 | 3,67 | 4,00 | 4,00 | **3,75** |
| F9 | 3,33 | 5,00 | 4,67 | 4,67 | **4,42** |
| F10 | 3,00 | 4,67 | 3,00 | 4,33 | **3,75** |
| F11 | 3,00 | 4,67 | 3,00 | 3,67 | **3,58** |
| NF12 | 4,33 | 5,00 | 4,33 | 5,00 | **4,67** |
| F12 | 3,67 | 4,00 | 3,67 | 4,33 | **3,92** |
| NF13 | 5,00 | 5,00 | 4,67 | 5,00 | **4,92** |
| F13 | 1,67 | 5,00 | 4,33 | 4,00 | **3,75** |
| F14 | 2,33 | 4,00 | 3,33 | 2,67 | **3,08** |
| **Average** | **3,160** | **4,605** | **3,605** | **3,975** | **3,836** |
| **St. Dev** | **0,87** | **0,52** | **0,67** | **0,58** | **0,46** |

## Appendix F - *Scores of the UTAUT2 questionnaire*

| **ID user** | **Effort Expectancy** | **Facilitating Conditions** | **Hedonic Motivation** | **Behavioral Intention** | **TA General** |
|---|---|---|---|---|---|
| NF1 | 4,00 | 3,33 | 3,67 | 4,00 | **3,75** |
| NF2 | 4,50 | 5,00 | 4,33 | 3,00 | **4,21** |
| NF3 | 4,50 | 4,33 | 5,00 | 4,50 | **4,58** |
| NF4 | 3,75 | 4,00 | 3,33 | 2,50 | **3,40** |
| NF5 | 4,75 | 4,00 | 4,00 | 4,00 | **4,19** |
| F1 | 3,75 | 4,00 | 4,00 | 3,00 | **3,69** |
| NF6 | 3,50 | 3,00 | 4,00 | 3,00 | **3,38** |
| NF7 | 5,00 | 5,00 | 4,33 | 3,50 | **4,46** |
| F2 | 3,50 | 4,67 | 2,33 | 3,00 | **3,38** |
| NF8 | 3,75 | 4,33 | 3,67 | 3,50 | **3,81** |
| NF9 | 4,00 | 4,00 | 3,00 | 2,50 | **3,38** |
| NF10 | 4,25 | 4,33 | 3,00 | 2,00 | **3,40** |
| F3 | 4,75 | 4,67 | 3,00 | 3,00 | **3,85** |
| F4 | 5,00 | 4,33 | 4,33 | 3,50 | **4,29** |
| F5 | 4,25 | 4,33 | 3,00 | 3,50 | **3,77** |
| F6 | 3,50 | 3,67 | 3,67 | 3,00 | **3,46** |
| F7 | 4,75 | 4,33 | 4,33 | 3,50 | **4,23** |
| NF11 | 4,50 | 5,00 | 4,00 | 4,00 | **4,38** |
| F8 | 4,25 | 4,33 | 5,00 | 5,00 | **4,65** |
| F9 | 5,00 | 5,00 | 4,33 | 5,00 | **4,83** |
| F10 | 4,75 | 4,00 | 4,00 | 4,00 | **4,19** |
| F11 | 4,75 | 4,33 | 4,33 | 4,00 | **4,35** |
| NF12 | 5,00 | 5,00 | 5,00 | 5,00 | **5,00** |
| F12 | 4,25 | 5,00 | 4,33 | 4,00 | **4,40** |
| NF13 | 5,00 | 4,33 | 5,00 | 5,00 | **4,83** |
| F13 | 5,00 | 4,67 | 4,67 | 4,00 | **4,58** |
| F14 | 3,75 | 4,00 | 3,00 | 2,00 | **3,19** |
| **Average** | **4,36** | **4,33** | **3,95** | **3,59** | **4,06** |
| **St. Dev** | **0,53** | **0,51** | **0,73** | **0,87** | **0,53** |