Taking to the (Virtual) Stage:

Developing a Questionnaire to Measure Immersion in the Mobile Augmented Reality Music Application 'Firstage'

Karl Knuttel

A research paper submitted to the University of Dublin,
in partial fulfillment of the requirements for the degree of
Master of Science Interactive Digital Media

2017

Declaration

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: http://www.tcd.ie/calendar

I have also completed the Online Tutorial on avoiding plagiarism 'Ready, Steady, Write', located at http://tcd-ie.libguides.com/plagiarism/ready-steady-write

I declare that the work described in this research Paper is, except where otherwise stated, entirely my own work and has not been submitted as an exercise for a degree at this or any other university.

Signed: ___________________

Karl Knuttel

8th May 2017

Permission to lend and/or copy

I agree that Trinity College Library may lend or copy this Research Paper upon request.

Signed: ____________________

Karl Knuttel

8th May 2017

Acknowledgements

I would like to thank my supervisor, Natasa Paterson-Paulberg, for all of her suggestions and guidance while writing this paper. Your help was invaluable.

I would also like to thank my mum, my dad, and my partner Gina for being eternally supportive throughout this master's programme.

Summary

This research paper is concerned with devising a questionnaire which could be used to measure immersion in *Firstage,* a mobile augmented reality (MAR) music performance application. *Firstage* overlays pre-recorded live music performances onto a user's smartphone screen and commingles it with the immediate environment captured through the device's camera in order to create a blended space. Thus music artists can appear to perform anywhere the user chooses, freeing them from physical stages and turning any surface into a potential 'virtual' stage.

As described in Kim et al. (2014), many MAR applications fail due to low uptake and success is often determined by their ability to exploit augmented reality's primary strength over other media, that of creating immersive experiences. A review of the relevant literature gives an overview of MAR technology, illustrates a number of theories of immersion applicable to mediated experiences, outlines the role of audio in immersion, and analyses a selection of relevant studies which measure for immersion in MAR. Through analysis of *Firstage*, it is determined that two models of immersion may be employed to measure the application's immersive performance: the models of mobile augmented reality immersion in Georgiou & Kyza (2017) and context immersion in Lee & Kim (2011) and Kim (2013).

Using the models identified, the paper goes on to create a potential questionnaire which could be used to measure immersion in Firstage. The theoretical approach to item generation, a proposed method of implementation, and the means through which results may be analysed are each described in detail.

Discussion suggests that while each model measures for different conceptions of immersion there is some overlap, and, in the particular case of *Firstage*, perhaps context immersion might be more indicative of the application's immersive performance.

Table of contents

List of figures

List of tables

Chapter 1. Introduction

1.1 Introduction

In *Sounding Out The City*, Bull (2000) argues that in its ubiquity, the Walkman revolutionised the way that people experienced their environment. It afforded users the opportunity to "aesthetically recreate their daily experience" (Bull 2000, p. 86) by soundtracking their lives and augmenting the mundanity of the everyday with music. With the development of smartphone technology that same portability and ubiquity now exists for all types of media. One need only look around oneself to see people checking email, watching TV or movies, reading the news, playing video games, and browsing the web, all through the use of a single personal device. Smartphones also allow for new types of media experiences. Augmented reality (AR), once restricted to use with desktop computers and bulky head mounted displays (HMDs), has now been made available to anyone who owns a smartphone. Real world environments can be blended with virtual objects and media to create richer, more multifaceted experiences of the everyday, with portability and ubiquity.

Through the use of mobile AR (MAR) technology, *Firstage* (Firstage Ltd. 2015) delivers pre-recorded live music performances to a user's smartphone and allows users to combine these performances with their immediate real world environment. Thus a sort of hybrid space—a 'virtual stage'—is created: a blend of the real and the virtual is formed in which these performances now take place. Just as a dependence on old media delivery methods such as newspapers, TVs, and computers was overcome by the smartphone, so too might the necessity for a live venue for music performances be overcome through MAR. Users can bring live performances with them and access them anywhere and at any time. *Firstage* is not just for users though. A stated aim is to allow artists to connect with potential fans by breaking down geographical and financial barriers. Why couldn't an Irish band play a gig in America, or a French band play a gig in Japan? "Emerging bands [without the money to tour worldwide] can now do that through a virtual stage. You can get loads of exposure, use the power of playing gigs and connect with people who really want to discover great music" (Murphy 2015).

That is not to say that such technology should replace the real thing, any more than video might replace live shows. However, one of AR's greatest advantages over other media, such as video, is in its ability to create *immersive* experiences. A convincing blend between the real and the virtual, as well as interactivity, can potentially create a suspension of disbelief in users such that the virtual elements become, in a sense, 'real' (Brooks 2003; Kim 2013). Furthermore, while more and more MAR applications are being released, many fail due to low usage rates; important for an MAR application's success is that it provides an immersive experience (Kim et al. 2014). An important question, then, becomes how well *Firstage* performs in terms of immersion and how this may be measured.

1.2 Background

There have been a number of studies conducted on immersion in mediated experiences, although until recently most have dealt with virtual environments (VEs) (Witmer & Singer 1998; Slater 1999; Slater & Wilbur 1997), non mobile AR (Azuma 1997; Azuma et al. 2001), or videogames (Brown & Cairns 2004; Cheng et al. 2015; Jennet et al. 2008). However, there are some recent literature and studies that deal with immersion in an MAR context, most notably those of Georgiou & Kyza (2017), and Lee & Kim (2011) and Kim (2013). In the former work, a model of immersion comparable to that found in other mediated experiences is tested, though adjusted such that it may be applied to an MAR context. In the latter two works, a model of immersion unique to the MAR context—context immersion— is proposed and tested. In these studies, the MAR applications investigated are 'location-aware' games or services, applications which use GPS to track the virtual content. Yet the application of focus in this paper is not location-aware, and indeed quite unique in the service it supplies.

1.3 Purpose of the paper

The aim of this paper, therefore, is to create a questionnaire which could be used to measure immersion in *Firstage*. There are two kinds of MAR application: marker-less, which includes location-aware applications, and marker-based, of which *Firstage* is a type (Prochazka et al.

2011). As most research in MAR immersion concentrates on marker-less applications, it seems necessary to study immersion in terms of *Firstage* and similar applications. Furthermore, it appears that both MAR immersion and context immersion are complementary and to some degree interdependent when it comes to MAR; indeed both appear suitable for measuring different aspects. Thus far they have never been measured for in the same study. The questionnaire in this paper aims to do just that.

It should be noted that it is beyond the scope of this paper—given time constraints—to conduct any qualitative data collection. Thus the questionnaire designed and the method proposed is theoretical. However, it is hoped that the theoretical approach could be employed to measure other marker-based MAR applications in the future; it is important for the developers of *Firstage*—and applications like it—to be aware of the immersive qualities of the application for future development with a view to increasing user retention and revenue.

1.4 Outline

In Chapter Two, relevant literature will be reviewed to identify the different ways in which MAR can be deployed so that a strong understanding of the technology can be achieved. Various theories of immersion will be introduced so that an appreciation of the complexity of the issue may be attained. Given that *Firstage* is an MAR application with a focus on blended musical performances, an overview of the role of audio in immersion will be given. Finally, relevant studies will be analysed to determine the appropriate models for use in measuring MAR immersion. In Chapter Three, the state of the art will be reviewed so that *Firstage* can be situated within the wider context of MAR applications. Firstage will then be analysed to ascertain the immersive characteristics of the application so that the relevant models of immersion can be applied. In Chapter Four, the relevant models of immersion will be applied and the process undergone for the generation of the questionnaire will be illustrated. A potential implementation of the study and the method by which the results of the questionnaire may be appropriately analysed will be described. Finally, in Chapter Five, the findings of the paper will be discussed and some potential concerns with and limitations of the proposed study will be outlined.

Chapter 2. Background

2.1 Introduction

As the intent of this paper is to develop a questionnaire which could be used to measure the immersive qualities of *Firstage*, this chapter will serve as a review of the appropriate literature. *Firstage* is an MAR application, and, importantly, one with a focus on music. The scope of the research thus includes several key areas. An overview of MAR implementations will be presented so that the technology used in *Firstage* can be understood. Then the various theoretical conceptions of immersion will be investigated, both in AR in general as well as in MAR specifically, as well as audio's role in the same. Finally, previous studies which have been conducted on MAR immersion will be explored so that a suitable structure for the questionnaire can be established.

2.2 MAR applications

In the past, AR applications required desktop computers, bulky dedicated hardware such as head mounted displays, and were thus out of the reach of most (Schmalstieg et al. 2011). With the rise of ubiquitous computing, AR has seen the creation of a new paradigm. Smartphones—with in-built cameras, large screens, extreme mobility, and access to the internet—provide for diverse and powerful implementations of AR and thus become an ideal platform for its development (Lee & Kim 2011). MAR is therefore characterised as AR which is created for and accessed by mobile devices and used within a mobile context (Olsson et al. 2013).

MAR applications can be divided into two main categories: markerless and marker-based (Georgiou & Kyza 2017; Kim et al. 2014; Prochazka et al. 2011).

2.2.1 Marker-less applications

The most common marker-less applications are location-aware games or services which utilise location data collected from the device's GPS and overlay appropriate context dependent information onto the device's display (Kim et al. 2014; Prochazka et al. 2011; Olsson et al. 2013). Reid et al. (2011) maintains that location-aware applications are so unique in the way that users interact with them that they may be considered a new type of interactive media entirely; they take place in a real space, and rely on spatial awareness and, importantly, movement on the part of the user.

The simplicity of these types of applications bears both negative and positive aspects. Due to the fact that they do not provide the user with any information about any specific objects shown in the camera picture, i.e., that all of the data is connected to GPS location, the richness of the experience may be limited. However, given the same simplicity, i.e., the lack of a requirement for complex calculations, such applications are quite easy to implement (Prochazka et al. 2011). Furthermore, it allows for the delivery of location and context-aware services with real time updating of content.

2.2.2 Marker-based applications

While location-aware applications use GPS to embed relevant information, marker-based applications analyse and process the input of a device's camera to identify a specific object and provide contextual information or media related to it (Johnson et al. 2010; Lee 2012).

This can be achieved in two ways. The first way uses artificial markers such as QR codes—or 2 dimensional barcodes—which are identified by the software. The detection of artificial markers is relatively simple due to that fact that the software will have the precise mathematical model of the object to hand and can calculate its presence—despite perspectival or morphological variances—based on its geometry. Indeed, this approach allows for the manipulation of virtual objects, which can be e.g., tilted or rotated by users given that objects are tied to the markers (Cheng & Tsai 2013). Nevertheless, it still necessitates allowing for variables such as low or saturated light conditions in which detection of the marker may become somewhat problematic (Prochazka et al. 2011).

The second way is through image recognition – identifying natural objects in the environment. In this case the software must calculate the presence of specific objects based on edge and corner detection, feature descriptor and matching, outlier removal, and pose refinement (Wagner et al. 2010). Thus the process is much more complex than artificial object identification and therefore requires more powerful devices.

The issue with a marker-based approach is that it does not allow for location and context-aware content delivery to the degree of a marker-less approach. Image recognition in some way overcomes this, however it is more costly to achieve in terms of processing power.

Considering the above, *Firstage* can be identified as a marker-based application which uses artificial markers to track the content it delivers.

2.3 Understanding immersion in MAR

In order to consider immersion in terms of *Firstage*, we must first understand what is meant by immersion and to separate it from the often conflated term, presence. Unfortunately, there is no consensus within the academic field as to precisely what immersion *is*. Slater (1999) and Slater & Wilbur (1997) regard immersion as something objectively measurable within a virtual environment (VE), the result of the fidelity of the experience that the system provides. Lara Rojas (2013) echoes this sentiment and holds that virtual systems should provide realistic graphics rendered in stereoscopic vision along with 3D sound, while also potentially giving haptic feedback so that the user can experience something closely related to real world experience; the greater the realism, the greater the immersion. According to Slater (1999), it is therefore possible to separate the concepts of immersion and presence due to presence being the user's subjective experience of that virtual environment, or their feeling of 'being there'. Indeed, the sense of 'being there' is found to be synonymous with the concept of presence in a number of studies (Kim et al. 2014; Reeves 1991; Steuer 1992). Witmer & Singer (1998) defines presence similarly, yet immersion too is understood as a psychological phenomenon which cannot be objectively measured. The further concept of involvement is introduced, and it is the role of involvement and immersion to bring about presence; both are necessary conditions for it. This is done through creating highly involving tasks within highly immersive environments. While the experience of immersion is a psychological state, it is nevertheless dependent on a technically advanced virtual environment to bring it about; the experience must have realistic virtual objects—and be continuous and undisturbed—to be immersive.

However, later studies have found that immersion does not depend on technically advanced systems; even applications with inferior graphics or audio can result in high levels of immersion (Georgiou & Kyza 2017; McMahon 2003). This could be the result of blocking out the real world environment through the use of goggles or headphones or through mental absorption in a highly engaging narrative (McMahon 2003). Weibel et al. (2010) argues that immersion occurs when people read, watch movies, or play video games; immersion here likewise depends on emotional involvement and absorption. Thus immersion is a natural human state which is neither new nor associated only with virtual environments (Georgiou & Kyza 2017). Brooks (2003) argues that a key component for immersiveness is observation of the rule '*Don't be boring*'. Storytelling and film has long adhered to this rule—humans are narrative animals—and we find that belief, or the suspension of disbelief, is essential for immersion. Murray (1997) argues that we are predisposed to believing; we do not so much suspend our disbelief as we actively look for things to believe in. Any technological additions should therefore aim to supplement this natural tendency (Brooks 2003).

Much recent research into immersion has been conducted in the area of videogames. It may be an area in which the concept may be quite appropriately investigated, given the above, inasmuch as videogames mix the narrative properties of film and storytelling with the interactive properties of digital media. Csikszentmihalyi (1990) conducts a study of the experiences of individuals doing activities they enjoy and defines *flow* as the extreme, optimal state of engagement during which "individuals are so involved in an activity that nothing else seems to matter" (1990, p. 4). In Cheng et al. (2015) the same concept of *flow* is recontextualised as immersion. Immersion, like *flow*, describes the degree of involvement with a computer game yet includes three distinct levels—*Engagement*, *Engrossment*, and *Total Immersion.*

Flow and *Total Immersion* experience some overlap here, particularly—as noted in Jennett et al. (2008)—as regards the experience of the distortion of time and involvement in doing a task. Yet immersion may be more appropriately applied, conceptually, to the suboptimal and non-extreme state of involvement present in the gaming experience; not all experiences bring about *Total Immersion* or *flow*, yet it is still correct to speak of immersion as an overarching concept (Brown & Cairns 2004; Ermi & Mäyrä 2005; Jennett et al. 2008). In this context, immersion retains its psychological quality in keeping with Witmer & Singer (1998) – it is not something that is objectively measurable in terms of the technology which provides for it, but rather a subjective state of relative involvement.

There is clearly much disagreement over an appropriate conception of immersion. Yet in the above we see that each of the conceptions appears in a particular context – be it games, VEs, or non-technologically-mediated media. Those varying contexts largely seem to determine what immersion is understood to be. Immersion in MAR is a relatively new phenomenon with quite little research having been done on it. As we are dealing with MAR with *Firstage*, and thus must locate an appropriate understanding of immersion, in the following we will first view the concept of immersion as it is applied to AR after which we will explore the concept of immersion as it may be applied to MAR technologies.

2.3.1 Immersion in AR

VEs and AR share many characteristics, yet stand apart in terms of experience. The most significant difference is that AR blends real world sensory information with virtual information (Azuma 1997; Azuma et al. 2001) while VEs exert considerable control over the entire sensory experience of the user. Thus the ultimate goal of VEs of attaining a sense of presence in the virtual world could be seen as too extreme for AR (Georgiou & Kyza 2017). To speak of presence in AR—the feeling of 'being there'—does seem somewhat tautological when we are talking about the augmentation of a real world environment with virtual objects. Thus any talk of presence in AR must relate to the way in which the virtual content appears as seamlessly integrated with the real environment, that "the 'machinery' of the AR experience should not be evident to the user" (Steptoe et al. 2014, p. 214). MacIntyre et al. (2004) argues that two main factors play a role in this illusion of non-mediation: technical factors, such as realistic graphics, and; consistency of content, which is related to the way in which virtual objects behave according to users' expectations.

Although the above literature are dealing with *presence* in AR, similar requirements could be said to apply to immersion, particularly if we consider immersion in terms of a graduated scale, as in Cheng et al. (2015) and Jennett et al. (2008), with presence or *flow* being reconceptualised as *total immersion,* the optimal, extreme level of involvement, below which sit *engagement* and *engrossment* respectively. Thus more immersive experiences in AR require applications with greater technical factors and effective consistency of content, though they might not be expected to reach the level of *total immersion* possible in VEs.

Figure 2.1. A HMD - the Microsoft HoloLens (source: www.microsoft.com)

Most implementations of AR, such as that of the *Microsoft HoloLens* (Microsoft 2016), rely on the use of HMDs (Figure 2.1). Such displays envelope most of the field of view of the user with a transparent lens onto which virtual objects or information are superimposed. Thus close to the entire sense of sight can be manipulated by the system. Early studies on immersion in AR, concerned mostly with the visual aspect, maintain that because virtual objects are being blended with real world information, significant graphical fidelity is necessary to create an immersive experience (Azuma et al. 2001; Höllerer & Feiner 2004). Photorealistic rendering presents a problem in terms of capturing the illumination information of the real world environment in order to model appropriate shading and highlights (Azuma et al. 2001). While VEs enjoy complete control over aspects such as lighting in order to create immersive experiences, AR is at a disadvantage due to the complexity in achieving a realistic blend.

Vividness is reported to be important for the immersive qualities of an AR system (Kim et al. 2014; Steuer 1992). Vividness is the ability of a system to create a "sensorially rich mediated environment" (Steuer 1992, p.90). It pertains to both the breadth—the number of different senses provided for—as well as the depth—the level of realism of the virtual information presented to those senses— of an AR system (Coyle & Thorson 2001). Thus the greater the breadth and depth of a system, the more immersive the experience.

Technical factors are not restricted to the sensory properties of virtual objects alone, interactivity is also important (Steuer 1992). Perceiving an augmented real world environment while being able to interact naturally with the virtual objects contained in it has been shown to positively affect the quality of user perception and immersion (Kim et al. 2014; Narzt et al 2006; Nee et al. 2012).

As regards the requirement for consistency of content (MacIntyre et al. 2004), we find in Azuma et al. (2001) four factors which affect a user's perception of such: latency, depth perception, adaptation, and long terms use. Latency is the delay a user experiences between interaction and response, and it is noted that even delays of 10 milliseconds make a marked statistical difference. Depth perception refers to the perceived distance of a virtual object from the user, which is considered difficult to achieve due to the many factors involved such as low resolution or dim displays and difficulties in rendering occlusion, the digital approximation of ambient light and shade. Adaptation relates to how users must adapt to an AR display—such as a HMD—and then readapt to real world conditions after removing it. Long term use refers to many AR displays historically tending towards being uncomfortable or causing eye fatigue during use.

Perhaps the most concise summary of AR immersion is found in Steptoe et al.:

> "immersive AR [...] relates to fostering in users a perceptual state of non-mediation, which arises from a high level of technologically-facilitated immersion and environmental consistency, and which in turn may give rise to realistic behavior and response" (2014, p. 214).

For AR to be immersive, it is essential that both virtual content and the consistency of that content be delivered effectively.

2.3.2 Immersion in MAR

Figure 2.2. MAR display (source: www.moremobilemarketingx.com)

AR technologies which utilise hardware such as HMDs can manipulate the entire visual field of the user. In the case of MAR, the capacity for the system to produce such a visually immersive experience does not exist to the same extent; users view the blended environment through the screen of their device (Figure 2.2). Nevertheless, it has been argued that while the screen of a smartphone is small, it does allow for wider peripheral view of the surrounding environment through the user's ability to move the device and scan the area (Billinghurst & Henrysson 2009; Kim 2013). Indeed, MAR might instead provide for a different kind of immersion than that of AR using HMDs (Georgiou & Kya 2017; Kim 2013).

Many of the same issues present in AR, such as the blending of the real with the virtual, and rendering lighting and occlusion realistically (Azuma et al. 2001), persist in MAR. However, MAR faces a further challenge: user distraction. Though speaking about VEs, Witmer & Singer (1998) argues that the 'distraction factor' present in a technology should be attended to when assessing its immersive potential. While one of MARs strengths is in its extreme portability, the range of environments it may be operated in increases the range of possible distractions that may arise (Georgiou & Kyza 2017; Reid et al. 2011). Traffic, general noise, animals or insects, weather, or unexpected events can distract the user from the task they are undertaking (Dunleavy et al. 2009; Georgiou & Kyza 2017; Reid et al. 2011). Furthermore, issues in MAR such as screen glare or the screen being too bright in dark environments are significant concerns, and the hardware on which MAR applications are run varies often quite dramatically in terms of graphical capabilities or the presence of GPS (Cheng & Tsai 2013; Dunleavy et al. 2009; Georgiou & Kyza 2017; Nilsson & Svingby 2009).

A preferred application of the concept of immersion in MAR is that of a graduated scale in which *Engagement*, *Engrossment*, and *Total Immersion* are the three levels (Cheng et al. 2015; Georgiou & Kyza 2017; Jennet et al. 2008). Georgiou & Kyza terms this the "operationalization of immersion as a continuum towards flow and presence" (2017, p. 26). Considering the tendency of MAR applications towards inviting distractions upon their users through real world environmental conditions, it has been argued that *Total Immersion*—or *flow*—is unlikely to be achieved in MAR application use (McCall et al. 2011; Reid et al. 2005). Nevertheless, many of the same immersion factors which come to bear on AR are present in MAR and due consideration should be given to them. In particular, accurately tracking and registering of virtual objects within the real world environment is considered very important, in keeping with the requirement of consistency of content (MacIntyre et al 2004; Papagiannakis et al. 2008; Regenbrecht et al 2002; Steptoe et al. 2014).

In MAR, the registration and tracking of content can be done in two ways, as illustrated in §2.2. Marker-less (using GPS) and marker-based (e.g., using 2D barcodes) approaches place the virtual information within the screen. However, the requirement for consistency of content means that the accuracy of the tracking must be high. Artificial markers are adept at this as they can establish a frame of reference for tracking (Schmalstieg et al. 2011). The virtual object can then be manipulated by the user by tilting, rotating, or moving the marker, which causes the object to move correspondingly (Cheng & Tsai 2013). Because of this, tracking through the use of an artificial marker appears to be more flexible than through the use of GPS which locks objects to geographical points (Cheng & Tsai 2013; Kim 2013).

2.3.3 Context immersion in MAR

Given that MAR does not enjoy the same level of control over the sensory immersion factors as AR using HMDs, a new approach to immersion in MAR—context immersion—has been theorised (Kim 2013; Lee & Kim 2011). Context immersion is derived from the notion of context-awareness, a term intimately tied to the idea of ubiquitous computing: static computer systems which adapt to a dynamic environment (Lieberman & Selker 2000). We find that context operates at several levels with the primary level being grounded in the elements in a given environment: the where, when, who, what, why, and how of people interacting with those elements (Kim 2013; Kirsh 2001). Smartphones serve as excellent devices for context-awareness as they possess cameras, good quality screens, GPS, compasses, and accelerometers, thus MAR applications can easily deliver personalised content to users according to their particular context (Kim 2013).

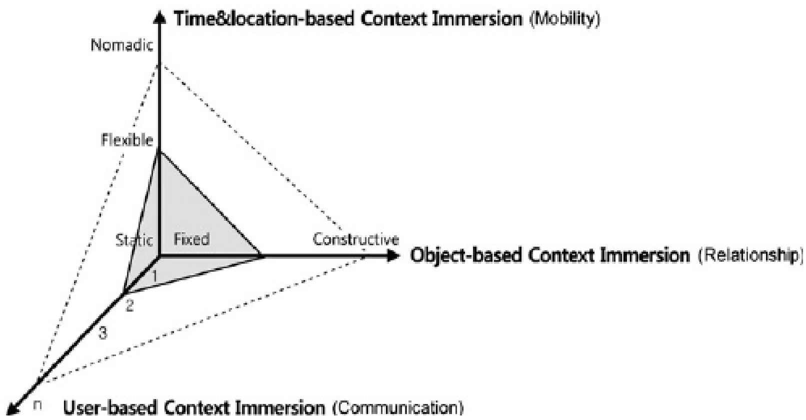

Figure 2.3. Dimensions of context immersion (source: Lee & Kim 2011).

Kim (2013) argues that through utilising context-aware services, MAR applications can provide for strong immersion, though of a type quite different to that of sensory immersion. Context immersion is the result of the relationship between user, object, and time and location. MAR applications which can exploit these relationships to augment a user's embodied interaction with the real environment with contextually-relevant information can be said to increase immersion by creating a sense of the virtual elements becoming real for the user (Kim 2013).

A three dimensional measure for context immersion is proposed (Figure 2.3), which is comprised of time and location-based context immersion, object-based context immersion, and user-based context immersion (Lee & Kim 2011). Time and location-based context, due to user mobility, requires that application use not be confined to small fixed areas or periods of use. Immersion, therefore, comes from the degree to which an application can provide itself for use regardless of time or location.

Object-based context is concerned with the relationship between the user and a given virtual or real object in the environment with the aim to create a sense of the virtual having a real relationship to the user (Lee & Kim 2011). For Kim (2013), immersion would begin with the application dynamically recognising objects or information in the environment. These objects could be real world objects registered by the application when users are nearby, or virtual objects such as media or user generated content tied to specific objects or areas within the environment. Immersion, then, depends on the degree to which the objects and information are reliably tracked by the software and hardware, creating a sense of the virtual objects being in some way grounded in reality. As seen previously (MacIntyre et al. 2004; Steptoe et al. 2014), the continuity of content is what is under scrutiny here. The better the registration and tracking of objects—and the higher the quality of information attached to those objects—the more the interaction with those objects becomes 'real' for the users (Kim 2013).

User-based context is concerned with the networking of users, the content they create, and how they can interact; it is the social networking capability which allows for personalization, sharing, and collaboration (Lee & Kim 2011). As we see in Figure 2.3, the number of users can range from one to potentially infinite, yet an ability to personalise the application remains important. The more an application caters for an individual user's personal needs, the greater the engagement and hence immersion. Communication between users is what promotes immersion in a user-based context (Lee & Kim 2011; Kim 2013). More immersive environments have many users rather than just one given that we want to interact and share our experiences with others (Lee & Kim 2011). Thus the extent to which users can share, interact, and even collaborate determines the degree of immersion.

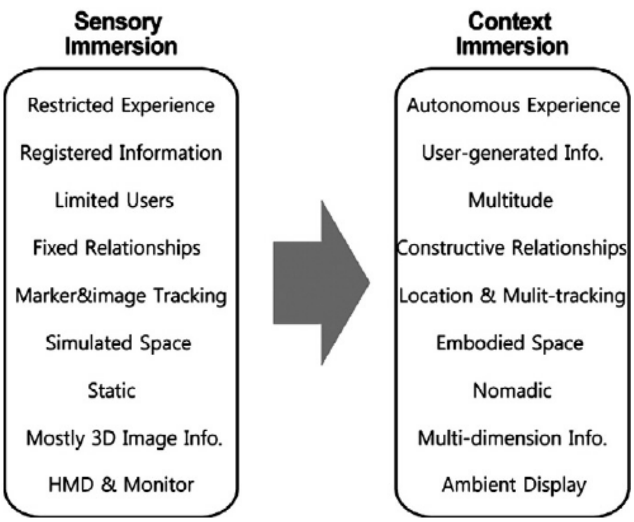

Figure 2.4. Characteristics of context immersion (source: Lee & Kim 2011)

In Figure 2.4, the differences between sensory immersion—which VEs and to a lesser extent AR excel at—and context immersion possible in MAR is displayed. The restricted experience of VEs—that the experiences may only take place in controlled and isolated environments with limited users—is replaced by autonomous experience and a multitude of users. User generated information, constructive relationships (dynamic content) and multi-dimensional information (information contained in many types of media) replace prerendered registered information and mostly 3D images. Location and multi-tracking, nomadic tendency, and ambient displays—display interfaces which provide context dependent information and draw attention only when needed—reflect the mobility and ubiquity of mobile devices and replace the fixed nature of and dedicated hardware present in VEs (Kim 2013). Finally, embodied spaces, spaces populated by objects which become 'real' for the users through information quality and relevance, replace simulated spaces which cannot be dynamically reshaped by users.

For Lee & Kim "context immersion in [MAR] is different from the sensory immersion in VE achieved by the conventional technologies such as visual and acoustic features" (2011, p. 207). Nevertheless, MAR affords users a meaningful connection with an augmented environment through building relationships between virtual objects and information and user's real world lives.

2.3.4 Immersive Audio

As we are dealing with an application which focuses on the delivery of audio to the user, it is necessary to consider the role that audio and acoustics plays in terms of immersion. For immersive AR, psychoacoustics—the science of sound and its physiological effects—plays an important role, particularly in terms of spatial audio and reverberation.

Spatial audio psychoacoustics affects our ability to locate sound objects within a 3D space, and depends on two primary mechanisms: the detection of i) timing differences, and ii) amplitude differences between the ears (Rumsey 2001). In Figure 2.5, a sound coming from an object on the right of the listener creates a delay between the right and left ears in reception of the sound. The listener thus perceives the sound as louder in the right ear than in the left ear and aids in an agent's perception of sound source localisation (Zhong et al. 2015).

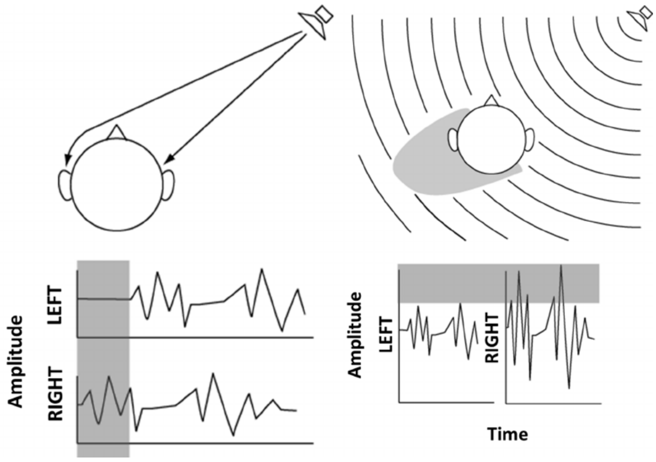

Figure 2.5. Interaural time and level difference (source: Zhong et al. 2015)

Very few natural environments allow for the propagation of sound in a direct manner, however. In most real world environments sound waves bounce off surfaces and objects on their way to the listener which results in many 'copies' of the same sound arriving to the ears with different time delays. This is known as reverberation (Figure 2.6). If I am standing in a room and a sound occurs in front of me, the sound waves will reflect off different surfaces in the room and reach my ears at different times. The direct sound will reach my ears first, followed by—in increasing time delays—the reflections of the same sound from nearer to further away surfaces respectively. This is what gives a space its acoustic characteristics (Rumsey 2001). In order to realistically place a sound object within a virtual space it becomes necessary to model the acoustic characteristics of that space using digital reverberation.

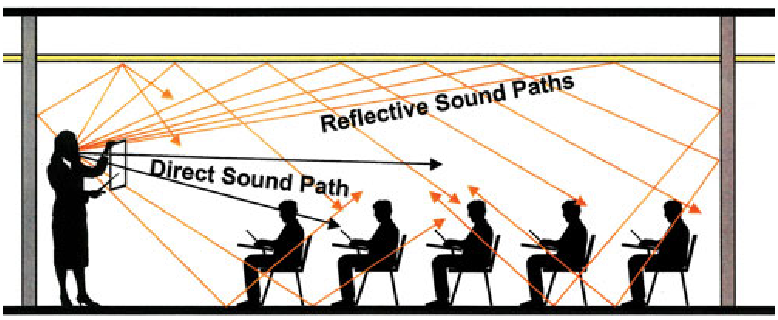

Figure 2.6. Reverberation (source: http://continuingeducation.bnpmedia.com)

For immersion, it is important that what is seen is corresponds to what is heard, and "a mismatch between the aurally perceived and visually observed positions of a particular sound causes a cognitive dissonance that can seriously limit the visualization enhancement provided by immersive sound" (Kyriakis 1998, p. 941). Thus accurately placing 3D sound objects within an augmented environment is important for consistency of content (MacIntyre et al. 2004; McMahan 2003; Steptoe et al. 2014). Implementing spatial audio in AR to place a sound object in a blended space can be done through binaural synthesis and the wearing of headphones by the user (Gardner 1999). Reverberation can be emulated through either algorithms or impulse response convolution. For the former, a digital signal processor uses an algorithm to model physical spaces. In the latter, by recording a space's impulse response (IR) and accurately capturing the its precise physical acoustic characteristics, the IR can be convolved with sound objects to virtually place them within the recorded space. Algorithmic reverberation generally requires less processing power, memory, and storage space while convolution reverberation is more accurately representative of a physical space.

In MAR applications, the use of both spatial audio and reverberation should be considered for their immersive qualities, though there are some shortcomings in the technology. Smartphones often only have one internal speaker which means that localisation—which requires a stereo sound playback source—may be difficult to implement. Nevertheless, many smartphone users utilise headphones which overcomes the problem. In the case of reverberation, the pre-rendering of reverberation onto sound objects is preferred as mobile technology is not yet powerful enough for real-time implementation with multiple sound sources (Paterson et al. 2013). It is therefore difficult to dynamically assign reverberation properties to sound objects, particularly if the same sound object should appear in several different spaces which each possess different acoustic characteristics. In this case, many different reverberation pre-renderings of sound objects could be included in the application yet this would add to the storage requirements on the hardware.

2.4 Measuring immersion

As noted in Georgiou & Kyza (2017), the majority of recent research into psychological immersion has been in the area of video games. As we have seen previously in Jennet et al.

(2008) and Cheng et al. (2015), the working model is that which Georgiou & Kyza (2017) terms “the operationalization of immersion”.

Both Jennet et al. (2008) and Cheng et al. (2015) base their research on the theory of Brown & Cairns (2004), which holds that immersion is a graduated psychological process in which the three levels are *Engagement*, *Engrossment*, and *Total Immersion*. The research in Brown & Cairns (2004) involves interviewing seven gamers about their gaming experiences. In this case, *Engagement* is the initial barrier to entry and can be measured in terms of access and investment, with access being the ease with which a user can control their avatar and that the type of game meets their preference, and investment coming as a result of this. In other words, if the means of control is intuitive and accessible and the type of game is one of the user’s preference, the user will subsequently invest their time and attention in playing.

Engrossment is the next level in which users have become more involved in—and thus direct more attention to—the game due to their emotional investment in it. In order for a user to become emotionally invested, the game must be constructed in such a way as to require more time, effort, and attention from the user. Brown & Cairns (2004) reports that gamers become less aware of the real world and begin to suspend their disbelief of the game world being virtual.

Total Immersion is equated to presence or *flow* (Csikszentmihalyi 1990), wherein the game world is all that matters to the user. Empathy as well as the game’s atmosphere are essential in this: the user must be able to intimately empathise with their character and the game environment must be sufficiently content-consistent to not produce any distractions (Brown & Cairns 2004).

Jennet et al. (2008) develops a questionnaire based on Brown & Cairns (2004) while incorporating aspects of both presence and cognitive absorption into its methodology. While acknowledging that gamers can reliably and intuitively identify immersion in themselves, one of the goals of the study is to identify the quantitative factors important in bringing about the feeling of immersion. Five factors to explore are identified: cognitive involvement, real world dissociation, emotional involvement, and two game factors–challenge and control. However, the study proves inconclusive .

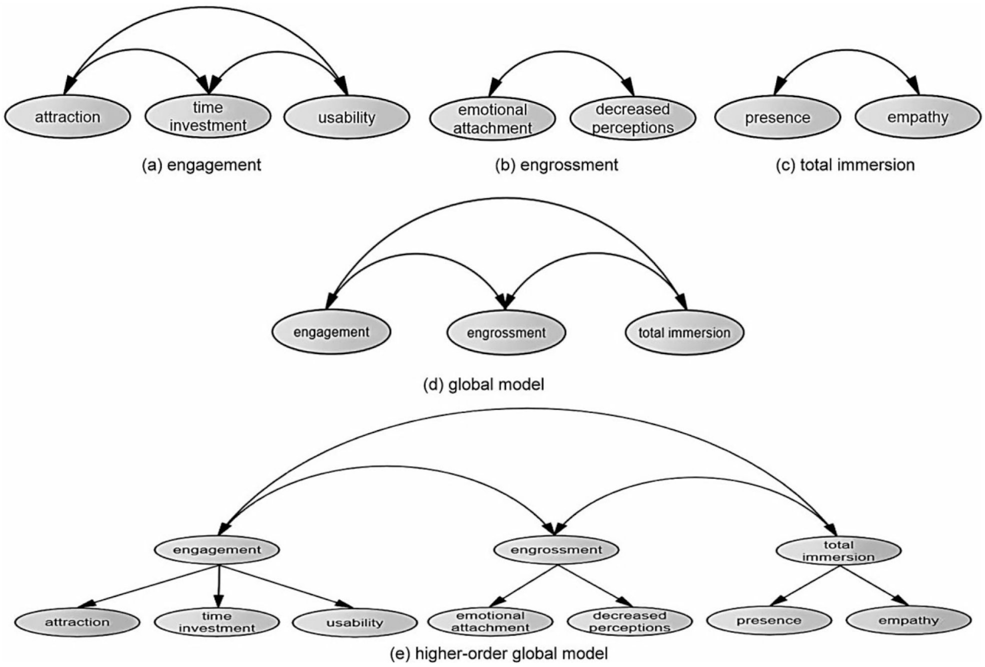

Figure 2.7. Multidimensional model of immersion (source: Cheng et al. 2015)

Cheng et al. (2015) designs the Game Immersion Questionnaire based on the graduated scale of *Engagement*, *Engrossment*, and *Total Immersion* of Brown & Cairns (2004), but explores and validates multidimensional factors within each level. The appropriate first order factors are found to be: within *Engagement* – “attraction”, “time investment”, and “usability”; within *Engrossment* – “emotional attachment” and “decreased perceptions”, and; within *Total Immersion* – “presence” and “empathy” (Figure 2.7). These factors are arrived at by interviewing a number of participants on their experiences of video game immersion and extracting relevant items for analysis based on the interviewees’ descriptions. Items are discarded or combined to remove overlap. A questionnaire employing a five-point Likert scale is created and deployed, and statistical analysis in the form of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) is performed on the results, which is found to confirm the proposed model (Cheng et al. 2015).

While observing the reliability of the model produced in Cheng et al. (2015), Georgiou & Kyza (2017) identifies the differences between MAR and video game experience and adapts the questionnaire to measure MAR immersion. The three level model of *Engagement*, *Engrossment*, and *Total Immersion* (Brown & Cairns 2004) as well as the multidimensional model of Cheng et al. (2015) are employed with inappropriate or irrelevant questions altered in order to make the questionnaire apt for determining immersion in two location-aware MAR games.

Georgiou & Kyza (2017) notes three main areas of difference. Firstly, MAR games employ mobile and location-aware interfaces, i.e., the users interact directly with the gameworld *through* the interface—which is itself part of the game world—and thus the user's action is not mediated through an in-game avatar. Secondly, real-world environments and virtual objects are combined, which is important in terms of feelings of total immersion; presence as in 'being there' is not possible in the same sense. Rather it is a feeling of being present in a blended space of the real and the virtual. Finally, MAR applications turn the real world into a game space while video game play takes place entirely within a virtual environment. The real-world therefore contains many more potential distractions, as we saw in section §2.3.2.

Bearing these differences in mind, Georgiou & Kyza (2017) develops a questionnaire with appropriate alterations to make it applicable for MAR immersion. This involves reconceptualising "empathy" as "flow", and combining "time investment" and "attraction" into one single factor, "interest". Following deployment of the questionnaire, both EFA and CFA, as well as Cronbach's α—a determination as to whether the items reliably measure what they purport to—show the alterations and implementation to be satisfactory.

The final study is that of Kim (2013) which measures context immersion. Employing a seven-point Likert scale, the questionnaire groups questions into six categories which relate to the characteristics we see in Figure 2.4 in section §2.3.3. *Interface* and *sensory* are interested in the control factors present in MAR, or how the user interacts with the technology. *Involvement* is interested in the user's experience of creating and contributing multidimensional information. *Motivation* relates to the constructive relationships which are forged between the multitude of users. *Mobility* is concerned with location and multi-tracking, the nomadic nature of users, and ambient displays. Finally, *reality* relates to how MAR allows users to experience embodied virtual space when using the technology.

Kim (2013) devises a number of questions for each category and the questionnaire is deployed after use of a location-aware MAR application, *Ovjet.* Context immersion is discovered to be moderate, and the paper concludes that the study suggests the presence of context immersion, that smart devices enhance accessibility of an AR system, that passive users become active in their environments, and that by exploiting context-awareness in MAR, constructive relationships and the creation of meaningful embodied virtual space can be achieved.

2.5 Conclusions

In this chapter we have discovered the two main approaches to implementing MAR applications: marker-based or marker-less. Artificial markers or image recognition can be used to achieve the former, while a device's GPS is used to achieve the latter. Both approaches have their strengths and weaknesses; marker-based provides for more accurate tracking of virtual objects, while marker-less allows for location-awareness and context relevant information delivery.

As we have seen, immersion in AR and MAR can be said to depend greatly on the illusion of non-mediation and a relatively seamless blend between the real and the virtual leading to a suspension of disbelief. This is achieved through consistency of content—accurate tracking of virtual objects in the environment—as well as media quality, be it graphical/audio accuracy or context relevant information. However, we have seen that rendering realistic graphic and audio objects is difficult in MAR: the former due to difficulties in simulating such things as realistic light and shadow, and the latter due to possible hardware limitations in mobile devices. Thus a realistic blend might prove difficult to achieve. Nevertheless, context immersion—the relationships between locations, users, and objects—provides for immersive MAR applications despite such technical limitations.

We have seen some key studies which measure immersion in digital media. For the purposes of this paper, the studies found in Georgiou & Kyza (2017) and Kim (2014) presenting models of multidimensional MAR immersion and MAR context immersion respectively are of particular interest.

In the next chapter we will explore *Firstage* in relation to other similar MAR applications, and present an overview of the potential areas of immersion to consider.

Chapter 3. Situating Firstage

3.1 Introduction

This chapter will first present an overview of some recent MAR applications in order to gain an understanding of the current state of the art. It will then analyse *Firstage* and compare it with the other applications so that it can be placed within a suitable context. Finally, the areas of immersion for which *Firstage* might be measured will be discussed, and some thoughts and concerns about potential strengths and weaknesses of the application in light of immersion will be offered.

3.2 State of the art

Greater processing power has enabled mobile technology to advance, which has allowed for more fully featured MAR applications. The most well known example of MAR is *Pokemon GO* (Niantic, Inc. 2016), having been downloaded approximately 550 million times within the first three months (Newzoo 2016). It is a location-aware 'exergame' (Kari 2016) which blends virtual characters with the real world environment. Players must search out the wider environment to hunt for the Pokemon characters which they can then catch by aiming and throwing a 'pokeball' at. Players can then use the Pokemon in in-game 'battles' or trade them with friends. As a location-aware game—and thus marker-less—it relies on GPS and compass, and registered content is geographically locked. Furthermore, since the game characters are cartoon-like, the realism factor is not of importance.

An example of a marker-based service is *INKHUNTER* (INKHUNTER, Inc. 2016), which allows users to 'try on' virtual tattoos (Figure 3.1). Users draw a marker symbol known as a 'square smile' onto the area of their body they would like to view the tattoo. They then choose a tattoo and the application erases the marker and places the tattoo in its place. Since it is marker-based, users can view the tattoo from different angles and distances by moving the smartphone around. Users can then capture an image of the scene, edit it with the in-built image editor, and share it on social media.

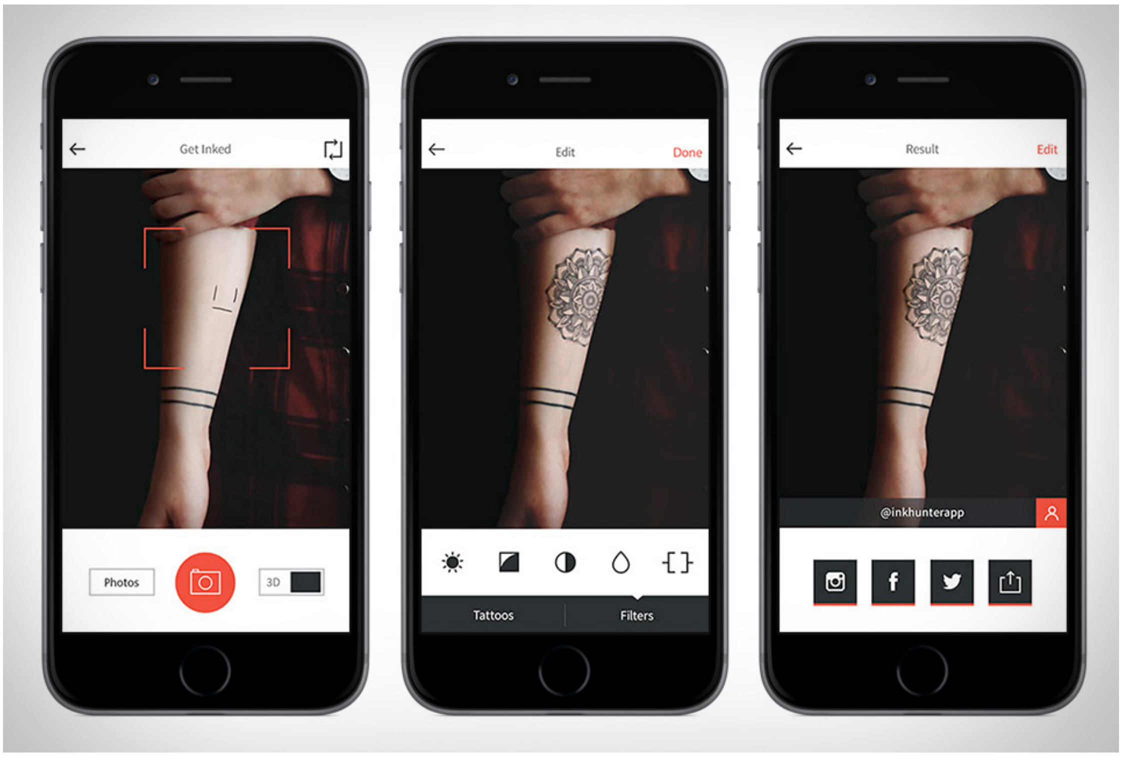

Figure 3.1. INKHUNTER application (source: http://inkhunter.tattoo/)

Blippar (Blippar 2011) allows users to tap on, or 'blip', objects in the environment and, using an image recognition marker-based approach, the application delivers content such as games, video, audio, or relevant information. For example, pointing one's device at the sky and 'blipping' might retrieve weather information, 'blipping' a movie poster might bring up the movie trailer or options to buy tickets, or 'blipping' a landmark might bring up tourist information about the area.

Both *Augment* (Augment 2011) and *Amikasa* (INDG 2014) allow users to place virtual objects in their real-world environments. *Augment* (2011) can import 3D models and can use user-defined markers or tracking. It serves as a virtual sales tool which can be used to demonstrate, for example, how products might look on a shelf or how a building might look once completed. Users can take screenshots of the augmented environment and share it over social media. *Amikasa* (2014) serves as a virtual interior designer and allows users to input the dimensions of their room and place 3D models of furniture from real brands int it within the application. They can then view their real-world room augmented with the new furniture in place.

While each of these applications augment the environment with digital information, none of them blends pre-recorded video and audio with real world locations. However, *The Lenz* (Deutsche Telekom AG 2017) does precisely this. Magenta, the Deutsche Telekom brand colour, is identified using image recognition in the form of what is termed Chroma Keying – the detection of a colour within a Pantone range. Any object in the user's surroundings can be scanned, be it flowers, a t-shirt, or a wall, and media will be blended with the environment with the media being entirely confined to the magenta coloured objects (Figure 3.2). The application is a partnership with music group *Gorillaz* and the media ranges from images of the band to video interviews to live performances.

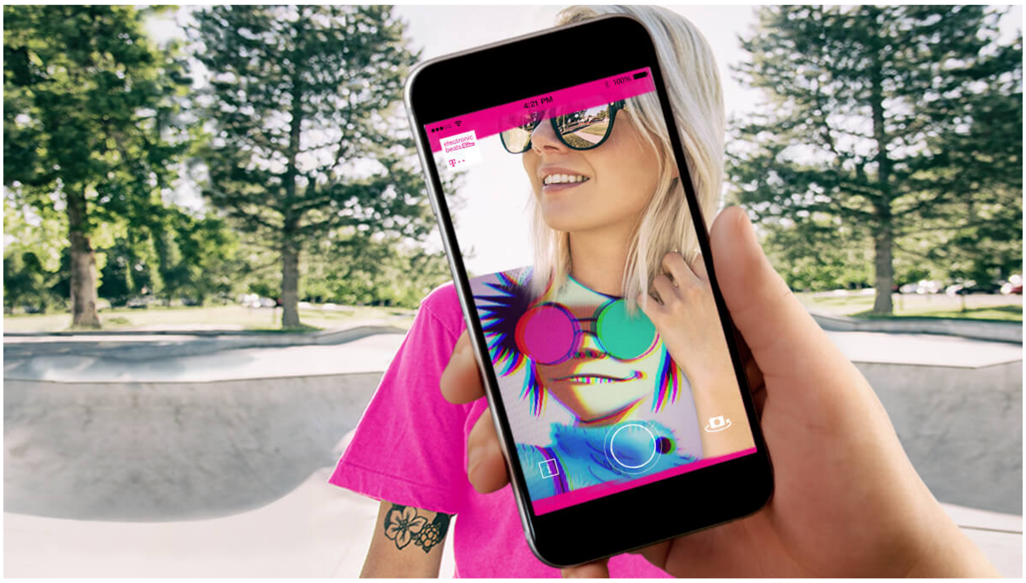

Figure 3.2. The Lenz application (source: http://www.electronicbeats.net)

Although a location-aware MAR game, *Bram Stoker's Vampires* (Haunted Planet 2012) involves having users search out their environment to hunt for ghosts and uses 3D audio to situate objects accurately within the environment. It incorporates both spatial audio and reverberation to achieve this which helps users find ghosts based on their location relative to the user, as well as convincingly place the sound and audio within the environment.

We have seen a selection of applications which blend digital media and information with real-world environments. Below, *Firstage* will be investigated to see how it can be compared.

3.3 Considering Firstage

Figure 3.3. A blended space created in *Firstage*. (Source: www.firstage.com)

Firstage is a marker-based MAR application which allows the user access to exclusive pre-recorded live video performances by music artists. The application uses both artificial markers as well as image recognition in order to deliver the experience. The artist's performance is recorded in the *Firstage* studio in front of a green screen with the performance subsequently being blended with the user's own environment. Unlike *Bram Stoker's Vampires* (2012), however, no use of spatial audio or reverberation is used to convincingly place performances within the environment. While similar to *The Lenz* (2017) in blending music with the user's environment (Figure 3.3), *Firstage* is nevertheless unique in that its focus is mainly on creating blended-space performances, while *The Lenz* (2017) introduces a variety of other media into the environment.

Figure 3.4. Firstage logo (Source: www.firstage.com)

The artificial marker—or 'stage'—used is the company's logo (Figure 3.4), which then serves to anchor the performance to a specific point. This way, similar to *INKHUNTER* (2016) and *Amikasa* (2014), the user can adjust their view of the area with their smartphone yet the performance remains situated in place. Anywhere a user may find the logo, be it in a magazine, on a coffee cup, or on a billboard or poster can serve as a 'stage' for a performance, just as products may be placed in any environment in *Augment* (2011). A more recent development is that by simple image recognition, in the form of pixel tracking, anything within the user's environment—be it a tabletop, a beer bottle label, or even a note of money—can be used as a marker (Holden 2016; Quinn 2016). This function is similar to *Blippar* (2011) inasmuch as any item in the environment can be enriched with digital information, yet it is applied in a more limited fashion.

Figure 3.5. Firstage (a) main screen and (b) performance mode (Source: itunes.apple.com)

A stated aim of *Firstage* is to create a promotional platform for musicians by delivering exclusive content in the form of said performances, sale of the exclusive audio from the performance, and various rewards for fans who donate money to the artist (Murphy 2015). It is also visualised as a way for potential fans to find out what an act is like before paying for ticket to see them by, for example, scanning the *Firstage* logo on a flyer or poster to view a performance from the artist (Holden 2016). Again, similar in function to *Blippar* (2011).

In Figure 3.5 (a), the app's main screen can be seen, which allows the user to select an artist whose performance they would like to watch. At this point, the app displays its performance screen (Figure 3.5 (b)) which offers the options of saving an image of a scene to share on social media, similar to *Augment* (2011), buying the track featured in the performance, or viewing the artist's profile. From the performance screen, it is also possible to forego the AR feature of the app and instead continue listening to just the audio from the recording.

A significant question is how the app separates itself from something like viewing an artist's performance on Youtube or other video sharing sites, or indeed simply listening to an audio recording through a music streaming application. This is a point of which the creators of *Firstage* are aware (Holden 2016). The ubiquity of live video performances on video sharing sites combined with their ease of production certainly presents a lower barrier for entry for emerging artists in comparison to *Firstage*, which requires an artist visit the company's studios in order to record a performance in front of a green screen. Furthermore, users must already have the app in order to view those performances. Thus *Firstage* must offer something beyond what is already available to both artist and user in return for their investment.

Of course, the blending of spaces is a selling point. However, while AR is a growing market, many AR applications fail soon after launch due to low usage rates (Kim et al. 2014). Novelty alone is not enough. Success in user retention is shown to rely on three major factors (Kim et al. 2014): information quality, interactivity, and visual quality. These are factors previously found to be integral for both immersive and context immersive MAR applications (Georgiou & Kyza 2017; Lee & Kim 2011; Kim 2013). Thus one way that *Firstage* can stand apart from the other media delivery methods is by offering a unique immersive experience to users.

3.4 Understanding Firstage in terms of immersion

The question then becomes what characteristics are present in *Firstage* and how they relate to the theories of immersion we have previously seen. Engaging users with relevant and interesting content and providing them with an easy to navigate interface is the first barrier to providing immersive experiences, and the quality of the delivery of that content then determines whether they will become more immersed (Georgiou & Kyza 2017). This is a basic consideration in measuring for immersion.

In terms of content delivery, *Firstage* blends performances recorded in one location with the immediate environment of the user. As we saw in Kyriakis (1998), there may be a degree of cognitive dissonance experienced by the user if attention is not paid to the spatial and reverberative characteristics of the space in which the blended performance takes place. However, as we have seen in §2.3.4, smartphones are not the ideal device for delivering such due to their single speaker and lack of processing power, though headphones can be used to overcome the former. Nevertheless, given that *Firstage* aims to deliver realistic performance experiences, the user's involvement will likely be affected by how well the blend is achieved. Similarly, how realistically visual elements such as lighting and ambient occlusion are rendered will have an effect on how convincing the blend between pre-recorded video and the live camera capture appears (Azuma et al. 2001; Kim et al. 2014; MacIntyre et al. 2004). These factors are important for immersion in MAR insofar as it affects a user's ability to engage emotionally with what they are experiencing (Georgiou & Kyza 2017). Likewise they are important for context immersion, particularly in terms of object-based context: creating embodied spaces where virtual objects come to feel real for users (Kim 2013).

Consistency of content delivered by the application in terms of its ability to accurately track virtual objects—the performances—within the augmented environment is considered essential for immersive experiences (MacIntyre et al 2004; Papagiannakis et al. 2008; Regenbrecht et al 2002; Steptoe et al. 2014). As *Firstage* uses a marker-based approach, and thus should provide reliable tracking of virtual objects, we might expect consistency of content to be high. However, that the application depends on connection to the internet may have a bearing on how well the application delivers an uninterrupted media feed, particularly since it is intended to be used without restrictions in a time and location-based context (Kim 2013).

Consistency of content also applies to how well virtual objects react according to users' expectations. Whether the performances can be interacted with and how well they respond to things like changes of viewing perspectives or distances are important considerations (Azuma et al. 2001; MacIntyre et al. 2004; Steptoe et al. 2014). How interactive the performances are—and what shape that interaction takes—will affect both immersion in general (Georgiou & Kyza 2017) as well as object-based context immersion (Kim 2013) by making the virtual performances feel more 'real'.

As we saw in §2.3.2, distraction factors play a role in immersion, and MAR is particularly susceptible to having external elements encroach on the experience due to its potential for use in a range of environments and environmental conditions (Georgiou & Kyza 2017; Reid et al. 2011; Witmer & Singer 1998). However, by delivering a more immersive experience through effective blending and consistency of content, users become less aware of their surrounding environment and any distractions can be largely filtered out (Georgiou & Kyza 2017).

Finally, Firstage offers users a way to share image captures of performances with friends on social media. The extent to which users can interact with one another is important for user-based context immersion (Kim 2013). Thus the ease with which Firstage allows users to interact and the extent of that interaction should be determined.

3.5 Conclusions

In this chapter some recent MAR applications which blend virtual media and information with real world environments in different ways were reviewed. We also discovered that *Firstage* bears many similarities with those applications, yet stands uniquely apart in that its focus is on delivering blended live performances, though it also allows users to share information with one another or to buy music from and invest in the artists they discover. Yet much of what *Firstage* offers can be achieved with non-MAR applications. Given that many MAR applications' successes are determined by their exploitation of the immersive potential in AR, the developers of *Firstage* may wish to determine how well the application performs in that regard. Therefore the characteristics of the application and how they relate to theories of immersion previously found in this paper were analysed. In the next chapter, a questionnaire informed by the characteristics uncovered here will be developed, with the appropriate models of measurement applied.

Chapter 4. Methodology and implementation

4.1 Introduction

This chapter considers previous research and questionnaires developed which are relevant to the immersive characteristics identified in *Firstage* in the previous chapter. Analysis of a number of studies in order to identify the structure of immersion as well as the conceptions of immersion relevant to the application was done in §2.4. This led to the development of a questionnaire to measure immersion in *Firstage* focusing on two relevant types of immersion: MAR immersion and context immersion. This chapter illustrates the models used and the process undergone for item generation. It then goes on to describe how the questionnaire should be implemented as well as how the results should be analysed.

As it is beyond the scope of this paper—given the time-frame allowed for the completion of the dissertation—no actual qualitative data has been collected at this point. Nevertheless, the purpose of this paper is to devise a questionnaire which could then be implemented by *Firstage* in order to determine the immersive qualities of the application. It is also hoped that the survey could be adapted by others wishing to measure immersion in MAR applications.

4.2 Models and item generation

As seen previously in §2.4 of this paper, there are two main models of immersion which may be applied to MAR applications. The first is the graduated scale model with three distinct levels of *Engagement*, *Engrossment*, and *Total Immersion* (Brown & Cairns 2004) expanded to include first-order factors within each level, e.g., the first order factors of "emotional attachment" and "decreased perceptions" within *Engrossment* (Cheng et al. 2015; Georgiou & Kyza 2017; Jennet et al. 2008). This will be termed the MARI (MAR immersion) model. The second is the three dimensional model of CI (context immersion) (Lee & Kim 2011; Kim 2013). Both of these models deal with different aspects of immersion in MAR. MARI deals more with 'general' immersion similar to the kind found in VEs, AR, and other media such as video games, films, or even books. CI deals with a kind of immersion that is unique to MAR and depends on the kinds of experiences ubiquitous computing affords. Thus the two might be considered somewhat complementary.

The immersive characteristics of *Firstage* illustrated in §3.4 show that both models may be applied to *Firstage*. Thus the questionnaire is composed of two sections, one for each model of immersion. Previous studies appear to have applied one model or the other, and while there is some degree of overlap in terms of qualities tested for, no study so far has measured for both.

4.2.1 MARI

Level	*Factor*
Engagement	Interest Usability
Engrossment	Emotional attachment Focus of attention
Total Immersion	Presence Flow

Table 4.1. The different levels of MAR Immersion and their first-order factors

The MARI model, based on the graduated model (Brown & Cairns 2004), identifies three levels of immersion: *Engagement*, *Engrossment*, and *Total Immersion*. Each of the levels are multidimensional; all have first-order factors which determine an application's performance for each level of immersion (Cheng et al. 2015; Georgiou & Kyza 2017). The structure (*Table 4.1*) is borrowed from the ARI Questionnaire (ARIQ) (Georgiou & Kyza 2017), which adapts the Game Immersion Questionnaire (Cheng et al. 2015) to allow for the differences between videogames and location-aware MAR games. The level of *Total Immersion* and its first-order factors in particular needed reconfiguration for an MARI context,

The question as to whether the structure of the ARIQ was applicable was considered given that *Firstage* is a marker-based service and the ARIQ is concerned with and location-aware games. Through analysis of the ARIQ, and recognising the immersive characteristics of

Firstage identified in the previous chapter, it was concluded that the similarities are such that the structure could remain. Thus the factors to be measured are: "Interest" and "Usability" for the level of *Engagement*; "Emotional attachment" and "Focus of attention" for *Engrossment*, and; "Presence" and "Flow" for *Total Immersion.*

Many items used to test the factors needed to be altered to account for the differences between the applications being investigated. In the case of "Interest", items needed to be reconfigured so that they related to the service *Firstage* provides—watching performances—as opposed to the experience of playing MAR games presumed in the ARIQ. Questions such as *"I liked the type of the activity"* or "*The topic of the activity made me want to find out more about it"* do not seem appropriate. Thus they become *"In general I like watching bands and music artists perform"* and *"Watching a performance encouraged me to watch more performances from other artists"* respectively. "Emotional attachment" too presented issues as games generally allow for more interactivity and challenges. Thus a question like *"I was impatient about completing the activity successfully"* is replaced by *"I felt a connection with the artist who was performing"*. Besides these significant divergences, many minor adjustments were made to all of the questions in order to make them suitable for *Firstage*, though the reasoning and theoretical considerations behind each item remained. In all, a total of thirty items for this section were developed: ten for *Engagement*, eleven for *Engrossment*, and nine for *Total immersion*.

4.2.2 Context Immersion

The CI model is composed of three dimensions: Time and Location-based, Object-based, and User-based context immersion (Lee & Kim (2011; Kim 2013). In the Context Immersion Questionnaire (CIQ) developed in Kim (2013), items are listed within a number of factors which broadly relate to the characteristics of context immersion we saw in Figure 2.4 in §2.3.3. These factors are: "Interface", "Sensory", "Involvement", "Motivation", "Mobility", and "Reality". However, the factors do not accurately correspond to the stated characteristics and there is much overlap between them, i.e., there are some characteristics represented in multiple factors. As a result, an entirely new questionnaire model was developed which is more consistent with the characteristics of context immersion identified in Figure 2.4. The factors are: "Nomadic" and "Location and Multi-tracking" for the dimension of *Time and Location-based* context; "Interface", "Embodied Space", "Sensory", and "Multidimensional Information" for *Object-based*, and; "Multitude" and "User Generated Information" for *User-based* immersion (*Table 4.2*).

Dimension	*Factor*
Time & Location-based	Nomadic Location & multi-tracking
Object-based	Interface Embodied space Sensory Multidimensional information
User-based	Multitude User generated information

Table 4.2. Measuring the different dimensions of Context Immersion

For the *Time and Location-based* factors, items which measure *Firstage*'s portability and ubiquity were generated such as *"I was able to use Firstage in a number of different environments"* and *"I felt that Firstage gave a unique experience depending on where I used it"*. For the *Object-based* factors, items which test how a user perceives and interacts with an embodied blended space were generated, e.g., *"Firstage made real world locations feel more meaningful"* or *"The performers stayed in place if I moved my smartphone around"*. For the *User-based* factors, items which investigate how *Firstage*'s users can interact were generated, such as *"I felt part of a community when using the app"*. In total twenty-six items were created: six for Time and Location-based, fourteen for Object-based, and six for User-based context immersion.

4.3 Implementation of the questionnaire

As this paper is focused on developing a potential immersion questionnaire for *Firstage* and does not carry out any empirical research, below will describe a potential approach. The proposed questionnaire can be found in the *Appendix* of this paper. As the questionnaire has not yet been implemented, it would be advised that exploratory and confirmatory factor analyses be conducted in order to determine the fitness of the questions.

4.3.1 Sample Group

Proposed participants would be members of the general public who have an interest in Alternative or Indie music and are aged between eighteen and forty. The sample size would be in the region of two hundred participants.

4.3.2 Method

The participants would be instructed to use *Firstage* for a week. Given that both MARI and CI may require extended use to assess, a controlled study is not advisable; assessing users' feelings of the application after a comfortable period of use is preferred. Participants should use the application in as close to their standard application usage patterns as possible, i.e., no more or less than they would usually give time to any other application unless they wanted to; it might be assumed that a more immersive experience would lead to greater usage. Given *Firstage*'s device independence, and assuming that this should be expected to impact any given user's experience of the application, no control would be in place in terms of preferred devices beyond a device's capability to run the application.

After a week's time, the users would be sent an email containing the questionnaire. They should answer all of the questions and also inform which device they used the application with. A five point Likert scale—in which 1 expresses Strongly Disagree, and 5 expresses Strongly Agree—would be employed in order to register the user's feelings.

4.3.3 Ethical considerations

Ethics approval is required for the questionnaire to enable participants to give informed consent. It should also be made clear that participants may submit a blank questionnaire if they wish and there is no obligation on the part of the participants to complete the questionnaire. The length of time that they should be expected to spend using both the application and undertaking the questionnaire should be made clear. Conflicts of interest must be considered, especially in terms of participants who may have a vested interest in the application, e.g., participants who might work for a rival company, bear ill will towards the makers of the application, or be friends or relatives of the application developers or the music artists featured in the application.

4.4 Analysing the results

Any incomplete questionnaires must be discarded. Each section of the questionnaire should be analysed separately.

The greater the item score, the more it is positively associated with its level/dimension of immersion. Participants rate along a 5-point scale for all questions: 1 is negatively associated, 3 is average, and 5 is positively associated. Any questions which are negatively related—that are phrased in such a way as 5 would be negatively associated such as *"I found Firstage confusing to navigate"*—should be calculated reversely, i.e., that a score of 5 becomes 1 and 4 becomes 2. The programme *IBM SPSS Statistics*, an analytics software application, may be used to analyse the data.

The mean of all the answers to each item should be calculated. This indicates the overall feeling of the participants in relation to each item. The means of all questions within a factor would then be summed and calculated as a percentage of the total score for that factor. For example, in the factor of "Interest" there are six items. In this case the total score is 30 (six items times a possible total score of 5 for each item). So the sum of the means of every item in "Interest" would be calculated as a percentage of 30, i.e., (100/30) x (sum of means of items within a factor). This will indicate *Firstage*'s performance in this factor overall.

In order to determine the application's performance in a given level/dimension, an item to total percentage must likewise be calculated. For example, in the MARI section, the level of *Engagement* contains two factors: "Interest" and "Usability". In this particular case there is a total possible score of 50: ten items (six in "Interest" and four in "Usability") times a possible total score of 5 per item. The means of every item within the level would be summed and the sum would be calculated as a percentage of 50. This shows Firstage's performance in the level of *Engagement* overall.

It is important to determine where strengths and weaknesses lie within each level/dimension. For example, the application may score highly in "Interest" but below the average of 3 for "Usability". This would indicate that *Firstage* is somewhat difficult to use and will affect its overall performance in the level of *Engagement*. Individual items can then be analysed to determine precisely why this might be and how it might be overcome. A frequency analysis to show the number of occurrences of each response chosen by the participants for each item could be undertaken to give more accurate insight into performance. For example, if the item *A7: "It was easy for me to start to use Firstage"* has an average score yet frequency analysis shows evenly dispersed answers across the scale, it might be determined that personal preference has come into play with a somewhat equal amount of users both agreeing and disagreeing with the statement. However, if answers are shown to concentrate mainly on a score of 3 (average), it might be determined that *Firstage*'s performance in this regard could be improved.

MARI is a graduated model with three distinct levels. Thus we might expect to see, for example, strong performance in *Engagement*, with gradually weaker performances in *Engrossment* and *Total Immersion* respectively. The theory behind the model holds that higher levels depend on lower levels to bring them about. Thus it would be unexpected to see a strong performance in *Total Immersion*, yet weak performances in *Engagement* and *Engrossment*. *Total Immersion* is the optimal state that users may experience.

However, in the case of CI, each dimension is equally important and they are not interdependent. High context immersion emerges from strong performance in every dimension. If some dimensions perform weaker than others, context immersion as a whole would be affected. Thus a relative balance between all dimensions is preferable. If a dimension indicates weak performance, it might be advised to investigate why using frequency analysis on individual items.

4.5 Conclusions

In this chapter the reasoning and methodology behind the creation of the questionnaire has been presented. By analysing a number of previous studies related to immersion in MAR, two relevant measures for immersion have been identified: MARI and CI. Using the appropriate theoretical approach for each model, the items for the questionnaire were been generated and adapted so that they are suitable to measure *Firstage* given its unique nature. As stated, the scope of this paper falls short of carrying out any qualitative data collection and thus a potential approach researchers may take to implement the study has been described. It should be noted that there is some overlap between MARI and the dimension of *Object-based* context immersion in CI. Indeed, it might be found that greater MARI could lead to or correlate with greater CI insofar as feelings of greater realism necessary for MARI might foster object-based context immersion necessary for CI.

In the literature reviewed for this paper, no conducted study measured for both MARI and CI. Given that both are important for MAR immersion—and might be considered complementary to one another—it is hoped that the questionnaire developed here might be adapted for use in measuring other MAR applications.

Chapter 5. Discussion and conclusions

5.1 Introduction

Firstage is an MAR application which blends pre-recorded live music performances with the user's immediate environment. In the same manner that smartphones overcome the restrictions of traditional media delivery, the aim of Firstage is to overcome geographical and financial restrictions to allow emerging music artists a chance to perform on virtual 'stages' across the world, with potential stages being any surface in the user's surroundings. Though video sharing sites can provide a similar service, what sets MAR apart from other media delivery platforms is its ability to create immersive experiences for users by convincingly blending the real with the virtual thereby making digital media and information become, in a sense, 'real' for the user (Brooks 2003; Lee & Kim 2011). Indeed, with many MAR applications failing due to low usage rates, the success of MAR applications is seen to depend on the degree to which they provide immersive experiences (Kim et al. 2014). The novelty of MAR is not enough. Users must feel that an application delivers high quality, contextually relevant information, permits interactivity, and provides convincingly blended video and audio content. Thus the purpose of this paper was to design a questionnaire which could be used to determine the performance of *Firstage* in terms of immersion, as well as to combine two models of immersion not yet measured for in a single study, MARI and CI, into the questionnaire.

5.2 Reflections

It is worth noting that an MAR application is not expected to perform extremely well in the level of *Total Immersion* in MARI (Georgiou & Kyza 2017). Indeed, as MARI measures for a kind of immersion similar to that found in VEs and videogames, the performance of *Firstage* in this regard is envisioned to be quite weak, especially as the ARIQ—which informs this section of the questionnaire—was used to measure immersion in location-aware games. Games, by their nature, are generally more immersive in an MARI sense due to the level of concentration and investment they command. Nevertheless, given that *Firstage* combines video and audio with a real world environment, there is a necessity for delivering effective consistency of content together with a realistic blend of that content with the real world environment. This is important for MARI, but so too is it important for CI—specifically in an *Object-based* context—due to its effect in creating embodied spaces where virtual objects come to feel real for users (Kim 2013). Thus *Firstage*'s performance in this regard should be strongly considered.

5.3 Limitations

Due to the scope of this research paper, no qualitative data was collected. As such, there are some considerations worth bearing in mind. As no pilot study has been undertaken, a Cronbach's alpha for each of the items could not be calculated in order to determine their suitability for measuring the factors they aim to measure. Thus it would be necessary for any potential undertaking of this study to attend to this. Furthermore, while every caution has been exercised in maintaining the methodical approach to item generation in the studies from which the questionnaire in this paper is derived, it would be advisable for any potential undertaking of the study to perform EFA and CFA on the items in the questionnaire due to their having been altered for suitability in measuring *Firstage*.

Using Likert scales may result in distortion of results. Participants often agree with statements as they are presented, i.e., positively worded statements result in positive responses. Often too do participants refrain from selecting extreme responses, i.e. either strongly agree or disagree. Thus any analysis of results should consider this.

5.4 Conclusions

The purpose of this paper was to devise a potential questionnaire to measure immersion in *Firstage* and similar MAR applications. This was achieved through a number of steps. First, to reach an understanding of the precise way that *Firstage* delivers its service, the various means by which MAR technology could be deployed was illustrated through a review of the appropriate literature. It then became necessary to establish an appropriate model to measure immersion in an MAR application. This was done by reviewing relevant literature and studies which deal with immersion in mediated environments, as well as MAR in particular. Next, *Firstage* was situated in relation to other MAR applications through a review of the state of the art, and an analysis of *Firstage* in terms of the immersive characteristics identified in the literature review was done. Finally, a questionnaire which could potentially be used to measure immersion in *Firstage* was developed, paying attention to the relevant research and applying the appropriate models found in Georgiou & Kyza (2017) and Kim (2013), and a proposed implementation of the questionnaire and a method of its analysis were described in detail. Two models of immersion, MARI and CI—which have not yet been measured for in a single study—have been incorporated into the questionnaire developed in this paper. It is hoped that the theoretical work done here could be used to both measure immersion in *Firstage* as well as in other similar MAR applications, though the limitations stated above must be attended to.

Further investigation as to the appropriate model to be used in measuring immersion in MAR must be done. While both MARI and CI have been validated in their respective studies, it appears that the nature of the application being measured—i.e., whether it is a game or a service—might play a role in this. It is envisioned that CI would be the area which would be of most interest in determining *Firstage*'s immersive performance. Given that CI is a unique feature of MAR, and that it applies equally to both services and games, it is believed that the focus of any future research into MAR services might concentrate here. Indeed, much of what is measured by MARI is covered by the *Object-based* dimension of CI, particularly considering MAR applications are not expected to bring about *Total Immersion* in MARI. Due to this, the most critical aspects of immersion in terms of *Firstage* will likely be in all three dimensions of CI.

List of Abbreviations

AR	augmented reality
ARIQ	augmented reality immersion questionnaire
CFA	confirmatory factor analysis
CI	context immersion
CIQ	context immersion questionnaire
EFA	evaluative factor analysis
GPS	global positioning system
HMD	head mounted display
IR	impulse response
MAR	mobile augmented reality
MARI	mobile augmented reality immersion
QR	quick response
VE	virtual environment

References

Augment (2011). Augment [mobile application software]. Retrieved from: http://www.augment.com

INDG (2014). Amikasa [mobile application software]. Retrieved from: http://www.amikasa.com

Azuma R. T. (1997). A survey of augmented reality. Presence 6(4), pp. 355–385.

Azuma R.T., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre B (2001). Recent advances in augmented reality. *IEEE Computer Graphics & Applications* 21(6): 34-47.

Billinghurst, M. & Henrysson, A. (2009). Mobile architectural augmented reality. In X. Wang, M.A. Schnabel (Eds.) *Mixed Reality in Architecture, Design and Construction*. Springer Science+Business Media B.V., pp. 93–104.

Blippar (2011). Blippar [mobile application software]. Retrieved from: https://blippar.com

Brooks, K. (2003). *There is nothing virtual about immersion: Narrative immersion for R and other interface*. Cambridge, MA: Motorola Labs/Human Interface Labs, MIT Media Laboratory. Retrieved from: http://alumni.media.mit.edu/~brooks/storybiz/ immersiveNotVirtual.pdf

Brown, E. & Cairns, P. (2004). A grounded investigation of game immersion. *Proceedings of the 2011 CHI Conference on Human Factors in Computing Systems*, pp. 1297–1300.

Bull, M. (2000). *Sounding Out the City: Personal Stereos and the Management of Everyday Life*. Oxford: Berg.

Cheng, K. H. & Tsai, C. C. (2013). Affordances of augmented reality in science learning: suggestions for future. *Journal of Science Education and Technology* 22, pp. 449–462.

Cheng, M. T., She, H. C., Annetta, L. A. (2015). Game immersion experience: its hierarchical structure and impact on game-based science learning. *Journal of Computer Assisted Learning* 31: 232–253.

Coyle, J.R. & Thorson, E. (2001). The effects of progressive levels of interactivity and vividness in marketing sites. *Journal of Advertising* 30, pp. 65-77.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. Harper Perennial, London.

Deutsche Telekom AG (2017). The Lenz [mobile application software]. Retrieved from: http://www.electronicbeats.net/the-feed/telekom-electronic-beats-releases-lenz-augmented-reality-app/

Dunleavy, M., Dede, C., Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology* 18, pp. 7–22.

Ermi, L. & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. In: *Proceedings of The 2005 DiGRA International Conference: Changing Views - Worlds In Play*.

Firstage Ltd. (2015). Firstage [mobile application software]. Retrieved from: http://www.firstage.com

Gardner, W. G. (1999). 3D audio and acoustic environment modeling. *Technical report*. Wave Arts Inc., Arlington, MA. USA.

Georgiou Y. & , Kyza E. A. (2017). The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings. *International Journal of Human-Computer Studies* 98, pp. 24-37.

Haunted Planet (2012). Bram Stoker's Vampires [mobile application software]. Retrieved from: https://www.hauntedplanet.com

Holden, J. (2016). *Start-up aims to rock live music with augmented reality*. [online] The Irish Times. Retrieved from: http://www.irishtimes.com/business/start-up-aims-to-rock-live-music-with-augmented-reality-1.2739189 [19 Apr. 2017].

Höllerer, T. & Feiner, S. (2004). Telegeoinformatics: location-based computing and services. In: Karimi, H. and Hammad, A. (eds). *Mobile Augmented Reality*. Haarlem, The Netherlands: Taylor & Francis Books, pp. 221-256.

INKHUNTER, Inc. (2016). INKHUNTER [mobile application software]. Retrieved from: http://inkhunter.tattoo/

Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66, pp. 641–661.

Johnson, L., Levine, A., Smith, R., & Stone, S. (2010). Simple augmented reality. *The 2010 Horizon Report*: 21-24. Austin, TX: The New Media Consortium.

Kari, T. (2016). Pokémon GO 2016: Exploring Situational Contexts of Critical Incidents in Augmented Reality. Journal of Virtual Worlds Research 9(3), pp. 1-11.

Kim, K., Hwang, J., Zo, H., & Lee, H. (2014). Understanding users' continuance intention toward smartphone augmented reality applications. *Information Development*.

Kim, M. J. (2013). A framework for context immersion in mobile augmented reality. *Automation in Construction* 33, pp. 79–85.

Kirsh, D. (2001). The context of work. *Human Computer Interaction* 16 (2–4), pp. 305–322.

Kyriakis, C. (1998). Fundamental and technological limitations of immersive audio systems. *Proceedings of The IEEE* (86), pp. 941–951.

Lee J. H. & Kim M. J. (2011). A context immersion of mixed reality at a new stage. In *Proceedings of Computer-Aided Architectural Design Research in Asia 2011: Circuit Bending, Breaking and Mending*, Newcastle, Australia, pp. 199–208.

Lee, K. (2012). "Augmented Reality in Education and Training", *TechTrends*, vol. 56, no. 2, pp. 13-21.

Lieberman, H. & Selker, T. (2000). Out of context: Computer systems that adapt to, and learn from, context. *IBM Systems Journal* 39(3-4), pp. 617-632.

MacIntyre, B., Bolter, J. D., Gandy, M. (2004). Presence and the aura of meaningful places. *Presence: Teleoperators and Virtual Environments* 6(2), pp.197–206.

McCall, R., Wetzel, R., Löschner, J., Braun, A.-K. (2011). Using presence to evaluate an augmented reality location aware game. *Personal Ubiquitous Computing* 15, pp. 25–35.

McMahan, A. (2003). Immersion, engagement, and presence: A method for analyzing 3D video games. In, Wolf, M.J.P., Perron, B. (Eds.) *The Video Game Theory Reader*. New York: Routledge, pp. 67–86.

Microsoft (2016). Microsoft HoloLens. https://www.microsoft.com/en-us/hololens

Murray, J. H. (1997). *Hamlet on the Holodeck - The Future of Narrative in Cyberspace*. New York: The Free Press.

Murphy, L. (2015). *A world's first: the app that brings live music straight to your mobile*. [online] The Irish Times. Retrieved from: http://www.irishtimes.com/culture/music/a-world-s-first-the-app-that-brings-live-music-straight-to-your-mobile-1.2347159 [12 Apr. 2017].

Narzt W, Pomberger G, Ferscha A, Kolb, D., Müller, R., Wieghardt, J., Hörtner, H., Lindinger, C. (2006). Augmented reality navigation systems. *Universal Access in the Information Society* 4(3), pp. 177–187.

Nee A., Ong S., Chryssolouris G., Mourtzis, D. (2012) Augmented reality applications in design and manufacturing. *CIRP Annals-Manufacturing Technology* 61(2), pp. 657–679.

Newzoo. (2016). Analysis of Pokémon GO: a success two decades in the making. Retrieved from: https://newzoo.com/insights/articles/analysis-pokemon-go/ [12 Apr. 2017].

Niantic, Inc. (2016). Pokémon GO [mobile application software]. Retrieved from: http://www.pokemongo.com/

Nilsson, E.M., Svingby, G. (2009). Gaming as actions: students playing a mobile educational computer game. *Human IT* 10, pp. 26–59.

Olsson, T., Lagerstam, E., Kärkkäinen, T., Väänänen-Vainio-Mattila, K. (2013). Expected user experience of mobile augmented reality services: a user study in the context of shopping centres. *Personal and Ubiquitous Computing* 17(2), pp. 287–304.

Papagiannakis G., Singh, G., Magnenat-Thalmann, N. (2008). A survey of mobile and wireless technologies for augmented reality systems. *Computer Animation and Virtual Worlds* 19(1), pp. 3–22.

Paterson, N., Kearney, G., Naliuka, K., Conway, F. (2013). Viking Ghost Hunt: Creating engaging sound design for location-aware applications. *International Journal of Arts and Technology* 6(1), pp. 61 - 82.

Prochazka, D., Stencl, M., Popelka, O., Stastny, J. (2011). Mobile Augmented Reality Applications. In *MENDEL 2011, 17th International Conference on Soft Computing*. Brno University of Technology.

Quinn, D. (2016). *Firstage - The First Virtual Stage*. [online] Bloom. Retrieved from: http://www.thisisbloom.com/2016/03/29/talk-firstage-first-virtual-stage/ [19 Apr. 2017].

Reeves B (1991). Being there: Television as symbolic versus natural experience. *Unpublished manuscript*, Institute for Communication Research, Stanford University, .Stanford, CA.

Regenbrecht, H., Schubert, T. (2002). Measuring presence in augmented reality environments: Design and a first test of a questionnaire. *In Proceedings of the Fifth Annual International Workshop Presence 2002*, Porto, Portugal. October 9–11.

Reid, J., Hull, R., Clayton, B., Melamed, T., Stenton, P. (2011). A research methodology for evaluating location aware experiences. *Personal Ubiquitous Computing* 15, pp. 53–60.

Rouillard, J. (2008). "Contextual QR Code," *Proceedings of the Third International Multi-Conference on Computing in the Global Information Technology*, Athens, Greece, July 27 – August 1, pp. 50-55.

Rumsey, F. (2001). *Spatial Audio*. Oxford, U.K.: Focal.

Schmalstieg, D., Langlotz, T., Billinghurst, M. (2011). Augmented Reality 2.0. In Brunnett, G. et al. (eds) *Virtual Realities*. Vienna: Springer, pp. 13–37.

Slater, M., Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments* 6, pp. 603–616..

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence: Teleoperators & Virtual Environments* 8(5), pp. 560–566.

Steptoe, W., Julier, S., Steed, A. (2014). Presence and discernibility in conventional and non-photorealistic immersive augmented reality. In *IEEE International Symposium on Mixed and Augmented Reality* (ISMAR), pp. 213- 218.

Steuer J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication* 42(4), pp. 73–93.

Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., Schmalstieg, D. (2010). Real-Time Detection and Tracking for Augmented Reality on Mobile Phones, *IEEE Transactions on Visualization and Computer Graphics* 16(3), pp. 355-368.

Weibel, D., Wissmath, B. (2011). Immersion in computer games: the role of spatial presence and flow. *International Journal of Computer Games Technology*.

Witmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence: Teleoperators and Virtual Environments* 7, pp. 225–240.

Zhong, X., Yost, W., Sun, L. (2015). Dynamic binaural sound source localization with ITD cues: Human listeners. *The Journal of the Acoustical Society of America* 137(4), pp. 2376-2376.

Appendix

MAR Immersion

Factor	*Category*	Item
Engagement	*Interest*	A1: Firstage captured my attention A2: In general I like watching bands and music artists perform A3: I found Firstage to be a novel way to experience a music artist's performance A4: I wanted to find out more about the artists whose performances I watched A5: Usually, I wanted to watch the whole performance A6: Watching a performance encouraged me to watch more performances from other artists
	Usability	A7: It was easy for me to start to use Firstage A8: I found Firstage confusing to navigate A9: Firstage was unnecessarily complex A10: I did not have difficulties in interacting with Firstage
Engrossment	*Emotional attachment*	B1: I felt a connection with the artist who was performing B2: I was excited by the experience B3: I found watching the performance to be a rewarding experience B4: I often felt like I was part of the performance B5: I felt that the app responded intuitively to what I wanted it to do
	Focus of attention	B6: If interrupted, I looked forward to returning to watching a performance B7: I was often more involved with watching a performance than anything else B8: I often forgot about the passage of time while watching a performance B9: Everyday thoughts and concerns faded out almost entirely when watching a performance B10: I was more focused on the performance than on any external distraction B11: During a performance, hardly anything could distract me
Total Immersion	*Presence*	Think back to the performance you most connected with C1: The performers felt so real that it felt as though they were in the room with me C2: The performance felt more like something I was experiencing rather than just watching C3: I forgot that I was watching an AR performance and believed it to be really happening C4: I was so involved that I wanted to interact with the performers directly C5: I was so involved that I felt like my actions could affect the performance
	Flow	C6: I had no irrelevant thoughts or external distractions while watching the performance C7: The performance became the only thought occupying my mind C8: I lost track of time and the only thing that I could think about was the performance C9: All of my senses were totally immersed in the performance

Context Immersion

Dimension	*Category*	Item
Time & Location based	*Nomadic*	D1: I was able to use Firstage in a number of different environments D2: I was able to use firstage at any time I wished D3: The quality of the media (audio and video) was consistent
	Location & multitracking	D4: Firstage lends itself to being used in many different locations D5: I felt that Firstage gave a unique experience depending on where I used it D6: I felt that Firstage provided performances which were relevant to my location
Object based	*Interface*	E1: I was able to interact with the performance in Firstage E2: The performers stayed in place if I moved my smartphone around E3: It was easy to look around the scene and inspect things E4: I felt like looking around the environment E5: I used both portrait and landscape view to look at performances
	Embodied Space	E6: Firstage made real world locations feel more meaningful E7: Firstage encouraged me to visit different locations to use the app E8: I felt that the real and the virtual blended together well E9: The performance felt consistent with the real world environment
	Sensory	E10: The performers looked like they were in the same location as me E11: The performances sounded like they were taking place in the same space as me E12: I was able to use the touch function to interact with the scene
	Multi-dimensional Information	E13: Firstage provided me with suitable information about the performers E14: I felt like there was a good mix of media available in the app
User based	*Multitude*	F1: I felt well connected to other users in my location F2: I felt broadly connected with other users of the app in general F3: I felt part of a community when using the app
	User Generated Information	F4: I felt like I could share information with other users easily F5: I felt like I could add my content to the app F6: I was able to view information from other users