Time's Running Out!

Exploring User Interaction of Location-Based, Time-Limited, Augmented Reality Experiences In Urban Environments

A Sound of Our Cities research contribution

Master Thesis

Mark Crispin Sandholt Dourado & Ingrid Germana Assis Lima

Medialogy & Sound and Music Computing
Aalborg University, Copenhagen
Department of Architecture, Design and Media Technology

AALBORG UNIVERSITY
STUDENT REPORT

Department of Architecture, Design and Media Technology
Aalborg University

Title:
Modeling User Interaction of Location-Based, Augmented Reality Experiences In Urban Environments

Project Period:
Spring Semester 2021

Participant(s):
Mark Crispin Sandholt Dourado
Ingrid Germana Assis Lima

Supervisor(s):
Cumhur Erkut

Page Numbers: 112

Date of Completion:
May 28, 2021

Abstract:

Location-based games, such as Pokémon GO, have leveraged Augmented Reality and real-life locations to great effect, both through marker-less and marker-based solutions. In cooperation with the European collaboration "Sounds of Our Cities", an AR location-based platform was created, with the purpose of integrating projects and installations from artists and architects around Europe to be experienced in Roeselare, Belgium and Barcelona.

This study investigates the relationship between theoretical concepts within Human-Computer Interaction such as user motivation and technology acceptance in mediated environments, by utilizing a marker-less prototype with augmented time-limited virtual content that cyclically changes throughout a session as well as providing points of interest on a digital map, for said content.

A correlational research method is employed to assess the aforementioned relationships and an exploratory approach is taken through the use of partial least squares structural equation modeling (PLS-SEM). 26 participants volunteered for testing this study's prototype and tests were conducted in urban areas around Denmark.
Results showed that five of seven hypotheses were statistically significant in explaining a relationship and associated variance between the respective constructs, most prominently between technology acceptance and presence as well as enjoyment and continuance intention.

Keywords: Location-based, Marker-less, Augmented Reality, PLS-SEM/PM, Geolocation, multimodality, Fear of Missing Out, Scarcity, Technology Acceptance Model (TAM), Self-Determination Theory (SDT), Presence, Continuance Intention

Contents

1 Introduction **2**

2 Analysis **4**

2.1 Augmented Reality 4
2.1.1 Devices and Displays 5
2.1.2 Marker-less vs. Marker-based AR 6
2.1.3 Geolocation and Location-Based AR 7
2.1.4 State of the Art - Related Works and Similar Applications 9
2.1.4.1 Fields - Soundzone App 9
2.1.4.2 Magic Leap - AMBEO Augmented Audio Lab 10
2.1.4.3 Pokémon Go 11
2.1.4.4 Ingress 11
2.1.5 Tools and Technology 12
2.1.5.1 Bose Wearables and Geo 12
2.1.5.2 Microsoft Soundscape 13
2.1.5.3 ARKit 4 13
2.1.5.4 Apple AR Glasses 2022 14
2.1.5.5 Microsoft Azure Spatial Anchors 15
2.1.6 Concluding Remarks on AR 15
2.2 Theoretical models for technology acceptance and motivation 16
2.2.1 Technology Acceptance Model (TAM) 16
2.2.2 Comparison to Similar Models 17
2.2.2.1 Theory of Reasoned Action (TRA) 17
2.2.2.2 Motivational Model (MM) 17
2.2.3 Further Evolvement of TAM 17
2.2.3.1 TAM for Augmented Reality 19
2.2.4 Continuance Intention and Motivation 21
2.2.5 Appropriation and Artifact Ecology 22
2.2.6 Concluding remarks of technology acceptance 23
2.3 Self-Determination Theory (SDT) 24
2.3.1 FoMO - Fear of Missing Out 25

2.3.1.1 Scarcity . . . 26
2.4 Presence . . . 27
2.4.1 Presence in Augmented Reality . . . 29
2.4.2 Auditory-Induced Presence in AR/VR . . . 31
2.4.2.1 Externalization . . . 31
2.4.2.2 Spaciousness . . . 32
2.4.2.3 Prior Expectations . . . 33
2.4.2.4 Localization . . . 34
2.4.2.5 Reverb Revisited . . . 36
2.4.2.6 Sound Quality . . . 37
2.4.2.7 Occlusion . . . 37
2.4.3 Acoustic Transparency . . . 38
2.4.3.1 Adoption . . . 39
2.4.4 Concluding Remarks for Presence . . . 39
2.5 The implication of COVID-19 on Location-Based Games and Applications . . . 40
2.5.1 Evaluating AR and VR experiments during COVID-19 . . . 42

3 Approach **44**
3.1 Motivational factors . . . 46
3.2 Presence factors . . . 46
3.3 Technology adoption factors . . . 46

4 Requirements **48**

5 Design and Implementation **50**
5.1 Phase I. . . . 50
5.2 Phase II. . . . 53
5.2.1 Covid-19 . . . 53
5.2.2 Alternative methods . . . 53
5.2.3 Prototyping, Limitations and Abandonment . . . 54
5.2.4 Revisitation . . . 55
5.3 Phase III. . . . 56
5.3.1 Application structure . . . 56
5.3.1.1 Engine . . . 56
5.3.1.2 Software Development Kits (SDKs) . . . 56
5.3.1.3 Occlusion (visual) . . . 58
5.3.1.4 User Interface (UI) . . . 61
5.3.1.5 Cyclicity . . . 62
5.3.2 Audio implementation . . . 62
5.3.2.1 Resonance Audio . . . 62
5.3.2.2 Ambisonics . . . 62
5.3.2.3 Reverb Probes . . . 63

5.3.2.4 Acoustic Materials . . . 64
5.3.2.5 Occlusion (auditory) . . . 65
5.4 Scenes . . . 65
5.5 Final look of the application . . . 66

6 User Study **68**
6.1 Participants . . . 68
6.2 Equipment . . . 68
6.3 Setup and Location . . . 69
6.4 Procedure . . . 69
6.5 Data Collection . . . 70
6.6 Taxonomy . . . 70

7 Results **72**
7.1 Defining the presence relationship . . . 73
7.2 The Complex Model . . . 75
7.3 Model Fit . . . 76

8 Discussion and key findings **79**
8.1 SoOC . . . 79
8.2 Design phases . . . 79
8.3 Evaluation . . . 80
8.4 Errors encountered . . . 82

9 Conclusion **83**

10 Future works **85**

11 Acknowledgements **87**

Bibliography **88**

Appendix **99**

1 Introduction

Augmented Reality (AR) is currently one of the most fast-paced technologies, sparking interest from both consumers and industry partners. Research in AR has been on-going since the 1990s, and has likewise been utilized in multiple industries for several purposes. Azuma [1] created a survey back then, investigating the field of AR and its purposes. In concise terms, AR can be described as superimposing virtual or digital content onto the real world. This can be achieved using different types of AR, most notably mobile AR [2]. Today, all newer smartphones and tablets have a number of AR features built in, meaning that AR can be experienced by anyone with access to such devices [2]. In order to be able to enhance an existing physical environment, AR applications can use certain types of sensors. Craig [2] lists three types commonly used in AR applications: gyroscopes, compasses and accelerometers. These are also some of the more common sensors found in mobile devices. Gyroscopes can detect the device's yaw, pitch and roll, i.e, information regarding the device's current rotation. Compasses facilitate determination of direction, which can be north, south, east or west. Lastly, accelerometers detect the device's change in movement and acceleration.

Similar to what is achieved visually with digital content, so can audio play a part in the experience of augmenting a physical environment. Audio as well as visuals, are based on virtual reality (VR) paradigms such as stationary, head-based and hand-based [3]. Sherman and Craig [3] mention subcategories to the aforementioned paradigms;

1 Stationary aural display -> speakers.

2 Head-based aural displays -> headphones (closed or open-ear speakers).

3 Hand-based aural displays -> controllers and mobile devices.

All types provide pros and cons for different purposes.
Applications created for e.g, entertainment as well as cultural purposes have been known to make use of location-based technology to present an altered environment through augmented reality. Several studies have investigated the effect of using Global Positioning Systems (GPS) in AR applications, e.g, for learning [4], cultural heritage [5][6], tourism, navigation (Google Maps AR navigation) and entertainment such as games. Some of the more popular location-based AR games are Pokémon GO and Ingress, which use points of interest on a map to provide the user information on real-life positioning of virtual content. Meaning, not only can

AR enhance a given real-life environment, but can also be targeted to enhance certain areas using GPS information. This study is based on a collaboration between Aalborg University (AAU) and Sound of Our Cities. Sound of Our Cities (SoOC) is a European collaboration, which seeks to combine art, technology and social space within the context of two cities - Roeselare, Belgium and Barcelona, Spain.

Figure 1.1: SoOC logo

The project includes 10 artists whose projects have been selected for further realization within the cities' social contexts, aiming to offer a fictional landscape upon which to project various realities and ideally foster debate about current issues. Furthermore, social maps created by the Dutch collective, Dear Hunter, may be used by artists as a reference for the different cultural points of interest.
Due to COVID-19, some of the SoOC work has been delayed and therefore this study aims to investigate and model the relationship between user behavior and motivation, technology acceptance, and augmented mediated environments. More specifically, the study will explore how users interact with audio-visually augmented location-based experiences in real environments, and how to interpret this. As part of the collaboration between AAU and SoOC, Aalborg University was initially tasked with designing a digital application and framework that could be used as a platform for the artists to form connections between locations and personal experiences through the modalities of sound and vision.
To achieve this, a marker-less Augmented Reality experience prototype will be designed along with an evaluation methodology allowing for exploration of the relationship between different aspects of the concept, such as: technology acceptance, presence, enjoyment, fear of missing out, scarcity, and continuance intention. Some of the objectives for this study are:

1. Develop an augmented reality application using GPS information to place digital audio-visual content.

2. Carry out statistical analysis of a complex model with theoretical constructs for user-behaviour, motivation and experience and identify their relationship with one another.

Based on what was tasked by the SoOC project together with our field of study and objectives, an initial research question can be formulated.

IRQ: *How can an interactive augmented reality application act as a tool for sound designers as well as provide immersive 3D content for users to experience in urban environments?*

2 Analysis

The following sections will cover multiple venues of research related to the posed initial research question, with the goal of acquiring knowledge on current technology and relevant theory which can be used to create an immersive audiovisual application using AR in urban environments, that can be evaluated in a significant way. The following sections of the analysis will cover different types of AR solutions, state of the art and theoretical models relating to both user-experience and human-computer interaction.

2.1 Augmented Reality

Head-mounted displays (HMDs) with head tracking, paired with controllers have been utilized for Virtual Reality (VR) in order to achieve a larger amount of freedom - specifically 6 degrees - and has seen an increase in affordable, consumer-grade HMDs being produced and sold in the past couple of years [7]. Formally, VR completely replaces the user's view of the real world - this is not considering the pass-through function of some newer VR headsets, allowing the user to interrupt an application to orientate themselves in the real domain before resuming, as that serves a tertiary purpose not relating to the respective experience at hand. However, where VR supplants the user's perspective, AR supplements it [8].
Augmented Reality (AR) refers to a spatio-temporal experience where physical objects are augmented with digital information, by superimposing virtual 3D objects onto a real environment, creating the illusion that those objects inhabit their respective space [8][9]. Azuma proposes that the potential of AR could be larger, due to its inclusion and mixing of both real- and virtual environments, and the inherent understanding and interactions afforded to the user by said inclusion. He also suggests that AR is the most likely route by which wearable systems will be adopted in place of smartphones, as it harbors the ability to provide a large visual- and/or auditory display in a compact form-factor [8]. In this context, form-factor refers to displays that would ideally already be acceptable in a social context, such as an HMD resembling sunglasses or eyewear [8]. However, socially accepted AR-capable technology already exists in the form of phones (see section 2.2.3.1.), albeit with some technical limitations.

The following subsections will elaborate on the different AR devices and their limitations and capabilities (see section 2.1.1.) as well as the main differences between marker-less and marker-based augmentation (see section 2.1.2.). Finally, Geolocation and location-based AR will be reviewed both from a technical- and design perspective.

2.1.1 Devices and Displays

AR devices can be categorized as either "optical see-through" (Magic Leap, Hololens) or "video see-through" (phones, tablets, VR HMDs) depending on their capabilities. Optical see-through displays use optical elements that are half-transmissive and half-reflective, which allows enough light in to perceive the real world while being subjected to the computer generated images projected from a separate display component, thus combining both the real- and virtual elements [10].
Video see-through displays utilize a camera, which captures a digital representation of the real world that is handled in real-time by a graphics processor, combining the camera feed with the virtual content. Because the final image is processed before being presented to the user, parameters such as contrast and brightness can be controlled for increased fidelity [10].

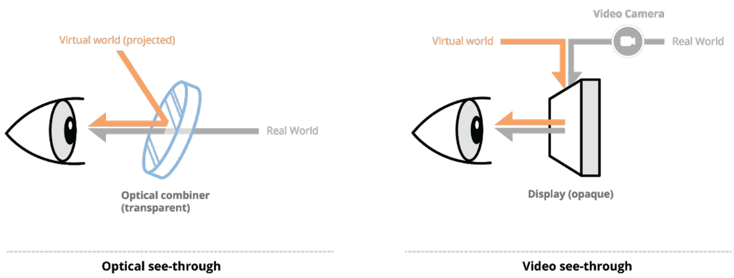


Figure 2.1: Two types of AR displays.

Mobile AR (MAR) devices span both video- (PhoneAR, ClosedAR and EasyAR) and optical (OpenAR) see-through devices [11]. In terms of wearability, these categories include hand held (smartphones), closely worn (head-worn, with phone as display), loosely worn (flip-on lenses) and worn (optical see-through HMD) [11]. However, each display type and subsequent devices utilising these have different capabilities and affordances, with an important difference being the resolution and Field of View (FOV). With current displays, either a small FOV with adequate spatial resolution can be provided or a wide FOV with low resolution [8]. For hand-held devices this resolution is measured in pixels per inch (PPI), but for HMDs, spatial resolution is measured in pixels per degree (PPD), the latter being heavily dependent on the resolution of the fovea. For a wide FOV in an optical see-through with an appropriate PPD, microdisplays with thousands of PPI higher than what is available in today's handheld devices would be required [8]. Research by Lin et al. [12] showed a correlation between presence (see 2.4. On presence), enjoyment and motion sickness as a function of FOV, in a virtual environment (VE); this was true from FOV ranges ranging from 60° to a plateau of 140°, with exceeding values showing negligible results. However, for a closed system (i.e, VR and ClosedAR), the overall FOV is limited to less than the average FOV of a person [12], while both optical see-through (OpenAR) HMDs and handheld devices support similar FOVs of augmentation (20°-30°) with no added limit to the overall FOV, i.e, the same as the person's own physical range [11][12]. Research by Choi et al. [11] suggested that phoneAR, when compared to ClosedAR, EasyAR and OpenAR) scored the highest in the usability and satisfaction categories: ease of use, comfort, suitability, fatigue (higher is better), social acceptance, general satisfaction and preference; However, the tests were carried out with a sample size of N=12, so the data should be regarded as suggestive.

2.1.2 Marker-less vs. Marker-based AR

Within the field of AR, there are two main types of augmenting real life: marker-based AR and marker-less AR.

Figure 2.2: (left) Example of marker-based AR. Business card used for recognition. (right) Example of marker-less AR (Pokemon GO). GPS and surface detection used for proper projection of augmented objects.

Marker-based AR works, as the name suggests, based on visual, physical markers for recognition. Business cards or images are examples of markers used to determine relative position of the virtual object in the physical space. Vuforia is one of the leading SDKs on the market and offers multiple features such as image targets (see figure 2.2.) and as of recent, area-, model- and multi-targets [13]. Vuforia works for multiple platforms and can be easily used in Game Engines such as Unity. The Marker-less AR does not require a physical prop to project virtual objects into real environments, but adjusts relative position based on e.g, surface detection, GPS services or gyroscope [14]. The Austrian-based company Wikitude claims that their popular SDK focuses on marker-less AR. However, this SDK is subscription-based and can be pricey depending on the intention of use (company or independent developer). Despite this, they offer a variety of features such as Geo AR, object tracking, scene tracking and more [15].

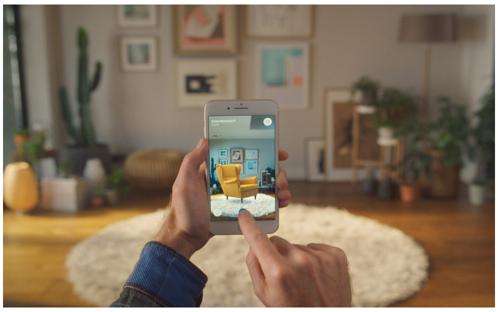


Figure 2.3: Ikea Place app. Surface detection used to place Ikea furniture in a physical space.

Examples of both types of AR being utilized could be Pokémon GO [16], which uses localization technology for projection of virtual objects. Similarly, Ikea Place is an AR app not using marker-based technology for placement of virtual objects. By scanning the floor for surface detection, the user is able to place virtual objects of different types of furniture to their living space.

	Marker-Based	Marker-less
Benefits	- More accurate positioning - Desktop supported - Mobile supported	- Does not rely on lighting - More stable - Mobile supported - Relies on localization technology and gyroscope
Disadvantages	- Is affected by lighting - Less stable - Relies on SDKs and markers for projection	- Less accurate positioning - Not usually supported on desktop

Figure 2.4: Table 1. Overview of benefits and disadvantages of using either marker-based AR and marker-less AR. [14]

2.1.3 Geolocation and Location-Based AR

Geolocation is the identification of a user's geographical location through a capable device, by using a variety of data collection mechanisms. Geolocation services tend to use network routing addresses or the internal GPS on a device to acquire a location [17]. There are different technologies for locating a device, such as: GPS, WiFi and BLE (Bluetooth Low Energy). GPS is the most precise, with an accuracy of up to 5 meters, but requires outdoor position for this to be true. However, WiFi and BLE can compensate for indoor (and outdoor) tracking, depending on the coverage, i.e, how many WiFi connections a device is in the vicinity of (with or without formal access), or how many BLE beacons are in the area [18].

As previously mentioned, geolocation is especially relevant for marker-less AR applications that seek to augment predetermined locations with digital content, as with the massively successful location-based PokeMon Go application, which focuses on entertainment and gamification. However, the concept has also been adopted in other fields such as navigation and tourism, with an example of the latter being the "KnossosAR" app, which seeks to promote

cultural heritage via the archaeological site of Knossos in Crete, Greece. Knossos utilizes geolocation for positioning of the site and raycasting to convey depth information by partially- or fully occluding points of interests behind existing modern, physical obstacles and geometry [19].
AR pioneer and Intel Labs researcher Ronald T. Azuma hypothesizes that new forms of media can be enabled through the use of AR, by making meaningful connections between the virtual content and surrounding real environment and by creating unique experiences where the power comes from the connection itself, rather than solely from either the virtual- or the real world [8][20]. Additionally, he proposes three strategies that can be categorized as *Reinforcing, Reskinning and Remembering.*

Reinforcing relates to the leveraging of a real location that is generally considered meaningful and by then adding appropriate content to said location [20]. An example of this is the 110 stories application where a silhouette can be seen of the twin towers that comprise the World Trade Center, but only if the user is within relative distance and is able to see the Manhattan skyline through their device. The silhouette itself changes depending on the time of day - while a pencil-like outline is presented during the day, a shimmering light outlining the buildings is visible at night [21].

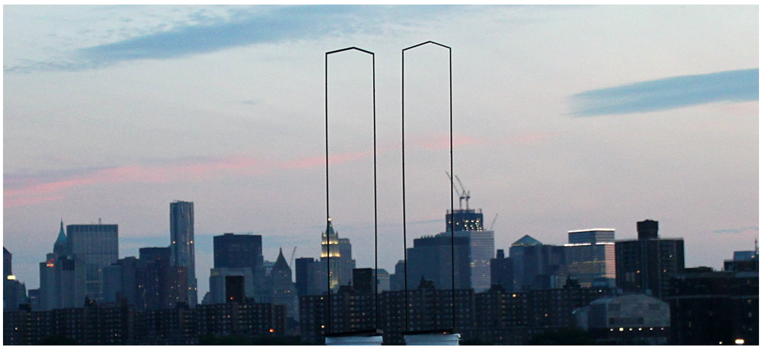


Figure 2.5: Example of reinforcing as described above

Reskinning puts less emphasis on the importance of the location itself and seeks to change the real world through embellishment, by making it fit the story being told through augmentation. An example of this is the aforementioned Pokémon Go, which effectively reskinned malls, monuments and other sights as POIs (i.e, training centres or Pokéstops). Another example of this could be experienced at the 2014 CES convention where Azuma and his team at Intel, in collaboration with the USC School of Cinematic Arts, created an AR experience of a flying whale from the fantasy series "Leviathan", that would fly out of the screen behind the speaker at the keynote speech, and over the heads of the 2500 people in the audience [8][20][22][23].

Figure 2.6: Leviathan experience

Remembering is similar to reskinning in the sense that it too utilizes locations that are generally considered mundane, but instead of redefining a location, the strategy instead seeks to leverage the personal memories that will have occurred in that space. This would take on some of the same affordances as mementos of past events, such as pictures from a wedding, graduation or other significant and personal events in an individual's life, but with the added benefit of being able to re-experience a version of said event at the same location where it happened, many years after the fact - an anachronistic rendering if you will [8][20].

2.1.4 State of the Art - Related Works and Similar Applications

2.1.4.1 Fields - Soundzone App

Fields is a light-weight app for iOS in which the user is able to "map" a local area with either recorded or imported sounds, or even experience predefined soundscapes by other artists. An audiosource is created at a given place and a short snippet of audio from the microphone is recorded and looped. The relative location of the soundsource is accompanied by a particle effect as a visual representation of the sound. While testing the application, the sound was spatialized and attenuated based on distance to the device, but both the audio and the visual effects had no interaction with each other in any tangible way; The sound would be almost as audible in an adjacent room and would only be attenuated by distance, while the visuals would clip through each other and barely diminish in size when moving away from the source. Testing the application outside, in an open area during the day, yielded better results but also introduced unwanted noise from the surrounding environment when creating recorded loops [24].

Figure 2.7: Screenshot from "Fields"

2.1.4.2 Magic Leap - AMBEO Augmented Audio Lab

Sennheiser's AMBEO Augment Audio Lab application, made specifically for the Magic Leap AR Head Mounted Display, is similar to "Fields" in both design and interaction. It features the same components of spatialized, looped audiosources with particle effects with the added benefit of being able to control a few effects (e.g, the centre frequency of a band-pass filter) through the affordances of the Magic Leap controllers [25]. Another benefit from using the Magic Leap is the motion tracking on the HMD itself and the 6 degrees of freedom (DOF) from the controllers - this allows for a more faithful implementation of binaural sound using e.g, Head-Related Transfer Functions (HRTFs) or Vector-Based Amplitude Panning (VBAP) as the motion tracking is directly related to the users head movement.

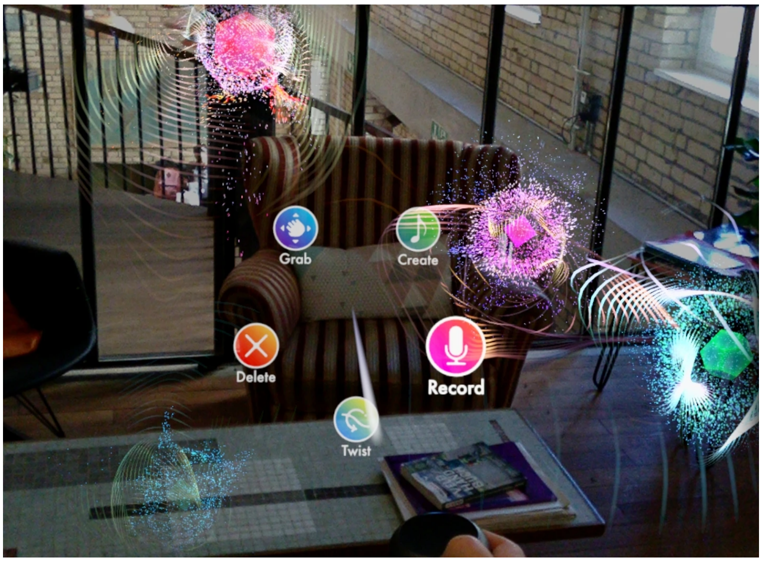


Figure 2.8: Image of some of the interactions available in the application, as well as some placed audiosources.

2.1.4.3 Pokémon Go

Pokémon GO is an AR mobile game developed by former Google startup Niantic Labs [16], which later became an independent company in 2015. The application makes use of geolocation as part of its core gameplay, and is credited as popularizing location-based games. The gameplay itself revolves around the player traversing an isometric representation of their geographical location, with Pokéstops that can be visited and wild Pokémon that can be captured; The types of Pokémon available in the area depends on the geography and climate (e.g, water Pokémon are found near a body of real water). Once a Pokémon is located, with AR enabled and using a video see-through approach, the environment is augmented with the animated Pokémon. From this screen, the user can take screenshots of the creature in the wild, or choose to capture it by flicking a Pokéball at it.

While the concept has garnered much positive feedback from increasing local business growth due to more foot traffic, some governments have expressed concerns regarding the safety of its usage, and some countries even regulate the allowed usage. In the month of the game's release (July, 2016), Facebook and Instagram reported 1.1 billion Pokémon GO related interactions from 231 million unique users [26].

Figure 2.9: Left) Virtual map using physical geographical location, featuring the player's avatar, Pokéstops and wild Pokémon. (Right) Player encountering a wild Pokémon with AR enabled.

2.1.4.4 Ingress

Ingress or Ingress Prime is a free-to-play location based AR game from 2013 (Android) and 2014 (iOS) by Niantic [27]. The game is based around the player's actual location seen on a map through an orthographic and isometric view and these 'portals' that are placed around the map (see figure 2.10.). The game is played solely through the use of the map which takes the player's position through GPS, similarly to the type of map in Pokémon GO (see section 2.1.4.3). However, Ingress Prime does not use the see-through camera approach as Pokémon

GO. The player's phone is referred to as a scanner for revealing nearby portals on the map. The player takes the role of an agent who needs to uncover some secrets surrounding an unknown force known as Exotic Matter (XM) that is posing a threat to mankind. The player must then choose a faction prior to beginning to play and capture these aforementioned portals. The player can choose to join the Enlightened faction or the Resistance. By deploying something called 'resonators', each faction can capture these portals placed on points of interest seen on the map (see figure 2.10.). When other members of your faction (other players) deploy a resonator on a portal, you can see who placed it by clicking on the portals seen on the map. The resonators help them grow stronger, and by deploying resonators from the same faction onto one portal, makes it more difficult for the opposing faction to capture it.

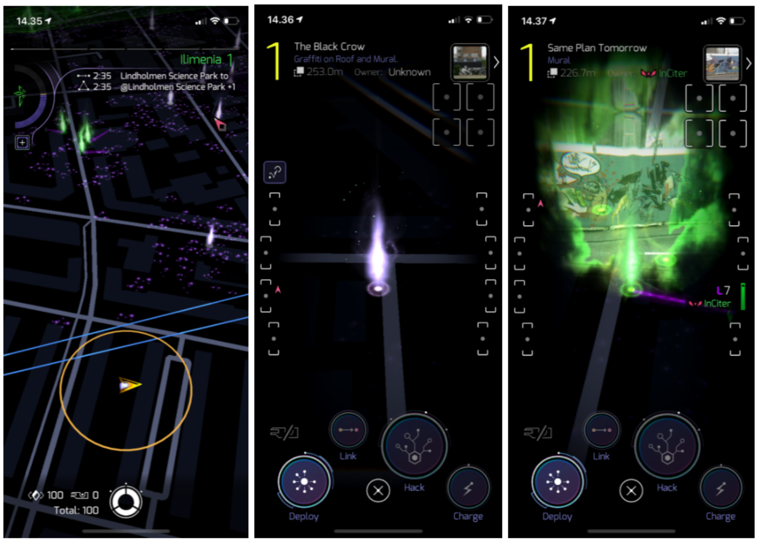


Figure 2.10: (left) Map revealing nearby portals. (middle) Action possibilities with a portal. Above is the geographical location and the current distance from the player. (right) A portal that has had one resonator deployed by a faction (hence the green color) as well as the name of the player that has deployed it.

2.1.5 Tools and Technology

2.1.5.1 Bose Wearables and Geo

In 2018, Bose announced a pair of sunglasses dubbed "Bose Frames", with speakers and motion sensors built into the temples of said frames, allowing for head tracking and AR sound. Paired with a mobile device using geolocation ("HERE" specifically), audio information could be relayed seemingly in physical space, by using virtual sound sources. One example featured in their promotional video is that of the "Golfshot" app that notifies the user of the direction of the nearest green, hazards and targets, through the use of directional sound.

Figure 2.11: (left) User attempting to locate the correct direction of the green through an obstruction, wearing the Bose Frames. (right) The Bose Frames. The speakers located in the arms of the frames visibly affect the form-factor.

2.1.5.2 Microsoft Soundscape

Microsoft Soundscape is a lifestyle navigation application that explores the field of spatial sound and audio-based technology to assist users who are visually impaired, by increasing ambient awareness of the surrounding environment. Paired with a pair of bone-conductive- or otherwise acoustically transparent headphones, Soundscape uses spatial audio cues and beacons to notify and navigate users safely to POIs in urban areas, by also recognizing pedestrian-crossings and other safety related traffic installations. Head-tracking integration has also been prototyped with the use of the Bose Frames, allowing for a hands-free approach, which in turn enables visually impaired users a higher degree of safety with the ability to have one hand reserved for a guide-dog, a blind walking cane or uninhibited usage.

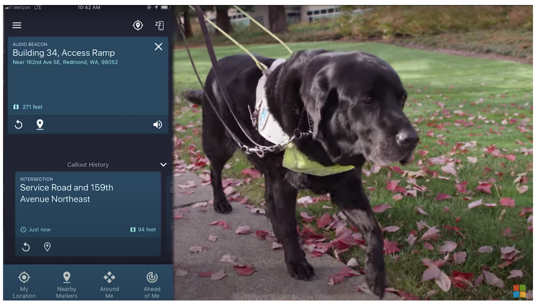


Figure 2.12: User with a guide dog following an audio beacon. The intersection is recognized by the application and the user is notified of its placement.

2.1.5.3 ARKit 4

With the release of ARKit 4 (beta version, 22nd of June 2020), a plethora of relevant features were introduced. One of the more relevant introductions is the inclusion of Apple's own location anchors, where objects can be placed at a specific latitude, longitude and altitude, and then be experienced on premise, in 3D. However, since Apple has to map the area, only

5 major cities in the US (New York City, Los Angeles, San Francisco, Miami and Chicago) support this functionality [28].

Another relevant inclusion is the new depth API which paired with the LiDAR (Light Detection and Ranging) scanner allows for much more precise people occlusion and faster calibration of geometry, when compared to what is currently available through AR surface detection. Additionally, this grants easy access to object occlusion at a much higher computational speed and fidelity, at a distance of up to 5 meters. The LiDAR scanner is essentially a time-of-flight sensor that emits low-power light at an environment, and by using the reflections of the different points in the real space, the distance is then calculated to create an accurate 3D depth map [29]. In this context, time-of-flight refers to the time from when a particle is emitted, travels through a medium and is acknowledged (received) [29][30]. As of May 2021, the LiDAR scanner is only available with the iPad Pro or iPhone 12 Pro, 2020.

Finally, yet another functionality that could prove beneficial for location-based AR applications is the contribution of "shared experiences", permitting users to experience the same AR scenario, with their individual user agencies intact and without being forced to follow each other around using a single device.

Figure 2.13: Reach and speed test of LiDAR scanner on iPad Pro 2020 [31].

2.1.5.4 Apple AR Glasses 2022

Supplementing the progress from ARKit 3 to ARKit 4, rumors and leaks from different sources as well as a statement from Apple CEO Tim Cook saying that the company sees AR as a "core" technology, seem to suggest that some degree of focus has been placed upon the development of ubiquitous AR technology, including wearables. While of a conjectural nature, the 53 patents recently won by Apple arguably strengthens the credibility of some of the information leaked. Of these patents, several of them relate to AR in the form of tangibility, maps, foveated HMDs, driving and wearables [32], with the latter including head-tracking for their true-wireless headphones (airpods) and optical see-through AR glasses equipped with LiDAR cameras and adjustable opacity [33][34][35]. According to some online tech sources [33][35], Apple seeks to utilize the user's own phone as the processor, rather than a designated unit, similar to the Magic Leap.

2.1.5.5 Microsoft Azure Spatial Anchors

As mentioned in section 2.1.5.4, ARKit 4 introduced a feature that allows for location anchors that can be used for placement of virtual objects in the real world. Another relevant software for this purpose is Azure Spatial Anchors. Azure Spatial Anchors are location anchors that can be detected by the device camera and uploaded through a cloud service. When being positioned at an anchor, one or more people can experience location-based 3D content in a precise manner. This software only emerged recently, during the finishing stages of the prototype, which is the reason for not having included this in this study's prototype implementation.

2.1.6 Concluding Remarks on AR

The previous sections offer some of the affordances available through the AR medium and are primarily presented to provide context and delimitation for the thesis. The main goal of using AR in the context of the thesis and SoOC, is to create an environment and a framework that can be used by artists across Europe to convey culturally significant content suited for the respective area, by means of reinforcing, reskinning or remembering.
Due to the inherent requirements of the SoOC project and the location-based nature of the culturally significant locations in Roeselare and Barcelona, we reviewed technological possibilities, limitations and accessibility of AR-capable devices that allow for remote manipulation of the 'virtual installations' by developers or content creators without the need for physical installations (markers). This makes it more sustainable to use marker-less technology for our thesis prototype and likewise, consider some of the audio-visual elements in other location-based AR applications. The next sections and subsections will give a perspective into users' attitude towards technology, such as AR, and how it is adopted and appropriated into the users' daily lives.

2.2 Theoretical models for technology acceptance and motivation

Within the field of user behavior studies, certain theoretical models regarding user attitude and motivation have been developed throughout the past decades. Relevant for this thesis are the ones relating to users' attitude towards using technology.

2.2.1 Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM), first used in 1986 by Davis [36][37], is a model used for predicting individual acceptance and usage of technology [38][39]. The model is based on the more generalized acceptance model Theory of Reasoned Action also known as the abbreviation, TRA (further elaborated in section 2.2.2.1).

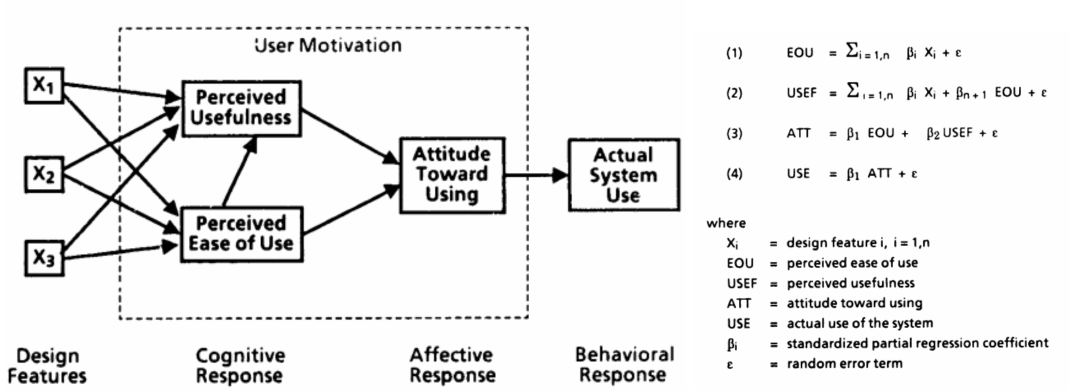


Figure 2.14: Technology Acceptance Model [36]

In his thesis, Davis [36] describes the TAM model and how the potential user's overall attitude toward using a given technology is determined by beliefs of the Perceived Usefulness (USEF also PU) and Perceived Ease of Use (EOU also PEOU). This study will adhere to the more recent abbreviations of these terms: PU and PEOU [40]. Davis [36] defined the beliefs PU and PEOU as being causally connected, as PEOU "has a causal effect on perceived usefulness" ([36] pp. 26). PU is related to the chance of which the user believes in using the system or application can enhance their job performance, where on the other hand, PEOU pertains to the degree of which the user would be "free of physical and mental effort" ([36] pp. 26) when using the technology or system at hand [36][37]. These beliefs are thereby correlated as if the system is easy for the user to use and requires minimal effort to use, the more useful the user will find the system. Allomary and Woollard [41] suggest that studies show how manipulation of these two beliefs, and their deterministic power on the system usage, can help developers predict user behaviour, intention and actual usage of the system.
Figure 2.15. shows the relationship between PU, PEOU and attitude towards usage (ATT also ATU). In short, the user's PU as well as PEOU result in the user's attitude towards using a given system. Attitude towards usage can be defined as the assessment of the user's positive

or negative reception of performing certain tasks and behaviour [41]. The success of a system can then be measured by these three factors, as a system that is not easy to use might not prove useful and then create a more negative attitude towards usage of the system [40].

2.2.2 Comparison to Similar Models

2.2.2.1 Theory of Reasoned Action (TRA)

As mentioned previously in section 2.2.1, the TAM model and framework was based on the Theory of Reasoned Action (TRA) model. The TRA framework was proposed by Fishbein Ajzen [42][43]. The person's attitude, subjective norms and perception of certain situations and contexts is linked to the person's behaviour intention which then provides a prediction of the person's actual behaviour based on the intention [41]. Again, the attitude is attributed to the concept of the negative and positive feelings toward a certain situation. Subjective norms are more related to a social aspect of behaviour, as it pertains to the perception of the person's peers thinking they should behave or not behave a certain way [38]. Later extensions to the TRA framework have also been used, such as the Theory of Planned Behavior (TPB) developed by Azjen [44]. This extension added the individual's perception of the ease of interaction or performing a certain behaviour [41]. Unlike the TRA, the TPB emphasized on mandatory situations whereas TRA was developed around voluntary situations. Studies have since determined that TPB is a better framework for determining intentional behaviour than the prior TRA framework [41].

2.2.2.2 Motivational Model (MM)

Developed by Davis, Bagozzi and Warshaw [45] for studying the relationship of PU and enjoyment of usage of technology and thereby the usage intentions. The study relies on the investigation behind extrinsic and intrinsic motivation for usage of computers in a work capacity. The extrinsic motivation is defined by Davis, Bagozzi and Warshaw [45] as being the user's will of performing a certain activity for the outcome and its gain to their e.g, job performance, promotion etc. The intrinsic motivation is on the other hand based on the user's will to perform a certain activity for no obvious reason or reinforcement other than just performing the activity itself [38][45].

2.2.3 Further Evolvement of TAM

Since the development of the TAM framework [36][37], further adjustments and evolution has happened to it. Based on the original TAM, the Technology Acceptance Model 2 (TAM2) was developed by Davis and Venkatesh [46] as an extension. The TAM2 framework adds new points to its theoretical construct, such as social influence processes and cognitive instrumental processes.

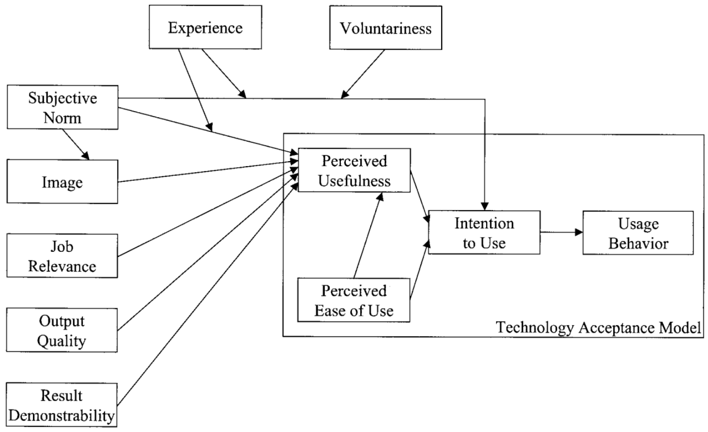
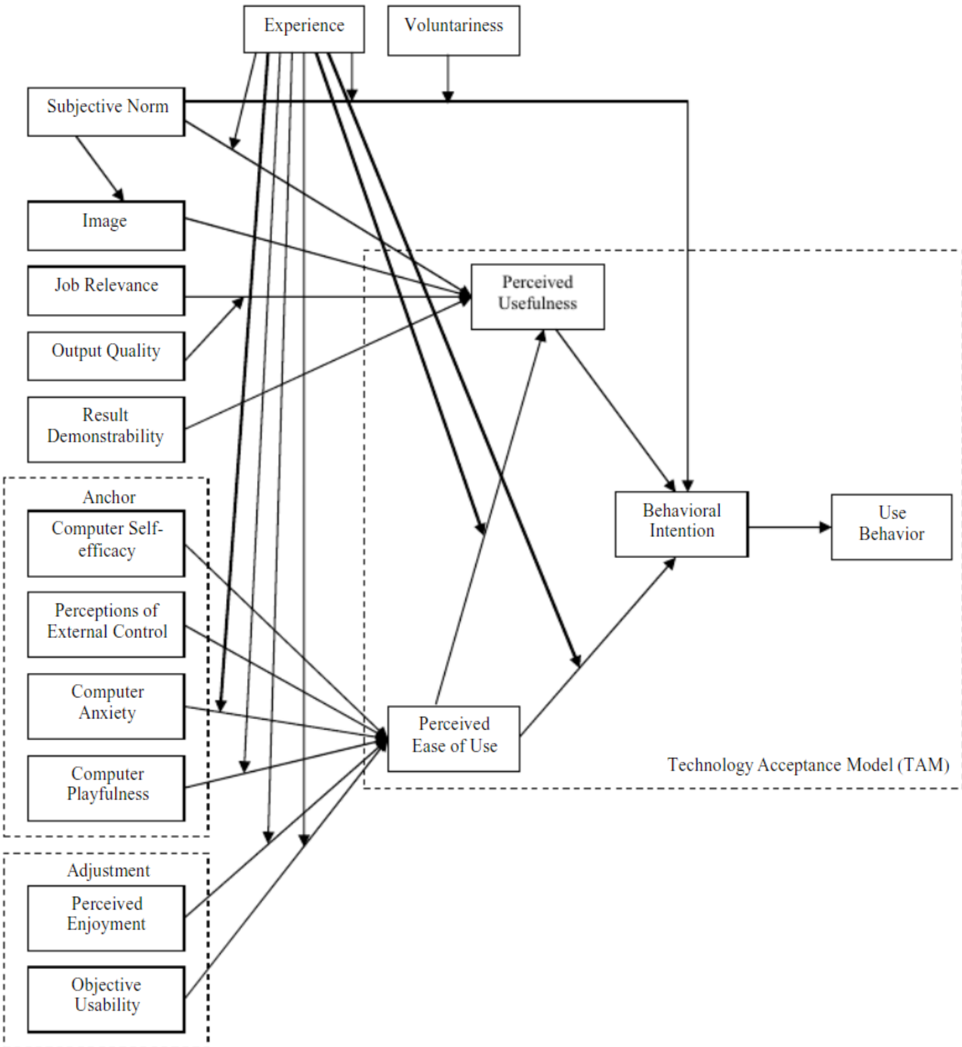


Figure 2.15: TAM2 model [46]

The social influence processes consist of three major concepts; subjective norms, voluntariness and image - all relating to the opportunity to accept or reject new technology [46]. Voluntariness pertaining to the aspect of the voluntary action opposed to the mandatory as well as the user's compliance when socially influenced. Subjective norms, like in the TRA and TPB models, refer to the user's behavioural intention based on what people important to them think they should do or not. Likewise perform actions which they would not have performed themselves, but might do if important peers think they should. Similarly, the user's image is based on the behaviour being favorable to the user's status within a social group. Davis and Venkatesh [46] suggest that the TAM2 model views the subjective norms to positively influence the user's image.
The other aspect of usage intention in TAM2 deals with the cognitive instrumental processes and are described by Davis and Venkatesh [46] as these four determinants: job relevance, output quality, result demonstrability, and perceived ease of use. Job relevance is the subjective judgement of the system's relevance and application to their work. Output quality refers to the system's ability to match the user's goals in a work capacity and how well the system works to reach these goals. Perceived Ease of Use in TAM2 adheres to the original TAM model conceptualization. Lastly, the Result Demonstrability is defined as "tangibility of the results of using the innovation" ([46] pp. 192) meaning that if the system can produce positive results from usage, the usefulness of the system can be enhanced.

In conclusion, the TAM2 differs from the original TAM model with more determinants for usage intentions as described above. Venkatesh and Bala [39] further evolved the TAM framework and proposed the TAM3 model. This framework is developed based on the TAM2 model and Venkatesh's [47] model for perceived ease of use, which added determinants such as Anchors and Adjustments. The anchors relate to the dimensions such as self-efficacy, computer playfulness and computer anxiety [39][47]. The adjustments are the result of perceived enjoyment and subjective usability. The TAM3 framework takes elements from both the TAM2 study as well as the PEOU study by Venkatesh [47]. Figure 2.17. shows the proposed model from Venkatesh and Bala's study from 2008. Determinants from both the TAM2 model as well as PEOU model have been united for this model.



[a]Thick lines indicate new relationships proposed in TAM3.

Figure 2.16: TAM3 model [39]

2.2.3.1 TAM for Augmented Reality

As this particular study deals with a relatively new technology (Augmented Reality), some of the same issues occur in terms of acceptance and adoption. Leue, tom-Dieck and Jung [48] proposed a theoretical framework for AR acceptance based on the original TAM framework by Davis [36]. The model differs from the original TAM model with its determinants. Leue, tom-Dieck and Jung's [48] study delves deeper into the application of such a model on AR usage in a tourism context. These external determinants are based on several studies, in which several concepts are established. This way the model becomes more context aware and tailored to the specifics of the study.

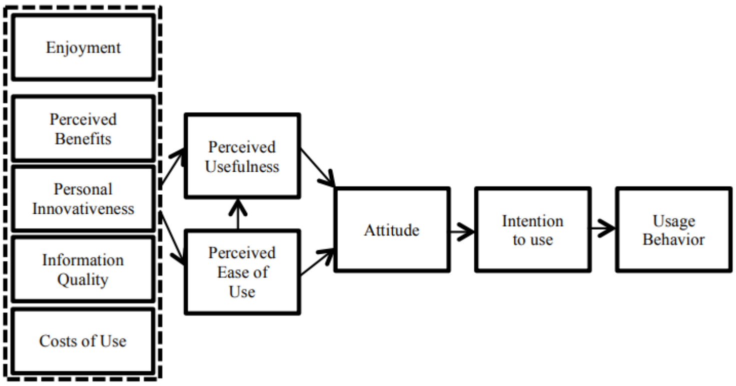


Figure 2.17: TAM for Augmented Reality Acceptance [48]

Enjoyment is a major factor for intention to use a technology [39][47][48]. Already in the third extension of the TAM model: TAM3, it was included as a part of the Adjustment relationship with PEOU (see figure 2.17). A study by Haugstvedt and Krogstie [5] investigated determinants in acceptance of such systems as AR. Their study was conducted in the context of cultural heritage and reported that perceived enjoyment has a significant impact on the intention of usage of an augmented reality application. Their research was also based on the TAM framework, although in an extended capacity (see figure 2.19).

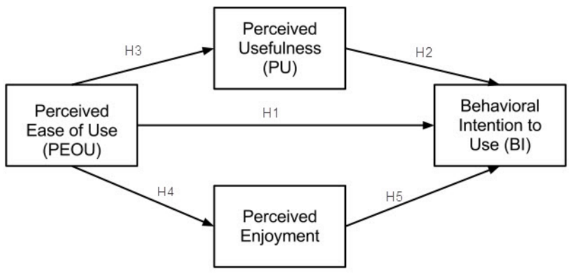


Figure 2.18: Haugstvedt and Krogstie's research model for AR acceptance [5]

Perceived benefits also play a role in the user's intention of usage of an augmented reality application. Leue et al. [48] included this part into their model (see figure 2.18) as previous studies such as Parra-Lopéz et al. [49], who investigated intention of use of social media and found perceived benefits directly influenced intention of usage of social media. Similar studies have likewise demonstrated that perceived benefits within the usage of AR technology can further enhance acceptance of the technology.
Personal innovativeness was likewise described as being an essential part of the AR Acceptance model proposed by Leue et al. [48]. Their reasoning behind it stems from other studies, likewise delving into AR usage, which when looking at the voluntary situations, allowed the users to be technological "pioneers" as well as their willingness to try new technology (inno-

vativeness) had a positive effect on intention of usage.
Information quality has also been studied, and in the context of AR usage, Leue et al. [48] added it to their framework as their research has suggested that e.g, early adopters of AR desire proper contextual information that is rich and high in quality.
Costs of Use refers to the idea of investigating the users' monetary and temporal cost when using the system. Parra-Lopéz [49] found that the users' perceived cost of use of social media did not negatively influence the predisposition of using the platform, seemingly suggesting that evaluating the relationship between cost and benefits for intentional use of a technology is important. The AR acceptance model provides guidelines in terms of creating an AR application, which incentivize users to use the application if successfully implemented.

2.2.4 Continuance Intention and Motivation

When it comes to mobile applications, especially games and free applications, player retention and "churn" are key metrics that are used to evaluate the perceived longevity of a virtual product; Churn refers to a user or player leaving the application behind, either for an extended period of time or for good, while churn-rate is the ratio of users leaving vs. staying as a function of time. The Uses and Gratification theory (UGT) is an audience-centered approach to understanding mass communication, focusing on how people interact with media and with what intent. The UGT has been applied to all manner of research on media usage, such as for mobile phones, internet and social media.
A study by Wu et al. [50] investigated several aspects of continuance motivation and "proactive stickiness" of players in online games, with motivation referring to continued play and the latter being their willingness to return and prolong their duration at each stay. The study showed that multiple gratifications, i.e, enjoyment, achievement and social interaction, had a significant effect on the player's continuance motivation, which in turn had a significant effect on proactive stickiness [50]; These findings are corroborated in a study by Li et. al [51], where results showed three types of gratification affected continuance intention to use a social-network game, namely: hedonic gratification (enjoyment and fantasy), utilitarian gratification (achievement) and social gratification (social interaction and social presence).

A study by Ghazali et al. [52] explores player intent through UGT, within the context of mobile AR gaming and the Pokémon GO application, with the purpose of identifying the motivational- and gratification factors that result in users choosing a specific medium, as well as what parameters stimulate continuance intention and loyalty (see proactive stickiness above for definition). The study showed that enjoyment had the strongest impact on continuance intention, with one of the key factors shown to positively affect said enjoyment being that of social interaction; Ghazali et al. [52] propose both network externalities and community involvement as important social factors for cultivating meaningful social engagement among players, and additionally suggest the frequent introduction of new elements be implemented, as that will elicit curiosity from the users, further fostering continuance intention and player involvement within the community.

2.2.5 Appropriation and Artifact Ecology

Similar to the aforementioned models for technology acceptance, another aspect of usage of technology and its reception is proposed as 'Process of appropriation' by Carroll et al. [53]. Their model is structured around the study of young people's appropriation of Information and Communication Technologies (ICTs) and further investigated in the Carroll et al. [54] study on young people's appropriation of Wireless Application Protocol (WAP) phones. The model is based on two different artifact types; Technology-as-designed and Technology-in-use. These types are combined by the process of appropriation in which the technology is adopted and transformed through time.
The first artifact type pertains to the infancy of the technology, meaning the state of the technology is as envisioned by the designer. The second artifact type pertains to the used technology, which means the integration of the technology by young people in their daily practices through the 'emergent process of change' ([53] pp. 3), also called the process of appropriation. The appropriation is meant to be understood as the part of which the technology is explored and evaluated and ultimately either adopted or rejected by the user. If the user rejects the technology and does not integrate it into their everyday life, disappropriation is determined to be the case. If the user ignores the technology altogether due to lack of interest in its features or functionalities, the non-appropriation happens and the process of appropriation does not happen [53]. See figure 2.20. for overview of the appropriation process.

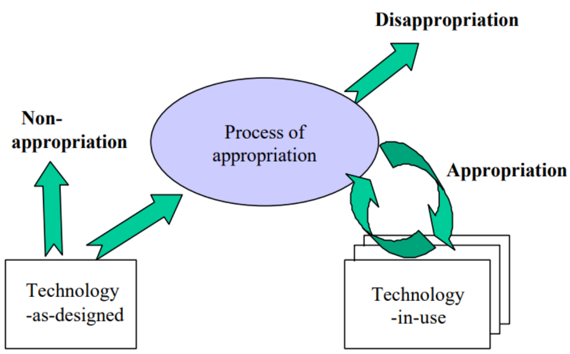


Figure 2.19: Appropriation model by Caroll et al. [53]

What this model takes into account which was not a major factor for the TAM framework to the same capacity is its perspective of the time domain and the technology's adoption and change over time for the users. Bødker and Klokmose [55] discuss the ecology of artifacts and their dynamic nature. Their ecology concept stems from Gibson's theory of ecology of visual perception [56]. For the study by Bødker and Klokmose [55], their usage of the concept is described as being dynamic and looks at the users' prospect of usage and the meaning of their daily activities. For their study, they rely on the Human-Activity Model stemming from basic Activity theory [57]. This model focuses on the why? (motivation), what? (goal-orientation) and how? (operation) of human activity. These are the three levels in the activity hierarchy. The model can be used to analyze the human experience along with the artifacts and also procure the clashes between skills and human capacity as well as the artifact's affordances and possible actions.

2.2.6 Concluding remarks of technology acceptance

The reasoning behind investigating earlier studies revolving different approaches and frameworks for understanding usage of technology as well as its adaption into the users' lives was simply to obtain an understanding of what to have in mind when using new technology with users that might not use AR on a daily basis or even explored it. Since this thesis looks into usage of AR in public areas, knowledge of user behaviour engaging with technology might be indicative of the behavioural patterns to be observed with this experiment. As mentioned in section 2.2.5, frameworks and models such as the original TAM, accounts for a lot of behavioural aspects for determining acceptance of technology, although, these models do not look into the aspect of acceptance happening over time and even being integrated into the users daily lives. Carroll et al. [53] proposed the model of Process of Appropriation as looking at the artifact as two separate ones, one used as designed and the other as integrated in the users' daily activities.
Some important takeaways from the studies explored are the factors of facilitating the acceptance of the technology, one being the perceived benefits to yield continuous use of a technology. The idea of having an augmented reality application, such as one proposed in this study, be accepted and adopted into users life could thereby be hypothesized as being facilitated by e.g, perceived benefits, ease of use and usefulness. For the sake of this thesis, continuance motivation offers some insight into some aspects of temporal user behaviour that will need to successfully be introduced in order to fully adhere to and utilize the concept of scarcity (see section 2.3.1.1.) and external FoMO appeals. Likewise, proper implementation of these concepts will be required for users to even encounter any cyclical (see section 3) aspects of the application, let alone fully realize a degree of loyalty that will have users come back to experience several epochs or locations.

The following sections and subsections will delve into aspects related to user motivation such as Self-Determination Theory (SDT) as well as the conceptualization of Fear-of-Missing-Out (FoMO) and scarcity.

2.3 Self-Determination Theory (SDT)

Self-Determination Theory (SDT), as conceptualized by Deci and Ryan [58], is a framework for understanding human motivation. According to Deci and Ryan [58], motivation comprises aspects that belong to activation and intention such as persistence, energy, direction and equifinality (reaching a goal can be done in many ways). Przybylski et al. [59] mention that three psychological needs are to be satisfied for obtaining effective self-regulation and psychological health. These needs are: autonomy, relatedness and competence. Autonomy refers to the 'personal initiative' (p. 1841), relatedness refers to how connected one is to others and lastly, competence refers to how effectively one can perform actions in the world. Seay and Kraut [60] defined self-regulation as "the ability of an individual to manage his own behavior through observation, evaluation, and consequation" (p. 831). They studied self-regulation in the context of harmful usage of online gaming. The concept of self-regulation is divided into three different classes: self-monitoring, self-evaluation and self-consequation. Self-monitoring is the player's ability to keep track of how much time they play and failing to do so is thereby also failing at self-monitoring. Self-evaluation is looking at the observed time spent playing and comparing it to either time spent on other activities or other's time spent playing. Self-consequation is referred to as the outcome of the evaluative process and the player either reinforces or punishes oneself accordingly, e.g, using playtime as a reward for completing neglected responsibilities caused by excessive playtime. This type of behaviour is linked to the controlled motivation in SDT. According to Hodkinson [61] SDT theory differentiates between autonomous- and controlled motivation. Autonomous motivation includes both elements of intrinsic and extrinsic motivation, while controlled motivation consists of external regulation and internalized factors (example above about harmful online-gaming). Specifically, intrinsic motivation is seen as the most positive form of motivation as it yields behavior that is done out of pure curiosity and in search of novelty and innovation [58]. Extrinsic motivation can be seen as a contrast to intrinsic motivation as a specific outcome is expected rather than the inherent satisfaction of intrinsic activity. These outcomes are based on external factors or regulations which can add pressure and undermine intrinsic motivations, e.g, monetary rewards, punishment can repel autonomous behaviour [58][62]. Controlled motivation can also be based on internalized factors such as self-esteem, approval or shame-avoidance [61]. Comparisons can be drawn between people whose motivation is internal and authentic (i.e, self-authored) and people who are purely controlled externally, as the latter typically express less interest, excitement and confidence which manifests itself in a detrimental fashion, as poorer performance, persistence, creativity, vitality and self-esteem than their internally motivated peers, even when the same level of competence is perceived [61]. Laato, Islam and Laine [63] correlated these components of SDT to further evolvement of 'Fear of missing out' (FoMO) in the context of usage of the location-based game (LBG) Pokémon GO. Namely, focusing on the rewards gained by play which can be argued to be external and internal pressure as mentioned above as part of SDT, where the controlled motivation of capturing e.g, rare pokemons somewhat suggests to undermine intrinsic motivation. The next section will elaborate on the term 'FoMO' and its relation to SDT and its usage in marketing.

2.3.1 FoMO - Fear of Missing Out

Having a fear of "missing out" or losing the opportunity to experience something deemed significant, either internally or externally is not necessarily a new concept. Przybylski et al. [59] defined FoMO as "a pervasive apprehension that others might be having rewarding experiences from which one is absent, FoMO is characterized by the desire to stay continually connected with what others are doing." (p. 1841), is not necessarily a new concept, but the acronym FoMO however, is a more contemporary idealization of the behavior that has become more ubiquitous and definable with the increased prevalence of social media. When seeing FoMO in relation to SDT, understanding the aforementioned psychological needs (see section 2.3.) become evidently important as FoMO is tied to self-regulation [59].
FoMO has been studied in multiple contexts such as online gaming [64][65], social media [66], journalism [67] and for marketing purposes [61]. The concept of FoMO has, in many contexts, a negative connotation as it is linked to e.g, unhappiness and loneliness [59].
Wegmann et al. [68] argue that FoMO is a 'bifactorial construct' (p. 35) with the two constructs: trait-FoMO and state-FoMO. Trait-FoMO refers to the individual predisposition of having a fear of missing out on something. State-FoMO refers to what the medium, in this case internet communication, can offer of possibilities to connect online and obtain information due to the ubiquitous nature of social media.
For measuring FoMO in several academic contexts, the scale (FoMOs) by Przybylski et al. [59] was used fully or to some extent. The scale offers ten items for quantifying the level of FoMO experienced. As previously mentioned, FoMO is also used in the field of marketing. Hodkinson [61] proposes a certain taxonomy for FoMO 'appeals'. Both commercial and non-commercial FoMO appeals are designed to motivate users to seize a particular opportunity and are therefore a "call to action" - a marketing term for a device that is designed to prompt an immediate response or purchase from a consumer [61]. These appeals can both be impersonal or in-person, referring to the mode of initialisation, and are defined as the following:

- Impersonal non-commercial: friend issuing an invitation through social media, that incorporates a FoMO appeal.
- Impersonal commercial: advertisement initiating a "missing out" appeal delivered by any impersonal communication mode (e.g, advertising).
- In-person commercial: FoMO appeal that includes general sales staff.
- In-person non-commercial: FOMO appeals that are typically seen as initiated by significant others [61].

	In-person	Impersonal
Commercial	**Sales staff**	**Advertisement**
Non-commercial	**Significant other**	**Social media**

Figure 2.20: Table showing examples of the different FoMO appeals.

FoMO appeals have already been shown to be effective for marketing purposes, in that they call on the consumer to directly address their internal hesitancy to assent to an action [61]. They have been used to stimulate demand in many different commercial areas, but one of the most common applications can be found in travel campaigns, suggesting a cultural omnipresence or even pervasion of the phenomenon.

2.3.1.1 Scarcity

The contrast to FoMO is scarcity appeals, which seek to facilitate an interest in the product, by informing the consumer (or user) of the limited time or supply available, be it perceptual or factual [61]. Scarcity refers to an insufficiency or shortage of supply. In this context, time is related to scarcity appeals as a user perceiving their time to be limited might opt for one activity over another. Marketers typically seek to induce a perception scarcity through claims of exclusivity (supply), excessive demand (first come, first serve) or offers that seem unlikely to be repeated in the future, typically limited to a short timeframe - examples of the latter which can be seen during Steam seasonal sales where popular video game titles are reduced in price for a limited time only, with some variant of a "seasonal" tag affixed to the title, to further stress the aspect of time; Let a more colloquial example be that of Lars Larsen's retail chain named Jysk, which would frequently offer specific household goods at greatly discounted prices but only for a single day - usually a weekday.
As mentioned, perception of scarcity can arise due to high demand or exclusivity with users being divisible into two groups - Individualists and collectivists. Individualists might reflect negatively upon something that is scarce due to popular demand, as it would not adequately satisfy the user's need for uniqueness in order to distinguish themselves from others, but are more likely to embrace exclusivity for the sake of said uniqueness [61].
Diametrically, collectivists' assessments of an object's value may increase when this is the case, as it indicates consensus through high popularity, and that the possession of said object would contribute to the feeling of inclusion [61]. For this thesis, the concept of scarcity will be employed as a positively-valent, external motivational appeal, with emphasis on generating an interest in the different locations chosen by SoOC, through non-commercial means, including both in-person- and impersonal modes of initialisation.

The previous sections have investigated human behaviour in the sense of motivation, intention as well as in correlation to usage of new technology and adoption of said technology into one's daily life. However, one aspect of adopting AR as the chosen type of technology has not yet been covered, namely user experience. While interacting with a virtual environment, several studies have tried to understand the users' experience and perceptions, and much research has gone into the concept of presence. Presence is a common term within human-computer interaction (HCI) and has been studied extensively in the realm of virtual reality. The next few sections will review some literature that discusses presence in augmented reality as well as auditory-induced presence due to the importance of audio in the thesis prototype's concept based on the overall vision and nature of the SoOC project itself.

2.4 Presence

Several studies have investigated the concept of presence in digital media and is often described as the sense of 'being there' [69][70][71]. According to Slater et al. [72], presence can be both behavioural and subjective, unlike immersion which deals with more objective factors such as Field of View (FOV), frame rate, interactivity and display resolution; Simply put, the objective circumstances that the equipment provides [71]. The subjective factor accrediting presence is described as how "an individual will express in response to questions about 'being there'" ([72] pp. 2), where the behavioural factor is the observable response to stimuli from the individual; arguably, presence is the observable response to immersion [71].
Witmer and Singer [73] contend that presence in a Virtual Environment (VE) can be sensed in degrees and determined largely by immersion and involvement. The degree of presence is also determined by the amount of attention shifted from the physical world to the virtual. Several studies suggest that attention is guided by the saliency of given information or stimuli in the environment [73]. The terms immersion and involvement are used in the study by Witmer and Singer [73] as core elements for achieving presence. As mentioned earlier, Slater et al. [72] categorized immersion as being an objective factor, however, this falls in disagreement with the definition by Witmer and Singer [73] who argue that immersion is, like presence, a subjective and individual experience where the individual feels included and enveloped in a stimuli-rich environment. Some of the factors which induce immersion include the seclusion of the physical environment, perceived self-movement, and control over interaction with the environment rather than being dependent on what the equipment can provide in terms of fidelity. Involvement is the state of consciousness experienced when one's attention and energy is directed on "a coherent set of stimuli or meaningfully related activities and events" ([73] pp. 227). However, these activities and events, and the degree of their significance determine how involved one can feel. As previously mentioned, presence can not be felt without immersion and involvement according to Witmer and Singer [73]. However, this is not agreed upon by Slater et al. [71], as their opinion on Witmer and Singer's definition of involvement is considered, by Slater et al., as being somewhat confounded with presence. Likewise, Slater

et al. [71] contradict the ideation of immersion and involvement being essential for presence, and distinguish this by arguing that one can be involved and not present (e.g, watching a TV show), and also be present while not being involved (everyday situations and scenarios).

For this particular study, we will adhere to the theories and definitions of presence by Slater et al. [69][71][72] to stick to a well documented and studied taxonomy.

Researchers have often referred to the experience and presence felt in VEs as mediated through stimuli in order to be sensed as nonmediated [74][75][76]. As mentioned by IJsselsteijn and Riva [75], the sense of presence is not exactly part of most people's conscious thought and is often reflected upon, unlike a mediated experience through certain technology. The perception of stimuli can be from either the virtual- and real world. What determines which world one will feel present in, is which provides the most dominant stimuli [75]. As previously mentioned, selective attention and the saliency of the percepts as well as the level of immersion in a mediated environment can lead to an enhanced sense of presence. This can ultimately lead to telepresence which happens when the majority of the attention is occupied by the technologically mediated environment [75][77][78] as Spatial presence. Spatial presence as referred to by Wirth et al. [78] is not necessarily reliant on stimuli rich environments with multimodal sensory inputs, as also more cognitive- or internal processes can provide the same sense of being somewhere else, e.g, as experienced through text. According to Wirth et al. [78], spatial presence is a two-dimensional construct where one dimension adheres to the sense of being somewhere else physically, through the medium, and the other dimension being the perceived possibilities of actions within said environment, meaning that the user will see action possibilities relevant to the mediated environment rather than from their real environment. Wirth et al. [78] defined spatial presence as a binary experience where the two dimensions of 'self-location' and 'realization of action possibilities', both of which adhere to a mediated environment, causing a person's mental capacities to shift to the mediated environment instead of reality ([78] pp. 4).
Amplifying and increasing sensory input and making the user feel some sort of feedback from actions can further enhance these two aforementioned dimensions. This was also similar in nature to what was suggested by Regenbrecht and Schubert [79]. Their emphasis was on the subjective perception of possible actions rather than objective. Schubert [80] has later argued spatial presence as a cognitive feeling in which unconscious spatial processes result in the subjective sense of spatial presence. He suggests that the "feedback of unconscious processes of spatial perception that try to locate the human body in relation to its environment, and to determine possible interactions with it" (p.170). He then suggests that conscious processes can happen if said body is located in the perceived environment and establish possible actions based on this.
In terms of measuring spatial presence, Vorderer et al. [81] created a specific questionnaire relating to spatial presence called MEC-SPQ (Measurement, Effects, Conditions - Spatial Pres-

ence Questionnaire). The structure of this questionnaire takes some of these aforementioned concepts such as attention, possible actions and sense of self-location as variables. Regenbrecht and Schubert [82] likewise tested a questionnaire pertaining to measuring presence in augmented reality. The emphasis was especially on the integration of the virtual objects into the real world through display-based technologies. Lombard and Ditton [74] discussed presence as transportation, where the three different types of presence as transportation were elaborated. These were denoted as; 'You are here', 'It is here' and 'We are together'. The first one is about being transported to a different place, the second one about having a place or objects being transported to the user, and the last type is about more than one person being transported together to a shared place. Regenbrecht and Schubert [82] highlight the second type ('It is here') specifically as what in reality happens when using augmented reality - the virtual objects are transported to the physical environment.

2.4.1 Presence in Augmented Reality

Much of the research conveyed in the previous section regarding presence, mostly pertains to studies of presence, specifically in Virtual Reality. This section will investigate some of the similarities and as well as differences within the study of presence in VR and AR.
Perhaps the most obvious difference between the two technologies is the level of projection of the real world to the user; with VR the user is fully exposed to a virtual environment with little notion of the real world, whereas AR superimposes virtual elements in real world environments (see section 2.1). The continuum of unmediated and mediated experience in regards to these two technologies is distinctive according to Tang, Biocca and Lim [76], as they suggest that AR is more of an unmediated experience than VR. As AR offers multiple platforms for experiencing an augmented environment (see section 2.1.1), putting it alongside the first person perspective as a regular HMD provides, the sense of being in the mediated environment is likely enhanced [76]. Two factors are relevant in this context: transparency and continuity of the interface. Transparency refers to the aspect of eliminating the mediation, meaning the gateway to the real world. Continuity of the interface refers to the disruption of the interaction or experience, also understood as the level of disruption of the experience or lack thereof [76]. This coincides with the similar theories by Slater and Steed [83], where disturbances count as breaks-in-presence (BIP), which they compared to being awakened during a dream in a Rapid Eye Movement (REM) sleep state.
Another aspect to consider in terms of presence in VR and AR respectively, is the user's sense of embodiment in the world. In VR it is possible to track the user's body through different usages of trackers and motion capture suits to provide the user with the sense of one's body in the mediated environment. Controllers have also served as trackers of the users' hands in the virtual world in which actions and interaction with the environment can happen. Issues with the projection of said user body are according to Steed et al. [84] currently difficult due to potential sensory mismatches of the proprioceptive perception of one's body parts in relation to your body, as well as creating believable and corresponding imitations of the user's

real body and attire, and ultimately affect the user's ability to interact with the virtual world. However, this does not pose the same issue in AR as the self-embodiment most often relies on the user's real world body [76].
The level of presence when using AR displays is still not well defined and difficult to properly investigate due to the nature of the medium along with the real world's effect on its perception [11]. However, their research showed the level of presence to be significantly higher using mobile phone AR rather than other types of AR displays (see section 2.1.1). As much of their experiment was rooted in virtual objects in real environments seen through four different MAR devices, the significant outcome for our particular study is the elements of user experience and usability. It was clear that virtual objects behaved the best using PhoneAR as they were not directly impacted by indoor/outdoor lightning as much as with some of the other MAR devices, especially Open AR (HoloLens). An important part of their research revolved around what they denoted as 'object presence' ([11] pp. 6). The concept of object presence was mentioned and defined in earlier studies, such as Stevens et al. [85] where the definition was based on Witmer and Singer's [73] definition of presence. Stevens et al. ([85] pp. 83) thereby suggest object presence to be "the subjective experience of being co-located with a set of objects, even when one is physically not in such a situation". This correlates to the types of transportation as mentioned in section 2.4, where the virtual world is transported to the user, and in this case through the use of a MAR device. The definition was then formulated by Choi, Kim and Kim ([11] pp. 6) as "how much the virtual augmentation feels to be realistic, physical, actually part of the real world, natural and harmonious." For this study, we will rely on the definition by Choi, Kim and Kim [11] as it coincides with the sentiment which is of importance for this prototype.
One last point to touch upon within presence, especially presence in AR, is the social aspect of the experience. What is known as 'social presence' is the part of presence that relates to the type of transportation that occurs when you are transported to a mediated environment with another actor (real or synthetic) [74][86][87]. When using and evaluating AR applications that are location-aware, it becomes problematic as the real world offers many random and confounding variables to the equation [88]. McCall et al. [88] studied how presence could be used to evaluate AR location-aware games, i.e, games that utilize an urban environment. They argue that users' presence is determined by the actions and interactions provided by the game or application. Essentially, the realities are not strictly divided but rather combined fluently, much like Milgram and Kishino [89] established in their continuum for Mixed Reality (MR).

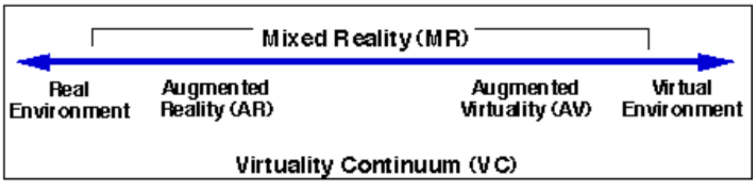


Figure 2.21: Virtuality Continuum by Milgram and Kishino (1994), showing there is no distinct line between the 'realities', but rather a bidirectional classification for MR.

Since the emphasis of this thesis is mostly focused on adding augmented sounds to urban environments (see section 1), it is therefore important to delve into the components of presence in the auditory modalities and investigate the affordances in terms of the medium. The next section and subsections will elaborate on auditory presence and its integration into virtual- and augmented reality.

2.4.2 Auditory-Induced Presence in AR/VR

It has long been recognized that head- and position tracking has a high degree of significance when attempting to induce a sense of presence within the user [90]. While the visual sense tends to be the dominant modality utilized for perception and discrimination of elements in the surrounding environment (virtual or otherwise) [91][92][93], Larsson et al. [94] note that the auditory modality has unique capabilities for assessing space and introducing presence in a different manner. They stress that while the visual system has a very high spatial resolution, it is limited by its field of view (FOV), requiring head movement in order to introduce and assess more data in a scene, when also being vulnerable to complete omission through obfuscation or occlusion (e.g, by simply closing one's eyes). However, the auditory system is less accurate in terms of spatial resolution, but offers the ability to assess spatial cues from the entire surrounding space - concurrently - primarily by localisation of objects' direct sound propagation and by attaining information of geometry and size of a location, through wall reflections and reverberation (or lack thereof) [94]. Furthermore, Larsson et al. [94] propose that the auditory modality possesses a higher temporal resolution than its visual counterpart, and research suggests that rhythm information across all modalities is encoded and based on "auditory code" [95]. Larsson et. al. submit several parameters they consider to contribute to auditory-induced presence, namely: Externalization, spaciousness, prior expectations, localisation, and sound quality [94].

2.4.2.1 Externalization

A study by Hendrix and Barfield [96] showed that the inclusion of sound as opposed to the complete omission of sound increased presence ratings, as well as showing that spatialized sound was favored in terms of presence when compared to its non-spatialized counterpart. However, Hendrix and Barfield attribute the (lack of) significance of the presence scores to their use of non-personalized HRTFs (head-related transfer functions), resulting in a lack of externalization of the sound, instead resulting in an internalisation. Data from studies conducted by Väljamäe et. al [97] corroborates the claim by Hendrix and Barfield, by presenting indications of HRTFs having a positive impact on presence in a virtual environment. With that said, externalization is considered important for MR applications that combine an auditory virtual environment (VE) with a real VE, since it potentially increases the perceived consistency between both environments, and thus it is not necessarily solely dependant on HRTFs to achieve the effect [94].

2.4.2.2 Spaciousness

Externalization can further be obtained or improved upon, by adding room acoustic cues, such as early reflections and reverberation [94]. This is corroborated by the findings from the study by Larsson et al. [98], which showed that the addition of room acoustics in a virtual environment resulted in higher presence ratings than the anechoic condition presented. The study attributes externalization as part of the result, as reverb is known to increase externalization in headphone-based systems [94][96][98][99], which additionally increases "naturalness" and presence (see section 2.4.3). As mentioned, acoustic cues also present the user with valuable spatio-temporal information regarding e.g, room size and geometry [100], with the ability to induce both object presence (see section 2.4), by eliciting the feeling of something or someone being located at a certain place in relation to one-self, as well as induce spatial presence (see section 2.4.), the feeling of being surrounded by- or being present inside a specific space [101]. The reproduction techniques concerning artificial reverberation in a VE are also a significant factor to be considered. A between-group study by Larsson et al. [102] showed that when comparing stereo sound with room acoustic cues in a VE, to a binaural rendering with acoustic cues in the same VE, the latter yielded significantly higher presence ratings [94][102].

Spaciousness Parameters

Gorzel et al. [103] express some of the challenges and features of note, regarding realistic sounding, synthesized reverberation in a virtual environment:

- Reflections and their reverberation time (RT60)
- Geometry and physical structure of the environment
- Material absorption coefficients
- Distance (depth cues)

Google's open source spatial audio SDK dubbed "Resonance Audio" [104] offers high fidelity audio simulations for different types of media, such as VR and AR, at a low computational cost. The virtual reflections of the SDK follow the same principles as their real world counterparts, in which reverberation is the persistence of sound after the occurrence and conclusion of its respective sound event. The time it takes for a signal to decay by 60dB (1/1000th SPL) is symbolized as RT60 and can be calculated using Sabine's formula:

$$RT60 = kV/A = 0.161V/A \tag{2.1}$$

Where:

- k is a constant that denotes the seconds per meter (s/m) = 0.161 s/m through air at 20°C (343m/s).
- V is room volume in m^3
- A is the equivalent absorption surface in m^2 expressed in Sabins = αS
- α is the absorption coefficient of the given surface material, normalized between 0 to 1, with 0 meaning 0% of the sound is absorbed (100% reflective efficiency), and 1 meaning 100% of the sound is absorbed (0% reflective efficiency).

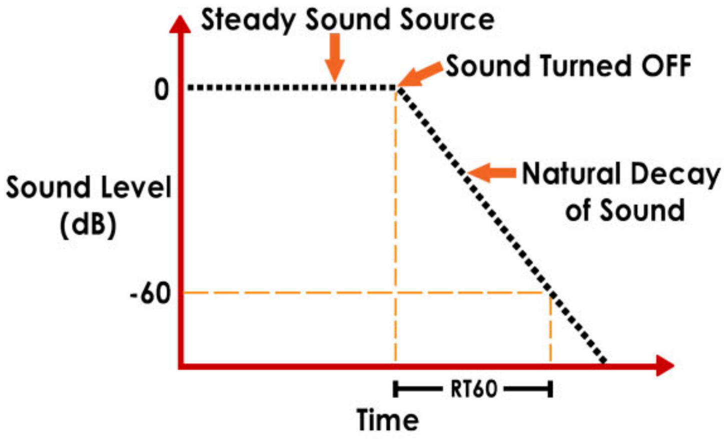


Figure 2.22: Basic diagram of the concept of RT60 reverberation time measurement.

Since spaciousness and auditory cues are shown to have a significant impact on presence, authentic representations of room acoustics are imperative for spatial audio and the perception of 3D space in both real-, virtual- and augmented environments. Larsson et al. [94] suggest that adequate real-time changes of the auditory environment can be crucial for inducing or maintaining a sense of presence, specifically in relation to localization of moving sound sources and the movement of the listeners themselves.

2.4.2.3 Prior Expectations

For virtual room acoustic simulations and subsequent auditory cues, personal heuristics and *intersensory discrepancies* [92][105][106][107] should be considered, as to best match the auditory feedback with the visual impression - be it real, augmented or virtual - otherwise, the user's sense of presence within the environment can be reduced or disrupted by the perceptually different environments [94]. The features stated in (section 2.4.2.2) are also considered important for the sake of conforming to the principles of personal heuristics.

Chueng and Marsden [108] proposed the notion of expectation and discrimination as factors related to presence, more specifically as "...the important perceived affordances of sound producing objects that contribute to a sense of place for users and contribute to a feeling of presence" [108]. Expectation is explained by Chueng and Marsden [108] as the degree of which a person will expect to hear a specific sound in a particular place, while discrimination relates to extent of which a sound will uniquely identify said particular place. Following this concept, the most effective sounds will be those that elicit high expectation and/or a high degree of discrimination from the user [108].

2.4.2.4 Localization

While the field of AR research has mostly revolved around the visual augmentation of physical- and digital space, other parts of the human-physiological methods of perception, have been less explored [9]. In the current tech climate, 3D- or spatial sound is utilized extensively in VR using methods such as vector-based amplitude panning (VBAP), Ambisonics or HRTFs to relay auditory information, i.e, direction and distance to an audiosource, on both the sagittal- (elevation) and transverse (azimuth, horizontal) planes; These are the binaural cues known as the Interaural time- and level differences (ITD and ILD respectively) [109]. ITD cues arise due to the spatial positioning of the ears with distance d, given by the head dimensions, which leads to a difference in the time for a signal to reach both ears, denoted Δt. The ILD cues occur when an incident wave hits the head, causing the sound to diffract and shadow, resulting in the perception of the sound having higher intensity in the ear closest to the soundsource, assuming the source is not directly in front or behind the listener.

The Duplex theory suggests that high frequency tones are localized by ILDs and low frequency tones by ITDs [109]. The ILDs are frequency dependant, due to the fact that wavelengths of low frequencies can be greater than the dimensions of the head, resulting in phase delays being the main attribute used to determine direction at low frequencies, while at high frequencies where d is greater than the wavelength, the level differences can be evaluated by the auditory system with less ambiguity.

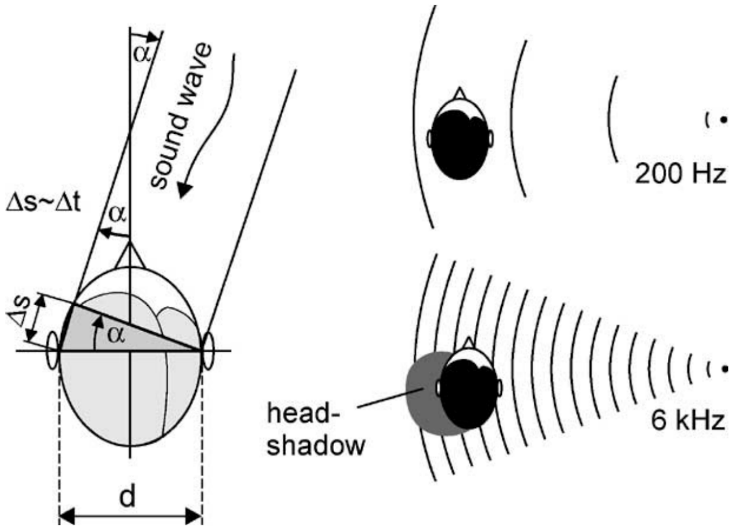


Figure 2.23: Left side showing an incident wave hitting a model of a head with causing a time delay due to the angle of the wave. Right side showing the principles of acoustic shadowing at different frequencies.

Furthermore, studies have shown that using generic HRTFs as opposed to personalized- or individualized HRTFs can result in increased front-back reversals and decreased performance in spatial localization [96].

Head Tracking

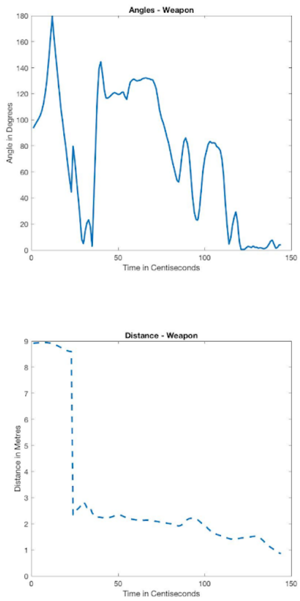


Figure 2.24: Plots showing distance and angle to audiosource as a function of time, for one participant during the test conducted by Dourado and Strübig [110]. See appendix A4 for more comparisons and participant data.

Front-back reversals (specifically classified as either front-to-back, or back-to-front) occur when the listener incorrectly judges a sound source to be flipped or mirrored on the horizontal plane (azimuth). However, head tracking has been shown to eliminate reversals in spatial audio applications, mainly due to the differential integration of interaural cues as a function of time, by which the cone of confusion is resolved by the motion of the head [99][111]. A test was conducted by Dourado and Strübig [110], in which participants were asked to localize a sound source in a virtual environment using ray tracing and a HMD for tracking head movement, efficacy and precision. Measurements were collected for both azimuth and elevation, with only 3DOF (degrees of freedom). This led to observations suggesting that participants with more initial head movement were quicker to resolve and estimate a guess for each new location, and were generally less prone to front-back confusion or azimuthal errors of more than 90 degrees. However, it is worth noting that time was not a condition for the test, as all participants were allowed an indefinite amount of time between each subsequent stimuli, as well as many occurrences of the sound source as they needed.

Furthermore, a study by Begault et al. [99] showed that the inclusion of head tracking for generic HRTFs resulted in a 3-degree localization improvement for azimuth error as opposed to the same conditions, but without head tracking.

2.4.2.5 Reverb Revisited

Another cue for perceiving distance is that of reverberation (see section 2.4.2.2) and energy dissipation. Regarding the latter, as the listener moves further away from an audiosource, the sound will be attenuated. In a free-field situation, the distance-to-SPL (sound pressure level) ratio follows that of the inverse square law, which dictates that every doubling of distance from a sound source will result in a -6dB change in intensity. Research by Begault et al. [99], on the impact of reverberation on spatial perception of a virtual speech source, indicated that auditory stimuli with reverberation yielded lower azimuth errors and higher externalization rates, albeit at the cost of some elevation accuracy. The presence of reverberation under two different conditions (early reflections only and full auralization), caused the accuracy of azimuth estimation to improve by approximately 5 degrees when compared to an anechoic condition [99].

2.4.2.6 Sound Quality

While hinted at in previous sections on sound (see section 2.4.2.4), quality and fidelity has been shown to have a significant impact on several aspects of presence [94][112][105]. Comparing 2-channel audio to a 5.1 (5 speakers, 1 subwoofer) surround setup, with and without high/low quality visual stimuli, Skalski and Whitbred [112] found that the 5.1 surround stimuli had a significantly more pronounced effect on player presence and enjoyment, even with the omission of visual stimuli. The parameters of sound quality in relation to auditory induced presence, as established by Ozawa et al. [113], also encompass sound information and localization. Their study employed binaurally recorded ecological sounds such as footsteps and vehicles etc., and found that the aforementioned parameters - sound information and localization - to be the principle factors for obtaining high auditory presence [113][94]. Larsson et al. [94] propose the implication that sound should be informative and enable listeners to imagine the original or augmented scene naturally, and that sound sources should be naturally or “easily” localizable. Other tunable parameters to consider are those of sound pressure levels (SPL), frequency content and dynamic range of the audio system. Ozawa and Miyasaka [113] observed that presence ratings increased with SPL without visual stimulus, and that ratings were generally highest for realistic SPL matching the auditory content, when paired with visual stimuli simultaneously [94]. This suggests that proper calibration of the overall SPL and frequency response of a MR audiosystem is relevant for producing a congruent auditory environment that is consistent with the visual stimuli presented.

2.4.2.7 Occlusion

In a real environment, the direct path between a given sound source and a listener will often be occluded by geometry, people or other obstructions that directly shape the sound. Some of this is attributed to the concept of diffraction (see section 2.4.2.4), as the sound still reaches the position of the listener by diffracting, or “bending”, around the occluding obstacle present in the line of straight propagation, hence why sound is audible from adjoining rooms with open doors if walking down a corridor, or why sound sources can be audible from around corners [114]. Diffraction is dependent on wavelength and the surface area of an obstacle and increases as the ratio between wavelength and obstacle size is increased, meaning diffraction will generally be greater for lower frequencies and smaller obstacles [114]. Since the human range of hearing is approximately from 20Hz to 20kHz, if not considering impairment nor decay caused by age or other factors, wavelengths (λ) within that frequency range (f_r), with a propagation speed (velocity - v) of 343m/s can be expressed as $\lambda = v/f_r$, and will correspond to wavelengths from 17m (20Hz) to 0.017m (20kHz). Wavelengths within this range match dimensions frequently encountered in physical environments, such as inside buildings and inhabited spaces, suggesting it is an important aspect of sound propagation, similar to reflection, absorption, diffusion and transmission [114].

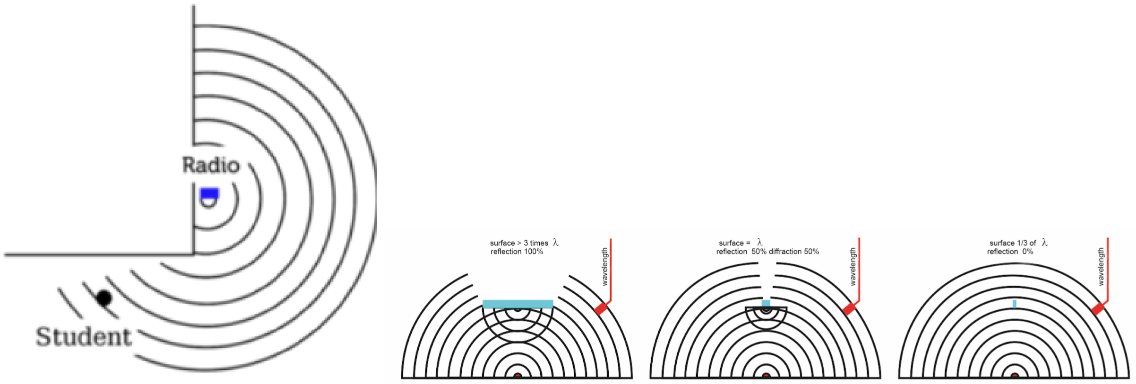


Figure 2.25: Example of a student who is able to hear the audio from the radio before having a visual (left). Degree of reflection and diffraction to size of an obstructing surface (right) [115][116].

Mixing occlusion between real and virtual objects adds credence to the existence of a virtual object existing in the physical space which it seemingly occupies, whereas incorrect occlusion can be detrimental to the user's perception of the object's existence and tangibility [117]. As with other auditory cues, an accurate representation of occlusion can further inform listeners of their surroundings by strengthening depth- and localization cue information [19][117]. Finally, omission of auditory occlusion and diffraction leads to a less congruent user perception of a virtual environment and has been shown to decrease presence and immersion [114][118].

The previous sections related to audio have mainly covered concepts and tools that need to be considered when designing and implementing audio, with the goal of soliciting a higher sense of auditory-induced presence in the user. However, the means through which the auditory information is conveyed could have reductive or even detrimental effects when considering an AR scenario.

2.4.3 Acoustic Transparency

Auditory awareness is an important feature when considering the usage of hearables and headphones in a public domain [101]. For some scenarios, a desired feature is to occlude the surrounding sounds through the use of e.g, passive noise-cancelling headphones, creating an auditory cocoon. However, conceding the ability to hear traffic or localize sound sources in a real world setting can be detrimental, not only for the users' safety, but also for their inclusion and adoption of virtual audio sources [101].

In terms of wearable audio, such as headphones and hearables, having the ability to intermix virtual audio with sounds from the real world surrounding is referred to as acoustic transparency. In this category, both passive- and active acoustic transparency exist, where passive transparency is when the ear canal is sufficiently open to hear both the virtual- and real sounds at the same time. For active acoustic transparency, active-noise cancelling (ANC) headphones can be configured to selectively filter unwanted elements (e.g, noise or surroundings), while boosting the desired output (e.g, speech). Additionally, inertial measurement units (IMUs - see Bose frames) can be considered a way to actively enhance spatial awareness and aspects such as localization and externalization [101].

Research by McGill et al. [101] strongly suggests that acoustic transparency positively affects the sense of externalization. Furthermore, acoustically transparent headsets showed a strong increase in **perceived** presence (spatial presence) as well as the sense of presence among other people (*social* presence) when tested outside for different media, with the adverse being the case for standard, more opaque headphones resulting in diminished scores and effects.

A study by Murray et al. [119] assessed the sense of presence in regards to auditory experience. In the study, the participants would be subjected to induced hearing loss through the use of earplugs, while within a mediated environment. Their results suggest that the use of closed- or inserted headphones lead to a higher sense of auditory self-awareness and self-representation, which had a detrimental effect on their overall sense of presence [94][119].

2.4.3.1 Adoption

Relevant to technology acceptance of wearables (see section 2.2.1), and more specifically adoption, integration and consumption of personal audio, ranging from ambient soundscapes and podcasts, through assistance (location-based audio, GPS, planning etc.), to music and drama content, McGill et al. [101] explored potential adoption of different types of media content with both acoustically transparent- and opaque headwear in public, professional and private scenarios. Given short exposure to the different technologies and media, participants showed a general inclination to isolation and auditory fidelity through the use of ANC headphones when it came to media and entertainment content, while functional audio, voice conversations and voice assistance were shown to be the most anticipated for adoption. It is worth noting that users generally favoured using ANC headphones for media requiring feedback or input from themselves, when in public, but with a large shift towards acoustically transparent audio for the same media, in a private setting (see appendix A1 for questionnaires and user responses)[101]; A possible factor for this lack of public adoption can be observed from almost half the participants expressing concern regarding overhearing of personal or sensitive information output by the headphones. Despite this, the vast majority (more than 75%) would still look for acoustic transparency as a key feature for future audio purchases as well as be more likely to use audio-based applications in their day-to-day lives [101].

2.4.4 Concluding Remarks for Presence

Investigating presence in its simple interpretation was included in our study to obtain a broader understanding of the user experience aspect of using technology such as MR, in which an environment is mediated for the user. Presence has multiple definitions and beliefs where some are closer to each other than others. Furthermore, when considering the vision and subsequent requirements established by SoOC, adhering to the principles and parameters presented in the sections on auditory-induced presence and acoustic transparency could feasibly serve to create an environment, where the artists hired for the project can convey their desired messages, ideally without reducing the end-user's sense of presence or immersion. Additionally, we have chosen to adhere to some of the theories mentioned in the sections above to establish some fundamental concepts for this particular project to build upon. Likewise, we have chosen to utilize some of the evaluative methods, and well researched and documented questionnaires from prior studies to investigate similar research questions our thesis poses.

The next section will delve into a new study (July 2020) which specifically investigated the usage of Location-Based Games and their social features during a global pandemic. The reason for including this section is to obtain an understanding of human behaviour during these circumstances, as it is arguably the worst global crisis in recent time with a lot of uncertainty and fear as a result of this.

2.5 The implication of COVID-19 on Location-Based Games and Applications

Location-based games (LBGs) or otherwise known as location-aware games, GPS games, location-based AR games and similar terms [63] are games that incentivize players to move around in the real world and play. This type of game, also called exergame, has a positive reputation of being helpful in activating people, who are not otherwise active, to do outdoor activities in which are fun and engaging [63][120]. Many of these games, Pokémon GO in particular, invite social interaction as a mechanic for obtaining rare Pokémon and other achievements. Some of these social gatherings happen in the shape of 'raids' in which multiple players meet up to fight a certain boss and then get some sort of reward, which results in several individuals gathering at a specific location. However, due to the renowned COVID-19 virus that has impacted the entire globe and called out for self-isolation and social distancing, these types of games fall in a category of being potentially harmful. This is due to the aspect of gathering up with many other players in close proximity and thereby exposing themselves and each other to the virus due to the risks of aerosol transmission [121]. Also, people are able to transmit the virus without experiencing symptoms and thereby not know they have the virus [121]. During the time where the virus peaked in many European countries, including Denmark, some governments issued several precautions which resulted in mass shutdown of most public institutions and social distancing. Developers of Pokémon GO reacted to this by changing certain features of the game to avoid gatherings and make rewards obtainable by solo play [63]. Laato, Islam and Laine [120] investigated how developers as well as players responded to said pandemic and how they adjusted accordingly. Since the games they used for reference all involved usage outside, the aspects of social distancing as well as quarantine did not fit the nature of the game per se. Research in this field is new territory, as this is the most severe case of pandemics in recent time, and ubiquitous devices are part of most people's lives which makes it rather effortless to use, e.g, for games such as Pokémon GO. To that effect, the same researchers conducted further studies on whether the implications of COVID-19 affect them playing Pokémon GO, socially [63]. The variables they discovered as salient in ultimately lowering social play among users were:

1. Player motivation
2. Health factors such as perceived severity and susceptibility
3. Organizational factors such as developers and government actions.

Their approach to this problem is threefold as they all affect the player's social play during a pandemic. All of these factors also play a role for our study as many of the same elements are at play. The structure of their study can be seen below:

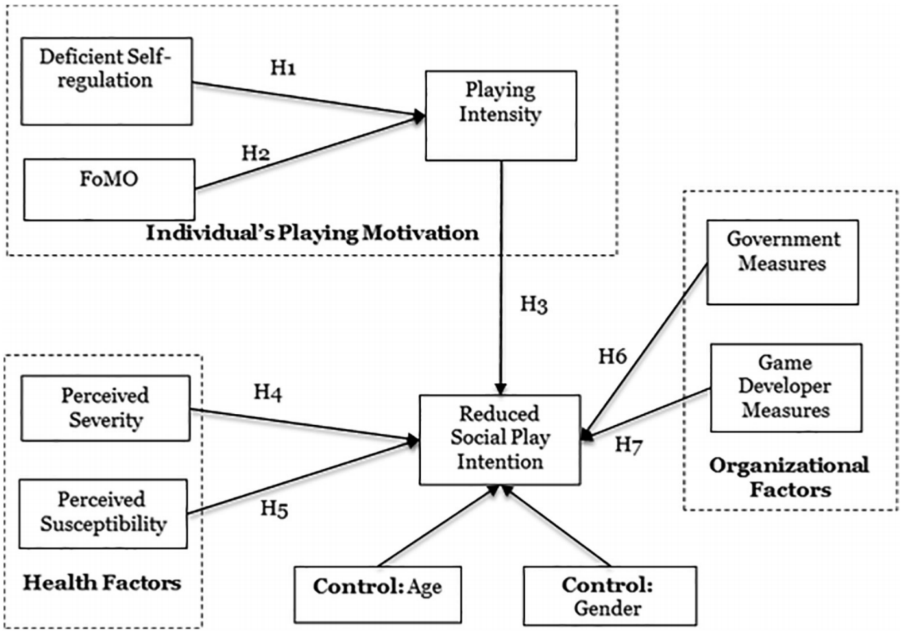


Figure 2.26: Model showing factors resulting in reduced social play intention [63]

Their study relies on many theoretical frameworks for studying intention and motivation. Self-determination theory (SDT) is used in this case to understand what lies behind the intrinsic motivation in players when playing LBGs in the current global situation. SDT is, as mentioned, a theory that tries to understand motivation and personality traits in terms of either being autonomous or controlled [62]. Autonomous encompasses elements from both intrinsic and extrinsic motivations, whereas controlled motivation is affected by external regulations as well as internal factors such as self-esteem and approval [61]. Deci and Ryan [58] conceptualized the intrinsic motivation combined with intrinsic regulation to yield regulative processes such as enjoyment and satisfaction. Hodkinson [61] conducted a study in which conceptualized a model for FoMO and specifically found SDT to be applicable. As also investigated as part of our study (see section 2.3.1) (due to the nature of providing content that is limited (Scarcity)), FoMO is part of user behaviour and rely on the intrinsic motivation for pursuing a certain action due to social pressure, e.g, usage of social media and in the case of the study by Laato, Laine and Islam [63], the FoMO deals with the unique rewards which yield materialistic gains that can happen when playing LBGs such as Pokemon GO, thereby adding pressure to continuously play to achieve these rewards. Another part of the user motivation for playing Pokemon GO according to Laato, Laine and Islam [63] is regarding Self-regulation or lack thereof, and is thereby denoted as deficient self-regulation (see section 2.3). Another theoretical concept they included as a factor for reduced social play intention was protection motivation theory (PMT) which is used to understand human behaviour during e.g, pandemics [63][120][122]. Perceived severity, as previously mentioned, is a component of PMT and is among others that appear in behaviours to protect oneself from infectious diseases. Many studies investigating human behaviour during epidemics or pandemics have been conducted based on hypothetical scenarios, to try and predict behavioural outcomes. Especially social distancing during a pandemic/epidemic was linked to all PMT components (e.g, perceived severity and self-efficacy), where specifically self-efficacy was a strong indicator for protective behaviour during a hypothetical epidemic/pandemic [122]. Their results indicated that governmental- and developer incentives significantly impacted reduced social play in a positive manner [63]. FoMO and self-regulation was seen to impact reduced social play in a negative fashion and in turn imply reluctance to comply with governmental- and health incentives such social distancing despite the efforts by the developers.

The main takeaway from these studies, especially the one by Laato, Islam and Laine [63] is the new perspective on how different factors affect the user's or player's motivation to procure social activities in the context of an augmented reality location-based game during a global pandemic. This provides food for thought in terms of which strategies are to be considered for our research.

2.5.1 Evaluating AR and VR experiments during COVID-19

Another implication, due to the current global health crisis, is mostly related to the field of research of evaluating such applications, i.e, augmented reality and virtual reality. As many institutions in Denmark have been shut down in order to prevent social gatherings and ultimately further spread the disease, conducting experiments and research normally has been challenging due to lack of access to facilities as well as the possibility to ethically conduct experiments with participants. Prominent researchers within the field of HCI conducted a study, published in July 2020, in which they investigated how COVID-19 affected evaluation of immersive experiences in VR and AR [123]. As previously mentioned, many of these types of experiments are conducted in accordance with e.g, university research labs, where access to both equipment such as HMDs and other high-end equipment is normally possible. However, due to the COVID-19 pandemic, access to these facilities have been denied for a while and thereby creating some uncertainty in terms of performing proper research. Steed et al. [123] proposed three solutions for conducting this type of experiment depending on the restrictions:

- **Short-term solution:** Remote experiments done with collaborating labs; using other researchers or lab associated people to conduct research. This however, poses a concern as the sample will be conveniently sampled as well as potentially yield biased results, as they might be considered expert users as they have extensive knowledge within the field. This then might not be indicative of generalized results for usability rather than having a mixed pool of participants to provide a more reliable outcome.
- **Medium-term solution:** Rather than recruiting participants that are associated with collaborating labs and colleagues, Steed et al. [123] propose the medium term solution to instead recruit participants with the required hardware for the experiment, e.g, virtual reality HMD. These participants might represent a more sustainable pool for reliable results. However, the lack of proper remote observation and conduction of the experiments are yet to be properly done, as well as ensure each participant gets the exact same experience, which can be problematic due to the nature of the varying equipment.
- **Long term solution:** This solution is as the name implies a more elongated process, yet effective in terms of getting a more diverse pool sample. Steed et al. [123] propose the process to be rooted in lending out equipment to individuals for testing. The amount of equipment could be of a smaller quantity, therefore making it a cheaper solution. They suggest that crowdfunding or governmental funding could alleviate the hardware restrictions and fund e.g, more equipment and better infrastructure to distribute to individuals from multiple regions. This would ultimately void the need for participants to meet up at a research lab to participate in an experiment and thereby be in concordance with health- and governmental incentives.

These solutions give a good understanding of the challenges that we face under these circumstances as well as seen in the perspective of the time constraints of this thesis. Given the circumstance, alternative methods should be used to ensure the experiment is conducted according to official precautions and laws and does not ensue ethical issues. The following section will elaborate on the study's approach to the initial problem statement as well as provide a list of hypotheses and a final research question. Section 3 will cover the requirements of the application, both in terms of the research as well as what was contributed from Sounds of Our Cities in the initial stages of the project (see section 5.1. on phase I of design).

3 Approach

As this is a study done in collaboration with the project Sound of Our Cities, the approach is obviously influenced by the early concepts and ideas provided by key individuals in the collaboration (Aalborg University included). The collective objective of the SoOC project is to create culturally sonic art for real locations (in this case, Roeselare, Belgium and Barcelona, Spain). What will be contributed from our side as students from Aalborg University is the technological integration of the aforementioned art and its ties to the location, hence making it a location-based application. Despite the SoOC project being mostly based on sonic art integration in urban environments, this thesis likewise presents the visual domain as a link to create virtual immersive environments in AR, which the user can interact with by both being a part of it, as well as experiencing it as an element of the physical space. The intention of making these environments is to give the appointed artists that are part of the SoOC project, the ability to mix the realities through visual and auditory stimuli.

In the perspective of other similar applications providing location-based AR content such as Pokémon GO, the aspects of usage, player behavior and motivation was investigated (see section 2.3 and 2.5), which in turn provide some ideas for continuously giving the user a reason for experiencing these environments. As the project is to be set to have installations in two specific cities - Roeselare and Barcelona - we propose creating environments changing cyclically (or after a given time) to encompass the changing physical environments through sound and visuals, similar as described by Ronald Azuma [8][20](see section 2.1.3). This stems from the theoretical concepts of scarcity, meaning that the virtual environments are only available throughout a given time, incentivizing users to procure these environments and experience them. This is similar to the strategies behind applications such as Pokémon GO that likewise have time-limited pokémons at certain locations for one to capture as well as social gatherings in the form of raids (see section 2.5), i.e, depending on the user and their motivations can experience FoMO (see section 2.3.1).

The choice of technology for this project is MobileAR, more specifically smartphones as this in its early stages lends itself more to being consumer friendly and easier adoptable (see section 2.2.3.1). In terms of research method, this study will be a correlational study as we seek to explore the relationship between certain factors of the analysis. More specifically, this study will have a more exploratory approach [124], as we do not aim to seek confirmation of existing theories but try to explore theoretical frameworks for a specific scenario (the prototype presented in this study). We will likewise utilize Partial Least Squares (PLS-PM or PLS-SEM) modelling or soft modeling as denoted by [125]. This type of path modeling is appropriate for exploratory frameworks unlike covariance based modeling, which aims to confirm existing theories and outputs a covariance sample matrix [125]. A different approach to the research design was considered, including that of an experimental setup, which would involve testing the application with and without the attached timers and cyclical behaviour, ultimately looking at scarcity as the independent variable and its causal effect on the dependent variable - continuance intention. Furthermore, designing under an experimental research paradigm would require a more controlled environment, reducing ecological validity, and would not allow for assessment of relationships between multiple variables in more realistic or "everyday life" scenarios; a correlational research design was thus chosen.

In order to formulate a research question and hypotheses for this thesis, an overview or structure for concepts and terminology is then set up in relation to one another. The research model will be pictured as followed:

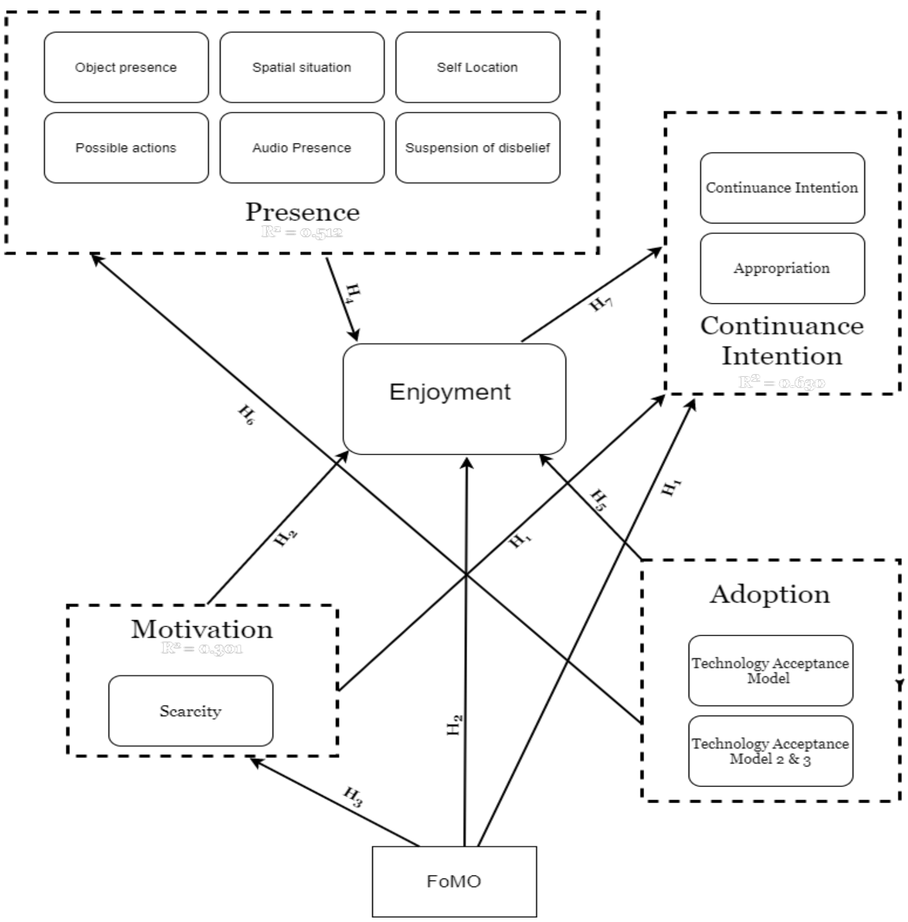


Figure 3.1: Research model

The model provides an overview of the relation between some of the core aspects from the analysis such as adoption (see sections 2.2 and subsections), presence (see section 2.4. and subsections) and motivation (see sections 2.3). These concepts are then hypothesized in relation to continuance motivation (see section 2.2.4) and the following hypotheses will indicate whether we are investigating directional or non-directional correlations between them. We therefore pose the questions that will serve as guidance for our evaluation of this thesis.

3.1 Motivational factors

FoMO appeals can be explained as directly addressing a users internal hesitancy to assent to an action and is characterized by the user's desire to stay connected with what others are doing. The concept has been applied in several contexts but is especially prevalent as a tool used in marketing and social media, and is typically manifested in LBGs as a "share" function allowing users to post images of experiences and achievements online [63]. Another trick employed by LBGs is to promote scarcity by creating time-limited events that feature unique rewards or experiences that players are likely to miss forever, unless acted upon. Thus FoMO and scarcity have shown to be effective external motivators when used in both commercial and non-commercial contexts. The following hypothesis can be made:

Hypothesis H1: FoMO and scarcity are positively related to continuance intention
Hypothesis H2: FoMO and scarcity decrease enjoyment
Hypothesis H3: General FoMO is positively related to scarcity

3.2 Presence factors

Enjoyment is the degree of which performing an activity is considered fun or pleasurable and has shown to have a positive influence on intention to play, as it leads to positive attitudes towards usage of a product. While intention to play is a positive outcome, intention to continue or play again could also likely be affected by enjoyment. The question is whether audio-visual environments can incur a degree of spatial-, object and social presence that ultimately, positively relates to the user's enjoyment. This leads to the third hypothesis:

Hypothesis H4: Presence increases enjoyment

3.3 Technology adoption factors

As mentioned in section (2.2.1) there are many factors which are to be considered when exposing users to technologies. Factors such as ease of use, perceived usefulness and perceived enjoyment have previously shown a significance in accepting technology and using it

[5][36][47][48]. In this case, we aim to investigate if there is a positive correlation between some of these factors and enjoyment for the purpose of continuance intention as well as the correlation with presence. We therefore propose these hypotheses:

Hypothesis H5: Adoption increases enjoyment
Hypothesis H6: Adoption is positively related to presence

Enjoyment has been a common denominator in this research and is tied to each of the factors: motivation [58][126][127], presence [12][74][73] however not in terms of spatial presence [80] and lastly, technology adoption which in many cases have enjoyment linked to the acceptance of a technology [5][36][47][48]. Our approach is therefore investigating the relationship between each factor and enjoyment as well as the relationship between enjoyment and continuance intention.

Hypothesis H7: Enjoyment is positively related to continuance intention

Based on the findings from the analysis as well as model structure, a research question can be formulated. Due to some of the unforeseen circumstances regarding the global COVID-19 issue, the final research question will divert from the initial research question as the relevance of designing for sound designers has been postponed for a later time. We will then focus on the experience and motivation of the user and their correlation to continuance intention and appropriation of the application.

FRQ: To what extent do factors such as presence, motivation, adoption, fear of missing out, and enjoyment contribute to the users' continuance intention and appropriation of a location-based augmented reality application?

The following section will list the requirements for the prototype in which the design and implementation process followed.

4 Requirements

Based on the analysis as well as formal conditions set up by SoOC and the nature of the project, a list of requirements was established:

Research Area	Design requirements	Section
Sound of Our Cities	• Must include a virtual map with the user's current position • Must be able to see POIs on the map corresponding to their real location, using coordinates and geolocation • Must allow users to experience POIs in AR mode • Must allow for implementation of culturally aware audio-visual content created by artists	Sections 1 2.1.3 5.1.
AR research Azuma (2017; 2018) Choi et al., 2019	• Must be location-based • Should utilize Reinforcing, Reskinning or Remembering • Should utilize AR hotspots • The placement of the environments should correspond to the POIs shown on the map, with a deviation in accuracy of max. ± 10m. • AR scenes/objects should be marker-less ○ The objects should be rendered without the need of physical markers, but rather rely on coordinates • Must be able to place objects in Roeselare, Barcelona and Copenhagen from anywhere in Europe • The application should be able to run on iOS devices, version 11 and up (MobileAR)	Sections 1 2.1.1. 2.1.2. 2.1.3.
Fear of Missing Out (FoMO) (Przybylski et al., 2013) **Scarcity** (Hodkinson, 2016)	• The AR scenes should employ non-commercial FoMO and scarcity as appeals • The content should be cyclical or time-limited to give the effect of scarcity ○ The AR environment(s) should change at least once after a certain amount of time during the test	Sections 2.3.1. 2.3.1.1.

Presence Wirth et al. (2003) Slater et al. (1997; 1998; 2009) IJsselsteijn and Riva (2003) Vorderer et al. (2004) Regenbrecht and Schubert (2002a; 2002b) McCall et al (2011)	• The virtual objects should visually convey a degree of physical tangibility ○ Virtual objects should be tetherable to the ground if desired ○ Virtual objects should be occluded by people ○ It should be possible to walk around the virtual objects	Sections 2.4. 2.4.1.
Audio Hendrix and Barfield (1995) Ozawa et. al (2002) Väljamäe et. al (2004) Lu and Smith (2009) Larsson et al (2008; 2010) Boneel et al (2010)	• The virtual sound sources should convey a degree of real-world auditory fidelity ○ Virtual sound sources should be spatialized ▪ Should allow for HRTF and Ambisonics usage ▪ Should allow for artificial reverb design that can mimic the surrounding augmented- and real space ○ Virtual sound sources should allow for head- or motion-tracking (wearables, mobile devices) ○ Virtual sound sources should be occluded by virtual objects ○ Virtual sound sources should retain an acceptable degree of quality	Sections 2.4.2. and subsections

Figure 4.1: Requirements for the application's design and implementation

5 Design and Implementation

In order to provide detailed documentation of the design and implementation of the prototype, we have chosen to document it in phases to encompass the entire process from concept to functioning prototype, including scrapped ideas and implementations. The design and implementation changed throughout the process due to changing input from key figures from the SoOC project as well as adapting to the precautions to the global pandemic that hit midway through the project and resulted in many things being put on hold.

Overview of chapter:

Phase I. Early brainstorming with SoOC and design ideas (section 5.1 and subsections)
Phase II. SoOC shutdown and adapting to COVID (sections 5.2. and subsections)
Phase III. Final implementation and design for the prototype (sections 5.3. and subsections)

5.1 Phase I.

This section will describe the first phase of the project, especially encompassing the early meetings with key persons from the SoOC project for establishing some general requirements.

Dear Hunter and map development In the beginning of the project, we were in tight collaboration with the architectural duo behind Dear Hunter. Their job was drawing and creating maps of the areas determined for the SoOC project in a carthopalogical fashion [128]. The idea was for Dear Hunter to situate themselves in the cities, in this case Roeselare, and gather information about the area, culture, people etc. to artistically create a map representing all of these. Their stay was appropriated through living in the area in a furnished office container (see figure 5.1).

Figure 5.1: Image of the Dear Hunter container on premise, in Roeselare

Our approach was to digitize the map and create a virtual version of their map to be used similarly to the map in Pokémon GO (see figure 2.9) with the isometric view of the virtual representation of the user's current location. This was implemented using the Mapbox SDK which provides location and navigation tools and services to e.g, Facebook. The reasoning for using this SDK was the compatibility with the Unity engine as well as its affordable and reasonable pricing. The SDK is free of use (with most functionalities available) until you reach 25k users of your platform. This was not something we deemed possible for the current status of our thesis. An initial map was created to conceptualize the integration of their work (see figure 5.3). We offered to create a virtual 3D model of their container and place it on a map as a demonstration. We were then provided with a drawing of each side of the container which could be applied as a texture to the 3D container. See figure 5.2. for the drawing and for the textured container model seen in Unity.

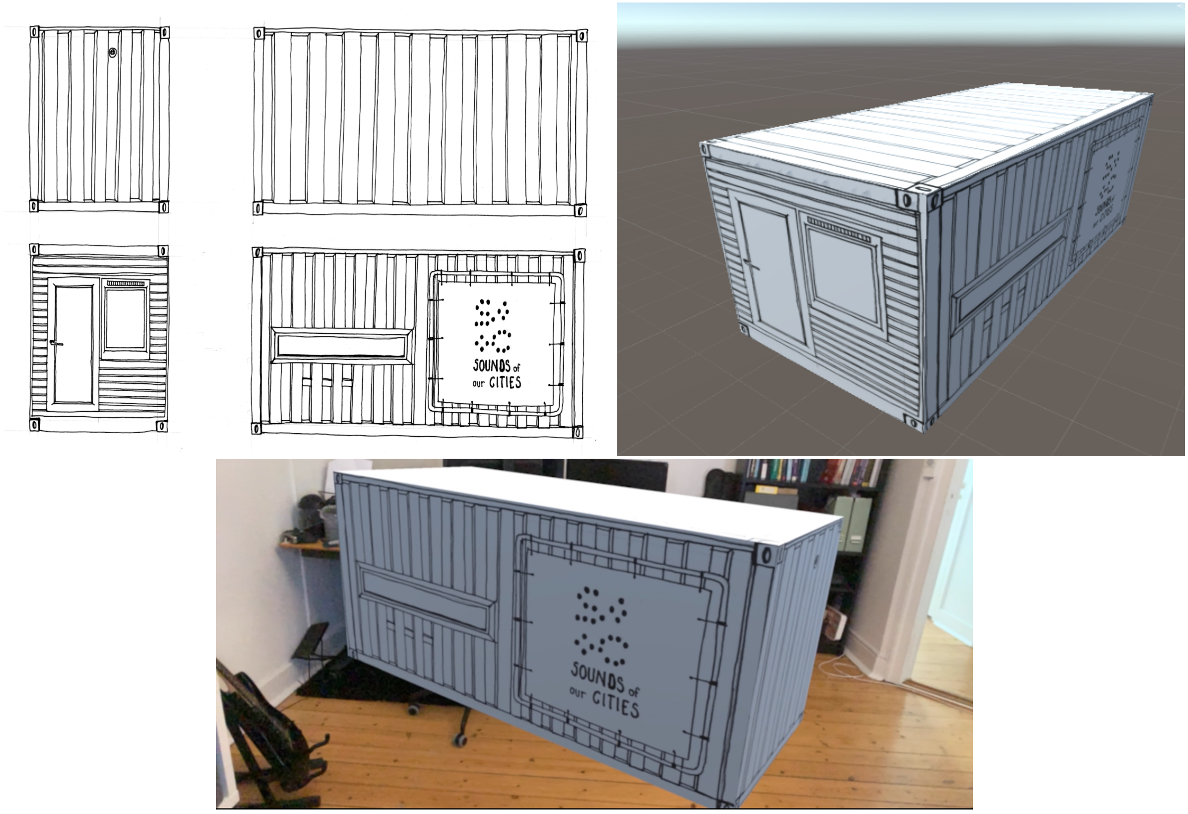


Figure 5.2: (left) Drawing by Dear Hunter of each of the container's sides.(right) a virtual container with the drawing as a texture. (bottom) The container seen in AR mode

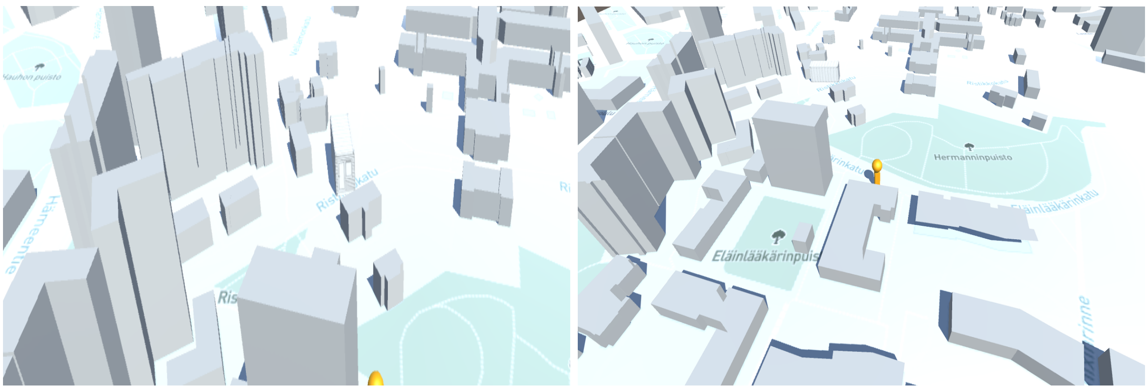


Figure 5.3: Screenshots from early map implementations. The Dear Hunter container is placed on the map among the buildings

5.2 Phase II.

This section will describe the second phase of the project, outlining some of the issues encountered during the initial months of the Covid-19 pandemic, and steps taken towards working around the situation.

5.2.1 Covid-19

In March of 2020, countries in Europe had severely inhibited or completely disallowed travel between countries, due to the pandemic. This resulted in the SoOC project being put on indefinite hiatus, with the communication and output of the different collaborative entities experiencing a severe reduction. Several events had been planned from the beginning, including a workshop that would have been held by the people affiliated with the project from Aalborg University. The workshop would have taken place in Barcelona with the recruited SoOC artists included, as a way to enable said artists to learn how to use the platform developed by AAU, as well as a method of gathering valuable data on usability and further requirements for the team at AAU to consider and implement. Because of the hiatus, the events were naturally cancelled or put on hold for the foreseeable future, which meant that any opportunity for testing the iteration at that point in time on the specific target group of users and creators (artists) would not be possible either. Additionally, testing end-users in a public space was also outside the realm of possibility, as it was clearly deemed unethical and irresponsible.

5.2.2 Alternative methods

Considering the limited options for testing a location-based AR game and the unavailability of a public space with regular social interaction during this period, a second approach was considered - namely that of simulating an augmented reality setting in a virtual reality environment, with similar affordances included. The VRChat environment was researched as a potential platform for recruitment and testing, since the platform allows users to interact with each other online through custom avatars, user created worlds and community-based events. An initial idea was to create an environment that players could join as they would see fit, with elements of the world changing cyclically, based on the time of day and day of the week. Another environment which would have been available through private invitation only was also considered, and would mainly be for the purpose of having a controlled environment for conducting tests with participants. Despite the obvious bias of not having access to any of the physical- or real properties of the AR concept, the environment itself could potentially allow for extraordinary capabilities for observation that a real environment would not. Furthermore, the virtual setting would allow for test conductors to be present but completely muted and invisible - a feature that would be hard to replicate in the real domain, as well as tap into the user's feed (camera viewport), meaning the test conductors could perform as observers as participants while retaining the anonymity of complete observers.

5.2.3 Prototyping, Limitations and Abandonment

A prototype environment was created in Unity along with the VRChat SDK. The environment featured a dynamic day/night cycle that matched the Danish timezone (GMT+1), with the idea that the scene would swap out assets, lighting, sound and change the overall "feel", depending on the (Danish) time of day.

Figure 5.4: (top left and top right) Different angles of the urban scene created in Unity. Building models were downloaded from TurboSquid. (bottom) Image of the city during night time for cyclical purposes and demonstrate the change in time.

The AR interaction was created by adding a second camera to the scene as a render-texture attached to a quad, simulating a tablet with video see-through capabilities. The camera attached to the tablet-object was then set to render an extra layer, namely the virtual objects (or conceptually speaking, the virtual-virtual objects), which were culled in the HMD-camera, i.e, the main camera.

Figure 5.5: (left) Snapshot of the tablet and showing how the environment is captured through a camera attached to the screen. (right) The researcher avatar to be used in VRChat standing next to the tablet.

Some limitations arose during the final stages of the prototype, since the VRChat SDK did not allow for developers to access the player camera due to it being a deprecated feature. This meant that the idea was ultimately abandoned as a way to get participants.

5.2.4 Revisitation

At the time of conception and prototyping (March 2020), relevant research relating to experimental design in social situations caused by similar unique events, i.e, the pandemic, was sparse, if not non-existent. This natural lack of information and precedence of the situation was arguably a major reason for abandonment, as the validity and reliability of methods was essentially undocumented, likely requiring much more rigorous development and testing of the applied methods than what had been imagined and prototyped up to that point. However, in late April, Saffo et al. [129] published a paper in which they tested the feasibility of using VRChat as a way of leveraging the large community by means of crowdsourcing. They similarly used Unity as the engine for developing the scene along with several SDKs, including the VRChat kit used for the AR/VR prototype, and were inhibited by similar limitations such as not being able to record the perspective of the participant or another fixed camera from within the test-scenario, but were forced to record the screen of the researcher running the experiment instead. They also realized that extracting and logging data through custom scripts such as time spent (checkpoints) or positional information (heatmap and HMD movement) would not be doable, since the only custom scripts allowed by VRChat were shaders. They even highlight the narrow scope of what can even be implemented for testing, due to the SDK relying on a basic trigger system (click, intersect, collide and timeout), ruling out stimuli with complex interactions or dynamic functionality, but rather only allowing static stimuli (ibid). Conversely, VRChat developers have, as of writing this, released an open-alpha SDK named Udon, which is a programming language built in-house by the VRChat developer team that works with Unity, allowing for variable-syncing between players and creation of custom behaviour, so the concept of collaborative or solo AR in VR experiences and experiments could potentially be carried out at a future stage, if required. As for reliability and validity of the data acquired, the research by Saffo et al. [129] did not include any actual conducted experiments and data-comparison, but could be considered exploratory in their methodological approach, as their focus was to document and reflect on the process of using a virtual, social platform as a way to crowdsource participants.

5.3 Phase III.

After a cumbersome design and development process due to COVID-19, this phase was the last and final, which was where we created the final prototype. The following sections will elaborate on the design and implementation.

5.3.1 Application structure

5.3.1.1 Engine

The Unity platform was used as the engine for the project as both Android and iOS were supported. Version 2019.3.0f6 was used throughout the entire implementation process as it supported many of the SDKs chosen for the project development.

5.3.1.2 Software Development Kits (SDKs)

The most central part of the application was the integration of augmented reality, but also the use of a map as part of localizing the virtual scenes. This was due to both testing purposes as well as from the early meetings with the Dear Hunter, which focused on the incorporation of map features. AR Foundation, along with ARKit (iOS) and ARCore (Android) were initially installed for cross-platform AR support, but ARCore was ultimately not utilized as some functionalities from ARKit were not supported and we had iOS devices, which then determined the platform. Beta-stage preview packages for AR Foundation and ARKit (version 3.1.0) were installed in lieu of the stable 3.0 releases, which supported people occlusion with an AR Raycast Manager script attached to the AR Session Origin (0,0,0), and an AR Occlusion Manager script attached to the AR Camera (Viewport: MainCamera). This meant that we were able to create 3D objects seen through the camera and that the camera could detect humanoids passing in front of the objects hence adding to the objects' more three dimensional appearance. Preview packages used AR Pose drivers, which were not compatible with people occlusion, as the pitch, yaw and roll would not track properly, so Tracked Pose Drivers from a legacy version were imported through the XR Legacy Helper package, attached to the camera and re-referenced in the script, so the camera could take the proper input from the device's camera.

For the map integration, we used the Mapbox Unity SDK. Mapbox is a cloud platform for real time mapping and location-based services [130]. The SDK works as building blocks for making software using their cloud services. We utilized the Location-Based Game prefab, which contains the most basic properties for an isometric map appearance and functionalities such as Points of Interests (POIs). Interaction with the map, such as drag map, zoom and re-centre functions were implemented by us for the specific use for the application.

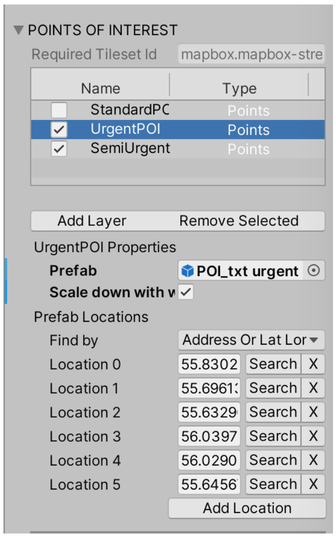


Figure 5.6: List of world coordinates for placement of some of the POIs.

In order to be able to place virtual objects to real-life locations, we utilized an SDK by Daniel Fortes called AR+GPS purchased from the Unity Asset Store [131]. At the time of implementation, placing AR objects using GPS was not widely accessible, especially free services and was only recently added to standardized AR SDKs such as ARKit and ARCore. The AR+GPS SDK provides easy prefab placement using geographical world coordinates.

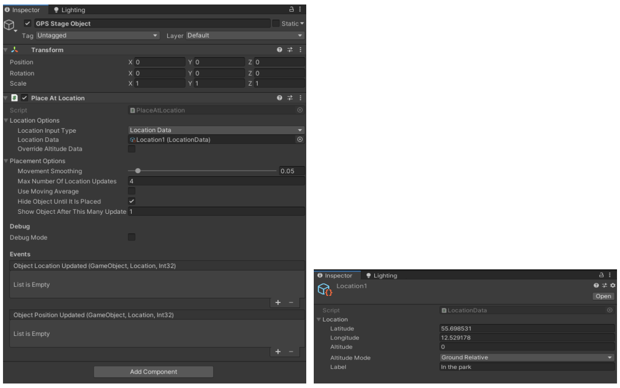


Figure 5.7: (left) Screenshot from Unity. Variable with location data that can be used to place objects. (right) Screenshot of component script that contains the location. Has to be put on the game object that will spawn on the location provided.

One of the SDK's key features is the implementation of hotspots, which is a distance-based detection for prompting the virtual objects when approaching the world location rather than having them visible at all times despite being far away from the actual location. However, this feature proved to be inconsistent and was scrapped. We then decided to create our own 'hotspot' script that worked in a similar fashion where it continuously detected the device's camera position in relation to the virtual object or scene's center. The distance was easily modifiable in order to accommodate the different urban location sizes. Code snippets can be seen in appendix A7.

Another SDK downloaded from the Unity Asset Store was the Native Gallery SDK by Süleyman Yasir Kula [132]. This SDK allowed us to save screenshots taken from the application to the device's gallery. The SDK contains several scripts that work natively to either iOS and Android. The original idea was likewise to implement dynamic lights and shadows corresponding to the time of day. We started out using a Unity SDK by Mark Hessburg [133], which takes the device's location and changes the lighting (angle of the sun, color etc.) dynamically in real time. This means also changing the shadows of the virtual objects. This SDK was however scrapped due to some inconsistencies with the shadows used in an AR setting.

5.3.1.3 Occlusion (visual)

With the release of the ARKit 3 preview package described in section 5.3.1.3. on SDKs, proprietary software from Unity in collaboration with Apple was made available for developers. This introduced the "Human Body Subsystem" allowing for motion capture and pose estimation, but more importantly the ability to recognize human bodies and provide apps with human stencil and depth segmentation images. The stencil segmentation identifies whether a pixel contains a person and the depth segmentation estimates the distances from each pixel that is recognized as a human - combined, these grant the option to occlude virtual elements with any real people present in a scene. A major limitation that factored into the development was that only iOS devices with the A12 Bionic chip and ANE (Apple Neural Engine) supported these features [134].

Figure 5.8: Early occlusion test-scenario using surface detection and a calibrated environment (transparent layer).

Further testing was performed with emphasis on textured virtual objects, lighting, geoprecision and audiovisual occlusion. A scene featuring an audio source and a textured wall was placed with geolocation in an urban courtyard during the daytime, by entering the coordinates given by manually clicking the location on Google Maps. The scene was located a few meters off center and did seem to drift slightly from its point of origin when moving around the objects, which was less of an issue when testing less enclosed environments (see appendices A3 for daytime build in a semi-open environment).

Future tests with the scene and the same device (iPad Mini) showed that getting the location from the device increased precision of the desired target origin (center of courtyard), suggesting locations that are indoors or in a more compact space could benefit from physically probing the location for coordinates with a similar device. This was however only using a single device in one location; further testing would be required for any additional assumptions on this matter. The application was restarted several times to infer the degree of deviation from the desired point of origin and empirically showed to deviate approximately +-3 meters from the two furthest spawn locations when opening the application; A single spawn point ended up in the wall of the nearest building, resulting in depth being displayed incorrectly when moving closer to the object.

Figure 5.9: Audiosource (capsule) and virtual object (wooden wall) occluding a person.

Figure 5.10: Person walking "around" objects and into the foreground. Notice the difference in stenciling, contrast and blurriness of the person vs. the wall between both images - Virtual objects retain clarity while the person is blurred when moving at close proximity.

A larger scene was placed at a grocery store parking lot, once again using Google Maps to get the coordinates of an approximated location, with the main purpose being to test moving sound sources, particle effects, relative distance between objects and mapping, and also interaction with larger objects. The scene consisted of several elements, including a moving particle system with a sound source attached that had been programmed to travel between patrol points, in a circle around the parameter. Another object in the scene was the virtual SoOC container with an occluded sound source (the sound of drawing and scribbling on paper) placed within the geometry. A few feet away, a tree and a floating, glowing sci-fi cube had been placed, with a voice reading an excerpt from Frank Herbert's "Dune". The last object was that of a relatively large labyrinth, which a player (or even observer, with guidance) could traverse. The labyrinth responded well to changes in perspective and was stable enough for a user to guide another person through, despite the latter's inability to visually perceive the object or its walls. The occlusion was also good for determining depth and location within the labyrinth.

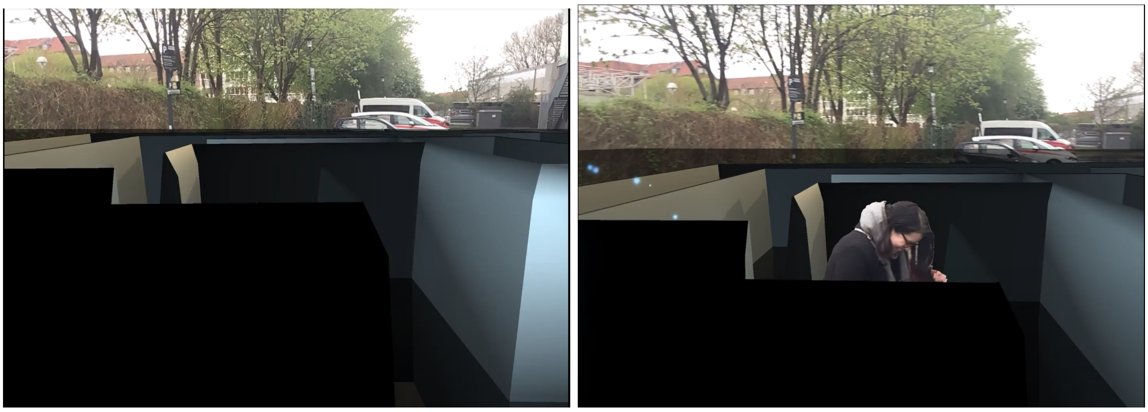


Figure 5.11: (Left) Labyrinth placed outside of a grocery store with a person crouched behind the wall, fully occluded. (Right) The person stands up and is only occluded by the front wall.

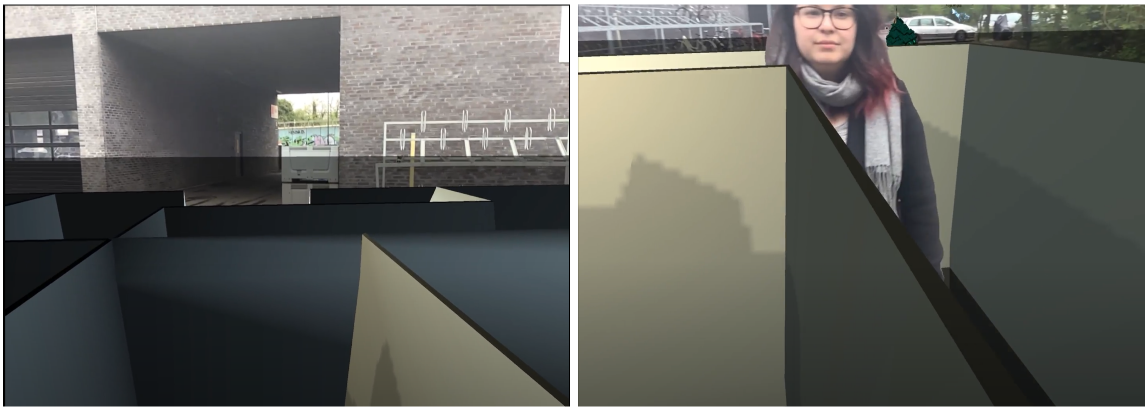


Figure 5.12: (Left) User's perspective from walking around inside the labyrinth. (Right) Person situated between two walls of the labyrinth - a single virtual object - with partial occlusion present.

5.3.1.4 User Interface (UI)

Since both components of the application, map and AR view, two different UIs had to be implemented leveraging the different functionalities of both. The map view had two buttons located in each bottom corner. The button in the left bottom corner was for re-centering the camera view back to the device's current location on the map. The current location was illustrated by a blue capsule seen moving inside the map relative to the device's location in real time. The button in the right bottom corner controlled the ability to switch between views, i.e, map- and AR view. This was done by changing the active camera from the one seeing the map and the camera with the AR properties and settings. Each respective camera was only scripted to see things of interest. The AR view had a similar UI to the map view. The button in the right corner had the same functionality as the right bottom corner button as in the map view, however switching to map view rather than AR view. Icons were self-made, however based on common iconography found from other applications.

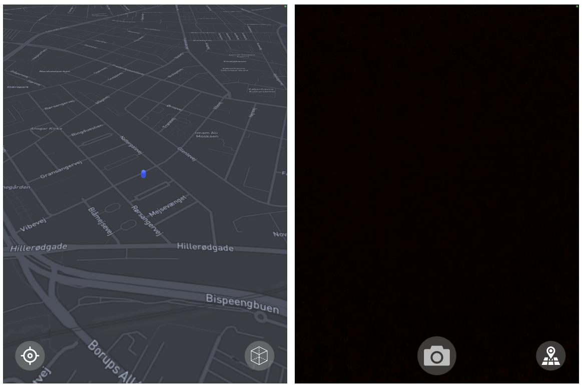


Figure 5.13: (Left)Screenshot of the map seen by the user. (right) Screenshot of the AR-mode with UI. The button in the bottom right corner switches from AR mode to map mode

A screenshot button was also implemented (bottom centre), where the user would be able to take pictures of what was seen on the screen when in the AR view. Each screenshot was saved to the device's camera roll and was designed to give the user the option to capture their experience. The photos taken by the participants did not include the UI.

5.3.1.5 Cyclicity

As mentioned in section 3, our approach to the sense of scarcity and appeal to FoMO was to implement virtual scenes that were time-limited and could only be experienced within a certain time. This cyclicity functionality was implemented so that two scenes were active at once, both having timers counting down, and after one of the timers would run out, the scene would change to something else and thereby give the possibility to miss out on it. All the scripts for this functionality were self-developed and snippets from the scripts can be seen in appendix A6.

5.3.2 Audio implementation

5.3.2.1 Resonance Audio

As mentioned in section 2.4.2.2. on spaciousness parameters, the Resonance Audio SDK was used for implementing spatial sound, which enabled sound source directivity customization, Near-field effects, sound source spread, geometry-based reverb and auditory occlusion efficiently enough to be scaled for mobile.

5.3.2.2 Ambisonics

Resonance Audio projects all sound sources into a global high-order Ambisonic soundfield, allowing for HRTFs to be applied a single time to the soundfield, instead of applying it to each single source within the field [104][135] The binaural reproduction of the Ambisonic sound fields is done by creating "virtual loudspeakers", where signals are generated by decoding the soundfield and then convolved with a Head Related Impulse Response (HRIR) corresponding to a loudspeaker location. The left and right channel signals are summed individually, to create the left and right output, respectively [135]. Furthermore, the order of the Ambisonic is adjustable, controlling spatial resolution and fidelity, as well as improved direct source localization. The cost of this optimization reduces the CPU load per sound source, allowing for playback of more sources simultaneously, making it a sturdy choice for mobile AR. The HRTF database used for the HRIRs are provided by SADIE Binaural Measurements, thus they are generic and not personalized.

5.3.2.3 Reverb Probes

A reverb probe is a term for a location where the reverb properties are sampled, through "baking", a term for processing and storing a precomputation of the properties for smooth transitioning between simulated environments [103]; The properties are baked using ray tracing to simulate the interaction of sound waves in the environment. The same technique can be used in runtime to include visibility of a probe as a parameter, e.g if placed around a corner or if several probes are in close proximity of one another, e.g, in a maze with different geometry, space and acoustic-material.

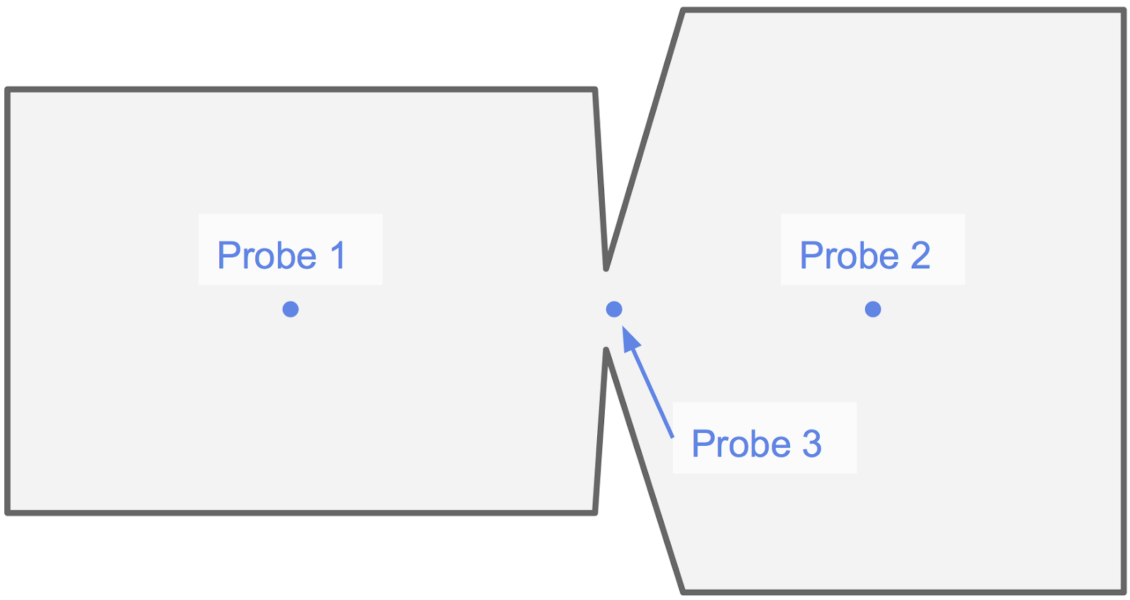


Figure 5.14: Visualization of a 3 probe setup, creating a smooth transition from one space to an adjacent space (e.g, outdoor to indoor, or two different rooms)

5.3.2.4 Acoustic Materials

Relevant to the probes is the ability to map virtual textures to acoustic materials, enhancing the ability to further increase fidelity, by making e.g, a marble texture have similar absorption and reflective capabilities to that of its real life representation. The textures of the assets used within the AR scenes were mapped to acoustic materials that would elicit a similar response in the real world; This was done by creating a material map and assigning acoustic properties to each material index in the array. After baking, the probes could be inspected for the RT60 values of different frequency bands (see section 2.4.2.2. On Sabine's formula).

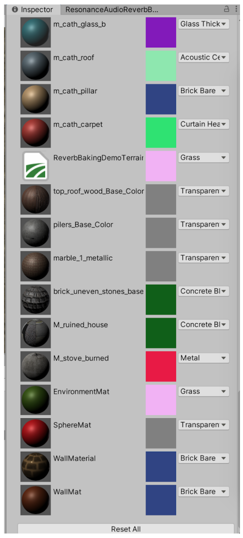


Figure 5.15: Screenshot of material map from the Unity scene

A virtual scene was built with different geometric configurations and acoustic materials, of which there were four distinct setups, i.e,:

- An alley-type configuration
- A ruined house with no roof or windows
- A spacious cube with highly reflective materials on each surface and protruding geometry
- Four walls overlapping, next to an audio source

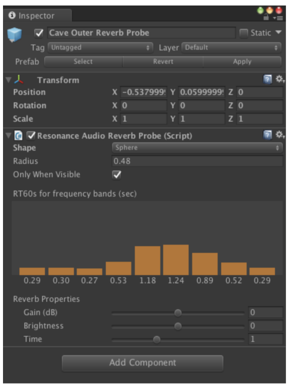


Figure 5.16: Screenshot from the Unity Engine

The first three setups (alley, house, cube) were used for testing reverb, reverb probes and absorption coefficients with textures and geometry, simply by moving around in the scene, baking audio and looking at the different RT60 values for the different probes. See appendices section A2 for screenshots of the three different setups.

5.3.2.5 Occlusion (auditory)

The fourth setup (four walls) was used for testing occlusion. A sound source was placed in close proximity to four overlapping walls that would occlude the sound to a higher degree for each added wall between the raycast attached to the camera view and the sound source, resulting in the most occlusion at the point where the back wall overlaps with the three walls closest to the audio source (towards the middle). From the Resonance Audio Github, the occlusion filter is presented as a monopole IIR low-pass filter with the difference equation:

$$y(n) = \alpha * (y(n-1) - x(n)) + x(n) \tag{5.1}$$

where is a coefficient that satisfies 0< α <1, $y(n)$ is the output at sample n, and $x(n)$ is the input at sample n [104].

5.4 Scenes

Three scenes were designed based on creating different types of tangibility as part of the requirements for the application (see section 4). Since there was a lot of focus on visual- and sound occlusion as well as emphasis on spatial audio (see section 2.4.2), it was important that each scene would provide different interactions in which to explore and experience. As mentioned in section (5.3.1.2), the implementation of dynamic lighting in the scenes (as well as shadows) was scrapped due to inconsistent appearance. Normal directional lights were then added to each individual scene to light up the objects and add slight shadows to emulate real physical objects as much as possible. The first scene was designed based on a relevant pop culture reference, the game 'Among Us', which is a multiplayer, social deduction game. The reasoning behind including this was to have a scene where an audio source (the red character), had the ability to walk around the participant utilizing spatial audio cues even when standing still. Audio from the actual game was used, including short leitmotifs and iconic sound effects, as well as the set of footsteps used ingame when a character walks around outside in grass and dirt. A custom walking cycle animation was created, where the character would stop at different locations and look around before resuming its path, and a script was created to relate the footsteps (which were randomized in pitch, within a range, to further differentiate each individual step), resulting in footsteps occurring only when the character was in motion.

The second scene was designed with simple shapes (cubes and cylinders) with varying colors. Initially, this scene was meant to be used during night time, hence the bright colors. We created a script that also changed the colors of the shapes every second in real time. The scene was a large and tall scene in order to ensure participants would be able to walk through or under the geometry and experience the visual- and auditory occlusion. The scene had an audiosource playing "Confusion" by "Pump Panel", i.e, the techno theme from the initial scene featured in the movie Blade (1998). The audiosource was located in between the two largest upright pillars and underneath the overarching cylinder, and featured reverb probes that would introduce reflections to the spatialized audio when the user stood inside the geometry. Additionally, occlusion was added, adding low-pass filtering if the ray-tracing could not reach the obstructed audiosource, which was not visible in the scene.

The third scene employed a generous amount of individual assets (planets) and inhabited a large physical area. The scene was designed as a Solar system consisting of the planets found in the system Sol. The planets were programmed to revolve around the sun as well as have individual rotations. The idea with the scene was to have multiple moving objects with different sounds, similar to the first scene, although with multiple sound sources. The audio used for the planets were audio recordings, or "space music", captured from each real planet in space and which were then added as individual audiosources for their respective virtual counterparts, with different attenuation depending on the planet's size. A "2D" (non-spatialized) audiosource was added for the prototype, which was an ambient soundscape featuring a cosmic aesthetic, but was later spatialized and placed inside the sun (without occlusion) as the 2D ambience clashed with the spatialized planets.

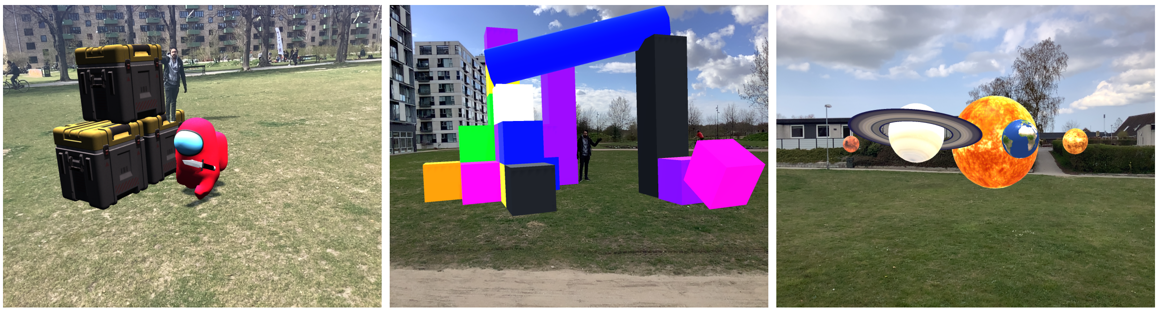


Figure 5.17: (left) First scene of a character from the video game 'Among Us'[136]). (Right) Scene two. (Bottom) Scene three of the solar system

5.5 Final look of the application

The following images are screenshots of the final application. These images represent the two 'modes' the user can see; map and AR-mode.

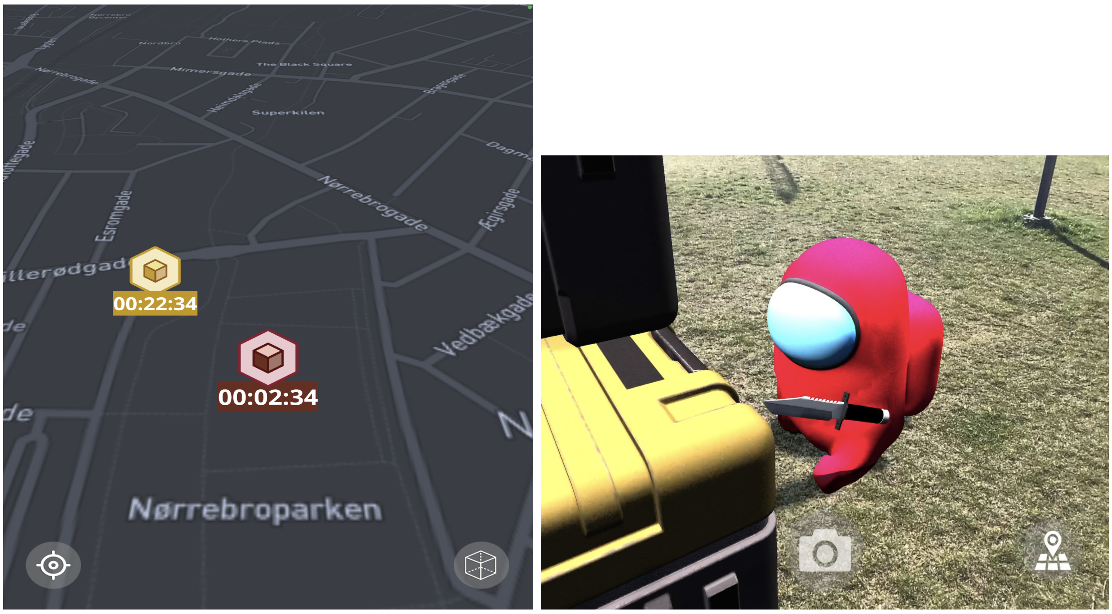


Figure 5.18: Images taken from the application. (left) Screenshot of the map with the two POIs. (right) Screenshot from the app's appearance in AR-mode (with UI).

6 User Study

The basis of the analysis provided a broad understanding of user behaviour and motivation when using technology as well as correlating it with what the SoOC project aims to develop for the urban environments of Roeselare and Barcelona. As mentioned in section 2.5.1, performing user studies during a global pandemic can be challenging, however not impossible. This section will cover the methodologies used for this thesis during COVID-19. We aimed to investigate the relationship between continuance intention, presence, motivation, FoMO and adoption. It was hypothesized that participants scoring high in motivation, presence and adoption would then have higher continuance intention. A more longitudinal approach to measuring these aspects was much preferred, however, due to the limited resources in terms of equipment and participants (social distancing), a more ethical, yet effective approach was chosen.

6.1 Participants

The sampling method used for gathering participants was convenience sampling. Convenience sampling is a non-probability sampling method which lends itself to easy and convenient sampling as participants gathered are the ones approximate and available [137][138]. Due to the restrictions during the COVID-19 pandemic in terms of limiting socialization as much as possible, the participants were recruited through personal relationships and acquaintances.

6.2 Equipment

The virtual content used for the study was developed in Unity using the Resonance Audio SDK (see section 5.3.1.2) Two iPads were used for the test: an iPad mini (5th generation) and an iPad 9.7” 5th generation, both using identical Apple Earpods for audio output. We chose to eliminate the possibility for participants to use their own earbuds due to many of them being wireless and (dis)connecting them several times a day was deemed unnecessary. Furthermore, using the Apple Earpods meant that the auditory transparency and the basic audio quality would be consistent throughout testing, as no ANC, sidetone or complete ear-canal occlusion would be present. Also, some newer apple airpods feature head-tracking, which was yet another point of bias that could be introduced through the earbuds alone.

6.3 Setup and Location

The participants were exposed to the same virtual content in a natural environment, at a predefined location, with at least one test-conductor present. By using a natural environment, it was possible to see the prototype in use as it would be in 'real life' rather than an artificial environment such as a lab [137]. The device was user-controlled and the headphones were placed in each ear at the start of the test. The experiment took place in Copenhagen, Helsingør and Gevninge and all were situated and conducted outdoors, in several urban areas with neighbouring structures such as buildings and roads etc.

6.4 Procedure

Each experience was initiated with a brief explanation of the interface and various functions of the application (i.e, snapshot, re-centre, isometric map, POIs), and were informed that any pictures they took during the test would be sent to them, if they wanted to have them.

Two POIs were available to the participant, placed approximately equidistant from the initial location of the user and were denoted as such in the isometric map-mode. The first POI utilized the first scene (Among Us character - see figure 5.17. in section 5.4.) had a 3-minute timer counting down to another prefab spawning in its place. After 3 minutes, this POI changed into the second scene (Solar system - see figure 5.17. in section 5.4.) which had a 2 day (48 hour) timer attached to it, for the rest of the session. The other POI (Cubes and cylinders - see figure. 5.17. in section 5.4.) had a 20 minute timer. However, despite the POI counting down, the scene did not change during the session, but solely represented the idea of it changing like the first POI. The timers would not be paused at any point during the test (even during the introduction to the interface), with the exception of encountering any major technical errors or crashes that would result in the conductors having to reset the application. The walking distance from the initial position to either point was estimated to take no more than 30 seconds to reach. Furthermore, a soft limit of approximately 15 minutes was set for the entire experience and the participant was informed of this, as well as the fact that they could stop at any given time for any reason, before the 15 minutes had passed. In terms of agency, the participant was allowed to visit each of the three scenes (2 POIs) within the preallocated 15 minute time-limit of the test itself, but would only be able to experience the time-limited event during the initial 3 minutes, meaning that potentially only 2 scenes out of 3 could be available, depending on the participant's initial decision to visit POI 1 or 2 first. Participants were not directed to any POI or asked to stay at one for a specific duration, nor were they asked whether they wanted to see the POI not initially visited. They were however informed, during the tutorial, that they could go to whichever POI they wanted to, during the experience.

For the scenes themselves, all three of them had at least one auditorily-visually occluded virtual object in the environment, with some type of navigation and movement required (i.e, ducking, sidestepping/walking, or clipping) to find, fully visualize and hear the unoccluded audiosources in the scene. No specific tasks were given or overtly explained in this regard.

6.5 Data Collection

Self-reported measures, through the use of an 84-item Likert type questionnaire, were collected. The questionnaire was designed to consist of individual constructs that were geared towards different areas, i.e, motivation, continuance intention and appropriation, presence, enjoyment, adoption, and FoMO, and was constructed using established questionnaires as well as self-made items (see taxonomy below and appendix A5 for list of items). The data gathered from these constructs were used to inspect possible relationships and effects on enjoyment and continuance intention. The questionnaire relating to Presence consists of 3 process factors from the "Measurements, Effects, Conditions - Spatial Presence Questionnaire" (MEC-SPQ) developed by Vorderer et al. [81] . 8-, 6- and 4-item scale setups with different weighting (1-3) have been designed, enabling the test conductor to reduce potential respondent fatigue, depending on the amount of process factors needed. The items of the following questionnaires have been placed in the order of highest to lowest weighting (mandatory to optional); Only the scale setups that will be used for the evaluation will be presented, and unused optional items have been omitted. The questionnaire by Regenbrecht and Schubert [82] on object presence has been utilized, but due to the format of the questions being binary, a reinterpretation to fit the Likert format has been proposed.

6.6 Taxonomy

The table below shows the different constructs of the questionnaire, and a short description of these, as well as the original sources from which they're adapted.

* denotes that the items in the construct are unaltered from their original state and are applied directly from its source.

Construct	Description	Items	Adapted from
General Questions	Questions related to demographics and generalized FoMO attributes	GQ: 1-14	Przybylski et al. (2013)*
Spatial Presence: Self Location	The user's sense of being present and situated within the virtual environment	SPSL: 1-6	Vorderer et al. (2004)*
Spatial Situation Model	The construction of a mental model of the (mediated) situation that includes space-related information.	SSM: 1-6	
Spatial Presence: Possible Actions	User's mental representation of their bodies and possible actions within the environment	SPPA: 1-6	
Suspension of Disbelief	An act of volition in which the individual actively pursues the goal of not being distracted from the current experience	SoD: 1-6	
Object Presence	The experienced realness of the virtual objects, and the spatial presence of these objects	OP: 1-6	Regenbrecht and Schubert (2002b)
Auditory Presence	The experienced presence as a function of specified auditory parameters	AP: 1-6	Witmer and Singer (1998); Ozawa et al. (2003); Regenbrecht and Schubert (2002b)
Adoption: TAM2&3	Questions relating to the user's attitude towards use of the application in a more social capacity as well as self-efficacy	MTAM2+3: 1-5	Davis and Venkatesh (2000); Venkatesh (2000); Venkatesh and Bala (2008)
Adoption: TAM1	Perceived ease of use and perceived usefulness of the features and functionalities of the application.	MTAM1: 1-10	Davis (1985); Davis, Bagozzi, Warshaw (1989); Shroff et al.,(2011); Allomary and Woollard (2015)
Motivation: Scarcity	The perception and motivation of general usage, of the application, when considering scarcity appeals	MS: 1-8	Hodkinson (2016)
Enjoyment	Perceived enjoyment and engagement of the audiovisual experience, interaction and general usage of the application	ENJT: 1-8	Ghazali, Mutum and Woon (2017)
Continuance Intention and Appropriation	The intention to interact with the experience or continue using the application, as well as future prospects	CIA: 1-7	Carroll et al. (2001); Carroll et al. (2002); Ghazali, Mutum and Woon (2017) Haugstvedt and Krogstie (2013)

Figure 6.1: Taxonomy for questionnaire.

7 Results

Age group	Total (%)
13-17	1 (3.8)
18-24	1 (3.8)
25-34	16 (61.5)
35-44	2 (7.7)
45-54	2 (7.7)
55-64	4 (15.4)
Total (%)	**26 (100)**

Figure 7.1: Age distribution

A total of 26 people (males = 18, females = 8) participated in the test in three different locations . Figure 7.1 shows a table of the participants' age distribution. Before we were able to create the complex model to answer the different hypotheses, we had to define the relationship between the different factors making up the construct of presence through its own model (7.2). For this study, we decided to process our data using Partial Least Squares (PLS), which is a regression algorithm using weight vectors. To calculate said data, we chose to use the SmartPLS-SEM (partial least squares structural equation modelling) software [139], with which we were able to determine both the convergent- and discriminant validity of the individual items for each construct that was observed. Convergent validity relates to the degree of which items in a scale measure the same construct [63] and the loadings of the items, along with composite reliabilities (CR) and average variance extracted (AVE), which were utilized in this regard [63]. Discriminant validity refers to the degree of which a scale measuring a single construct measures other related constructs as well, with the ideal measurements showing little or no relation between one another. To determine discriminant validity, we utilized the Heterotrait-Monotrait Ratio (HTMT) matrix as recommended by the SmartPLS documentation, as older calculations such as the Fornell-Larcker criterion [140] is deemed unreliable and ineffective in some circumstances [139][141]. Discriminant validity is meant for ensuring that a construct has the strongest relationship with its items (indicators). By creating a table with an overview, it can be determined if there is some overlap in the relationship between latent variables and indicators. HTMT values under 0.9 are considered adequate for discriminant validity [139]. The following sections will refer back to these essential assessments and adherence to the thresholds. Despite removing items, the values could not get any lower than this.The lower the value, the higher discriminant validity [141].

Discriminant Validity

Fornell-Larcker Criterion | Cross Loadings | Heterotrait-Monotrait Ratio (HTMT) | Heterotrait-Monotrait Ratio (HTMT)

	Adoption	Continuance i...	Enjoyment	FoMo	Motivation	Presence
Adoption						
Continuance in...	0.880					
Enjoyment	0.994	0.875				
FoMo	0.986	0.534	0.658			
Motivation	0.478	0.292	0.219	0.637		
Presence	0.874	0.867	0.888	0.689	0.406	

Figure 7.2: Table with HTMT values.

7.1 Defining the presence relationship

As the presence category encompasses a bigger scope, taking multiple studies and questionnaires into account, a parsimonious model [142] was created to establish the most salient aspects of presence in order to congregate the scores to be used as a construct in the final model. The parsimonious model helped identify which items from the questionnaire yielded insignificant factor loadings or suffered from multicollinearity issues, resulting in the removal of these items from the presence construct included in the final model. Multicollinearity refers to the case when an independent variable is correlated to another independent variable in a regression model. However, this is not seen as a positive thing, as independent variables should be independent. If independent variables are highly correlated with one another, it can yield a large standard error [143][144][145].

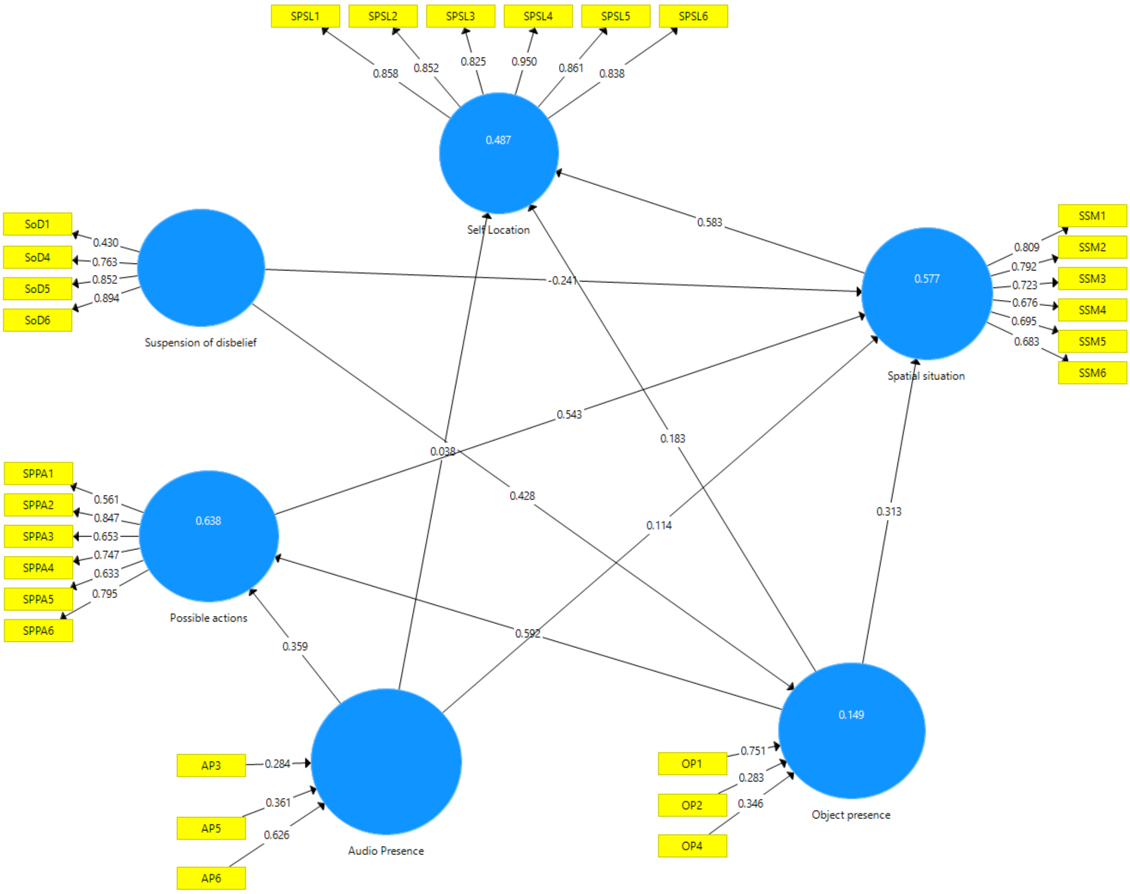


Figure 7.3: Presence model with all the subcategories and their respective items.

After running the PLS algorithm, certain weights were given to each indicator from the latent variable (in this case: object presence, audio presence, spatial situation, self location, suspension of disbelief and possible actions), which should reach the weight of >0.7. Some of the items from each latent variable had to be removed due to either negative or low weight. A few indicators with low weights were kept due to the construct's integrity [146]. After removing the items that caused some issues, an aggregate score was made for each participant for each latent variable, so that they could represent the indicators for presence in the complex model.

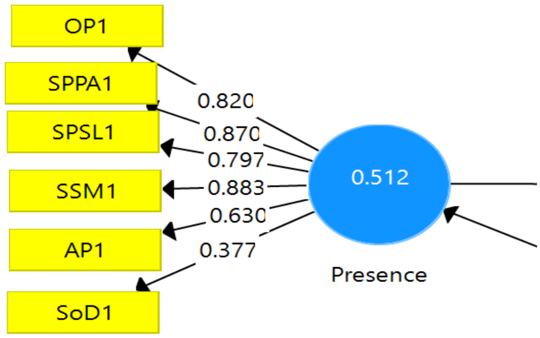


Figure 7.4: The indicators' aggregates used in the complex model.

7.2 The Complex Model

To measure the reliability of the complex model, Chronbach's alpha, AVE and CR were calculated for each construct within. The Chronbach's alpha values for each construct met the threshold of 0.7, ranging from 0.726 to 0.906. However, AVEs for three of the six constructs, namely Adoption (0.427), Enjoyment (0.469) and FoMO (0.457), reported scores below the threshold of 0.5 specified by Fornell and Larcker [140], albeit with the added notion that AVE-scores may be considered more conservative estimates of the validity of the measurement model [140][147], and that the composite reliability with a value of >0.6 are enough to reach an acceptable level of convergent reliability [140][147]. CRs for all six constructs were above the 0.6 threshold, ranging from 0.814 to 0.920, reaching acceptable levels of convergent validity [147]. See appendix A5 for survey loadings, CR and AVE values. In terms of discriminant validity, we utilized the Heterotrait-Monotrait Ratio (HTMT) as measurement. As mentioned in section 7, the threshold for adequate discriminant validity is <0.9. For our data, the values range from 0.219-0.994, which means there are two constructs that go over the acceptable threshold.

After running the PLS algorithm, we bootstrapped the data to find the statistical significance of the results (Cronbach's alpha, R^2, HTMT etc.). The PLS-SEM does not assume the data to be normally distributed, and requires a non-parametric procedure such as bootstrapping to find the significance of path coefficients [139]. When bootstrapping, subsamples are generated based on samples drawn from the original observations. The most typical number of subsamples created are 5000, which we adhered to for our calculations. These subsamples are then used to estimate the PLS model [139].

7.3 Model Fit

The primary objective of PLS is to demonstrate the significance of an alternative hypothesis, allowing for a potential rejection of the null hypothesis through significant t-values and a high R^2, with the sample data provided [148]. Results showed that the constructs of Continuance Intention and Appropriation (t = 7.302, R^2 = 0.631), Enjoyment (t = 9.939, R^2 = 0.764), and Presence (t = 4.926, R^2 = 0.514) were all statistically significant with $p<0.001$, but with the construct of Motivation (t = 1.640, R^2 = 0.281) being non-significant with p = 0.051.

f^2 scores showed that five of the moderation effects displayed a large effect size ($f^2 >$ 0.350), with the largest effects measured between the effect Enjoyment had on Continuance Intention and Appropriation (f^2 = 1.022), and between the effect Adoption had on Presence (f^2 = 1.058). Two effects showed a medium effect size (0.150 < f^2 < 0.350), namely between the effect FoMo had on Motivation (f^2 = 0.390) and Motivation on Enjoyment (f^2 = 0.170). One effect showed a small effect size (0.020 < f^2 < 0.150), specifically between the effect FoMo had on Enjoyment (f^2 = 0.025). The remaining two effects were shown to be non-significant, that is between FoMo and Continuance Intention and Appropriation (f^2 = 0.014) and between Motivation and Continuance Intention and Appropriation (f^2 = 0.004). Effect size thresholds for significance were established by Benitez et. al. [142].

f^2	Adoption	Continuance Intention	Enjoyment	FoMO	Motivation	Presence
Adoption			***0.461***			***1.058***
Continuance Intention						
Enjoyment		***1.022***				
FoMO		0.014	0.025		***0.390***	
Motivation		0.004	*0.170*			
Presence			***0.433***			

Figure 7.5: The indicators' aggregates used in the complex model.

The research model was presented in section 3 (figure 3.1), as well as accompanying hypotheses that were the foundation for this study and its research question.

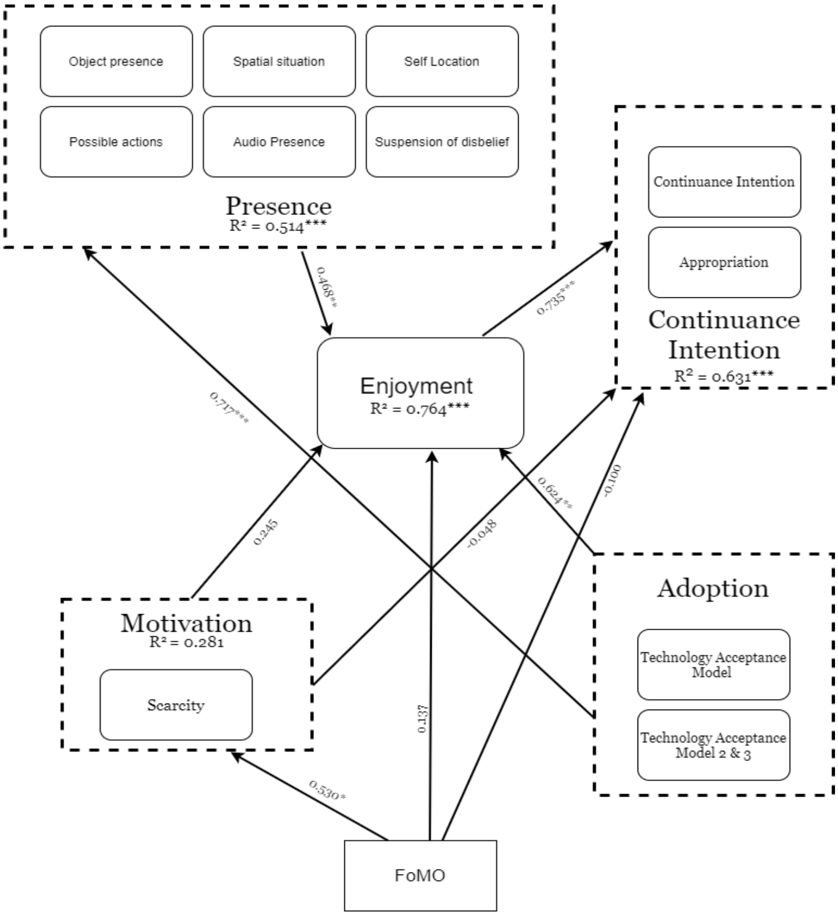


Figure 7.6: Model with path coefficients and R^2 values

Based on the model above as well as the values retrieved from the bootstrapping, the data shows that:

H1. The first hypothesis focused on the participants' general FoMO and the scarcity provided by the application with the intention of continuous use. This hypothesis does not statistically show any relation between either FoMO (M = -0.126, SD = 0.187, p = 0.297) or scarcity (M= -0.001, SD = 0.171, p = 0.390) with continuance intention.

H2. The second hypothesis checks for the participants' general FoMO and scarcity from the application and if these decrease enjoyment. This hypothesis does not show a negative relationship between both FoMO (M = 0.101, SD = 0.244 , p = 0.288) and Scarcity (M = 0.214, SD = 0.166, p = 0.069) with Enjoyment.

H3. The third hypothesis checks for FoMO scores and their possible positive relation to scarcity. General FoMO is positively related to scarcity (M = 0.527, SD = 0.222, p = 0.009).

This hypothesis shows a significant positive relationship between FoMO and scarcity. The relationship between FoMO and scarcity explain 53% of the variance.

H4. The fourth hypothesis focused on the sense of presence and its effect on enjoyment. Presence was found to increase enjoyment (M = 0.400, SD = 0.206, p = 0.012). This hypothesis shows a positive relationship between presence and enjoyment and explains 46.8% of the variance.

H5. The fifth hypothesis looked at adoption having an effect on enjoyment. It was found that adoption increases enjoyment (M = 0.630, SD = 0.253, p = 0.007). This hypothesis shows a positive statistical relationship between adoption and enjoyment and explains 62.4% of the variance.

H6. The sixth hypothesis covers the relationship of adoption and presence. The data shows that adoption had a positive effect on presence (M = 0.756, SD = 0.070, p = <0.001). This hypothesis shows a very significant relationship between adoption and presence and explains 71.7% of the variance.

H7. Lastly, the seventh hypothesis focused on enjoyment and its relationship to continuance intention and appropriation. Based on the data, it was found that enjoyment is positively related to continuance intention (M = 0.729, SD = 0.130, p = <0.001). This hypothesis shows a strongly significant positive relationship between enjoyment and continuance intention and explains 73.5% of the variance.

8 Discussion and key findings

8.1 SoOC

Early in the process, the entire SoOC project was put on hold due to COVID-19 lockdowns around Europe. This resulted in all other aspects of the project coming to a stop, however we made the decision to carry on and implement as much as possible from previous meetings, by primarily focusing on our thesis and research. This had a great impact on the workflow and implementation, as decisions about the content and functionalities were altered midway to accommodate the research. The reasons for having some seemingly "redundant" paragraphs (e.g, in the introduction regarding SoOC, mentioning the artists) is due to not knowing, during the different phases of the project, when/if Corona restrictions would be lifted and if we would be able to meet with the artists and integrate their assets into our environment/research. Out of respect for the initial vision, we opted to document it in this report, despite the aspects of the project which were caught in limbo throughout the process.

8.2 Design phases

The scope of the design and implementation phase was initially set to be about a month, however, due to the unforeseen consequences of COVID-19 and the shutdown on the SoOC project, we tried to create backup prototypes depending on the prospect of the situation (see section 5.2). Even though we did not utilize the implementation done in that short period, it was still somewhat of a solution to some of the issues faced at that moment. In a perfect world, the prototype had been finished in due time for hosting workshops and tailoring the environment to the artists' needs and projects, and no time would have been diverted to designing alternatives as a knee-jerk reaction to the initial pandemic panic. In lieu of the artists' inputs and assets, the design choices for creating the objects and scenes were mostly based on research and functionality, to serve a purpose, whereas the inclusion of professionally made art installations and concepts could have potentially added an extra layer of depth, albeit likely at the cost of more noisy data.
In terms of the program itself (application), certain issues arose during development which then postponed the initial testing period for the application. This was however not a detriment as the application ended up being stable and functioning as intended. Even though we had

to piece together different SDKs and self-made code, creating a grand puzzle, to make the application work at an acceptable level, the application ended up being relatively polished. During the implementation process, a lot of new important updates to SDKs were released, essentially updates that would facilitate the implementation further had we not had our own version. However, due to compatibility issues and time constraints, we were not able to utilize it for the testable prototype. In section 10, we elaborate on these new technologies and how they could be used for expansion and upgrade of the application.

8.3 Evaluation

Data from 26 participants was collected from 4 different locations, with each participant tested separately. Apart from the differing locations, some bias was introduced through the fact that the participants consisted mostly of friends and family, or their colleagues and co-students, as setting up an installation or a location and approaching people during lockdown was deemed unethical. The location bias would have been reduced if this was not the case and a single "setup" would have been used instead, but it can arguably be recognized as a positive addition to the research, as one of the main focal points is that of location-based experiences, which ideally would have been conducted in both Roeselare and Barcelona.
The choice of data collection for this study was deemed sufficient with only a questionnaire. However, further data collection might have strengthened the research such as mechanical observations, i.e, video footage or screen captures. Screen captures did however already have an adverse effect on the application, as the device was already allocating plenty of memory to the application itself, and would struggle to distribute it to recording the screen as well, which is why this particular method was avoided. Behavioural observations (filming the participants) could have strengthened the data in terms of technology adoption, presence and other, by running a data session in which researchers deliberate and determine behaviour from set criteria related to the research. The only mechanical outputs from the application were the participants' photos that were taken during the session, which was optional and not forced upon the participants (see figure 8.1).

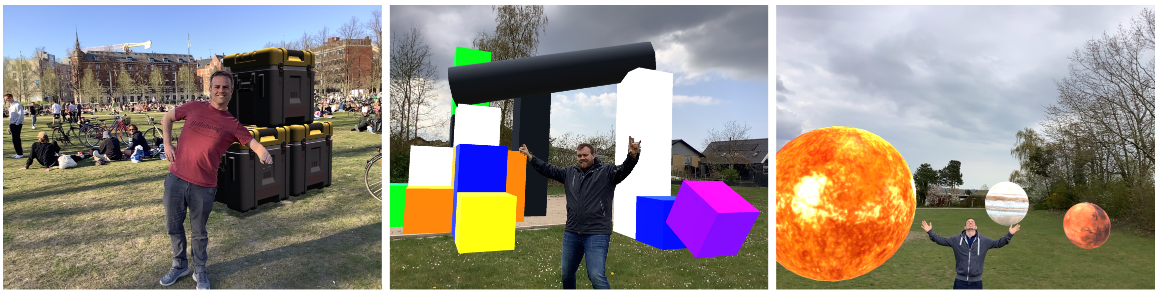


Figure 8.1: Three different participants getting their photos taken during the test - one from each scene

However, this feature was very popular among the test participants as several of them

heavily used it to capture themselves in the environment, similar to a postcard (in this case, a virtual postcard).
Further potential bias could be attributed to respondent fatigue, as the test encompassed several constructs, employing aspects derived from multiple questionnaires, resulting in participants having to answer 84 questions after using the application. Likewise, was it observed that some participants struggled to understand the wording of some of the items in the questionnaire, most of which were from the premade questionnaires from other studies. Some of these items use a more esoteric rhetoric, by including specific terms and expressions that might not be common; these had to be explained and elaborated upon. The size of the final questionnaire could have been reduced slightly by testing the self-made questions in a similar fashion to Laato et al. [63], through the use of card sorting and association. Additionally, this would in turn have strengthened the discriminant validity of some of the constructs and factors of the final model. In terms of the collected data, a larger sample size would undoubtedly be beneficial, specifically to allow for more variables and less parsimony in the model(s), as well as new findings in areas of the research that showed a smaller effect size. With that said, PLS-SEM models take more of an exploratory approach and less of a confirmatory, meaning they do not require a large sample size to infer consistency or relationships between different trends or phenomena [124].

The majority of the data gathered from the test, fit the hypotheses posed in this study. Only two out of seven hypotheses were not statistically significant. Both of those were relating to the motivation- and FoMO constructs. These were initially hypothesized to have an effect on continuance intention, however, this did not end up being the case. This could be explained by not making the point of scarcity clear enough through the time-limited scenes in the application.
A more longitudinal approach to the test could have made the implication of scarcity much clearer for the participants as they would have to be motivated and make an effort to "catch" the scenes during the right time. Especially, if the participants would use the application through a longer period for themselves rather than a short test with only a few scenes to experience. The FoMO items used for the questionnaire were taken from the study by Przybylski et al. [59], which relates mostly to a person's sense of FoMO in a more social media-esque capacity. These items were mostly used to get an understanding about the level of FoMO each participant experiences in their daily life and not necessarily in terms of the presented application. Another point to make in regards to the FoMO scores would be more of a personal one, as many of the participants used for the study were known to not be particularly active on social media, and thereby not be inclined to the type of FoMO portrayed in the items by Przybylski et al. [59].

8.4 Errors encountered

During the tests, an error occurred for four of the participants, which was only fixable by restarting the application. This is likewise suspected to be a factor leading to a lesser true sense of scarcity as the timers also reset when rebooting. However, all four of these participants had already picked a location prior to resetting the application. Another issue that occurred during the test was what could be deemed as GPS interference. Some of the testing locations were either far from cell towers or in a crowded location with many surrounding devices. This led to some tracking issues, where the virtual object was difficult to locate despite being in the right position. Also, the map also had some tracking issues when being surrounded by a lot of other devices, including locations such as Kronborg Slot (our initial testing spot in Helsingør), but struggled to locate the device due to lack of cell towers and conflicting signals from Sweden, across Øresund.

9 Conclusion

The basis of this study was the collaboration of the European Art Project Sound of Our Cities (SoOC), which aims to combine art, technology and social space in two urban cities, namely Roeselare, Belgium and Barcelona, Spain. Our contribution to the project, as students from Aalborg University, was to provide a platform in which sound artists from all over Europe could use to artificially place augmented audio in these aforementioned locations. This platform was specifically a mobile application that utilizes augmented reality- and GPS technology. An initial research question was posed based on the task and role given in the SoOC project:

IRQ *How can an interactive augmented reality application act as a tool for sound designers as well as provide immersive 3D content for users to experience in urban environments?*

During the project's inception, a global pandemic caused many facilities to go into lockdown. This included the collaborative work between partners in the SoOC project. Since the thesis was aimed at providing a platform for artists, the direction had to change as work was shut down, hence the thesis had to shift direction to investigate another aspect of the work such as human-computer interaction (HCI). After thorough research within areas such as AR development, state-of-the-art, technology acceptance (TAM), presence, fear of missing out (FoMO), enjoyment, and motivation, a final research question was formulated to encompass the slightly changed direction of the thesis:

FRQ *To what extent do factors such as presence, motivation, adoption, fear of missing out, and enjoyment contribute to the users' continuance intention and appropriation of a location-based augmented reality application?*

A prototype was developed in the form of a mobile application using AR technology as well as GPS for map positioning. The concept of the application was to provide the user with points of interest on a map that would represent real-world locations for which they could experience virtual scenes in AR. Each of the scenes were time-limited, i.e, only available to the user for a given time before either changing or disappearing. This meant that users would have to decide whether they wanted to 'catch' a certain scene in the time frame. A correlational research framework was established, looking into many of the concepts mentioned above and their relationship to the intention of continuous use of the application in the future. A user study was set up resulting in gathering 26 participants for testing the application. The test was conducted at a moment, where it was deemed appropriate for social contact. The data was processed using PLS-SEM, as a complex model was created with many constructs and relationships. Results showed that general FoMO (mostly in regards to social media) was found to have a positive effect on the sense of scarcity in the application, i.e, the timed virtual events. The sense of presence while experiencing the virtual scenes was found to have a positive effect on the sense of enjoyment. The level of technology acceptance or adoption likewise contributed positively to the sense of enjoyment. Technology adoption was also a prominent factor contributing to the sense of presence. A strong positive relationship was found between the sense of enjoyment while using the application and the intention of continuous use as well as appropriation was found. It was initially assumed that FoMO and motivation (scarcity) would positively affect the intention of continuous use, however this was not found to be the case. Likewise did these not decrease the sense of enjoyment while using the application. This means that 5 out of 7 hypotheses could be statistically accepted. Further testing and development would be needed to strengthen the research and findings from this thesis, ideally at a time where there are no social restrictions. Longitudinal tests as well as additional data collection methods could further validate the current findings.

10 Future works

Based on the acquired knowledge and the current state of the art, a future iteration of the prototype would likely feature the ARkit 4 SDK, as both human- and object occlusion have become baseline, with even greater capability, in part due to the integration with LiDAR technology available on newer devices. Another introduced feature is the ability to add location anchors at specific coordinates, possessing essentially identical affordances when compared with the prototype presented in this report, albeit with better tracking and triangulation (see section 2.1.5.3 and 2.1.5.5). Additionally, incorporating marker-based techniques, e.g, AR projections onto buildings or mapped out surfaces in an urban area, alongside the location anchors and marker-less techniques cloud server as an additional point of triangulation, increasing the stability of the marker-less scene once within the vicinity of a POI. A more extensive audio setup could also be considered that features head tracking (e.g, Apple Airpods Pro 2021), adding the ability to attach the AR camera's rotation and position within a scene, to that of the user's head instead of the handheld device, potentially increasing several aspects of the auditory induced presence. As a consequence of the lessened acoustic transparency stemming from wearing in-ear headphones instead of open-back headphones or earbuds that sit in the concha, in front of the ear canal, real time binaural mixing of the real world (audio passthrough) and augmented sound, with simulated reflections ideally matching the physical space, would be worth considering in this context. Further steps towards closing the gap between the auditory- and the visual stimuli would be by possibly taking advantage of the capabilities of LiDAR and the ARKit4 occlusion algorithms, to make physical objects and people also occlude virtual sounds. This together with an elaborate evaluation scenario with longitudinal data collection could be further supportive of the research hypotheses and strengthen the results.

Since the data from the conducted tests showed positive relationships between technology adoption and acceptance of the medium used (mobile device), and the sense of both presence and enjoyment, it is clear that continuing with this type of medium could be beneficial to reach other markets than those focused on art and culture. Applying the research to other fields and contexts, such as leveraging the fact that physical space can be occupied by virtual objects (internally dubbed Virtual-Real Estate). Physical billboards or promotional materials situated in a public place can be very expensive, but the cost of virtual objects in the same space would likely be attributed to the platforms through which they're conveyed (Facebook or Instagram etc.). Furthermore, the affordances and implementation of physical promotion differs vastly from what is possible through virtual means, as remote-access for the developer would be constant and the material would be hot-swappable or perhaps even cyclical.

11 Acknowledgements

We would first and foremost like to express our sincere thanks to our supervisor Cumhur Erkut, who has been incredibly patient and helpful throughout the entirety of the project. Even through times where the direction of the project was unclear due to the pandemic, he provided support and guidance.

We would also like to thank our collaborative partners of Sound of Our Cities. Not only was it exciting trying to conceptualize the ideas of the project, but also meeting interesting and intelligent partners working on completely different aspects of the project.

This study has been a learning experience for both of us and we have ventured out of our comfort zones in so many ways, which has made this project special to us.

Another thank you to our family and friends that supported us through our work on this thesis, and for being available for testing whenever needed, even when social distancing was a strict concern. We are very thankful for each and every one's time and effort to help us complete the study.

Bibliography

[1] Ronald T. Azuma. *A Survey of Augmented Reality*. 2005.

[2] Alan B Craig. *Understanding augmented reality: Concepts and applications*. Newnes, 2013.

[3] William Sherman and Alan Craig. *Output: Interfacing the Virtual World With the Participant(s)*. 2019, pp. 258–396.

[4] Te-Lien Chou and Lih-Juan Chanlin. "Location-Based Learning Through Augmented Reality". In: *Journal of Educational Computing Research* 51 (Oct. 2014), pp. 355–368. DOI: 10.2190/EC.51.3.e.

[5] Anne-Cecilie Haugstvedt and John Krogstie. "Mobile augmented reality for cultural heritage: A technology acceptance study". In: *2012 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE. 2012, pp. 247–255.

[6] Gheorghe-Daniel Voinea et al. "Exploring cultural heritage using augmented reality through Google's Project Tango and ARCore". In: *International conference on VR Technologies in Cultural Heritage*. Springer. 2018, pp. 93–106.

[7] VIAR. *Virtual reality market size in 2018 with forecast for 2019*. URL: https://www.viar360.com/virtual-reality-market-size-2018/. (accessed: 22.05.2021).

[8] Ronald Azuma. "Making Augmented Reality a Reality". In: Jan. 2017, JTu1F.1. DOI: 10.1364/3D.2017.JTu1F.1.

[9] Jing Yang and Friedemann Mattern. "Audio augmented reality for human-object interactions". In: Sept. 2019, pp. 408–412. ISBN: 978-1-4503-6869-8. DOI: 10.1145/3341162.3349302.

[10] Niteesh Yadav. *Understanding display techniques in Augmented Reality*. URL: https://blog.prototypr.io/understanding-display-techniques-in-augmented-reality-c258b911b5c9. (accessed: 22.05.2021).

[11] Hyeonah Choi, Youngwon Kim, and Gerard Kim. "Presence, Immersion and Usability of Mobile Augmented Reality". In: Aug. 2019.

[12] James Lin et al. "Effects of Field of View on Presence, Enjoyment, Memory, and Simulator Sickness in a Virtual Environment". In: *Virtual Reality Conference, IEEE* 0 (Jan. 2002), p. 164.

[13] Vuforia. *Develop AR Experiences With Vuforia Engine*. URL: https://www.ptc.com/en/products/vuforia/vuforia-engine. (accessed: 22.05.2021).

[14] Jack Cheng, Keyu Chen, and Weiwei Chen. "Comparison of marker-based AR and markerless AR: A case study on indoor decoration system". In: July 2017. DOI: 10.24928/JC3-2017/0231.

[15] Wikitude. *Wikitude Augmented Reality SDK*. URL: https://www.wikitude.com/products/wikitude-sdk/. (accessed: 22.05.2021).

[16] *Niantic*. URL: https://nianticlabs.com/en/products/. (accessed: 22.05.2021).

[17] Google. *Geolocation: Displaying User or Device Position on Maps*. URL: https://developers.google.com/maps/documentation/javascript/geolocation. (accessed: 22.05.2021).

[18] Sensolus. *Geolocation: The what and how behind it*. URL: https://www.sensolus.com/comparison-geolocation-technologies/. (accessed: 22.05.2021).

[19] Panagiotis Galatis et al. "Mobile Augmented Reality Guides in Cultural Heritage." In: *MobiCASE*. 2016, pp. 11–19.

[20] Ronald Azuma. "The road to ubiquitous consumer augmented reality systems". In: *Human Behavior and Emerging Technologies* 1 (Feb. 2019), e113. DOI: 10.1002/hbe2.113.

[21] Brian August. *110 Stories: Augmented Reality Twin Towers iPhone App*. URL: https://www.kickstarter.com/projects/110stories/110-stories-augmented-reality-twin-towers-iphone-a. (accessed: 22.05.2021).

[22] Ronald Azuma. *110 Stories: Augmented Reality Twin Towers iPhone App*. URL: https://ronaldazuma.com/Leviathan_at_CES2014.html. (accessed: 22.05.2021).

[23] Chanel Summers, Vangelis Lympouridis, and Cumhur Erkut. "Sonic interaction design for virtual and augmented reality environments". In: Mar. 2015, pp. 1–6. DOI: 10.1109/SIVE.2015.7361290.

[24] *Fields*. URL: https://fields.planeta.cc/. (accessed: 22.05.2021).

[25] Felix Faire et al. *AMBEO Augmented Audio Lab*. URL: https://felixfaire.com/work/ambeo-lab. (accessed: 22.05.2021).

[26] David Kirkpatrick. *Pokemon Go led to over 1B interactions on Facebook and Instagram in July*. URL: https://www.marketingdive.com/news/pokemon-go-led-to-over-1b-interactions-on-facebook-and-instagram-in-july/424187/. (accessed: 22.05.2021).

[27] Jon Russell. *Google's Ingress augmented reality game quietly launches for iOS devices*. URL: https://thenextweb.com/news/googles-ingress-augmented-reality-game-quietly-launches-ios-devices. (accessed: 22.05.2021).

[28] Ethan Saadia. *AR Location Anchors in ARKit 4*. URL: https://medium.com/@ethansaadia/ar-location-anchors-in-arkit-4-9a7ca19652f5. (accessed: 22.05.2021).

[29] Mike Peterson. *What you need to know about Apple's LiDAR Scanner in the iPad Pro*. URL: https://appleinsider.com/articles/20/03/18/what-you-need-to-know-about-apples-lidar-scanner-in-the-ipad-pro. (accessed: 22.05.2021).

[30] Apple. *Introducing ARKit 4*. URL: https://developer.apple.com/augmented-reality/arkit/. (accessed: 22.05.2021).

[31] *Apple iPad Pro Lidar Scanner for Augmented Reality - Superba AR*. URL: https://www.youtube.com/watch?v=ipcq02PEX5c&ab_channel=SuperbaAR. (accessed: 22.05.2021).

[32] *Augmented Virtual Reality*. URL: https://www.patentlyapple.com/patently-apple/augmented-reality/. (accessed: 22.05.2021).

[33] Gerald Lynch and James Peckham. *Apple Glasses release date, price, news, leaks and what we know so far*. URL: https://www.techradar.com/news/apple-glasses. (accessed: 22.05.2021).

[34] Mark Gurman. *Apple's Secretive AR and VR Headset Plans Altered by Internal Differences*. URL: https://www.bloomberg.com/news/articles/2020-06-19/apple-team-working-on-vr-and-ar-headset-and-ar-glasses. (accessed: 22.05.2021).

[35] Kate Kozuch. *Apple Glasses: Release date, price, features and leaks*. URL: https://www.tomsguide.com/news/apple-glasses. (accessed: 22.05.2021).

[36] Fred Davis. "A Technology Acceptance Model for Empirically Testing New End-User Information Systems". In: (Jan. 1985).

[37] Fred D Davis, Richard P Bagozzi, and Paul R Warshaw. "User acceptance of computer technology: A comparison of two theoretical models". In: *Management science* 35.8 (1989), pp. 982–1003.

[38] Viswanath Venkatesh et al. "User acceptance of information technology: Toward a unified view". In: *MIS quarterly* (2003), pp. 425–478.

[39] Viswanath Venkatesh and Hillol Bala. "Technology Acceptance Model 3 and a Research Agenda on Interventions". In: *Decision Sciences - DECISION SCI* 39 (May 2008), pp. 273–315. DOI: 10.1111/j.1540-5915.2008.00192.x.

[40] Ronnie H Shroff, Christopher C Deneen, and Eugenia MW Ng. "Analysis of the technology acceptance model in examining students' behavioural intention to use an e-portfolio system". In: *Australasian Journal of Educational Technology* 27.4 (2011).

[41] Azza Alomary and John Woollard. "How is technology accepted by users? A review of technology acceptance models and theories". In: (2015).

[42] Martin Fishbein and Icek Ajzen. *Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research.* 1975.

[43] Lara LaCaille. *Theory of Reasoned Action*. Ed. by Marc D. Gellman. Cham: Springer International Publishing, 2020, pp. 2231–2234. ISBN: 978-3-030-39903-0. DOI: 10.1007/978-3-030-39903-0_1619. URL: https://doi.org/10.1007/978-3-030-39903-0_1619.

[44] Icek Ajzen. "The theory of planned behavior". In: *Organizational Behavior and Human Decision Processes* 50.2 (1991). Theories of Cognitive Self-Regulation, pp. 179–211. ISSN: 0749-5978. DOI: https://doi.org/10.1016/0749-5978(91)90020-T. URL: https://www.sciencedirect.com/science/article/pii/074959789190020T.

[45] F Davis, R Bagozzi, and P Warshaw. "Extrinsic and Intrinsic Motivation to Use Computers in the Workplace". In: *Journal of Applied Social Psychology* (1991).

[46] Viswanath Venkatesh and Fred Davis. "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies". In: *Management Science* 46 (Feb. 2000), pp. 186–204. DOI: 10.1287/mnsc.46.2.186.11926.

[47] Viswanath Venkatesh. "Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model". In: *Information systems research* 11.4 (2000), pp. 342–365.

[48] M Leue, D tom Dieck, and TH Jung. "A Theoretical Model of Augmented Reality Acceptance". In: *e-Review of Tourism Research* 5 (Mar. 2014).

[49] Eduardo Parra-López et al. "Intentions to use social media in organizing and taking vacation trips". In: *Computers in Human Behavior* 27.2 (2011). Web 2.0 in Travel and Tourism: Empowering and Changing the Role of Travelers, pp. 640–654. ISSN: 0747-5632. DOI: https://doi.org/10.1016/j.chb.2010.05.022. URL: https://www.sciencedirect.com/science/article/pii/S0747563210001615.

[50] Jen-Her Wu, Shu-Ching Wang, and Ho-Huang Tsai. "Falling in love with online games: The uses and gratifications perspective". In: *Computers in Human Behavior* 26 (Nov. 2010), pp. 1862–1871. DOI: 10.1016/j.chb.2010.07.033.

[51] Hongxiu Li et al. "Modeling hedonic is continuance through the uses and gratifications theory: An empirical study in online games". In: *Computers in Human Behavior* 48 (July 2015). DOI: 10.1016/j.chb.2015.01.053.

[52] Ezlika Ghazali, Dilip Mutum, and Mei-Yuen Woon. "Exploring player behavior and motivations to continue playing Pokémon GO". In: (Aug. 2018). DOI: 10.1108/ITP-07-2017-0216.

[53] Jennie Carroll et al. "Identity, Power And Fragmentation in Cyberspace: Technology Appropriation by Young People". In: (Jan. 2001).

[54] Jennie Carroll et al. "Just What Do the Youth of Today Want? Technology Appropriation by Young People (PDF)". In: *Hawaii International Conference on System Sciences* 5 (Jan. 2002), 131b. DOI: 10.1109/HICSS.2002.994089.

[55] Susanne Bødker and Clemens Nylandsted Klokmose. "Dynamics in artifact ecologies". English. In: *NordiCHI '12 : Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*. Nordic Conference on Human Computer Interactions, NordiCHI ; Conference date: 14-10-2012 Through 17-10-2012. Association

for Computing Machinery, 2012, pp. 448–457. ISBN: 978-1-4503-1482-4. DOI: 10.1145/2399016.2399085.

[56] James J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, 1979.

[57] Susanne Bødker and Clemens Klokmose. "The Human–Artifact Model: An Activity Theoretical Approach to Artifact Ecologies". In: *Human-Computer Interaction* 26 (Dec. 2011), pp. 315–371. DOI: 10.1080/07370024.2011.626709.

[58] Richard Ryan and Edward Deci. "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being". In: *The American psychologist* 55 (Feb. 2000), pp. 68–78. DOI: 10.1037/0003-066X.55.1.68.

[59] Andrew K. Przybylski et al. "Motivational, emotional, and behavioral correlates of fear of missing out". In: *Computers in Human Behavior* 29.4 (2013), pp. 1841–1848. ISSN: 0747-5632. DOI: https://doi.org/10.1016/j.chb.2013.02.014. URL: https://www.sciencedirect.com/science/article/pii/S0747563213000800.

[60] A. Seay and Robert Kraut. "Project massive: Self-regulation and problematic use of online gaming". In: Apr. 2007, pp. 829–838. DOI: 10.1145/1240624.1240749.

[61] C. Hodkinson. "'Fear of Missing Out' (FOMO) marketing appeals: A conceptual model". In: *Journal of Marketing Communications* 25 (Oct. 2016), pp. 1–24. DOI: 10.1080/13527266.2016.1234504.

[62] Edward L Deci and Richard M Ryan. "Self-determination theory." In: (2012).

[63] Samuli Laato, A.K.M. Najmul Islam, and Teemu Laine. "Did Location-based Games Motivate Players to Socialize During COVID-19?" In: *Telematics and Informatics* 54 (June 2020), p. 101458. DOI: 10.1016/j.tele.2020.101458.

[64] li li et al. "Fear of missing out (FoMO) and gaming disorder among Chinese university students: Impulsivity and game time as mediators". In: (Apr. 2020). DOI: 10.21203/rs.3.rs-25628/v1.

[65] Shaun Ansari. "Video Game MicroTransactions - Engineering desire in digital experiences". PhD thesis. Jan. 2019. DOI: 10.13140/RG.2.2.24788.35207.

[66] M. Perrone. "FoMO: Establishing Validity of the Fear of Missing Out Scale with an Adolescent Population". In: 2016.

[67] *Feel Like a Wallflower? Maybe It's Your Facebook Wall*. URL: https://www.nytimes.com/2011/04/10/business/10ping.html. (accessed: 22.05.2021).

[68] Elisa Wegmann et al. "Online-specific fear of missing out and Internet-use expectancies contribute to symptoms of Internet-communication disorder". In: *Addictive Behaviors Reports* 5 (2017), pp. 33–42. ISSN: 2352-8532. DOI: https://doi.org/10.1016/j.abrep.2017.04.001. URL: https://www.sciencedirect.com/science/article/pii/S235285321730007X.

[69] Mel Slater and Sylvia Wilbur. "A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments". In: *Presence: Teleoperators and Virtual Environments* 6 (Dec. 1997), pp. 603–. DOI: 10.1162/pres.1997.6.6.603.

[70] B Goldiez and Jeffrey W Dawson. "Is presence present in augmented reality systems". In: *Proceedings of Presence 2004. VII. International Workshop on Presence*. 2004, pp. 294–297.

[71] Mel Slater et al. "How we experience immersive virtual environments: The concept of presence and its measurement". In: *Anuario de Psicologia* 40 (Sept. 2009).

[72] Mel Slater, John McCarthy, and Francesco Maringelli. "The Influence of Body Movement on Subjective Presence in Virtual Environments". In: *Human Factors* 40.3 (1998). PMID: 9849105, pp. 469–477. DOI: 10.1518/001872098779591368. URL: https://doi.org/10.1518/001872098779591368.

[73] Bob Witmer and Michael Singer. "Measuring Presence in Virtual Environments: A Presence Questionnaire". In: *Presence Teleoperators amp Virtual Environments* 7 (July 1998). DOI: 10.1162/105474698565686.

[74] Matthew Lombard and Theresa Ditton. "At the Heart of It All: The Concept of Presence". In: *Journal of Computer-Mediated Communication* 3.2 (Sept. 1997). JCMC321. ISSN: 1083-6101. DOI: 10.1111/j.1083-6101.1997.tb00072.x. URL: https://doi.org/10.1111/j.1083-6101.1997.tb00072.x.

[75] Giuseppe Riva, Fabrizio Davide, and Wijnand A IJsselsteijn. *Being there: Concepts, effects and measurements of user presence in synthetic environments.* Ios Press, 2003.

[76] Arthur Tang, Frank Biocca, and Lynette Lim. "Comparing Differences in Presence during Social Interaction in Augmented Reality versus Virtual Reality Environments: An Exploratory Study". In: *Teleoperators and Virtual Environments - Presence* (Jan. 2004).

[77] Frank Biocca. "The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments." In: *J. Computer-Mediated Communication* 3 (Jan. 1997).

[78] Peter Vorderer et al. "Constructing Presence: Towards a two-level model of the formation of Spatial Presence". In: *Report to the European Community, Project Presence: MEC (IST-2001-37661)* (June 2003).

[79] Holger Regenbrecht and Thomas Schubert. "Real and Illusory Interactions Enhance Presence in Virtual Environments". In: *Presence* 11 (Aug. 2002), pp. 425–434. DOI: 10.1162/105474602760204318.

[80] Thomas Schubert. "A New Conception of Spatial Presence: Once Again, with Feeling". In: *Communication Theory* 19 (May 2009), pp. 161 –187. DOI: 10.1111/j.1468-2885.2009.01340.x.

[81] Peter Vorderer et al. "MEC spatial presence questionnaire (MEC-SPQ): Short documentation and instructions for application". In: *Report to the European Community, Project Presence: MEC (IST-2001-37661)* (June 2004).

[82] Holger Regenbrecht and Thomas Schubert. "Measuring presence in augmented reality environments: design and a first test of a questionnaire". In: *arXiv preprint arXiv:2103.02831* (2021).

[83] Mel Slater and Anthony Steed. "A Virtual Presence Counter". In: *Presence* 9 (Oct. 2000), pp. 413–434. DOI: 10.1162/105474600566925.

[84] Anthony Steed et al. "The Impact of a Self-Avatar on Cognitive Load in Immersive Virtual Reality". In: Mar. 2016, pp. 67–76. DOI: 10.1109/VR.2016.7504689.

[85] Brett Stevens et al. "Putting the Virtual into Reality: Assessing Object-Presence with Projection-Augmented Models". In: *Presence* 11 (Feb. 2002), pp. 79–92. DOI: 10.1162/105474602317343677.

[86] Martijn Schuemie et al. "Research on Presence in VR: a Survey". In: (Mar. 2001).

[87] Timothy Jung et al. "Effects of Virtual Reality and Augmented Reality on Visitor Experiences in Museum". In: (Feb. 2016).

[88] Rod Mccall et al. "Using presence to evaluate an augmented reality location aware game". In: *Personal and Ubiquitous Computing* 15 (June 2011), pp. 25–35. DOI: 10.1007/s00779-010-0306-8.

[89] Paul Milgram and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays". In: *IEICE Trans. Information Systems* vol. E77-D, no. 12 (Dec. 1994), pp. 1321–1329.

[90] Tina Wu et al. "The Effect of Head Tracking on the Degree of Presence in Virtual Reality". In: *International Journal of Human-Computer Interaction* 35 (Dec. 2018), pp. 1–9. DOI: 10.1080/10447318.2018.1555736.

[91] David H Warren and Wallace T Cleaves. "Visual-proprioceptive interaction under large amounts of conflict." In: *Journal of experimental psychology* 90.2 (1971), p. 206.

[92] Mark Dourado and Ingrid Lima. "Perceived Density of Virtual Objects in VR: A study on the Effect of Multimodality on Density Perception". In: (2019).

[93] David Hecht, Miriam Reiner, and Avi Karni. "Multisensory enhancement: Gains in choice and in simple response times". In: *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale* 189 (June 2008), pp. 133–43. DOI: 10.1007/s00221-008-1410-0.

[94] Pontus Larsson et al. "Auditory-Induced Presence in Mixed Reality Environments and Related Technology". In: Oct. 2010, pp. 143–163. ISBN: 978-1-84882-732-5. DOI: 10.1007/978-1-84882-733-2_8.

[95] Sharon Tenenholz, Lee Gilroy, and Randolph Blake. "Hearing What the Eyes See". In: *Psychological Science* 16 (Apr. 2005), pp. 228–235. DOI: 10.1111/j.0956-7976.2005.00808.x.

[96] Claudia Hendrix and Woodrow Barfield. "The Sense of Presence Within Auditory Virtual Environments." In: *Presence* 5 (June 1996), pp. 290–301. DOI: 10.1162/pres.1996.5.3.290.

[97] Aleksander Väljamäe et al. "Auditory Presence, Individualized Head-Related Transfer Functions, and Illusory Ego-Motion in Virtual Environments". In: *Proceedings of 7th Annual Workshop of Presence* (Jan. 2004).

[98] Pontus Larsson, Daniel Västfjäll, and Mendel Kleiner. "Effects of auditory information consistency and room acoustic cues on presence in virtual environments". In: *Acoustical Science and Technology* 29 (Mar. 2008), pp. 191–194. DOI: 10.1250/ast.29.191.

[99] Durand Begault, Elizabeth Wenzel, and Mark Anderson. "Direct Comparison of the Impact of Head Tracking, Reverberation, and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Speech Source". In: *Journal of the Audio Engineering Society. Audio Engineering Society* 49 (Nov. 2001), pp. 904–16.

[100] Priscilla Chueng-Nainby and Phil Marsden. "Designing Auditory Spaces to Support Sense of Place: The Role of Expectation". In: Dec. 2002.

[101] Mark Mcgill et al. "Acoustic Transparency and the Changing Soundscape of Auditory Mixed Reality". In: Apr. 2020, pp. 1–16. DOI: 10.1145/3313831.3376702.

[102] Pontus Larsson, Daniel Västjäll, and Mendel Kleiner. "On the quality of experience: A multi-modal approach to perceptual ego-motion and sensed presence in virtual environments". In: *ISCA Tutorial and Research Workshop on Auditory Quality of Systems*. 2003.

[103] Ian Kelly, Marcin Gorzel, and Alper Güngörmüsler. "Efficient externalized audio reverberation with smooth transitioning". In: (2017).

[104] *Developing with Resonance Audio*. URL: https://resonance-audio.github.io/resonance-audio/develop/overview.html. (accessed: 22.05.2021).

[105] Nicolas Bonneel et al. "Bimodal perception of audio-visual material properties for virtual environments". In: *ACM Transactions on Applied Perception (TAP)* 7.1 (2010), pp. 1–16.

[106] Martin Giesel and Qasim Zaidi. "Frequency-based heuristics for material perception". In: *Journal of vision* 13 (Dec. 2013). DOI: 10.1167/13.14.7.

[107] Robert Welch and David Warren. "Immediate Perceptual Response to Intersensory Discrepancy". In: *Psychological bulletin* 88 (Dec. 1980), pp. 638–67. DOI: 10.1037//0033-2909.88.3.638.

[108] Priscilla Chueng and Phil Marsden. "Designing auditory spaces to support sense of place: the role of expectation". In: *CSCW Workshop: The Role of Place in Shaping Virtual Community. Citeseer*. 2002.

[109] V. Willert et al. "A Probabilistic Model for Binaural Sound Localization". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 36.5 (2006), pp. 982–994. DOI: 10.1109/TSMCB.2006.872263.

[110] Mark Dourado and Daniel Strübig. "Enhancing the Player's Efficiency in Narrative-driven Search Tasks in Virtual Reality through Auditory Icons". In: (2018).

[111] Jan Pedersen and Torben Jorgensen. "Localization Performance of Real and Virtual Sound Sources". In: *Proceedings of the NATO RTO-MP-HFM-123 New Directions for Improving Audio Effectiveness Conference* (Apr. 2005).

[112] Paul Skalski and Robert Whitbred. "Image versus Sound: A Comparison of Formal Feature Effects on Presence and Video Game Enjoyment." In: *PsychNology Journal* 8 (Jan. 2010), pp. 67–84.

[113] Kenji Ozawa et al. "Psychological factors involved in auditory presence". In: *Acoustical Science and Technology* 24 (Jan. 2003). DOI: 10.1250/ast.24.42.

[114] Brent Cowan and Bill Kapralos. "GPU-based real-time acoustical occlusion modeling". In: *Virtual Reality* 14 (Sept. 2010), pp. 183–196. DOI: 10.1007/s10055-010-0166-6.

[115] *Short Notes: Diffraction of Waves*. URL: https://spmflix.com/courses/form-5-physics/lectures/4913389. (accessed: 22.05.2021).

[116] *Acoustic Lessons*. URL: https://carusoacoustic.com/us/magazine/first-lesson-of-acoustics-the-sound-wave/#intro. (accessed: 22.05.2021).

[117] Yuzhu Lu and Shana Smith. "GPU-Based Real-Time Occlusion in an Immersive Augmented Reality Environment". In: *Journal of Computing and Information Science in Engineering - JCISE* 9 (June 2009). DOI: 10.1115/1.3130144.

[118] Nicolas Tsingos et al. "Modeling Acoustics in Virtual Environments Using the Uniform Theory of Diffraction". In: (June 2001). DOI: 10.1145/383259.383323.

[119] Craig D Murray, Paul Arnold, and Ben Thornton. "Presence accompanying induced hearing loss: Implications for immersive virtual environments". In: *Presence: Teleoperators & Virtual Environments* 9.2 (2000), pp. 137–148.

[120] Samuli Laato, Teemu H. Laine, and A.K.M. Najmul Islam. "Location-Based Games and the COVID-19 Pandemic: An Analysis of Responses from Game Developers and Players". In: *Multimodal Technologies and Interaction* 4.2 (2020). ISSN: 2414-4088. DOI: 10.3390/mti4020029. URL: https://www.mdpi.com/2414-4088/4/2/29.

[121] *Coronavirus disease (COVID-19): How is it transmitted?* URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted. (accessed: 22.05.2021).

[122] Lynn Williams et al. "Protection motivation theory and social distancing behaviour in response to a simulated infectious disease epidemic". In: *Psychology Health and Medicine* 20 (May 2015). DOI: 10.1080/13548506.2015.1028946.

[123] Anthony Steed et al. "Evaluating immersive experiences during Covid-19 and beyond". In: *Interactions* 27 (July 2020), pp. 62–67. DOI: 10.1145/3406098.

[124] Vincenzo Esposito Vinzi, Laura Trinchera, and Silvano Amato. "PLS Path Modeling: From Foundations to Recent Developments and Open Issues for Model Assessment and Improvement". In: Jan. 2010, pp. 47–82. DOI: 10.1007/978-3-540-32827-8_3.

[125] V. Esposito Vinzi et al. *Handbook of Partial Least Squares*. 2019, pp. 449–485.

[126] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. "The motivational pull of video games: A self-determination theory approach". In: *Motivation and emotion* 30.4 (2006), pp. 344–360.

[127] Ron Tamborini et al. "Defining media enjoyment as the satisfaction of intrinsic needs". In: *Journal of communication* 60.4 (2010), pp. 758–777.

[128] *We are Dear Hunter.* URL: https://dearhunter.eu/. (accessed: 22.05.2021).

[129] David Saffo et al. "Crowdsourcing virtual reality experiments using vrchat". In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–8.

[130] *Maps and Location For Developers*. URL: mapbox.com. (accessed: 22.05.2021).

[131] *AR + GPS Location*. URL: https://assetstore.unity.com/publishers/40190. (accessed: 22.05.2021).

[132] *Native Gallery SDK*. URL: https://github.com/yasirkula/UnityNativeGallery. (accessed: 22.05.2021).

[133] *SunLight – Location based Time of Day*. URL: https://assetstore.unity.com/packages/tools/particles-effects/sunlight-location-based-time-of-day-66399. (accessed: 22.05.2021).

[134] *AR Foundation support for ARKit 3*. URL: https://blog.unity.com/technology/ar-foundation-support-for-arkit-3. (accessed: 22.05.2021).

[135] marcin gorzel et al. "efficient encoding and decoding of binaural sound with resonance audio". In: *journal of the audio engineering society* (2019).

[136] *Among Us*. URL: https://innersloth.com/gameAmongUs.php. (accessed: 22.05.2021).

[137] Thomas Bjørner. *Qualitative methods for Consumer Research*. 2015.

[138] Justin Jager, Diane Putnick, and Marc Bornstein. "II. MORE THAN JUST CONVENIENT: THE SCIENTIFIC MERITS OF HOMOGENEOUS CONVENIENCE SAMPLES". In: *Monographs of the Society for Research in Child Development* 82 (June 2017), pp. 13–30. DOI: 10.1111/mono.12296.

[139] *SmartPLS 3*. URL: https://www.smartpls.com/. (accessed: 22.05.2021).

[140] Claes Fornell and David F Larcker. "Evaluating structural equation models with unobservable variables and measurement error". In: *Journal of marketing research* 18.1 (1981), pp. 39–50.

[141] Jörg Henseler, Christian M Ringle, and Marko Sarstedt. "A new criterion for assessing discriminant validity in variance-based structural equation modeling". In: *Journal of the academy of marketing science* 43.1 (2015), pp. 115–135.

[142] Jose Benitez et al. "How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research". In: *Information Management* 57.2 (2020), p. 103168. ISSN: 0378-7206. DOI: https://doi.org/10.1016/j.im.2019.05.003. URL: https://www.sciencedirect.com/science/article/pii/S0378720617309394.

[143] John P Meyer and Natalie J Allen. *Commitment in the workplace: Theory, research, and application*. Sage publications, 1997.

[144] *Multicollinearity in Regression Analysis: Problems, Detection, and Solutions*. URL: https://statisticsbyjim.com/regression/multicollinearity-in-regression-analysis/. (accessed: 22.05.2021).

[145] Wynne Chin. "How to write up and report PLS analyses". In: *How to write up and report PLS analyses* (Jan. 2010), pp. 655–669.

[146] *Categorical Predictors via Dummy Variables SmartPLS 3*. URL: https://www.youtube.com/watch?v=L-uynRF8_cg&ab_channel=JamesGaskin. (accessed: 22.05.2021).

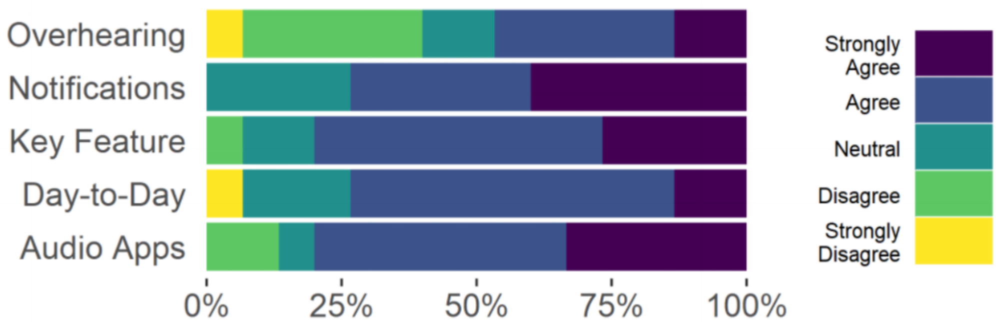
[147] Long W. Lam. "Impact of competitiveness on salespeople's commitment and performance". In: *Journal of Business Research* 65.9 (2012), pp. 1328–1334. ISSN: 0148-2963. DOI: https://doi.org/10.1016/j.jbusres.2011.10.026. URL: https://www.sciencedirect.com/science/article/pii/S0148296311003705.

[148] James Gaskin and Paul Lowry. "Partial Least Squares (PLS) Structural Equation Modeling (SEM) for Building and Testing Behavioral Causal Theory: When to Choose It and How to Use It". In: *IEEE Transactions on Professional Communication* 57 (June 2014), pp. 123–146. DOI: 10.1109/TPC.2014.2312452.

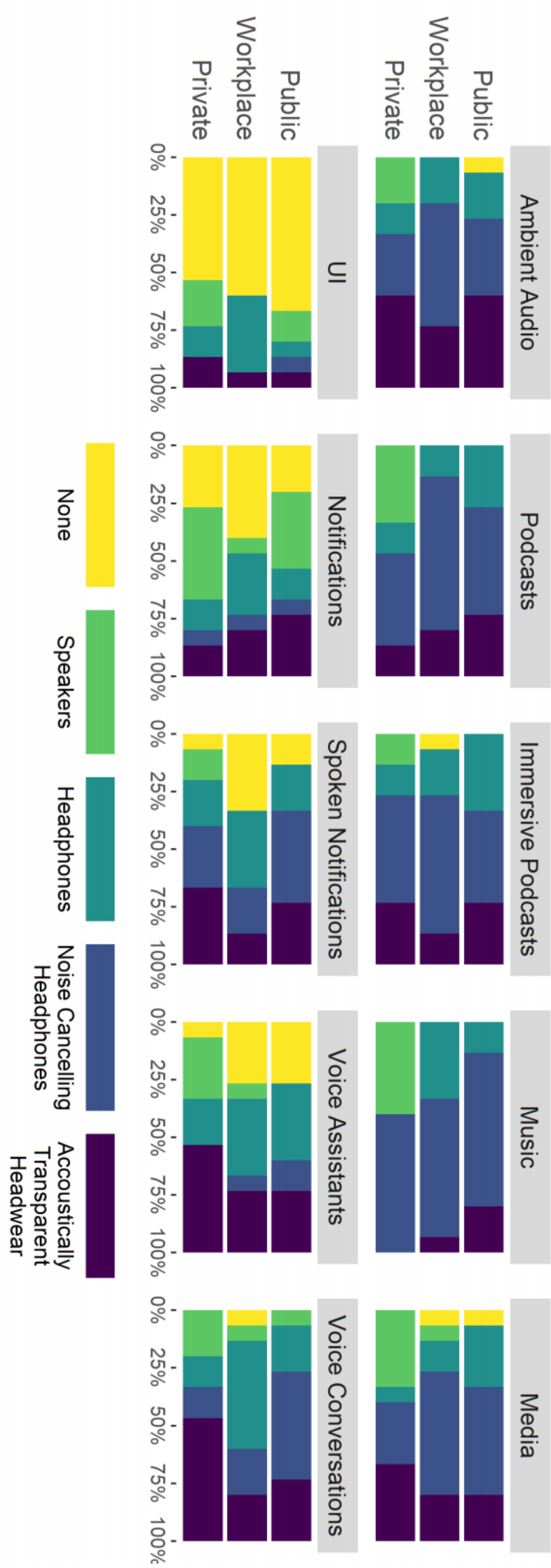
Appendix

Appendix A1. Attitudes to Acoustic Transparency as a Feature [101]

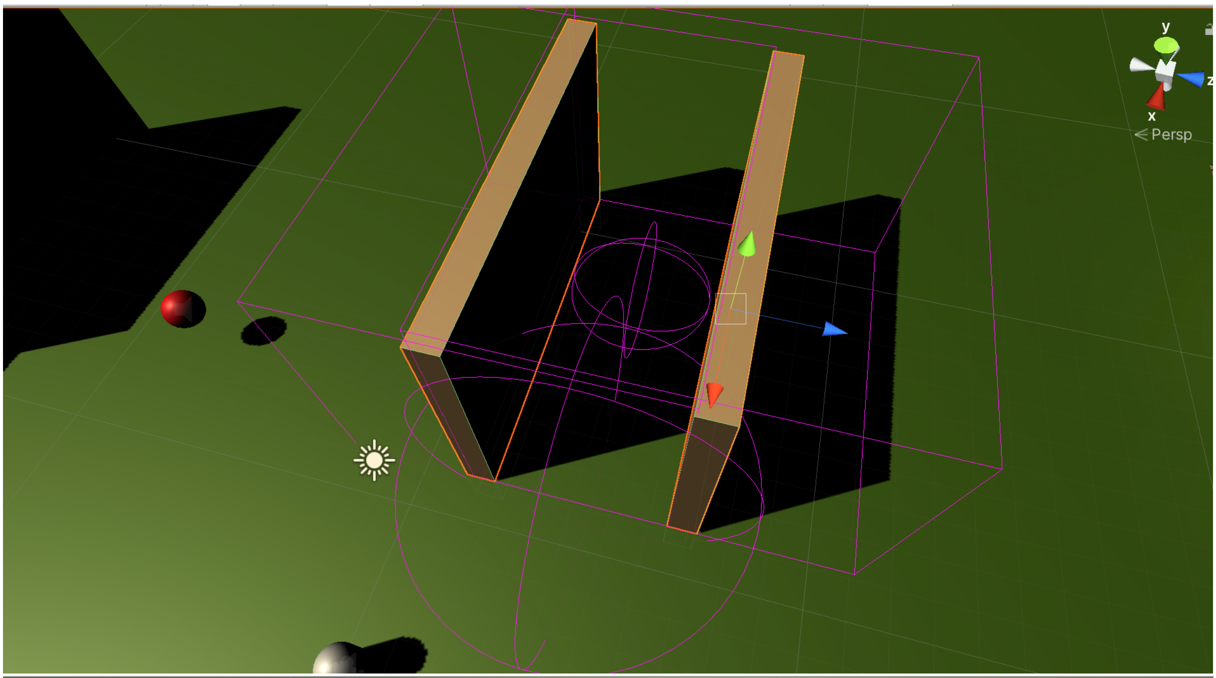
"The figure below shows participants almost universally agreed that acoustic transparency would be a key feature to consider in future headwear purchases, making it more likely that they would wear an audio device throughout the day and use audio based applications. Using a headset such as the frames was also seen as preferable to smartphone speakers with respect to auditory notifications. However, based on the experience wearing the frames, just under half of participants were concerned about other's being able to overhear their audio". [101]. "For each audio content type , participants were asked to indicate how they would prefer to listen to it in private (e.g, in the home or a car), in public (e.g, public transport, parks, restaurants) and in the workplace (e.g, office, library)."

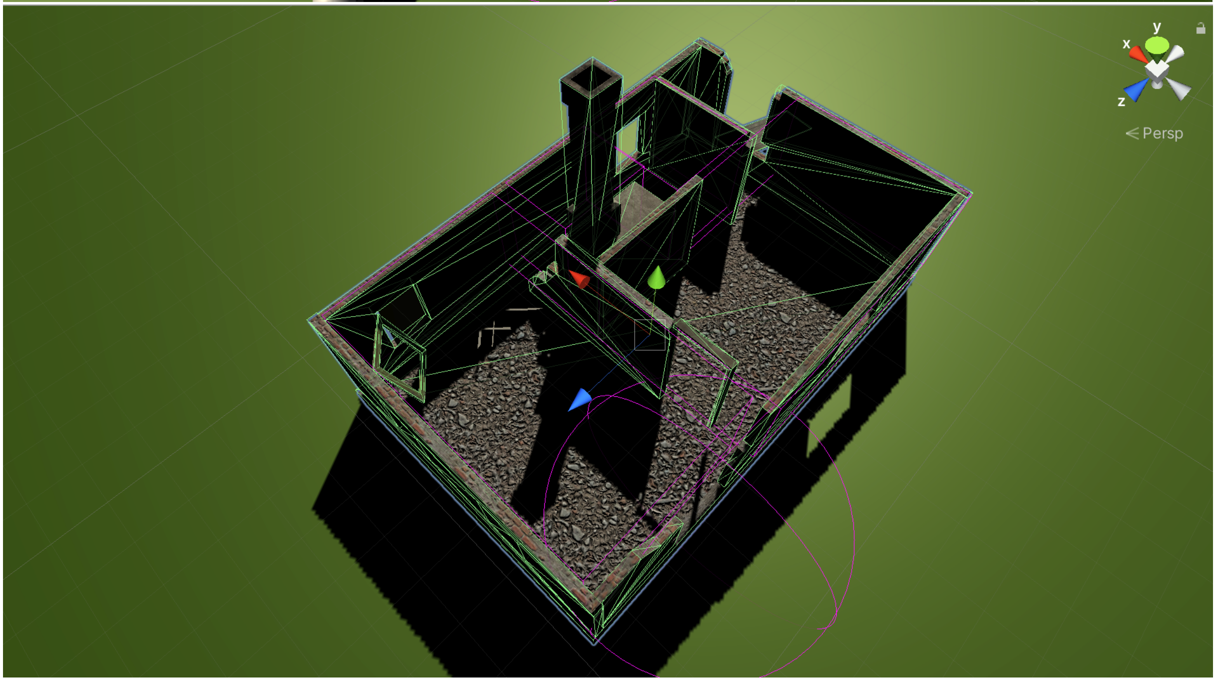


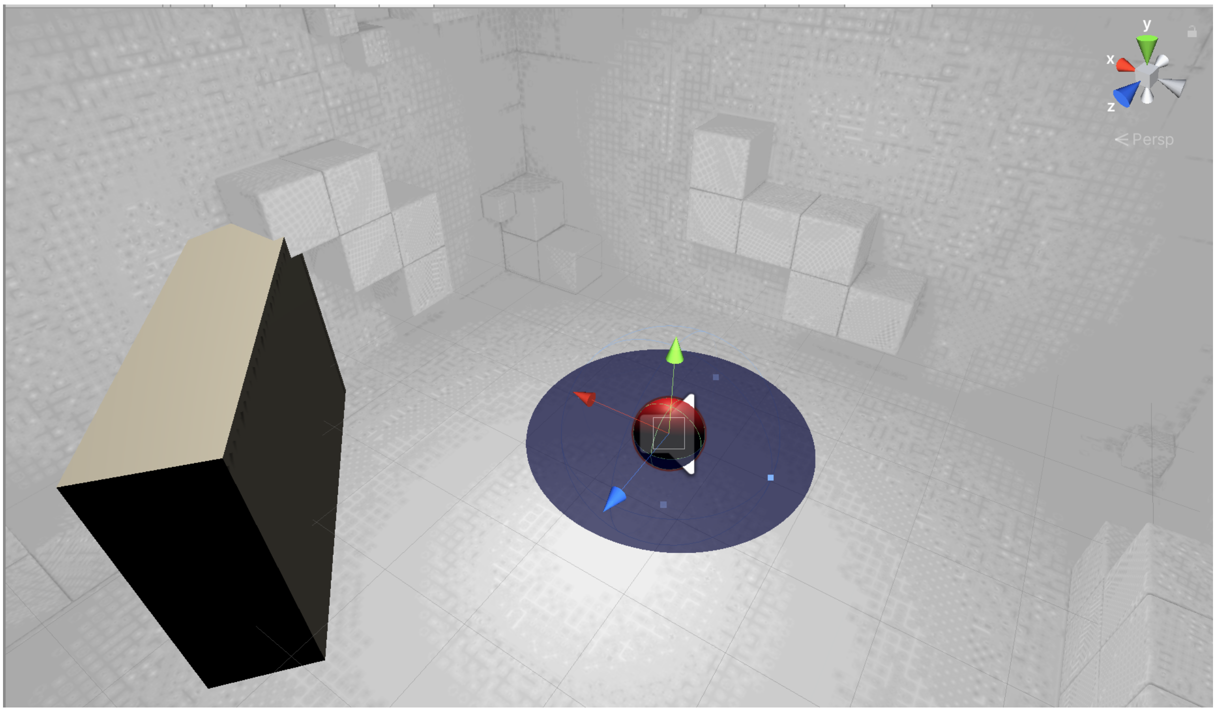
"Responses to questions specific to acoustic transparency. Overhearing: "I would be concerned that others could hear what I was listening to"; Notifications: "It would be more socially acceptable to hear audio notifications over acoustically transparent headsets instead of from my smartphone"; Key Feature: "Acoustic transparency would be a key feature I would look for in future audio purchases"; Day-to-Day: "Acoustic transparency would make it likely that I would wear such an audio device throughout my day-to-day life"; 'Audio Apps: "I would be more likely to use audio-based applications in my day-to-day life given an acoustically transparent headset."[101].


Ambient Audio
Podcasts
Immersive Podcasts
Music
Media
UI
Notifications
Spoken Notifications
Voice Assistants
Voice Conversations
Public
Workplace
Private
0%
25%
50%
75%
100%
None
Speakers
Headphones
Noise Cancelling Headphones
Accoustically Transparent Headwear

Appendix A2. Screenshots of three different reverb setups in Unity

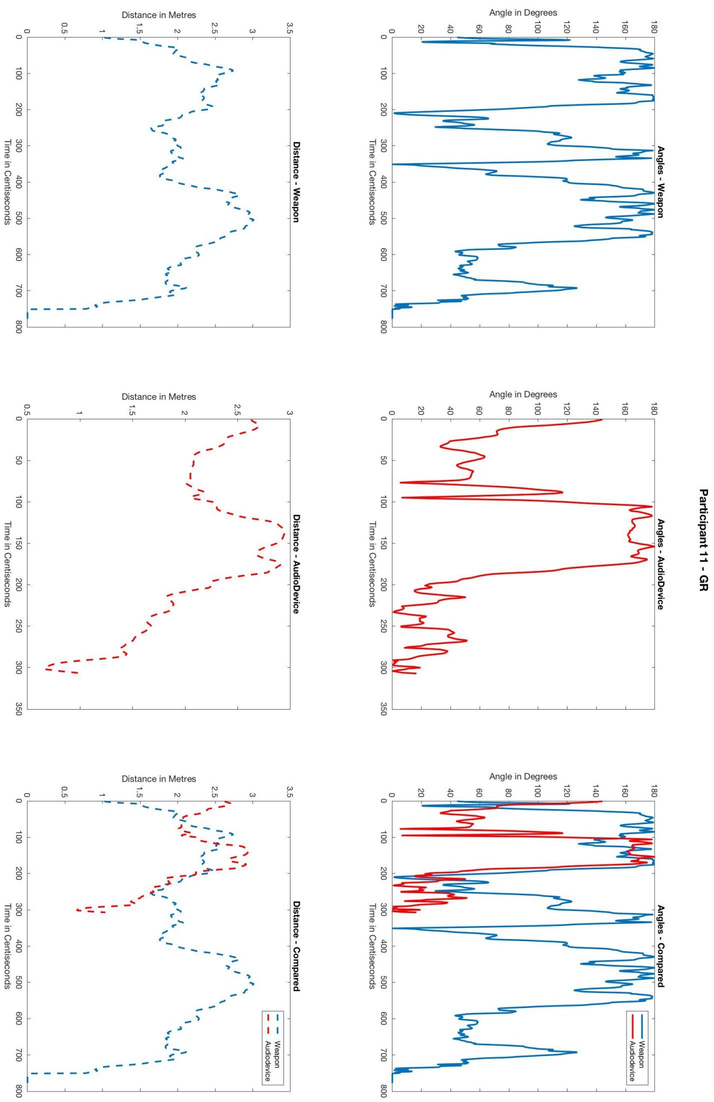


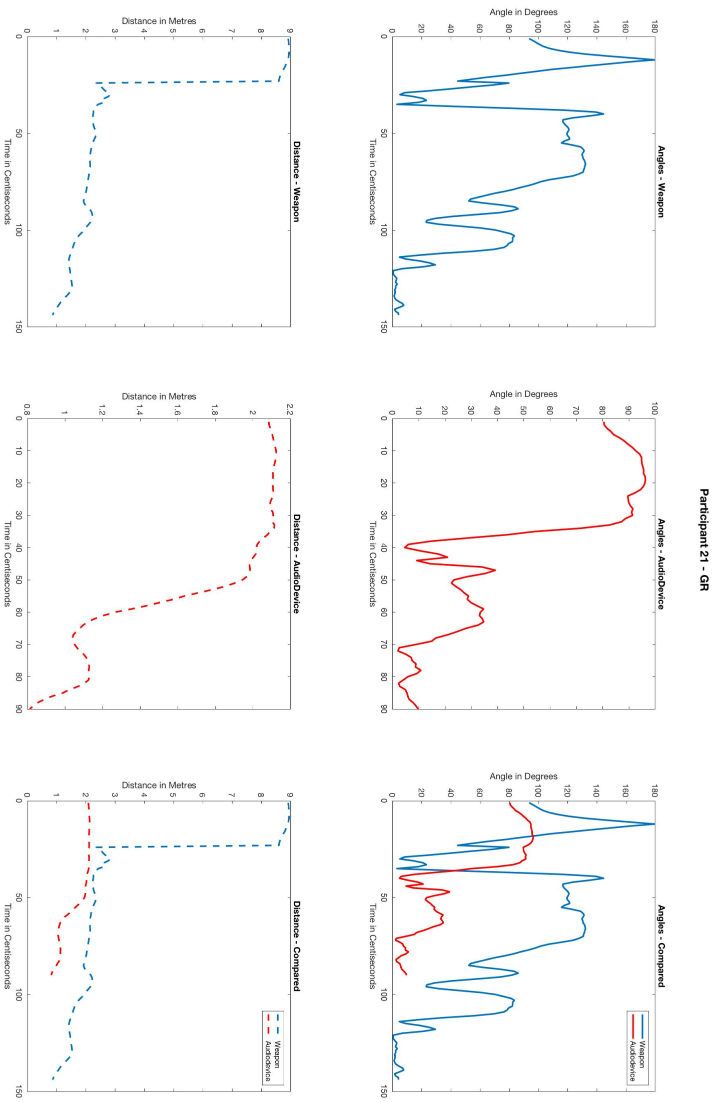



y
x
z
Persp

Appendix A3. Tree build (Day/Night)

Appendix A4. Distance and angle as a function of time, for multiple participants [110].




Participant 21 - GR
Angles - Weapon
Angles - AudioDevice
Angles - Compared
Distance - Weapon
Distance - AudioDevice
Distance - Compared
Angle in Degrees
Distance in Metres
Time in Centiseconds
Weapon
Audiodevice

Appendix A5. Questionnaire items with loadings and each category with average variance extracted (AVE) and composite reliability (CR).

Items denoted with (R) are reverse-scored

Construct	Items	Loading
Presence **Vorderer et al. (2004)** Spatial situation (SSPL) Self Location (SSM) Possible Actions (SPPA) Suspension of Disbelief (SoD) **Regenbrecht and Schubert (2002b)** Object presence adapted Audio presence (self-developed) **AVE: 0.564** **CR: 0.880** *Loadings for presence are from the parsimonious model in section 7.1.*	SSPL1: I felt as though I was physically present in the environment of the presentation	0.858
	SSPL2: It was as though my true location had shifted into the environment in the presentation	0.852
	SSPL3: It seemed as though I actually took part in the action of the presentation	0.825
	SSPL4: I felt like I was actually there in the environment of the presentation	0.950
	SSPL5: I felt like I was part of the environment in the presentation	0.861
	SSPL6: I felt like the objects in the presentation surrounded me	0.838
	SSM1: I had a precise idea of the spatial surroundings presented in the augmented environment	0.809
	SSM2: I was able to make a good estimate of the size of the presented space	0.792
	SSM3: I was able to make a good estimate of how far apart things were from each other	0.723
	SSM4: Even now, I could still find my way around the spatial environment in the presentation	0.676
	SSM5: I was able to imagine the arrangement of the spaces presented in the augmented environment very well	0.695
	SSM6: Even now, I still have a concrete mental image of the spatial environment	0.683
	SPPA1: I felt like I could move around among the objects in the presentation	0.561
	SPPA2: The objects in the presentation gave me the feeling that I could do things with them	0.847
	SPPA3: It seemed to me that I could have some effect on things in the presentation, as I do in real life	0.653
	SPPA4: I had the impression that I could act in the environment of the presentation	0.747
	SPPA5: It seemed to me that I could do whatever I wanted in the environment of the presentation	0.633
	SPPA6: I had the impression that I could be active in the environment of the presentation	0.795
	SOD1: I didn't really pay attention to the existence of errors or inconsistencies in the augmented environment.	0.430
	SOD2: (R) I wondered whether the augmented presentation could really exist like this.	Removed
	SOD3: (R) I took a critical viewpoint of the augmented presentation.	Removed
	SOD4: (R) I thought about whether the action or the augmented presentation was plausible.	0.763
	SOD5: (R) I directed my attention to possible errors or contradictions in the augmented environment.	0.852
	SOD6: (R) I concentrated on whether there were any inconsistencies in the augmented environment.	0.894
	OP1: I felt that I could touch or grasp the virtual objects	0.881
	OP2: (R) I paid attention to the difference between real and virtual objects	0.630
	OP3: I made an effort to recognize the virtual objects as being three-dimensional	Removed
	OP4: Watching the virtual objects was as natural as watching the real world	0.464
	OP5: (R) The objects appeared to be visualized on a screen rather than giving the impression that they were located in space	Removed
	OP6: (R) The objects appeared as flat images rather than three-dimensional	Removed
	AP1: It was easy to distinguish between real world- and virtual sound	Removed
	AP2: The sound seemed to interact realistically with the virtual objects in the environment	Removed
	AP3: (R) The sounds were hard to locate in the environment	Removed
	AP4: The virtual sounds fit the physical space	0.366
	AP5: The sound quality was good	0.853
	AP6: The virtual sounds seemed to be located in the physical environment	0.940
Motivation Scarcity adapted by (self-developed) **AVE: 0.692** **CR: 0.900**	**If I knew the AR experience was time-limited…**	
	MS1: I would find it more appealing than if it was a recurring event	Removed
	MS2: I would find it hard to decline or pass up	Removed
	MS3: And that everyone else had tried it, I would be more encouraged to experience this application	0.897
	MS4: And my friends also experienced it, I would feel pressured to engage in the same activity just to not miss out	0.868
	MS5: And that it was unique (i.e. only I could try it), I would be more encouraged to experience this application	0.788
	MS6: I would regret if I missed out on it	0.762
	MS7: (R) I would not necessarily go out of my way to experience them	Removed
	MS8: (R) I would only experience it if I happened to be in the area	Removed

Adoption TAM1 adapted by (self-developed) TAM2 & TAM3 adapted by (self-developed) **AVE: 0.428** **CR: 0.814**	**After having tried the AR experience...**	
	TAM1-1: I found the application was easy to use	0.657
	TAM1-2: (R) I found it hard to navigate around the virtual objects	0.632
	TAM1-3: I found that using the application did not require a lot of mental effort	Removed
	TAM1-4: I can see the map being useful in locating other virtual environments	0.832
	TAM1-8: I would use the application as a way to navigate specific areas while also exploring any virtual environments nearby	0.642
	TAM1-9: I found the sound quality of the headset to be sufficient for the application	0.537
	When learning about the AR experience...	
	TAM2-2: (R) I initially felt insecure about my ability to interact properly with such technology	Removed
	TAM2-3: I saw no problem initiating the activity required in front of people, on my own	0.583
	TAM2-4: (R) I felt insecure about using the technology when strangers were around	Removed
	TAM2-5: I would prefer using my own device to perform the activities required, in a public space	Removed
Fear of Missing Out (FoMO) adapted by Przybylski et al. (2013) **AVE: 0.457** **CR: 0.851**	Q1: I get anxious when I don't know what my friends are up to.	0.559
	Q2: I fear my friends have more rewarding experiences than me.	0.779
	Q3: When I have a good time it is important for me to share the details online (e.g. updating status)	Removed
	Q4: When I go on vacation, I continue to keep tabs on what my friends are doing	Removed
	Q5: I fear others have more rewarding experiences than me.	0.759
	Q6: I get worried when I find out my friends are having fun without me.	0.830
	Q7: It is important that I understand my friends' "in jokes."	0.558
	Q8: It bothers me when I miss an opportunity to meet up with friends.	0.531
	Q9: When I miss out on a planned get-together it bothers me.	Removed
	Q10: Sometimes, I wonder if I spend too much time keeping up with what is going on.	Removed
	When learning about the AR experience...	
	FOMO1: I felt more compelled to engage with such technology if my friends also did	0.649
Enjoyment (self-developed) **AVE: 0.469** **CR: 0.853**	ENJT1: I enjoyed using the application as a way to experience the surrounding urban area	Removed
	ENJT2: I enjoyed experiencing the virtual environment (audio and visual)	0.895
	ENJT3: Finding the virtual objects was fun	0.672
	ENJT4: I enjoyed engaging with the objects	0.915
	ENJT5: It was interesting how the sound matched the visual setup of the objects	0.643
	ENJT6: (R) I found it boring that I could not interact with the virtual objects	0.518
	ENJT7: I was excited to see how the environment would change at a later point in time (e.g. during the day or night)	0.593
	ENJT8: I'm excited to share any photos I took of the experience with my friends	0.402
Continuance intention and appropriation (self-developed) **AVE: 0.523** **CR: 0.920**	**If you had to imagine the events were more extensively built and placed in multiple areas, consider the following questions:**	
	CIA1: I intend on using this application on a regular basis to be able to explore the changing environments	0.905
	CIA2: I would be interested in finding and exploring the rest of the augmented areas	0.822
	CIA3: (R) I don't expect to want to continue experiencing augmented locations and activities in the future	0.654
	CIA4: I am willing to try this experience in the future	0.657
	CIA5: (R) I can not see myself using the application in the future	0.587
	CIA6: (R) Time-limited AR events would not incentivize me to use this application in the future	0.497
	CIA7: I can see myself using the application to explore new areas of my city I haven't explored before	0.834
	TAM1CIA-5: I would be able to use it in the future to discover different areas around me	0.764
	TAM1CIA-6: I would venture outside more to explore virtual environments	0.834
	TAM1CIA-7: (R) I can not see myself going out of my way to explore the virtual environments on the go	0.372
	TAM1CIA-10: (R) I can conclude that this type of application is not for me	0.837

Appendix A6. Part of script controlling cyclicity of the objects

```
public void ChangePrefab(GameObject prefab, GameObject previousPrefab)
{
    prefab.layer = 8;
    previousPrefab.layer = 9;
}
```

Main function for changing between active objects. The layer determines whether the camera can see it or not. The camera's culling mask is set to only see objects of layer 8, and layer 9 is not is not visible.

```
public IEnumerator ChangeToFirst()
{


    obj1.SetActive(true);
    prefabChildren[0].SetActive(true);
    currentPrefabFomo = obj1;
    Debug.Log("AmongUs Active");
    obj2.SetActive(false);
    prefabChildren[0].layer = 8;
    prefabChildren[1].layer = 9;
    prefabChildren[1].SetActive(false);
    isAmongUs = true;
    isSolar = false;
    ChangePrefab(obj1, obj2);

    if (isSolar == true)
    {
        obj1.SetActive(true);
        prefabChildren[0].SetActive(true);
        obj2.SetActive(false);
        prefabChildren[1].SetActive(false);
    }

    obj1.transform.GetChild(0).gameObject.SetActive(true);

    obj2.transform.GetChild(0).gameObject.SetActive(false);
```

```
yield return new WaitForSeconds(fomoTimer);

StartCoroutine(ChangeToSecond());
```

Screenshot of the first coroutine ensuring that only the first scene is active. The change happens when the timer (fomoTimer) runs out and initiates a new coroutine for the next scene (ChangeToSecond()).

Appendix A7. Hotspot activation based on distance

```
public void Update()
{
    distanceToAmong = Vector3.Distance(distCamera.position, distAmong.position);
    distanceToSolar = Vector3.Distance(distCamera.position, distSolar.position);
    distanceToNeon = Vector3.Distance(distCamera.position, distNeon.position);

    if (amongScene.tag == "Hide")
    {
        amongScene.SetActive(false);
    }
    else if (amongScene.tag == "Active" && CyclePrefabs.GetComponent<CyclePrefabs>().isAmongUs == true)
    {

        amongScene.SetActive(true);


    }


    if (solarScene.tag == "Hide" && CyclePrefabs.GetComponent<CyclePrefabs>().isSolar == true)
    {
        solarScene.SetActive(false);

    }
    else if (solarScene.tag == "Active")
    {
        solarScene.SetActive(true);

    }


    if (neonScene.tag == "Hide")
    {
        neonScene.SetActive(false);

    }
    else if (neonScene.tag == "Active")
    {
        neonScene.SetActive(true);

    }
```

Distance is measured from the object to the device's camera and activates/deactivates it

```
            if (distanceToAmong >= range)
            {
                //amongScene.GetComponentInChildren<Renderer>().enabled = false;
                muteOne(true);
                amongScene.tag = "Hide";
            }
            else if (distanceToAmong < range && mapActive == false && CyclePrefabs.GetComponent<CyclePrefabs>().isAmongUs == true)
            {
                //amongScene.GetComponentInChildren<Renderer>().enabled = true;
                muteOne(false);
                amongScene.tag = "Active";
            }
            if (distanceToSolar >= range)
            {
                muteTwo(true);
                solarScene.tag = "Hide";
            }
            if (distanceToSolar < range && mapActive == false && CyclePrefabs.GetComponent<CyclePrefabs>().isSolar == true)
            {
                muteTwo(false);
                solarScene.tag = "Active";
            }
            if (distanceToNeon >= range)
            {
                muteThree(true);
                neonScene.tag = "Hide";
            }
            else if (distanceToNeon < range && mapActive == false)
            {
                muteThree(false);
                neonScene.tag = "Active";
                //TODO:::: ADD CULLING (Layers)
            }

        }

}
```