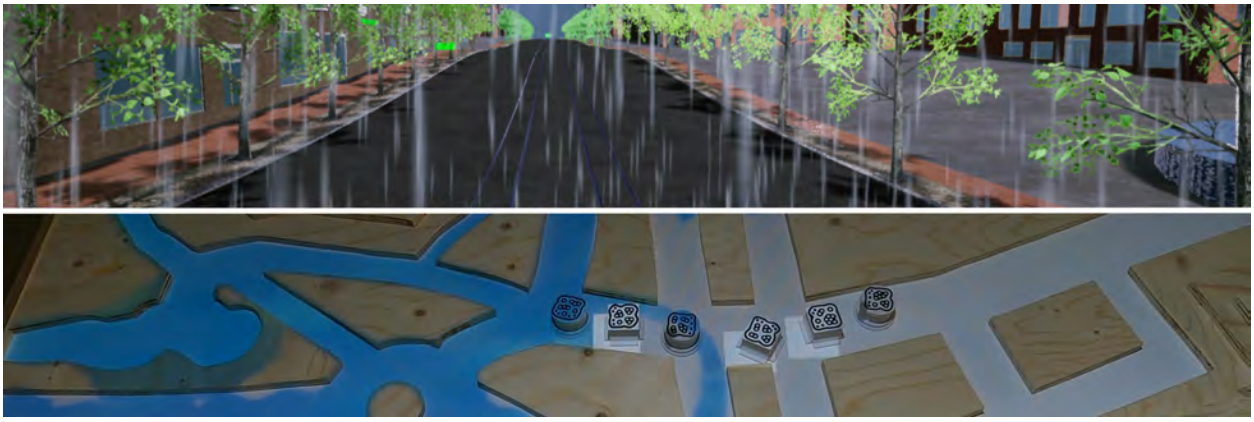


Exploring tools for interactive visualizations in physical and virtual space

An evaluation study between visualization technologies used in urban planning and development

Master's thesis in Computer science and engineering

VICTOR LÖFGREN NORRMAN

HENRIK NILSSON

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2020

Master's thesis 2020

Exploring tools for interactive visualizations in physical and virtual space

An evaluation study between visualization technologies used in urban planning and development

VICTOR LÖFGREN NORRMAN & HENRIK NILSSON

UNIVERSITY OF
GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2020

Exploring tools for interactive visualizations in physical and virtual space
An evaluation study between visualization technologies used in urban planning and development
VICTOR LÖFGREN NORRMAN & HENRIK NILSSON

Supervisor: Josef Wideström, Computer Science and Engineering
Advisors: Kristina Knaving & Peter Ljungstrand, RISE Prototyping Society
Examiner: Staffan Björk, Computer Science and Engineering

Master's Thesis 2020
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Gothenburg, Sweden 2020

Exploring tools for interactive visualizations in physical and virtual space
An evaluation study between visualization technologies used in urban planning and development
VICTOR LÖFGREN NORRMAN & HENRIK NILSSON
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

Abstract

Citizen dialogue in urban planning projects does not guarantee a successful process and can only be considered a valuable instrument if used intentionally, rather than as a goal in and of itself. The interface between large-scale projects and the citizens they aim to involve is essential to create a common ground for dialogue, where factors like structure and design of the interaction are often overlooked.

Present thesis attempts to outline the most significant aspects to consider when designing interactive visualization interfaces to support citizen dialogue. By adhering to a compound evaluation process of user experience and cognitive science, the results are grounded in data while substantiated by relevant factors for human information processing. Through an iterative design process, two interfaces were created, Virtual Reality and Projection Mapping, to mediate and facilitate an evaluation process from which fifteen design guidelines were derived. The study revealed a number of differences between the technologies, but also between user groups of citizens and urban planners. The gap between the users' desires and the experts knowledge of the optimal conditions of interaction pose a notable challenge to design for. The significance of this study appear in this discrepancy, where the mismatch is likely where dialogue continuously fails. Conclusively, the main contribution of this thesis is the novel outlook on the design of interactive technologies for urban development through a compound evaluation of data and scientific theory.

Keywords: Urban planning, information visualization, citizen dialogue, virtual reality, projection mapping, cognition, evaluation, design guidelines.

Acknowledgements

We want to express our deepest gratitude to the people who have supported us throughout the work with this theses. Thank you Josef Wideström, our supervisor and always so insightful support in discussions and at design crossroads. To Per Ottosson - the Touchdesigner genius, who was an invaluable asset in the development phase of this thesis. To our advisors at RISE, Kristina Knaving and Peter Ljungstrand, for giving us the opportunity to spend our time at your playful and creative studio - working in such an inspiring environment has contributed significantly to the quality of this thesis. Also, a thank you to Eric Jeansson and the city's department of urban planning, for sharing your 3D model of Linnéstaden with us.

Finally, we want to thank our family and friends for the love, patience and support throughout the past five years, without which none of this would have been possible. Sorry for all the cancelled plans and weekends, we promise we will make it up to you.

Victor Löfgren Norrman & Henrik Nilsson, Gothenburg, May 2020

Contents

List of Figures

1

Introduction

Adapting to a more digitized society comes with many challenges. Among them, the question of how to involve citizens in the urban planning processes to a higher degree, as well as a desire to work more efficiently across sections within city management and concerned industrial actors in order to better analyze and make well-informed decisions. Gothenburg City together with Visual Arena at Lindholmen Science Park have been assigned the task of creating a testbed with the purpose of exploring, developing and implementing various digital technologies intended to support different aspects of urban development and planning through the means of visualization. The testbed is intended to pave the way for new possibilities of cross-sectional collaboration and co-creation, a consolidating force between the public and business sector, research and civil society, and hopefully facilitate dialogue and create opportunities for democratizing urban planning processes [5]. This project is a part of an initiative within Gothenburg City to foster innovation and ability to reform. The city stands before significant challenges within urban development and new methods of interdisciplinary effort are required to manage the changes to come.

The development of infrastructure, institutions, buildings and communities is a complex process that can have large and often unforeseeable environmental, societal and political impact on its inhabitants. Implementing projects within urban development is a process in which the majority of all planning needs to be accounted for before the execution. Once in this phase, the process is often static, without much room for structural changes or an iterative workflow. Measures have been taken in several urban planning projects to minimize the risk of undesired consequences that are noticed too late to adjust. For example, the urban planning community has in large started to adhere to the concept of participatory design, in which citizens are involved in early phases of the projects in order to create a solid information base upon which well-informed decisions can be made that are data-driven rather than based on assumptions [20]. After all, it is the citizens that are going to live with the results from these projects, and it is therefore wise to consider their needs and requirements when changing the environment they live in.

Urban development is a creative, technology-driven research field in which new visualization technologies are embraced and constantly developed to cope with the technological expectations that exist today. Conventional technologies such as 3D sketches have in large migrated into the digital space, and modern 3D visualizations can nowadays be complemented with digital data layers that adds information and complexity to the visualization. According to Wideström and Eriksson [24], virtual models are among the most common technologies used in urban development to stimulate citizen involvement. It is clear that visualizations are important for mediating information about urban development since it concerns the physical space in which a city and all its inhabitants operate. The question is, what context provides the optimal preconditions for successful and meaningful use of a certain type of visualization technology, and how would you go about evaluating it. Visualization evaluation can be done in many ways, and the method should depend on the goal and purpose of it [50]. No single method of evaluation can or should be applied to every project, since the objectives of different projects are never identical [36].

As the levels of complexity and knowledge domains differ profoundly between and within concerned sectors of urban development, and it is well-known to be a top-down process [79], attention needs to be brought to the significance of understanding the user. This vacuum can be accommodated by user experience studies, an iterative design process and application of theories on human cognitive processing. Interactive technology tools possess the ability to create interfaces for acquiring and distributing information in novel ways, facilitating participation and generating engagement to a larger extent than traditionally used methods [67]. The iterative interaction design process is important to apply in projects of this magnitude, wherein wicked problems are a commonly appearing type of challenge. With the user in focus, this is countered to the extent it can be, since the process becomes data-driven and oriented after the stakeholder. Public officials, decision makers and civilians have differing qualifications that need to be taken into account in order to create a successful interaction. By extension then, the type of visualization technology also needs to differ in order to match their mental model of whatever the goal of their particular user group is.

The ambition of this thesis is to address these issues from an interdisciplinary perspective of interaction design and cognitive science. The interaction design approach is fundamental to all fields concerned with research and design of information technology systems. The core resides in the understanding and shaping of behavior, and the task of matching that between user and system. Establishing requirements, designing concepts to fulfill them, building prototypes as conceptual representations and evaluating the prototypes, with cascading iterations, is a resilient approach to wicked problems such as those in urban development. The process starts in the user, includes the user during, and ends with the user, unlike urban planning, which often starts in the grasp of decision makers and public authorities but still ends with the user. There is a tendency of letting the engineering perspective prevail, of prioritizing cost, durability, structural aspects, environmental restrictions or matters of construction. All of absolute relevance, but void of methods for managing the user experience. Countless examples exist in the software engineering world where the expert, or engineer, designs the interaction and end up presenting an interface based on their intricate conceptual understanding of a system, completely lacking connection to the actual user's mental model [15]. The user's conceptual understanding is generated by the minimal path of spent cognitive resources to enable them to interact successfully with an interface. A user-centered approach coupled with continuous evaluation ensures that the direction is aligned with the goal of creating something appealing, engaging, useful, usable, valuable, or what other requirement the end-user may express.

Involving human information processing in urban development is key to understanding the social dynamics of participation, the value of distributed cognition in interactive artefacts, how attention and memory affect knowledge acquisition, the benefit of rotating information between sensory modalities, heuristics and biases, among other factors. Urban planning and citizen participation is an innately social activity. Acknowledging and utilizing learning elements like shared attention, imitation, play and embodied interaction is principal to creating a favorable learning environment [35]. Likewise, multisensory integration of information has a major influence over attention, perception, memory, learning [63], meaning that experiences shaped to include several senses will be processed to a greater extent. This premise is compatible with Gestalt laws, which claims that we perceive the world with a holistic lens where we strive to organize bits into meaningful wholes. From this theory, design principles have been derived that are important to the shaping of a user experience since it dictates how we patternize our perception [80]. Incorporating aspects of cognitive science into urban planning will provide perspectives on human information processing that are not captured by the interaction design process, serving as an ample complement.

1.1 Project Background

In order to improve the capacity for innovation of Gothenburg city, the project Virtual Gothenburg Lab is intended to provide new opportunities to work across sections within urban management, develop new work procedures, and create the right conditions for establishing digital and virtual technologies as instruments. The planned outcome of the project, running from January 2020 to January 2022, is a testbed, wherein development and implementation of digital visualization technologies that supports urban planning and development can be carried out. It will serve as a means of unity and cooperation between public sector, industry, the science and research community, as well as civil society from a regional scale to an international scale. Within these parties, stakeholders and target groups vary from regular citizens to public officials and decision makers, each with different needs. The project is owned by Visual Arena at Lindholmen Science Park, together with a range of different actors from above mentioned sectors, and is intended to provide a basis for various development projects concerning urban planning. To support future use cases, the project will make use of themes and pilot areas in order to assure process accuracy and knowledge of using the testbed. The different themes will handle different valuable aspects, such as how to initiate citizen dialogue and increase participation to the benefit of democratization and learning.

The focus toward urban development was determined because of a pre-established need, and the fact that other projects within the city are working in parallel with a similar scope, such as the Gothenburg Digital Twin [14, 42]. In preparation of the project, an analysis of existing comparable projects was carried out. The findings revealed that there have been no previous attempts at creating a testbed of the described type, although similar efforts exist. Singapore has successfully created the first complete digital twin of the city [60], a dynamic three-dimensional model and collaborative data platform. The platform will be used for long-term planning and decision making, such as handling the infrastructural impact of Singapore's aging demographic or improving the wireless network coverage of the city, in order to make simulations and projection for analyzing costs and feasibility. Another ongoing large-scale project in the field that emphasizes the importance of participatory development processes is Augmented Urbans [6], within which actors from both academy and industry from five EU countries in the Baltic region are involved . There is also a growing trend towards the elicitation of new technologies to better support the urban planning process, both as a platform for guiding discussions, and for creating immersive experiences to better understand different perspectives in different urban development phases. In Augmented Urbans, Extended Reality (XR; which includes virtual, augmented, and mixed reality) is explored as a set of technologies for urban planning, with a partial aim of revealing the potentials and limitations of XR as a technology for urban development [6]. Gävle Innovation Arena (GIA) is another example, working to create a collaborative platform for innovation and research based solutions for smart and sustainable urban development. They focus on developing a data lake, where data, normally exclusive to a specific organization or source, is made accessible to projects within urban development [31].

With a focus on exploration, simulation, collective investigation and understanding, Virtual Gothenburg Lab has the intention to investigate the use of interactive technologies such as virtual reality, augmented reality and projection mapping. There is also a desire of visualizing and simulating effects and consequences of decisions to support critical user analysis. Through these methods, the testbed is intended to provide support for, insight, planning, decision-making, learning, co-creation and participation. The project proposal mentions that the project should strive to enable users to learn and understand through shared cognition, attention and collaboration. It is assumed that a successful implementation of these things in a design would enrich user experiences, since mental resources would be free to focus on such things instead of actively having to establish shared cognition.

In detail, by project end the aim is to have established, tested and evaluated the testbed. This will be carried out through careful examination of the visualization technologies, the methods of citizen dialogue, the competences necessary for operation and development as well as having the right infrastructure in place. It will do so by execution of five tests, where visualization technologies and methods are implemented and evaluated, each adhering to one of the following pilot areas. City Core - The core of the city should be characterized by a rich street life with a focus on culture and green areas. This case looks into how the city environment is experienced from a point of view of safety, noise, and overcrowding, by visualizing or simulating interactively how such an experience might unfold. CoExist - Autonomous vehicles are and will become a more common occurrence in the city. A core issue concerns how these vehicles will interact with their environment and other means of transport such as bicycles and pedestrians. By the use of simulations, urban and traffic planners together with vehicle manufacturers are able to communicate to make sense of data, in order to optimize for the well-being of the city. Rains - There is a need and desire to develop new tools for visualizing the negative effects of heavy downfall through key numbers, evaluation of countermeasures and making well informed decisions. Also, a will to compare these decisions to costs and effects of noise, mobility in the city and the value of green areas. Social sustainability - Segregation of living communities has been brought to attention as a major challenge for sustainable urban development. This case involves visualization of experiences of community squares in participation with its inhabitants, to create sustainable and equal environments. Lastly, City Skyline - As the city grows, the right conditions for experimentation, analysis and impact regarding the city silhouette need be present in order to preserve its history and characteristics. With the right tools for visualization, analysis and communication can facilitate this process to act and plan accordingly. Out of these five pilots, present thesis will focus on *Rains*.

1.2 Approach

The thesis will be carried out as a part of the Virtual Gothenburg Lab project, but only attending to one of the pilot areas and with focus on evaluating visualization technologies. This will serve as an important precondition to the development and subsequent implementation of the parts of the testbed that address interactive technology as a medium for dialogue.

Initially, a literature review of the previous research will be carried out, covering the core issues of urban development and a combined theory section of cognitive science and user experience methodology, to provide grounds for analysis of the results. An overhaul of the commonly occurring visualization technologies within the field as well as the ones that have potential will be presented. Frameworks used for evaluation of previous work will be examined, as they could provide new methods for the current project to follow in the design process further on. Meetings will be set up with stakeholders to create a clear picture of what has been accomplished so far in pilot studies and previous connected studies. This will serve as the foundation.

The thesis will focus on carrying out a user study of the target group citizens, to map out their behavior, thoughts, and needs. The user research phase will be iterative, and approached with methods learned and practiced at the program such as focus groups, workshops, and interviews. This phase will be conducted with consideration of the objectives of the Virtual Gothenburg Lab project and with an outlook of identifying and mapping out visualizations technologies. Once these areas have been investigated, a demarcated path can be decided on to which the rest of the project will follow. Consecutively, visualization prototypes will be created, depending on the results of the user study, targeting different technologies that fall in line with the project's expected outcome without sacrificing insights from the study. The technologies will be developed in an iterative low-to-high fidelity format and subject to formative and summative evaluation. The summative evaluation will be based on a user experience perspective as well as a cognitive benchmarking in order to discover how they perform from a subjective as well as an objective point of view. That is, it is desirable to understand how the user will receive the experience since that is required for them to actively want to engage with it; but, it is of equal importance to determine how and if the information that is trying to be communicated is received. A mixture of qualitative and quantitative data often tends to give a nuanced picture of users' experiences and will be gathered to generate a base for conclusion and discussion of results.

Insights from user research and analysis of literature will be combined with concrete insights from evaluations of the visualization prototypes, to lay ground for the proceedings of the main project. The authors share a background in cognitive science and see the potential of supporting the interaction design approach with theories on human information processing. Its role will be to motivate aspects of the design that might not be self-evident from interaction design alone, such as the importance of multimodal interaction or the social interplay within a group of participants. Simply, utilizing the human capacities that go beyond the actual interface and experience. In addition, perspectives on what and how to evaluate the collected data will be of use when interpreting the results. By combining these perspectives, the thesis aspires to deliver a guiding document for the continued Virtual Gothenburg Lab project as well as for future projects of urban planning.

1.3 Stakeholders

Besides the main parts of the Virtual Gothenburg Lab project relevant to the thesis, Gothenburg City, Lindholmen Science Park and RISE Interactive, a number of other actors are involved. These actors represent different industry sectors and also academia and will naturally have various internal expectations on the project outcomes. Finding a balance as well as a common ground among actors of the project will constitute a challenge of its own. For example, representatives of the automotive industry, such as Volvo Cars, will naturally take interest primarily in aspects of the project that will contribute to car infrastructure and things that are relevant to their research field. Other actors might be more concerned with education and learning, such as Chalmers University of Technology, Universeum, and Ordrum, and might thus focus primarily on pedagogical aspects, and how to communicate findings that results from the testbeds. Lastly, several actors involved in the project are primarily focused on how to work with technologies for information visualizations to support the context of urban planning and city development, and will thus not interfere with plans regarding the purposes of the testbeds. However, these actors might have internal interests regarding software and hardware to be used, as they might enter the project with varying competencies and experiences regarding different visualization technologies. Actors involved with technologies for information visualizations are for example Berge, OutHere, Insert Coin, Atvis, and HiQ. Being able to distribute activities to the "right" actor will thus pose a challenge that requires a thorough understanding for their respective interests, for the project to gain most from their competencies and knowledge.

Other sectors represented in Virtual Gothenburg Lab are architechture (Tengbom), infrastructure (ÅF, Göteborgsregionen), and IT- and innovation consulting (Unicornsulting and Ramböll).

1.4 Scope

Whereas Virtual Gothenburg Lab is an economically resourceful large-scale project with a long timeline - present master thesis is not. Consequently, some limitations must be established in order to be able to deliver results that will prove useful for the community at large, and contribute with formative information regarding visualization technologies to Virtual Gothenburg Lab for their future progress.

Firstly, present study will only cover one out of several use cases, meaning that the prototypes developed will only visualize information relevant for that use case. The prototype process will have a vertical focus where narrow functionality is prioritized before a broad representation of features, in order to successfully be able to target and evaluate the essential properties. The aim is not to develop a prototype that aims to solve a multitude of visualization challenges, but rather to create one that can be compared to commonly used technologies from a cognitive, as well as an interaction design perspective.

The evaluation will not be exhaustive on technologies. Urban planning is a technology-driven field wherein a plethora of technologies are used for visualization purposes. The preceding literature study will serve to narrow and delimit the scope to include a few relevant ones. Thus, the aim is not to treat all technologies from the field, but rather to 1) present a new way of evaluating visualization technologies that combines cognitive processes with values of usability and user experience, and 2) outline the premises for when to use a visualization technology before others. A challenge in comparing technologies with few shared attributes in terms of physical properties and interaction opportunities is to find a common ground upon which they can be judged on the same basis. Present thesis will therefor follow a data-driven approach in which user interests inform the design process, combined with a theoretical grounding to support rationales regarding cognitive aspects of the developed interfaces.

1.4.1 Research objectives

What do we want to achieve?

- Present thesis aims to contribute to the field of Urban Planning by applying a data-driven user-centered process derived from the field of Interaction Design to evaluate the capacity for different visualization technologies to support citizen dialogue.
- Develop an evaluation method that can be used to determine the efficiency and value of a visualization technology in relation to another, irrespective of their level of similarity.

Research questions

1. How can projects within urban planning be improved in terms of citizen interest, by adhering to a compound evaluation process of user experience and cognition?

2. Which are the key factors for creating engaging, valuable and meaningful interactive visualization interfaces to support citizen dialogue in the context of urban planning?
3. How do the visualization technologies of focus relate to each other, what attributes set them apart, when do they prosper, and under what circumstances do they underperform.

2

Background and Previous Research

Urban Planning (UP) is a large and versatile research field, homogenous by no means. The interdisciplinarity of the field brings methods and insight from several related scientific areas such as architecture, interaction design and behavioral sciences to be combined in an infinite number of combinations. Consequently, no project or study within UP is like the other, and the accumulated knowledge within the field stems primarily from case studies, as in many other design fields. In UP projects, there are typically several different groups of people involved, each with their own needs and demands. It is often not easy to even generalize about the intended end user, since end results are designed to remain physically present for people of the future, and to be used by various target groups.

Present section makes an effort to narrow down the research field to the previous work within UP that the authors deem relevant to the thesis. The presented research covers topics relevant for different phases of the thesis' timeline, such as visualization technologies, methodologies in large-scale projects, participatory design, ways of evaluating designs and more. Doing so will provide the reader with the research context to which the authors wish to contribute, and the presented previous research will serve as a rationale for the design decisions made along the way.

2.1 Involving Users

Having its roots in Scandinavian working environment philosophy in the 70s', where workers were provided the opportunity to influence their own workspace and the systems they worked with, Participatory Design (PD) have developed into an important field in Human-Computer Interaction [36], and influenced the way designers think of their users. Although PD has been a well-used method within interaction design for decades, it is not yet common to apply to large-scale projects such as urban planning [18, 62]. Having seen the great effects of PD within interaction design, many have raised their voices in favour of the implementation of PD in urban planning as well [24, 74].

There are several cases to be made in favor of PD in urban planning, as will be presented below, but there are also bad examples when large-scale projects have been lacking citizen participation, that clearly demonstrates the downsides from not having a participatory approach.

Being called "an 'extravagant folly' unfit for its purpose" by French President Jacques Chirac, the multibillion-dollar 'Biblioteque National' in Paris has been widely criticized for its design that put spectacular aesthetics before functionality [56]. The planning process of the national library has turned into a prototypical case of how not to carry out large-scale design projects, embodying the traditional "top-down" process, that follows a sequential "waterfall model", exclusive of user participation in the planning phase [18]. The library was built without consulting librarians, something that led to the designers not understanding the target group, and the design being dysfunctional for their needs [56]. Neither had the designers considered the needs of all types of visitors, as the library was not sufficiently accessible to wheelchair users.

In another case, Jan Gehl [32] describes the " people landscape in Brasília ... [as] an abject failure. City space is too large and utterly uninviting, paths are too long, straight and uninteresting, and parked cars prevent pleasurable walking in the rest of the city.".

Gehl discuss urban design and city planning as work that involves planning on different scales. At the large scale, a holistic approach is required, since infrastructure, traffic and quarters are considered here. At the middle scale, individual areas of the city are considered, and the positions of buildings are decided on. At the small scale, the human scale, citizens are considered, and the human landscape is to be designed. On this scale, the city needs to be experienced and approached at eye level, and urban designers must be sensitive to human values in order to design landscapes that will add to citizens' well-being. Gehl proposes that urban planning procedures should follow the principle: "first life, then space, then buildings", in which the planning process occurs after the desired social atmosphere and way of life has been established [32].

> “*Working with the small scale is the key to ensuring better conditions for the human dimension.*”
> - Jan Gehl, 2010

2.2 Participatory design in large-scale projects

When a new culture house was planned for in Swedish municipality Lundby, an early approach to meet user involvement was to develop a Virtual Culture House, that was supposed to: "realize a set of virtual spaces where cultural activities and expressions can take place and later complement the physical culture house, inform and support the ongoing design process of the culture house, and develop an identity for the culture house..." [25].

The Virtual Culture House was a web-based portal onto which communication between the municipality and citizens regarding the culture house could take place, encouraging visitors to engage in media-sharing and open communication at the site. The website was organized in different categories, including "Events", where cultural events taking place were marketed, "Projects", where local cultural organizations were presented, and "Rooms", which included thematic virtual exhibitions, in the form of pictures, text collections, video, and sound files. In these exhibition rooms, visitors were able to post their own media files, or share their opinions about already posted files, in a forum-like place, visible to other visitors. This platform for cultural exchange and, more importantly, its social constituent, generated data on the level of appraisal of different exhibitions, along with other opinions valuable for decision makers in later design phases of the physical culture house. Another way of informing the design phase of the physical house on the website was through "The Culture House Creator", in which visitors could create their own Culture house in an application, based on cultural activities. Visitors could define the proportions in terms of percentages that should constitute concert halls, exhibition areas, library, gardens etc. [25].

Involving citizens in activity planning, as opposed to detail planning of the culture house, goes in line with the process advocated by Dalsgaard & Eriksson [18], who found that users and architects/planners had different foci, with users focusing more on activities and services enabled by the design, whereas architects prioritized the building and visual details. The development of the new public library 'Urban Mediaspace Aarhus' was a decade-year long project that consequently followed a participatory approach [18]. Throughout the years, many participatory activities were organized, and interfaces were created to better engage and inform the citizens and future users of the public library building. One of these was "City Voices", which was an interactive table with integrated auditory scenarios that users could listen to. The scenarios were related to various projects in the city, and the interface allowed users to record their opinions of the projects directly. The recordings were later analyzed and incorporated into the project's core values, that upcoming activities were bound to follow [18].

Interviews with stakeholders such as architects and librarians revealed that an issue with the activities was the large amount of data generated that was hard to utilize [18]. Planning and working with participatory activities must follow a systematic procedure, with a clear aim for the activities, and analytical measures must be taken to translate raw data into meaningful pieces, comprehensible to the stakeholders and decision makers.

2.2.1 Challenges with participatory design

Participatory design in urban development projects faces many challenges that arise due to their innately, and highly, heterogenous target group - citizens. Dalsgaard [17] points to the challenges that come from working with people from different backgrounds, and with varying levels of insight and background knowledge in a project, but also to the fact that people's opinions change over time. What is perceived to be a problem today might not be an issue in the future, and user needs that are identified early in the process might prove insignificant even before the project has ended. Also, involving users is not a recipe for success, and requires extensive work in order to contribute to a project. As pointed out in Dalsgaard & Eriksson [18], activities with heterogeneous participants require that a shared language is established in order to obtain mutual understanding for project agendas and ideals. However, establishing a shared language is time-consuming and requires effort from all people involved. For most projects, common language is insufficient for handling technological or project-specific subjects, whereas the technical language of experts is too specialized to be used in diverse user groups [62].

Having a clear purpose for every activity that is being held will enable participants to focus on the right things and stick to the subject, regardless of their own agendas. Participatory design is both time consuming and costly, and needs to be performed with a solid plan in order to contribute to the success of projects in urban development. Oostveen & Van den Besselaar [62] argues that clear restrictions have to be established regarding the influence of citizens in participatory projects. In a case study, they utilized user participation to expand their insights into civic perspectives, and restricted participatory activities from treating subjects about functionality in an IT system.

A noticeable attempt to systemize participatory processes can be found in the AELIA-model, which is an acronym for attention, experience, learning, influence and action [20]. The model aims to address five phases of a participatory urban development process, from the first initial planning phase to implementation, in chronological order. The first and last phase is concerned with high-level strategies for communication between experts and citizens, such as getting citizens attention in the first place, making them interested in participating in activities, and extracting data from the activities, to support action by relevant actors in the project. The three middle phases are all concerned with ways of engaging the citizens, by providing them with novel experiences, introducing them to new information and thus stimulating learning, and by giving them real influence, and the chance to play an important role in the development [20].

2.2.2 Critique against participatory design

Critique against citizen participation has been raised, stating that the concept is inflated and politicized, and often subject to abuse [45]. The argument goes that the concepts is to loosely defined, and thus easy to tie to one's own process in an effort to embellish it. Other critique regards the uneven representation of citizen groups that attend to participatory activities, leading to several groups being left out of the process [41]. In addition, of the people that do attend to the activities, it is often true that those with the most extreme opinions are the one that makes themselves heard, affecting the results of the activities[41].

The complexity of involving citizens in urban planning projects must be dealt with by a pragmatic approach. Juarez & Brown [45] argues that participation does not have an intrinsic value, but its value should rather be weighed against its costs. Only then can the participatory process be evaluated. Clear goals for participatory activities must be articulated, to enable successful outcomes that match the goals. Otherwise it will lead to the further inflation of the concept, instead of being regarded as the instrumental tool it can be, when used efficiently.

2.3 Technological Exploration in Urban planning

In Augmented Urbans (Augmented Urbans), a large-scale EU project, Extended Reality (XR; which includes virtual, augmented, and mixed reality) is explored as a supporting tool for UP processes, with a partial aim of revealing the potentials and limitations of XR as a technology for urban development (Augmented Urbans News). For example, 3D models have been used to visualize in VR planned projects at their intended locations in order to make planning processes more transparent and accessible to citizens (Augmented Urbans: Riga; Augmented Urbans: Helsinki). In another project, AR technology is explored as a tool for transferring knowledge and skills in regard to the maintenance of plants and ecosystems, with the aim of educating citizens in biodiversity awareness in populated areas (Augmented Urbans: Gävle). Other related projects that take a holistic perspective on urban development are the so-called digital twin cities that are currently under development in several large-scale projects (Virtual Singapore; Virtual City Chalmers).

A digital twin city is a digital replica of a physical city that in addition to 3D-mirroring of the city's physical attributes include sensory data, visually represented in the virtual space (Virtual City Chalmers). The digital twin is scalable in a way that enables visualization activities in different phases of an urban development project, meaning that digital twins have the potential to satisfy all three scales defined by Gehl [32], if utilized properly. Making a digital twin city compatible with XR technology can even enable urban planners to switch between the scales in a single interface, something that is not possible with conventional technologies. In addition, a digital twin city that is updated in real-time with data from active sensors can let urban planners analyze the physical city's state in four dimensions (space and time; Digital Twins Chalmers), and conduct predictive analyses of future states, using historical data.

2.4 Interaction Technologies as Design Material

Urban Planning is a technology-driven field and, as such, can make use of IT as a design material. The degrees of freedom of IT are vast, revealing the inherent wickedness in such projects [70]. Limitations are more often due to computing power, server size and human creativity than lack of potential in the material. Löwgren & Stolterman [54] call IT the "material without qualities", and refers to the absence of affordances, that results in split standards in the field and differences in opinions of what methods and technologies should be common practice. Consequently, the field has become lush with various technologies that can be applied in different situations, from analogue clay models to highly technological tabletop interfaces, wherein graphical tools are combined with tangible devices, utilized through sensors and computational devices.

2.4.1 2D tools

When designing urban environments, spatio-cognitive properties must be taken into account, something that is challenging to do using 2D interfaces [51]. When representations of true, and proposed urban environments must be altered in order to fit 2D interfaces, spatial information and qualitative perceptions related to the lack of a dimension are lost, and users have to rely on static analyses of the environment, and numeric data instead. The most common 2D tools are screen-based visualization programs that combine maps with dynamic analyses that can be used to predict and explore consequences of alternate planning decisions. Presenting information in two dimensions is a large part of every project that is conducted, with information types spanning from text, sketches, charts, diagrams and pictures, to more complex 3D-rendered physical models and animations. In urban planning processes, several dialogue-supporting 2D tools have been implemented using map based interfaces. Urban Mediator [13] is a web-based interface that enables location-based discussions among citizens and from the community. The interface lets users create, obtain and share information that is categorized into categories, and tagged and portrayed on maps. The interface is accessible to anyone with access to the web and does thus not require physical meetings, much like the Virtual Culture House [25].

Community PlanIt [34] is an interactive online game platform for urban dialogue in which the user gets to play minigames such as trivia questions, and engage in social activities such as contributing to forum discussions and giving "likes". Doing this, the user earns different badges depending on the type of activity he has made. Game mechanics such as the point system will encourage users to be more active on the platform and thus contribute to the community in the form of suggestions, solutions and attitudes. 2D interfaces are good for the sharing and gathering of location-based communication, and can favorably be done at physical gatherings as well as from a distance [9]. Having shared attention enables seamless communication, and non-verbal communication facilitates understanding among the participants.

2.4.2 GIS

Geographic Information Systems (GIS's) are used to "computationally model, simulate, and analyze [the morphology of the terrain, the structure of vegetation, the built form, erosion by wind and water, gravitational forces, fire, solar irradiation, or the spread of disease] and their impact on the landscape" [64]. Needless to say, GIS's are innately complex, and learning to interpret and extract valuable information from them requires user skills. GIS programs are powerful tools for studying, predicting, and reshaping environments digitally, as part of the planning in urban development or landscape architecture. GIS interfaces are graphical, which adds an element of abstraction to the tool, since 3D forms have to be translated to a 2D interface and manipulated through non-natural artifacts such as mouse and keyboard. Petrasova et al. [64] argues that this unnatural interaction is unfit for designers, as it restricts creative thinking, spatial navigation, and learning.

2.4.3 Tangible interfaces

Tangible user interfaces (TUI's) draws upon the notion that physical affordances, rather than digital mimicries, utilizes human kinaesthetic intelligence and situated cognition, thus constituting a better starting point for natural and intuitive interactions [64, 76]. The tangible approach is "motivated by the desire to retain the richness and situatedness of physical interaction, and by the attempt to embed computing in existing environments and human practices to enable fluid transitions between 'the digital' and 'the real'" [76]. By equipping physical artifacts with digital properties through the use of physical-digital feedback loops, digital representations can be directly manipulated. This is manifested in Ullmer & Ishii's [78] early contribution to the TUI field: 'Tangible Geospace', in which small physical models of buildings are used to manipulate a 2D graphical map, that repositions itself accordingly with the models placed on it. This coupling of models to the graphical map allows the user to manipulate information presented in both 3D and 2D simultaneously. Although using physical models is a conventional technique within urban planning [24], digitizing them creates vast opportunities interaction-, computational- and visualization-wise.

Kobayashi et al. [48] created a similar setting as Ullmer & Ishii, but applied in a different context, namely disaster planning. With the aim of supporting collaborative planning in time-critical disaster situations, they projected a GIS generated map onto a tabletop that contained disaster measures based on predictive simulations. Users were able to control parameters of the disaster simulation such as: type of disaster (eg. earthquake, fire, tidal waves etc.), scale, location, capacities of shelters etc. using physical tokens. LC (inductor-capacitor parallel circuit) tags embedded in the physical tokens were detected and located through antenna arrays on the table, enabling it to register location, rotation and button-presses on the tokens. Users received instant graphical feedback of their actions as evacuee movement data was displayed, that dynamically adjusted to the current state of the map [48].

The Luminous Planning Table [7] made use of projections on a tabletop to represent aspects otherwise hard to analyze in urban planning such as wind, sun/shadows, and traffic flow during different times of the day. The projections dynamically changed as buildings were placed on the tabletop, enabling the users to explore how nature correlates with different buildings' height, placement, relative position, rotation, etc. [7]. Usability tests showed that the interface excelled at things that are normally cumbersome in the profession by "bringing the site 'to life'", but at the same time, concerns were raised that an overemphasis on physical aspects of urban areas redirects focus from qualitative aspects such as economy, society and political.

In CityScope [4] LEGO bricks constituted the modelling material being used. The bricks were tagged with chips readable to an overhead projector so that they, once placed in the modelled environment, provided input to the interface. GIS data was projected onto the interface, and could be manipulated through the replacement of bricks, or addition of new ones, making the interface highly interactive and changeable.

Other notable modelling materials that have been used in TUI's are: sand (Sandscape)[69] and clay (Illuminating Clay)[66]. In these projects, sensing technologies have been implemented to capture the surface geometry of the modelled environment, in order to map it to GIS data such as height, thus being able to simulate water flow etc. In Sandscape, an infra-red (IR) camera was used, whereas the Illuminating Clay used a laser scanner [69].

2.4.4 3D tools

Virtual Reality (VR) is a technology in which a user enters a 3D-modelled environment and experiences it from a first person's perspective. Being a technology that lets users experience rather than passively view the virtual environment [67], it draws upon rules and conventions from the physical world, providing the visitor with an intuition of how to interact with it. Through the concatenation of GIS and VR, urban planners are now able to visit and explore data-generated real environments on a larger scale, in which 3D views of images and 3D models are combined [57].

In a study examining the effectiveness of VR in participatory urban planning, van Leeuwen, Hermans & Jylhä [79], presented participants with prospective design variants of a public park in either VR or on a laptop computer rendering the same 3D model. Participants were not able to move in the virtual space, but could turn their heads 360°. Participants in the VR condition experienced significantly more engagement than those watching the computer screen, something that the authors argued is an important factor in participatory processes, wherein it is a challenge to recruit citizens to participate in activities.

In Augmented Reality (AR), the user is perceptually located in a continuum between Reality and Virtual Reality. The user experiences the real world, but with integrated digital information that is mapped onto it in real time, either through a digital screen, or AR-glasses. The product used for depicting the digital information must be an input-output device, since the recognition of physical objects or regions is important for the device to overlay digital information realistically and accurately [67, 65]. Ways of doing this is through using visual markers such as QR codes, GPS coordinates, or real time object recognition [65].

Calabrese & Baresi [65] aimed to elicit AR technology to communicate ideas and proposals of urban transformation to designers, students, and interested citizens. Using a 3D tracking software that recognizes objects in a three-dimensional space, a mobile application was developed, able to map objects even with the user moving the phone, by utilizing common mobile sensors such as oscilloscopes and compasses. Although stable to movement once calibrated, the start-up phase with the application was not accurate, which led the researchers to place external bluetooth transmitters around the area of interest, providing coordinates to the application that helped with the navigation and mapping of the environment. Once calibrated, users of the application could experience the physical environment through their device's camera, complemented with digital 3D models of buildings proposed to be built at their respective locations [65].

2.5 Evaluating Technologies

The most common technologies used for citizen involvement in urban development projects are virtual models, web services, social media, physical models or public hearings [24]. Out of these, social media and public hearings are the only innately interactive methods. Neither, however, is designed to mediate visual properties, although vision is a useful modality for urban planning. Physical and virtual models are methods that excel at this but instead lack innate interactivity, while web services possess the fundamentals for interaction and displaying visual media. The same reasoning can be applied to evaluatory features of these methods. Social media and public hearings are channels dedicated to receiving feedback but are disconnected from richer interactive media sources that could provide common ground for discussion, models are used for demonstrating purposes, and web services are used to distribute information. The distinct properties and availability of these methods make them a convenient choice, but rarely are the technologies themselves evaluated and studied in comparison to other methods of representing information. Reviewing the literature on visualization evaluation, a significant part appears to focus on tasks presented in 2D, e.g. graphs, pictograms, images, color, shape etc, in strict and controlled environments [12, 11]. In addition, the studies that seek to make statistical claims about visual and cognitive processing aspects have clear and defined limitations, making it challenging to formulate comparisons between technologies. While visualizations of any type is the most common technology used in efforts of urban dialogue, few of them are accompanied by evaluations of their usability [27].

Lam et. al [50] recognizes the complexity of evaluating visualizations and argues that the primary reason for this is that every visualization tool will support a variety of internal information processing of the user. Their comprehensive study of over 800 publications within the field of information visualization provides a framework of scenario categories and their typical evaluation strategies, rather than specific methods, in order to obtain a broad understanding of the alternatives that could match each individual study. They identify seven categories of these scenarios: visual data analysis and reasoning, user performance, user experience, environments and work practices, communication through visualization, visualization algorithms, and collaborative data analysis. More often than not, these processes overlap to some extent, which creates additional requirements that a single evaluation method rarely is able to manage on its own. On the other hand, clear research objectives ensure only some evaluation scenarios will match each case. To effectively be able to find the right objectives and define and delimit the evaluation scope, questions about what aspect of the interaction is valuable to the researcher should be formulated. In addition, there is need for a clear focus on whether the study mainly aims to evaluate the visualization itself (e.g. encoding and processing), or cognized products of the interaction (e.g. exploratory data analysis, user experience).

Billger and Wästberg [9] conducted a literature study of digital dialogue tools in contemporary urban planning processes and found that the majority are either 2D maps, 3D environments or game-based. They identified challenges of representing appropriate levels of detail and realism in virtual models, integration difficulties of qualitative and quantitative data during evaluation, and how to properly evaluate the user experience of the tools. Their study revealed an overrepresentation of evaluations being performed in simulated environments, while a few were carried out in real planning processes. Unsurprisingly, studies on the latter had a tendency to involve and engage users to a higher degree by letting them influence the real planning process, something the authors suggest ought to be an important factor for creating participation in urban dialogue. The evaluations performed in simulated settings focused primarily on usability and pragmatic characteristics, while the few real life studies focused on interaction and level of immersion of the experience.

2.5.1 User experience vs knowledge acquisition evaluation

Evaluating the user experience, in terms of pragmatic and hedonic qualities, of a visualization or technology is more common in the literature than evaluating the facilitation of learning or memorability of that experience [73]. From the ISO standard, a User Experience is defined by: “User’s perceptions and responses that result from the use and/or anticipated use of a system, product or service. Users’ perceptions and responses include the users’ emotions, beliefs, preferences, perceptions, comfort, behaviours, and accomplishments that occur before, during and after use.” Employing this definition of UX would mean that besides sheer usability qualities, internal goal states of the user is included, such as what they expect to achieve from an interaction. However, evaluation of internal and subliminal cognitive processes affected by attention, motivation, emotion, mental state, beliefs and so on, is not as straightforward as typical UX aspects like efficiency or ease of use. In order to do so, the components of what constitutes a knowledge experience have to be classified.

2.5.2 Information retention and comprehension

It is a complicated task to evaluate the extent to which knowledge or insight is obtained from experiences of one visualization tool to another. Visualization technologies differ in their compositions, features, pro’s and con’s, making them suitable for different situations (Lam et al. 2016). For example, comparing technologies that have access to interactive learning elements (e.g. VR) and regular visualizations that do not (e.g. video), can be questionable from an empiric perspective beyond simply stating their differences. Also, if a method depends on several participants interacting while another prevents it, then social factors may be the underlying cause. That being said, in order to draw conclusions about when to use certain technologies, some type of evaluation must be conducted. Conclusions from these types of studies often report results that tend more toward a correlatory nature, where the exact causes are not self-evident - Is the technology “VR” causing the effect, or is it the fact that the interface is embodied? Then again, if a technology or method innately supports this type of behaviour the results can be attributed to it, but a critical discussion should follow debating possible coefficients.

Depending on the objective of a visualization, a range of evaluation methods exist. Kuliga et al. [49] suggest that adjective based evaluation tools, such as semantic differential scales, can be used in order to assess how users experience the differences between a virtual reality environment and its actual reality counterpart. They mean that how participants respond and describe the experience depends on perceptual concepts that are subconsciously important, which enables a researcher to determine the causes of their subjective rating. With a similar evaluation approach, Van Leeuwen et al. [79] present a case of comparing the experiences of a virtual immersive 3D model and a screen-based rendering of the same 3D model in a civic participation effort to redesign a public park. Following the two scenarios, a twofold questionnaire was distributed to the participants, tuned to measure memory and recall of objects and differences in the scenes. The first part was a free recall task where the instruction was to note down all objects they could remember, and the second part was a recall accuracy task where items were described verbally and the participants would check off whether they had seen them. The results from the free recall task showed a significant difference in favor of the immersive scene, although the recall accuracy remained equal.

In Allcoat and Mülenens [2] study, knowledge acquisition was compared in three learning conditions, VR, video and regular textbook-style. The VR condition presented a fully interactive 3D model of a plant cell, which the participants could manipulate by highlighting, resizing and rotating, accompanied by a floating menu with detailed information about the different parts. The video condition used a recording of the very same view from within the headset that previous participants had experienced, ensuring that the components between those conditions would not differ. In this condition, the participant could navigate the video as they saw fit. The textbook-style condition used screenshots of the 3D model with all the information presented to them in separate views. The VR condition facilitated recall to a larger extent than the other conditions, while both VR and video conditions facilitated comprehension better than the textbook condition. These measures were evaluated using a pre and post test knowledge questionnaire. A discussion follows where they point out that active and passive learning may be the cause to the results, and that VR simply capitalizes on that fact.

In order to determine what information was encoded into memory and what aspects of a visualization attracted attention, Borkin et al. [11] investigated how and to what extent visualizations were recognized and recalled, by analyzing gaze patterns and text descriptions generated by their participants. To what extent a visualization can be remembered, they claim, has direct implications for how well it can be comprehended. A previous study of the same authors concluded that elements that are visually distinct and utilize properties such as color, visual complexity and familiar objects will improve overall memorability [12]. Their quantified results were in line with conventional visualization guidelines such as improved recall if the essence of the visualization was displayed in text, pictograms improved recognition, and redundancy facilitated the message transmission. Eye tracking data revealed that visualizations that are memorable “at-a-glance” are also more easily retrieved from memory. Concurrent with this retrieval, details about the visualization were also activated, concluded by participants descriptions being of higher quality for these cases. In the same vein, Saket et al. [73] measured time and recall accuracy between node-link diagrams and map-based visualizations and found a significant difference in performance accuracy favouring the latter. The effect was proven to persist several days past the original exposure. Although controlling for confounding factors such as font-sizes, color and recall format, the authors open up for further research of the possibility that the map-based visualizations were more memorable due to engagement factors rather than more specific features of the method.

A large scale visualization study [12] on what properties of a visualization make it memorable arrived at the conclusion that unique visualization types had a significantly higher memorability score. The study employed Amazon's Mechanical Turk to process 2070 single-panel visualizations involving infographics, charts, graphs, pictorials, grids/matrixes and more, with sources ranging from news media sites to scientific publications. The task was constructed as a simple image matching format, where a sequence of images were shown to the participant and once an image reappeared more than once the participant would flag it. The study finds that unique visualization types with high visual density, as opposed to clear and more conventional ones, were more memorable to the participants. They reason that this may be due to the fact that we have a predisposition to visually process natural and unstructured scenes well and pick out the important features of our environment. In order to cover themselves, the authors grant that what makes a visualization memorable does not necessarily make it comprehensible, but argue that it may be an important precursor to answer higher level questions of what makes visualizations effective and engaging.

2.5.3 Engagement

Results from Van Leeuwen et al. 's [79] experiment reported that more vivid memories were retained from an immersive virtual 3D environment compared to the same model presented on a two dimensional screen, but they also reported a statistically significant difference in the level of engagement from the participants. More specifically, the self reported measures where this effect was present were immersion, translocation and concentration. However, they don't make any claims about what exactly caused this effect, but speculate that it might be caused by innate qualities of the VR format - immersion and interactivity. The authors argue that it is valuable from a planning perspective to be able to isolate the feeling of engagement in order for civic participation in urban planning to work well. Similar results were reported in Allcoat and Mülenen's [2] knowledge acquisition study where an evaluation scale questionnaire was distributed to gauge participants' opinion on their level of engagement between an immersed virtual scene, a video recording of the same scene, and a text-book style presentation of screenshots of that same environment.

A strategy for framing the evaluation of engagement is by employing Delman et al's [20] AELIA-model. An approach to participatory urban development where they declare that attention, experience, learning, influence and action are necessary features in order to achieve successful citizen engagement.

2.5.4 Mixed method approach

User experience evaluation traditionally relies on self-reported measurements, observations and usability studies. However, these methods may not always be able to reveal aspects of subliminal cognitive processes like emotion and attention [40]. Since self-reported measurements rely on intro- and retrospection they are subjective by nature and therefore subject to distortion by users' mental state, biases and other human factors. Observation is a method that avoids this to some degree, but since it is still performed by a human, all the same factors apply to the observer instead, only from an exocentric perspective. A way to counter or complement qualitative evaluation is by employing a mixed method approach. That is, combining UX evaluation methods with biometric sensors that can measure physiological reactions which are hard to self observe. Eye tracking is a commonly used method that tracks corneal reflection and pupil dilation to identify gaze movements and fixation, thereby giving information about visual attention, fatigue, arousal, engagement and interest in general. Electrodermal activity or galvanic skin response, typically measured by a wristband, measures skin conductivity in order to determine stress and arousal. Electroencephalography is a non-invasive technique used to record electrical cortical activity, by which emotional responses and motivation of the user can be inferred.

Hussain et al. [40] explains that each biometric sensor can identify a part of a user's behaviour. The strength of one method is the weakness of another, and combined, they provide a more complete mapping of the experience. Then, when coupled with UX evaluation that reveals conscious aspects of how the user is feeling about the experience, the researcher gets a step further to understanding the user.

3

Theory

Design research is an academic field that is sensitive to cultural conventions, subjectivity, and technological advances. This results in rapid shifts in paradigms, and makes the field ever so interesting to act in. The "goodness" of design lies in the eyes of the observer, and thus a design will often appeal more to some than others, and it can be hard to compare between designs with inherently different properties - even when they address the same problems. Hence, design research needs to be grounded in, and evaluated through frameworks with the ability to make more objective assessments of design qualities.

Cognitive Science is the interdisciplinary research field interested in the human mind and the underlying processes to cognition such as the driving forces to decision making, emotions, and reasoning. Knowledge of how the mind works and what drives humans to act on their environment is essential for understanding what constitutes good design. Good design is design that is made for humans, and developed with human cognition in mind. Having this perspective will enable designs that are not solely subjectively grounded, but instead created to best suit the cognition of the intended user, with regards to their cognitive as well as contextual prerequisites.

This chapter will present relevant theoretical frameworks from Cognitive Science as well as Interaction Design with the intention to provide a substantial base for analysis of the collected data. Moreover, the treatment of these subjects is intended to facilitate an understanding of the actions and opinions of the participants in the user-centered activities, and thereby theoretically ground the resulting guidelines, something that will elevate their relevancy and contribution to the field of Urban Planning and Interaction Design.

3.1 Embodied, Situated, and Distributed Cognition

The embodiment thesis relies upon the foundation that many features of cognition are determined by aspects of an agents physical body, such as knowledge representation, arranging and conveying information, and perceptual processing [63, 81]. Robert and Foglia [81] explains that at the most general level, cognition can be considered to rely on the body, either as a constituent or a causal role, in three distinct functions. 1) as a Constraint, where the body of an agent constrain the character and content of a mental representation and thus significantly affect the way in which the cognitive system will process it. 2) as a Distributor, where it acts more as a part of the cognitive system and distributes processing tasks between brain and body in order to balance computations and representations between neural (internal) and non-neural (external) structures. 3) as a Regulator, in terms of allocating cognitive activity in the spatial and temporal dimension by coordinating real-time execution of complex actions in concert with a continuously changing environment. In this role, the body is highly engaged in feedback-driven processes of cognition, and is considered to not only be responsible for receiving input and producing behavioral output, but also an essential real time regulator of cognition. Together, these three perspectives encompass embodied, situated and distributed cognition which are central to understanding the cognitive processing of multimodal interactive technologies.

3.2 Social Cognition

Humans are innately social animals with an exceptional ability to cooperate and communicate. The human ability for cooperation is believed to be due to our well developed intersubjectivity (a.k.a Theory of Mind), i.e our "capacity to judge the beliefs, desires and intentions of other agents ... [as well as] emotions and attentional foci" [29]. When interacting with other individuals, we are skilled at interpreting and utilizing non-verbal communication that often cannot be verbally translated [68]. This information is crucial for our sense-making of other individuals' actions and intentions; it constitutes the information from which we derive our situational and peripheral awareness, i.e. our conceptual understanding of our current physical and social context, our awareness of presently unfolding events, and the impact that one's own actions will have on ongoing and future events [68]. The ability to utilize individual strengths on a group level has enabled human communities to survive, and societies to form, within which many individuals can co-exist and mutually benefit.

Socially Distributed Cognition is a sociotechnical system-based cognitive approach that challenges the notion of cognition being an isolated phenomenon in the brain, to include human-human interaction as well as interaction with artifacts and materials in the environment [38]. By extending the boundaries of what constitutes a cognitive system, sociotechnical constitutions can be better understood, and technologies aimed to support collaborative interactions can be developed. The social distribution of cognition states that multiple users of a system can not only cooperate, by sharing their own mental states through various communication channels, but also create new cognition, through the social activity of cooperative interaction [38, 16]. Socially distributed cognition takes into account the ideas and mental states that arise in the interaction between two humans, that would not occur in any other context.

3.3 Gestures

Human communication is a complex interplay between verbal as well as non-verbal conventions and expressions. Our ability to mentally depict representations of real-world objects as well as abstract concepts distinguishes us from other animals [29]. This ability has given us a rich inner language, that we efficiently share between individuals through language conventions [10].

With verbal language conventions, we are able to convey indexicality, and thereby establish mutual understandings with receivers regarding referred objects. However, decoupled from its context, conversations can appear nonsensical and unclear, since indices are often contexually bound and can thus only be understood by attendants of the conversation [3]. Additionally, humans can communicate indices non-verbally through gestures, for example by pointing to objects, and thereby providing an index to a verbal statement. Gestures are an integral part to human communication and thus involved in cognitive processes associated with language and thinking [1]. As a linguistic tool, gestures possess the ability to activate mental concepts in the mind of a receiver [77], thereby also causing any motoric responses coupled to the same concept.

Pointing gestures are used to index objects or locations, however, these must not be represented in the real world, but can refer to abstract concepts, such as the metaphoric locations for choices ("on the one hand - on the other hand") [1], or directions for time concepts ("back then, or ahead in the future"). The embodiment of cognition can be understood through gestures, since they seem to prove that the environment we perceive the world through is integrated in our cognitive system, as we can index it to communicate abstract concepts not necessarily existent in it.

Gestures are not only used to index objects external from the body, but can also represent physical as well as simulated actions in communicative contexts. It lies in the nature of action gestures to not exert any impact on the environment, but rather communicate the form of the represented action [1]. Hence, in natural communication, people can simulate actions without consequences through gestures, since the environment will remain in its present state. Simulating an action will affect the cognitive state of the executor and observer though, as a simulation will cause cortical activation patterns similar to the represented action [39]. When activating a mental concept through simulation, all cortical pathways coupled to that concept will activate simultaneously, including activations in motor cortex that are connected to the concept. Thus, when an action is simulated through a hand gesture, the motor cortex will not only respond to the movement patterns of the hand, but also the represented action that the hand represent [39]. Consequently, when simulating an action, not only is the represented action communicated, but also *experienced* in a sense, and therefore, knowledge can be obtained regarding its effects on the environment, if performed. Several studies, primarily on monkeys, indicate that mirror neurons give rise to the cortical activation following observations of the behavior of other individuals [71, 28, 39]. With mirror neurons, interacting individuals are able to approach each others' cognitive states through the mirroring of cortical activation patterns. This cortical connectivity is independent of mental models for interpretations, and instead occur automatically, without mentally loading the participants of the social interaction [53].

3.4 Conceptual Models

Conceptual models is a strong concept within Interaction Design that refers to the abstract representation of an interactive system or interface that a user has. This representation is often a simplification of the actual system, and not necessarily coherent with the system's actual functionality and underlying structures, but will regardless determine the way that users interact with its interface [59]. More often than not, it is simply not meaningful for a user to understand a potentially complex system structure in order to interact with it, which is why a simplified model will suffice to provide enough understanding for interaction. Using a conceptual model for interaction allows the users to predict the system's behavior and generalize how the system will respond in similar usage contexts [44]. Therefore, it is important to be aware of the intended users' background knowledge and preconceptions of the system's constituents, as to match these in the interface design.

By developing interfaces with human cognition and perception in mind, chances are that they will be better understood by the users, and that the users' conceptual models will suffice for them to successfully interact with them. Donald Norman's famous principles for successful designs are all grounded in psychology research, and constitute a good base upon which designs can be created with the end-user in mind [59].

Affordances
Every attribute of an object that gives a clue about ways to interact with it is has an affordance, meaning that all objects possess several affordances, aggregating to a compound knowledge of how to interact with it. For example, a tennis ball might afford being hit on with a racket - an affordance probably inherited through convention. For a person unfamiliar with tennis though, a stronger affordance might be the ball's bounceability, that enables the user to use it as a bouncy toy. Apparently, the origins of affordances can vary, as we interact with objects in our environment for different reasons. A door handle might possess a strong enough physical affordance for it to be understood as a pullable object to most people, but in less obvious situations, conventions - and thus objects' perceived affordances - might differ between users.

Signifiers
Signifiers are the tools that designers can use to communicate the affordances of an object [58]. These are the attributes of interfaces that signifies the possible interactions. Whereas affordances constitutes the actual possibilities with an interface from the user's point of view, signifiers points to these affordances through their perceivable features.

Discoverability
The principle of discoverability states that the possible interactions available to a user must be easy to detect, as well as the current status of the system.

It is self-explanatory that interface details that are hard to identify will be hard to interact with, but this is also why important functions must be easier to discover than functions not used frequently.

Feedback
The system must provide feedback that clearly communicates the effect that the users' action has in order to avoid confusion. If the user is unaware of the effect of an action she will continue to try and achieve her intended goals unnecessarily.

Constraints
The ways to interact with a system always has some constraints that will limit the user, and narrow the possible actions. Constraints can frustrate - if the user can not do what she wants in a system - but they are also powerful tools for guiding the user towards the intended actions and system states. Norman [59] identifies four different types of constraints: Physical, Cultural, Semantic, and Logical - each working as guiding tools for interaction possibilities. Physical constraints are simply the physical attributes of an interactive system that is preventing certain actions. Cultural, semantic, and logical constraints are all constraints that takes place in the users mind, from which she can derive what actions are possible based on previous knowledge of the world, or the society in which she lives. Cultural constraints are similar to conventions, apart from the fact that they are not descriptive of how to behave but rather how not to behave, or what actions are inappropriate. Semantic constraints are not tied to any specific culture but rather to the intended meaning of actions. Logical constraints might be the least complex constraint that is located in peoples' minds, as some actions are constrained to certain contexts, or bound by other actions to have been taken place, such as pressing the gas pedal only after the motor is running.

Mapping
Mapping refers to the connection between controls and the resulting effect from a user action. In order for a user to apprehend a system, it must be clear what actions preceded a new system state. Mapping can be especially important in interfaces where a system receives several inputs in parallel during a short period of time, and must communicate their respective effect with clarity. Physical mapping can for example refer to how well a knob or button - though its spatial location - communicates where and how the system will respond when interacted with.

4

Methodology

Following a design process can facilitate the work of a researcher, as it will provide tools to employ at different phases of the process. Having a clear goal with one's design activities is crucial in order to plan them so that they generate relevant data. Articulating these goals can be facilitated by the process, as it can help identifying what is needed from one phase, before moving further to the next.

Present thesis will employ an iterative and data-driven process structure with loosely defined stages, as to enable flexibility to the process. This strategy will provide clues as to what phase the project is in, and thus what design activities are relevant to execute to inform the process and move forward to future phases. However, the loosely defined process will not prevent the return to a previously visited phase, if newly acquired information requires a revisit and remake of a previous step.

Present chapter will describe the process that the project will follow, as well as a discussion regarding the purpose of following a process. Following, an examination of methods deemed relevant to the project will be described, and placed in their respective design phase for a simpler overview of their chronological order in the process.

4.1 Research through Design

Research through design is an approach to knowledge acquisition that dictates pragmatism, in which exemplars of artifacts are what constitutes the empirical body within the field, and the community judges the success of the methods and theories applied in given research by either adhering to similar methods and rationales, thus justifying them, or denying their presence in their own process. Gaver [30] argues that since design research revolves around artefacts, insights generated throughout the design process should be "considered as annotations of [the artefacts]" - and not more than the researchers' subjective experiences from the process.

There is an ongoing debate within the field of Human-Computer Interaction (HCI) regarding the justification of the design through research approach and its validity as a scientific method [30]. Opponents of the approach argue that it impedes true progress in the field due to the lack of standards and criterias for what constitutes "good" design research [26]. However, since designers often deal with wicked problems, design studies are not replicable, and their outcome will depend on situational factors and the judgements of the researchers. Therefore, argues Gaver [30], methodological standards and research criteria will not be able to cover all research problems needed to be addressed, and use cases and design examples should be the means through which designers share and learn from experiences, and extend their knowledge and designerly intuition. Adhering to this notion will encourage explorative and creative ways of approaching problems of a wicked character. Hartson & Pyla [36] stresses the need for designers to follow a process in their practice, or else they will only apply their favourite methods, and ignore activities with which they do not have experiences with. This will lead to important activities being neglected, causing unwanted consequences. Following a process must however not be a limiting or rigid commitment, and they argue that experienced practitioners can take liberties within a process, and trust their design intuition instead of following the process blindly, if these are conflicting. Hartson & Pyla's argument in favor of a process is thus not so much a call for the standardization of methods within the field that Gaver opposes, but rather a collegial advice to follow a plan that is well-tested. Their process model "The Wheel" [36] is an outline of activities that they argue should be part of every design process. The activities are: Analyze, Design, Prototype and Evaluate. What is important to point out is the iterative workflow advocated by the authors, which is characteristic to the field of Interaction Design. Hartson & Pyla advises practitioners to iterate, not only the whole process, but also within each activity, and between activities. As an example, Prototyping is by no means a linear activity by itself, but rather several attempts to explore a design space through the means of prototypes. This exploratory process might require a revisit to the previous activity, namely Design, within which data-driven planning of prototypes occurs and conceptual designs are created.

Conclusively, following an established design process can provide you with a toolbox of activities, methods and techniques that can be useful throughout a design project. These can support thinking and guide workflow, releasing time for you to work goal-oriented. However, following a process blindly can lock you into a rigid workflow that does not fit your specific project, disabling you from working creatively and following your design intuition.

4.2 Design Methods

Present section will go through the design phases of "The Wheel" and describe the methods associated with the respective phases. The described methods are all frequently occurring methods in the Interaction Design literature, and possess different qualities that makes them usable in various phases of the design process.

4.2.1 Analyze: Understanding user needs

Activities found in the first and analytical phase of the design process are concerned with collecting domain-specific data and information about the users' work and needs. In order to design for a user group, one must first get to know the context in which the end-product will be used, and the users themselves [36]. User requirements should be grounded in real data, as to not set the project off in the wrong direction by making potentially false assumptions about the user group.

Contextual inquiry. Being an ethnographic approach used to outline and understand the tacit knowledge and activities performed by a target group, contextual inquiry is a method used in early stages of the design process, before the designer has enough knowledge about the intended user to design for them. By observing and interviewing people in their natural environment of interaction, the researcher can draw conclusions about the environment and people interacting in it that would not emerge from interviews or observations alone. It is important that the contextual inquiry collects data from various user segments and different users, in order for it to provide a versatile ground for future analysis. The focus of the method should be that the researcher tries to shift perspectives to the user's, as to not fall victim to researcher biases that color the data [36].

Literature review. Commonly, an early act towards exploring a problem space is to review the literature that is thought to be relevant to the current project, something that will provide the researcher with knowledge about related work that has been made in the chosen area of research. Knowing the area of research to which you aim to contribute provides a more solid understanding of whatever research gaps that may exist and thus inspires to fill them with new knowledge. Also, it will help the researchers to plan their studies with proper demarcations since they can rely on the findings from previous work and will not have to reinvent the wheel. The purpose of a literature review, apart from the insights gained by the researchers, is to present to readers of the future work the essence of previous research within the area, such that it is clear as to why the presented work came about [36].

Interview. As part of the process of getting to know your target group, interviews are good for meeting the people who will be affected by your design, either directly through the usage of it, or indirectly. In interviews, you get the chance to gain insights on the attitudes and conceptions of your target group through guided conversations. The conversational format will not only provide the verbalized answers to the questions you ask, but sometimes, more importantly, it will provide the implicit information communicated through body language and personal expressions through tonal changes [36]. The observant interviewer will notice these, often subtle, cues, and extract well-needed information that might prove valuable in his future work. Interviews can be more or less structured, something that will alter their conversational topic and outcome [68].

In unstructured interviews, open-ended questions are asked in a natural conversation-like manner around a topic. The interviewer must follow an inner sense of what is relevant in order to obtain useful information, or transcribe and analyze the answers in a later phase when he knows what information is sought for, although by then, he may realize that the right questions were not asked. In structured interviews, the researcher knows what questions he wants answered, and asks them in the exact same manner for all participants. The questions are scripted, as to remain consistent between participants, and counteract the risk of obtaining biased data. Structured interviews are often conducted when the interviewer has certain expectations of what answers he will obtain, and has clearly defined goals with the interview [68].

Semi-structured interviews are, as they sound, a mix between structured and unstructured interviews and thus contain both scripted questions, and a pre-established topical path, but with an open-ended structure wherein interviewees are allowed, and encouraged, to elaborate on their answers.

Focus groups. Groups interviews such as focus groups are convenient to conduct when the purpose is to gather qualitative data from a sample that is representative of a community, rather than individual experiences [68]. If the moderator succeeds with creating a socially comfortable peer context, participants tend to open up, and share their personal opinions, feelings and attitudes openly [55]. Having a heterogenous group can stimulate discussion and lead to topics that would not have come up in individual interviews.

4.2.2 Design: Interaction design concepts

In the design phase, conclusions are drawn from the processed results from analysis phase. This is the time where design concepts are introduced and sketched as potential solutions to identified problems from the user studies. However, the design solutions produced should be kept at a low resolution, and sometimes at a conceptual stage only, as to not become focused on details. It is important to understand the mental models of the intended users, as to match those with the proposed design concepts [36].

Mental models. The concept of mental models refers to a person's understanding and reasoning about how a system works, what components it consists of and how one would interact with it. At opposing ends, the interaction designer and the user often have profoundly different mental models of a design, but it is the designers mission to match the mental model of the typical user with the interaction sequences of a product [36]. The designer's mental model is grounded in the vision of how the system would ideally work and be interacted with, based on what was uncovered in the analysis and contextual inquiry. The conclusions that can be drawn from that data should constitute the building blocks of the design. The users mental model is more complex, being built on previous experiences with similar encounters as well as general knowledge about the world, such as context, cultural conventions, their individual differences of expertise and so on [36].

Concept design. The design of the concept is how the designer connects the mental models of their own and the intended users by establishing metaphors, themes, system transparency in the form of clear feedback mechanisms and adherence to intuitive and conventional interaction methods [68].

4.2.3 Prototype: Realization

A prototype is an artifact aimed to communicate one or several aspects of an end-product [52]. It is an externalization of the designer's thought of what the end-product might be, and must thus not be similar to any existing product or sketch. The prototype is a useful tool to explore conceptual designs in a cheap way, communicate design ideas to stakeholders, test technical feasibility, and even conduct tests of usability - all without having to develop expensive, fully functional products [68]. In addition, the process of building prototypes can be mentally stimulating for the designers, and encourages reflection in the design, thus leading the process forward [72, 75].

Prototype categories. Prototypes can look radically different due to their pragmatic purpose, and there is no "prototypical" prototype. There are, however, ways of dividing prototypes into categories descriptive of their complexity and purpose, as will be presented next.

Horizontal vs. Vertical. Since prototypes are not finished products, they do not contain the functionality nor content of the thought end-product that they represent. The designer must decide what aspects of the end-product should be represented in the prototype. A prototype that has a finished look and contains many design features is called a horizontal prototype, due to the broad representation of features that comes on behalf of its functionality that is close to non-existent. On the contrary, the vertical prototype has a narrow focus with a high degree of functionality for only small parts of the design concept, making them useful for communicating specific functionalities without putting them in a broader context [36].

Level of fidelity. The level of fidelity in a prototype refers to how closely it resembles a finished product [36]. When evaluating hi-fidelity prototypes, users might not notice that they are not interacting with a finished product, and will thus anticipate a reasonable response to every action they make. Hi-fidelity prototypes are therefore often used for presentational, or evaluative purposes [8]. Low-fidelity prototypes do not aim to fool users of their functionalities as they do not resemble a finished product. Instead, they are quick and cheap to create, often out of paper, and aim to explore broad design ideas "- such as content, form and structure, the 'tone' of the design, key functionality requirements and navigational structure" [8], thus informing the process with directions of interest to explore in the design space.

4.2.4 Evaluation: Verify and refine

The process of evaluating the design of an interaction is a matter of deduction and inference. Aspects such as usability or user experience are not directly measurable. Instead, indicators that are connected to those concepts are measured, such as task time or error count, and the resulting value is hinted at the level of usability [36]. In the design cycle, evaluation is the last step before it resets. In reality, evaluation takes place during every step of the process, in a range of different places and in different shapes [68]. Formative evaluation is often quick and informal, typically occurs early on in a diagnostic manner and that lets the researcher uncover clear flaws in a design. Summative evaluation is a quantitative and often, but not necessarily, formal procedure where performance metrics and end-user opinions are collected in order to assess and improve the user experience [36]. The latter tend to be connected to the concluding evaluation of a higher level prototype that has improved beyond iterations of obvious issues.

Much like the dichotomy between formative and summative evaluations complement each other, so does quantitative and qualitative data. Early in the design process, when a prototype is being shaped and evaluated, there is no use in measuring quantitative aspects like error count, since there likely will be many of them and they might not necessarily be able to be captured numerically. In this case, the descriptive nature of qualitative data suits the stage of the process. When evaluating user experience, qualitative data is always relevant since it's a product of the subject of the interaction, which is often the user. However, when evaluating cognitive aspects of a user experience, such as attention, learning or decision-making, many of them are subconscious and unfavorable to capture qualitatively. A user's subjective opinion of their attentiveness to a certain element of an interaction is bound to be influenced by a range of human factors and biases. Thus, a more accurate way of capturing this data is by measuring it quantitatively with the help of devices such as an eye-tracker, since eye movements and gaze is a direct coefficient of attention. In conclusion, evaluation should be a part of all the stages of the design process. In time for summative evaluation, it is beneficial to consider how subliminal aspects of the user experience can be captured. A mixed method approach of UX research methods coupled with quantitative tracking methods may provide a more comprehensive account of the experience.

Wizard of Oz. This technique is used when a non-functional prototype is to be evaluated, and the functionality is simulated by a researcher from behind the scenes. The researcher then controls the interactivity of the artifact, as the user interacts with it, leading the user to think that the prototype is in fact working [55]. The technique enables researchers to evaluate "real" user interaction instead of having to guide them through the product, with narrow interaction possibilities available, or by using a narrative to describe a usage context, rather than a real context.

Observation. Collecting data that can be used at any time during the design process can be achieved through user observations, in certain contexts or during interactions with artifacts. Early on, it can serve as a way of contextual inquiry where researchers are interested in understanding and mapping out the user behavior and goals [68]. It can take place as a part of an uncontrolled field study in an exploratory manner, sometimes referred to as semistructured observation, where the researcher has an idea of what to look for but is primarily observing with an open mind [55]. Alternatively, it can be carried out in a more controlled environment, such as a laboratory setting, in a structured manner where aspects are well defined and more exact causations are investigated. Either the observation is direct, where the user is aware of being analyzed and that a researcher is gathering data about their observable interactions, or it can be indirect, where data instead is gathered in log format and often over time [68]. The method works well for uncovering aspects of an experience that might be unknown to the user. On the contrary, if not coupled with other data collection methods, subjective experiential aspects, mental states and the fact that a user is being observer, can affect their performance and behavior, sometimes referred to as the "Hawthorne effect" [36].

Think-aloud protocol. The protocol is an evaluation method used when the researcher is interested in the underlying thoughts and feelings that goes through the user's mind as they complete a set of tasks [36]. As many other evaluation methods focus on results and retrospectives in the analysis of interactions and usability, the think-aloud protocol extracts the explicit underlying cognitive analysis that precedes and co-occur with the interactions of interest. Since the user is asked to verbalize their thought process as they complete a task, aspects that cause enjoyment or frustration come to light and are labeled by the one experiencing them. The method is flexible in that it can be used in whatever situation a user is able to verbalize their thoughts, and can be a good compliment to observations in controlled or uncontrolled environments [68]. The focus of the method should be oriented toward what is happening, rather than why, in order to keep the task burden light so that the user can focus on the interaction [55].

Usability testing. Testing the usability of an interface is typically aimed at isolating aspects that are frustrating or confusing to the typical end user, as they perform a predetermined set of tasks [55]. The method is specific in the sense that it focuses on the expected goals of the end user, and therefore tasks are often coupled with scenarios to provide a context to the testing session. A think aloud protocol is commonly used in conjunction with test scores to reveal problems with the interaction as well as determining the efficiency of the interface. Usability testing is often restricted to a controlled environment in order to be able to determine that the developed product is usable for the tasks it was designed for and by the intended user population [68]. It is therefore most commonly applied in a summative evaluation setting.

Affinity diagram. Affinity Diagramming for Contextual Inquiry is a bottom-up, generative process in which qualitative data from interviews is clustered and categorized based on shared affinity [55]. Observations from the interviews are written on individual sticky notes and then placed on a large surface. The observations are then coupled if they express similar problems, issues, or subjects and can thus be categorized differently depending on the researchers' purpose with the analysis. The clusters that emerge compose inductively derived categories that can inform the design process, and act as support for a continued data-driven approach to prototype development.

Semantic differential scale. In semantic differential scales, adjectives of different semantic poles compose a scale on which a participant judges the conformity to their attitudes and perceptions regarding an object or a concept. For example, a scale with the semantic poles *Good - Bad* will force the user reflect about their attitudes to the object's properties solely in terms of their goodness and badness, thus encouraging them to inhibit attitudes related to other aspects than those treated in the scale [37]. Unlike Likert scales, that are often used to measure the amount of congruity between the perceptions of a user and statements about an object, semantic differential scales measure the semantic connotations associated with the object.

5

Execution and Process

Following section describes the process of the thesis, as it unfolded throughout the project, covering the four design phases from "The Wheel": Analysis, Design, Prototyping, and Evaluation. All methods used as well explorations of the design space are presented and described, to provide the reader with a context within which the end results in terms of both design solution and theoretical conclusions can be understood.

5.1 Planning

Entering a project of such an extent as a master thesis, it is essential to set up a time plan to ensure that the project will finish within the time frame of the thesis course. However, when setting up a time plan before the project begins, flexibility must be embedded in the plan as it will change throughout the course of the project. The initial time plan looked as following:

Week 1-5: Literature research and Planning report. The initial phase is dedicated to exploring the design space of focus, both by reviewing relevant literature on the field, but also by getting acquainted with Virtual Gothenburg Lab, and look into the various interests that are represented in the project. Examples of such activities are meetings with stakeholders as well as reviewing project documents such as the project plan.

Week 5-8: Planning methodology and user studies. Informed by the exploration of the design space, a more concrete plan for the thesis project and methods to use can be formulated. Initial user studies are conducted to get acquainted with the intended target group, as well as gaining empirical data that can be used as design rationales in forthcoming phases.

Week 9-16: Prototyping and iterating The prototyping phase compose the biggest part of the project, and involves several iterations, with various levels of fidelity. Parallel prototyping will be acquired in order to develop more than one prototype to be evaluated and compared.

Week: 15-16 Evaluation. Parallel to the development phase, an evaluation session is to be planned and finally executed when the prototypes are functional. The evaluation sessions will generate data that will be processed and compose the results of the study, as well as the topic for analysis and the guidelines.

Week 16-17 Analysis. The data is analysed and interpreted through theoretical frameworks. The analysis is then compressed and refined into concrete guidelines,that can be easily applied to future projects within urban planning.

Week 14-20: Report writing. Working with data analysis is time consuming and accumulates large amounts of content to discuss in the report. Hence, writing on the report must be done parallel to the development and evaluation phase. When the practical activities are finished, the report will attain the authors undivided attention, and be focused on until the thesis project ends.

5.2 Analysis: User and Field of Research

There are several aspects to take into account when designing a design project. Firstly, in order to contribute to a research field, one must be well-informed with previous work and methodologies used in that field, as for the conducted project to contribute with novel information or fill research gaps. Also, it is important to understand one's target group, to enable the creation of meaningful and useful solutions that meets their expectations and needs. This section will present the approach used to become acquainted with previous work and commonly used methods within the research field, as well as how the researchers collected empirical data that would inform the design process and constitute support for design rationales formulated throughout the design process.

5.2.1 Literature review

In the early phase of the project, the research area within which the project was to be conducted was traversed through a literature review. A semi-structured snowball search for relevant previous research was applied to the exploration in which keywords sought for included but was not restricted to: urban planning, visualization technologies, tangible interaction, tangible interface, citizen inclusion, participatory design, evaluation methodology, geovisualization, and GIS interface. Previous case studies were reviewed, and methodology as well as technological implementation was analyzed for a better understanding of both existing knowledge within the research field but more importantly of knowledge gaps that need to be dealt with. The literature review resulted in a theoretical ground upon which research questions could be formulated, and a broad outline of the study proposed.

5.2.2 User study

Collecting data from a user group as diverse as "citizens" is problematic mainly because of issues concerning sampling. Ensuring that the data represents the opinion of the average citizen means either situating the study where a random selection is present or performing studies in several different geographical locations, the latter being a very time and resource consuming option. With this in mind, locations for contextual inquiry, semi-structured interviews and observation were determined to be held at two different locations: Älvrummet at Lindholmen Science Park and the Gothenburg City Library. Both are public spaces where a significant amount of people pass through. The difference between them is that Älvrummet gets more spontaneous visits from IT industry and academia, sacrificing some of its representativeness, but allows for performing contextual inquiry since it is the city's main display of plans for urban development. The city library however, is presumably the space visited by the most diverse user group available to the public.

Figure 5.1: Overview of the physical model at Älvrummet. Scale 1:400.

Figure 5.2: Perspective view of the model at Älvrummet.

Figure 5.3: Zoning plan posters and bulky binders. An example of poor information visualization targeting citizens, according to the Älvrummet guides.

The purpose at this phase of the design process was to map out the user group by uncovering their current relation to citizen dialogue, their preferences for how they want to find and receive information provided by the city and what their attitudes are towards novel ways of digitally visualizing this information. A conscious decision was made not to ask direct questions about different visualization technologies but rather let the users mention it if they felt it fit into the interview, in order to collect more generalizable data that could later on be interpreted and used as a foundation for the coming evaluation. In total, 27 participants were interviewed, with an age span ranging from 17 to 95 (median age = 39).

In addition to targeting regular visitors to the sites of data collection, the researchers interviewed two of the guides located at Älvrummet, whose task is to inform visitors about ongoing urban development projects and also to organize guided tours for larger groups of visitors.

5.3 Designing the interaction

Following the user studies, an affinity diagram was used to extract useful insights from the raw data and identify themes in the opinions of the interviewees. All comments from the annotated interviews were written on post-it notes and then categorized due to their shared attributes. Since the categorization process is subjective by nature as the categories are determined by the researchers' preconceptions of what defines a (relevant) shared attribute, there is not one but several possible categories to cluster data nodes in. Therefore, the categorization process was iterative, and several categories were explored before settling for the final ones.

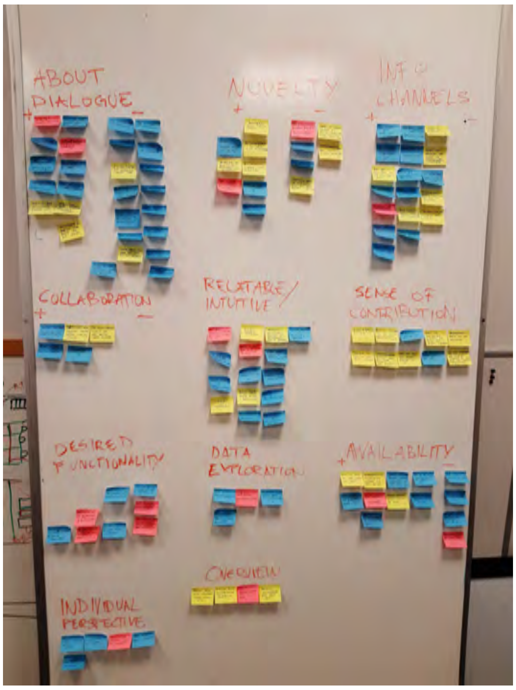


Figure 5.4: Using affinity diagram to sort and identify themes in the data.

The categories that were distilled from the affinity diagram were of different kinds; some described general attitudes towards urban dialogue or were descriptive of the interviewees primary information channels about urban development projects, whereas other addressed concrete suggestions of what features the interviewees appreciate in interfaces and models depicting urban development projects. The categories can thus be further categorized into higher-level categories, as depicted below.

Information channels

About dialogue. The majority of the participants emphasized the importance of citizen dialogue but at the same time, many stated that they are not interested in actively participating in dialogue meetings or hearings. Several stated that they did not think that the city's current methods for citizen dialogue were working well and called for better ways to reach a majority of citizens.

Info channels. The majority of the interviewed received information about urban development projects through digital media such as newspapers and social platforms. A few participants were members of forums on which urban planning was discussed. Only a few received information from the municipality's official communication channels. Individual perspective - The large majority of participants were interested in ongoing and planned projects in their vicinity and wanted to be able to have influence over their realization.

Attitudes towards participation

Sense of contribution. Many participants expressed dejection towards current tools and methods for citizen dialogue, and argued that the real influence of citizens through these channels is non-existent. Participants with experience from citizen dialogue activities expressed that they did not feel listened to, and that there was no feedback on their suggestions.

Physical/geographical availability. The majority stated that geographical availability is of great importance for them to participate in dialogue activities. Interfaces that are accessible from one's own technical devices do not require the effort of getting to a physical place, or schedule for a meeting, but can instead be used whenever appropriate. If an interface would require physical presence, many participants preferred it to be located at places they visited frequently.

Desired functionality

Novelty. Many participants expressed positive attitudes towards novel technologies and interfaces such as AR, VR, holograms and 3D projections but at the same time, concerns were raised that new technology was excluding the senior citizens.

Collaboration. There was a large agreement regarding the positive aspects about collaboration in general, and participants agreed that accumulated knowledge was better than knowledge kept individually. It was argued that it is easier to become blinded by an interface when interacting single-handed, and that collaboration erases that effect.

Relatable / Intuitive. Several argued in favor of physical models, as they are easy to understand and simple to relate to. Physical models are "..concrete and clear", and provide an easy-to-grasp perspective in a transparent manner. Some participants stated that they wanted more interactivity to the physical model standing in Älvrummet, and that interactive physical models are engaging and interesting. One participant suggested that a physical model made of LEGOs would engage people and promote interaction.

Data exploration-drill-down functionality. Some participants expressed a desire to see more information than what is displayed. Drill-down functionality is desired, where additional information can be explored that does not interfere with the overview perspective. Two participants suggested a temporal dimension to an interface, where progress and development over time could be visualized, and even historic perspectives could be displayed.

Information overview. Most participants agreed that an overview perspective is important in visualizations. Both because it makes orientation and navigation easier, but also because more information can be presented simultaneously without cluttering the information. In physical models, the participant decides for herself what information to attend to, and what perspective to experience the visualization from.

5.4 Prototyping: Projection Mapping

Conclusions from the analysis phase were crucial to extracting the outlines for the next stage - prototyping. Informed by the values expressed by the target group, a discussion ensued of the propriety of the different dialogue tools reviewed in the body of previous research. Looking back into definition, contexts and applications of each of them, coupled with key facilitating cognitive functions, it became more clear which technologies had the most potential. Essential to this process was asking the question of to what extent the themes from the user study could be matched with the most prominent features of each technology, thereby allowing them to prosper. In order for the evaluation to be valid it was established that the technologies needed to be able to display their full potential, while acknowledging the the fact that it would likely mean sacrificing comparability. Discussion over this trade-off really defined the project, and would follow through the iterations to come.

Projecting onto a designated space in which users may interact is a feature that can be achieved in several different ways. Deciding which alternative would suit the objectives of the project better was difficult to envision. Therefore, the initial focus lay on exploring the different options and existing solutions to similar problems, and look back into the details of the technology from where it was encountered in previous research. Limiting the search to solutions within a reasonable resource range narrowed down the alternatives to a handful different solutions. Further conclusions could be drawn from the themes that emerged from the user study. *Novelty*, *collaboration* and *intuitiveness* all characterize this technology. Thus, the interface needed to embrace and reflect these features. A number of approaches to projection mapping focus on the visual effects rather than the interaction, such as art installations or stage lighting for live music concerts. Consequently, these also tend to lack the fundamental dynamic elements that lets explore an interface or environment. This realization led toward computational models where the user was able to provide input and in one way or another alter the behavior of the system. Similar solutions have previously been used successfully in urban planning [4, 69, 7], which led us to study their methods and see how and if similar functions could be integrated with the objectives of this project. What they all have in common is a tracking system; where either body, depth, color, shape, infrared light, or fiducial markers are utilized to provide input data about the changes taking place within a tracked scene in order to alter the projected output. In addition, the human body is somehow involved, either by being the tracked object or by manipulating the objects being tracked, generating a natural and embodied interaction.

After discussing these aspects with supervisors from RISE, questions regarding technical aspects to support the interaction were raised. For instance, whether the project required real time or pre-rendered projections, and whether to simulate the projected content - water, or implement a particle system using a physics generator. As the interface would be experienced similar to that of a game, in terms of interactions and problem solving in a simulation, the decision partly rested on the choice of whether to implement it using a game engine or more of a visual development platform. At this point, Unity or Touchdesigner [21] were the main contenders, as they both provided solid prerequisites for tracking as well as for the reason that there was in-house expertise of these platforms at RISE. Besides that, both perform well in terms of 3D or immersed experiences, they have well-documented implementations using motion tracking input devices (initially considered to be the tracking solution) such as RealSense or Microsoft Kinect, and handling virtual geometrical objects with physical properties. Ultimately, what separated them was Touchdesigner's tools for managing projections, making it the better choice.

The first two weeks of prototyping consisted in large of learning how to use Touchdesigner. It is a node based procedural and visual programming language for real time interactive multimedia content, used for 2 and 3D production. The main features important for this project being its 3D and physics engine, projection mapping, and the procedural programming approach. Workshops were held by RISE supervisors for practicing the fundamentals while simultaneously lo-fi prototyping to allow the design to take shape.

5.4.1 Components

In order to provide an overview as well as introduce the following sections, a summary of the design components will be presented. No component had a linear path of development and fidelity was deliberately kept at low resolution until the next phase of the design demanded an improved capacity.

> *Program.* The actual program itself and the behavior of the design, created in Touchdesigner. Containing the virtual environment, the virtual actors, the particle system and the code. Responsible for producing the output image projected onto the physical space.
>
> *Tracking system.* The final version of the tracking system was comprised out of three components: camera, computer vision framework, and Touchdesigner. The camera continuously sent real time video to a computer vision application that scanned the scene for identifying features from a library of patterns. If any patterns were found, positional data about them was sent to input nodes of the Touchdesigner network. The program interpreted and translated the data in order for it to be received by acting operators in the network.
>
> *Virtual Space.* The design ended up utilizing a physics generated virtual environment to host a particle system - water. This environment was built in 3DS Max as an image of the Linnéstaden city district. The desire to create a dynamic and natural interaction required a virtual environment in

order for water to flow, collide and behave in a legitimate way. The virtual scene provided the base 3D geometry for the Touchdesigner network, in which particles and other virtual actors were contained.

Physical Space. The physical scene was created as a materialized copy of the virtual environment, making the user able to interact with it in a tangible fashion. Its only function being that the projection was contained within it, behaving as if it was interacting with the projected content.

Virtual Actors. Within the virtual scene in Touchdesigner there were two actors performing different geometrical operations. First, the "barrier", a 3D cube acting as a wall within the virtual scene preventing particles from flowing freely. Second, the "well" a roaming 3D cylinder creating a hole in the bottom plane of the virtual scene, thereby letting particles trickle through it. Both virtual actors were connected to a physical actor counterpart that could be manipulated by the user.

Physical Actors. The physical actors were 3D printed tangible representations of the virtual actors - a barrier and a well. On top of each actor a fiducial marker was placed, enabling the camera to track each object individually. The physical actors could be placed anywhere in the physical scene considered to be ground level. This position in space corresponded to the same location in the virtual space. Thus, wherever the physical actor went, so did its virtual equivalent.

As all components where refined gradually and at different stages of development, the below account will process each component's prototyping phases in full.

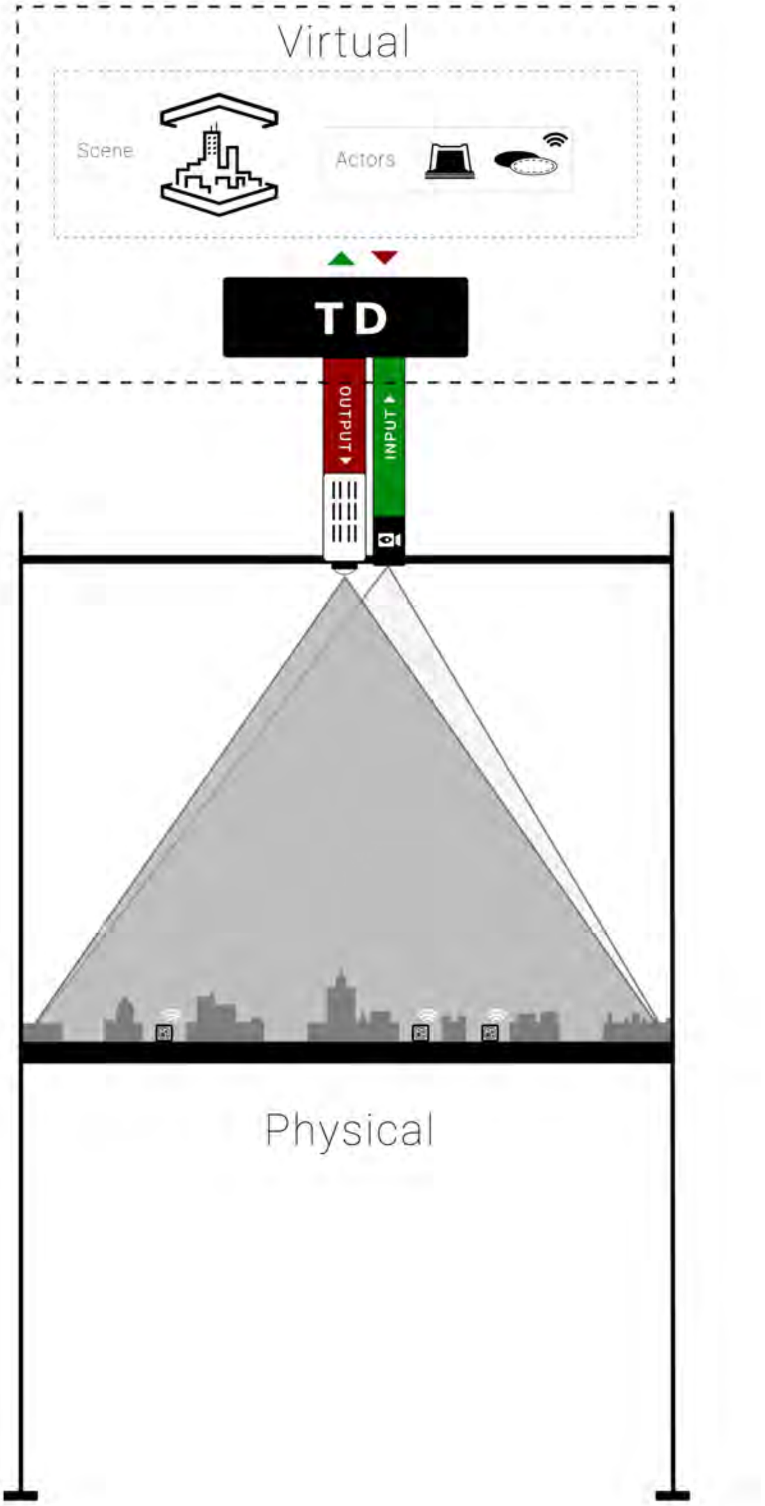


Figure 5.5: Infographic of all the components, relations and information flow within the projection mapping interface.

5.4.2 Tracking

Throughout the development of the PM interface, several tracking solutions were explored to find the most stable one, eventually settling for for overhead tracking of fiducial markers. The issues of tracking affected the development of this prototype to a large extent, which will reflect in the sections below, but was crucial to solve for the sake of functionality.

5.4.2.1 Reflective surfaces and color

The initial tracking attempt focused on color tracking using the RGBA camera feed from Microsoft Kinect v2. Touchdesigner has built in support for capturing video, depth and IR data from the Kinect cameras, which made it ideal for quick prototyping to explore tracking alternatives. By isolating distinct color values through the application of measurement thresholds, red, green and blue could be filtered out and separated from each other. The program tracked each color with relative ease as long it could be clearly distinguished from the background. However, setting the color value thresholds in such a way that not one of the channels would pick up skin color or color of garments of users proved to be challenging. In addition, the lighting conditions in the workshop greatly affected the tracked scene, causing color values of the tracked items to vary by extension. Though ceiling lights and other environment lights could likely be shut out, colored light from the actual projector would inevitably affect the scene. This interference raised the question of whether it was plausible to position the camera to track from the same top-down perspective as the projector at all.

Discussion ensued about other tracking methods, such as tracking from below, a solution often used for tangible tabletop user interfaces [78, 46]. However, that would require the surface of the interface to be see-through which seemed incompatible with the presence of other untracked physical objects in the scene, such as buildings or terrain. The second attempt instead focused on utilizing the infrared camera feed from the Kinect. Discovering that different surfaces reported different values depending on their level of reflectivity, simulating an effect of an object appearing closer than it is, a new prototype took shape with tracking highly reflective surfaces. In practice, this meant that reflectors could be at the same distance as the surface of the base, but report max vicinity values to the Kinect, making them easy to set apart from the rest of the tracked scene. This solution turned out to not scale well as the code established tracking position by calculating the average of all distances in the video stream, within specified thresholds, and could therefore not discriminate between several reflectors of the same proximity values. Since the network needed to be able to track and separate a number of objects there was an investigation into how the reflective objects could be manipulated to report discrete values. The first speculation was to use IR emitting LEDs that, if coupled with resistors of different strengths, could output stable values at varying levels, making it possible for the network to separate them from each other. This path was quickly scrapped due to the realization that Kinect only tracks the flight time of the infrared beams emitted by itself and filters out all other sources of infrared light.

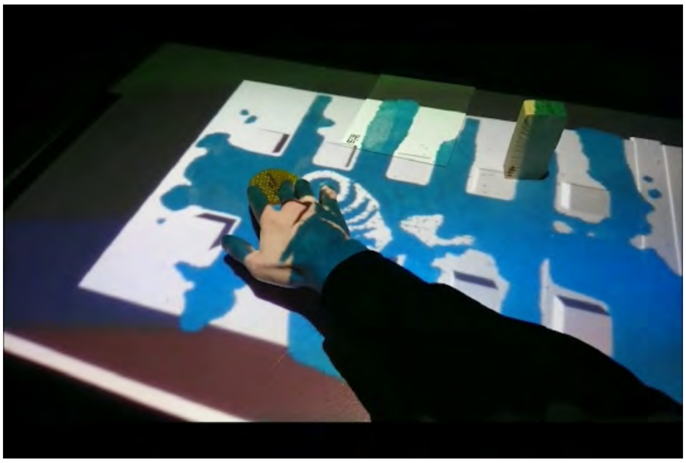


Figure 5.6: A low fidelity prototype without physical objects except one reflector, tracked by its perceived proximity to the camera.

5.4.2.2 Computer vision

Acknowledging that tracking IR or color would not establish a stable foundation for the network, other methods were searched for. Tracking fiducial markers or QR codes was another approach encountered previously in the literature, meaning computer vision data would need to be provided to Touchdesigner. One of the options looked into was OpenCV, an open source library of programming functions used for real-time computer vision and image recognition, used specifically for object and motion tracking among others. While this would likely have provided the most reliable solution as it could be tailored to our needs, creating the code would take a considerable amount of time not directly dedicated to the actual purpose of the thesis, and therefore ruled out.

Next, the open source computer vision framework reacTIVision [46] was tested because of its tracking capabilities of fiducial markers. It is a toolkit for rapid prototyping of interactive surfaces tracked by video in real time. The application uses the TUIO protocol to send messages through UDP to a specified network port, in this case sent on localhost; for any client application to read that can decode the protocol. TUIO is encoded using OSC (Open Sound Control), a well adopted protocol for communication between multimedia devices, because of its efficient encoding of discretionary controller data as well as its wide use. In the case of reacTIVision, the controllers are the fiducial markers, and the transmitted data contains presence, orientation, identity and location of the markers. Unbeknownst beforehand, Touchdesigner, being a platform dedicated to real time production of interactive multimedia, proved to have components capturing input on both OSC and TUIO protocols. This realization caused a breakthrough to the tracking issue, as reacTIVision provided a relatively stable input to Touchdesigner. Although, the fact remained that the scene was continuously interrupted by lighting from the environment and directional light from the projector, occasionally causing the tracking of the markers to be lost and with it the input to Touchdesigner.

Figure 5.7: A picture showing the tracking of fiducial markers on a plane surface. The markers are coupled to their virtual correspondents, deciding their positions in the virtual scene.

5.4.2.3 Cameras and lighting

A range of different cameras were tested besides the Kinect, primarily web cameras because of their often convenient form factor. The minimum requirements to keep tracking stable appeared to be 1920x1080 resolution and 30 frames per second, something most cameras could match. Parallel to assessing camera performance, other aspects affecting the tracking where looked in to. The primary issue was caused by the mismatch in frequency of different light emitters in the environment. The power distribution grid in Sweden runs on 50hz, meaning electrical lights will flicker 50 times every second. The projector being used had a default image output frequency at 60hz which unfortunately could not be changed. Any camera can typically be set to capture images at 50 or 60 hz, though, by necessity negating the other. Despite matching projector output frequency with the video capture frequency, interference was found, likely because of sporadic drops in framerate caused by Touchdesigner's demands on CPU. As a result, the image input to reacTIVision was continuously flickering.

In order to see how the tracking performed, each marker in the scene was paired with a graphical representation - a virtual cube - projected onto the physical surface. The tracking problems then became apparent as the cubes reset their position every time the tracking was lost, skipping between their actual position and the reset value. To cope with this, one of the more significant implementations was a function created for the Touchdesigner network wherein the last recorded position of a marker would be cached; the effect being that if the tracked position was to be lost, Touchdesigner would behave as if the marker was still left at the position it was last seen. This feature caused considerable stability to the interaction and avoided user confusion, since the virtual cube would not leave the scene if a marker was somehow hidden.

Despite the network optimizations a few issues of tracking remained. Unless reacTIVision started picking up the fiducial patterns from the video stream after losing them, the interactive connection of the bond between marker and virtual cube would be uncoupled.

5.4.2.4 In the infrared spectrum

Several solutions to the lighting issue were explored during the course of developing the prototype. One of them concerned the idea of tracking the scene completely in the infrared light spectrum. To achieve this, visible light would have to be filtered out with an infrared lens and the tracked scene lit up by infrared spotlights. Since the Kinect v2 has an infrared video stream and its own infrared light emission, this was the first solution to be explored. A downside to reactTIVision was that in order for it to find a video input source it would have to be recognized by the operating system as a camera. The Kinect then, primarily being a motion sensing device and not recognized as a camera, had the be transformed to a virtual camera for the connection to work. This was accomplished through the use of OBS Studio, which accepts a wider range of input devices and can output anything held in its scene as a virtual camera stream. During testing of this solution it was discovered that the infrared video stream had a resolution of 512×424 pixels and only the color stream had a full 1920×1080 pixel resolution, the former not being high enough to distinguish the features of the fiducial markers.

Figure 5.8: IR camera together with two infrared spotlights, assembled with a Raspberry Pi during an iteration as part of a solution to the tracking issues.

With no infrared cameras with enough resolution at our disposal, the decision was made to build one. An IR camera module for Raspberry Pi was purchased from Amazon and set up with a Raspberry Pi 4 Model B. The camera module had infrared spotlights attached to a shared board, making it capable of lighting the scene on its own, a field of view of about 150 degrees, and 1920×1080 resolution, making it ideal for the task. As there is no possibility of connecting the RPi directly to another computer, a few different methods of sharing the media were investigated. First, a web interface was set up that could be accessed and controlled from any browser over local network. With a live-preview, low latency and high framerate capacity, the setup essentially became a wireless camera sending real time video to any client on the network, creating a very flexible solution. For reacTIVision to access the video feed, the web interface window had to be transformed to a virtual camera through OBS Studio. Despite making sure no image compression was taking place at any stage in the sequence, the output quality was subpar and could not be used to track the markers. An alternative version was tested where the RPi desktop was accessed remotely, through local network using the SSH client PuTTY, by the computer running reacTIVision. By previewing the camera feed through the built in interface of the RPi OS, no compression took place, producing a clear image. Instead, the framerate at which the desktop image was transmitted was well below 10 frames per second, making it equally dysfunctional for tracking. In both solutions the clients were connected over ethernet as to eliminate concerns regarding connectivity.

Ultimately, the RPi solution was abandoned due to time constraints. A search for cameras began again, testing new and old to see if something might have been overlooked. Testing the 4k Logitech Brio webcam for the second time came with the realization that it required USB 3.0 data transfer speeds to perform optimally. While USB ports and cables it had previously been tested with where 3.0, the necessary extension cable had likely not been. Switching this part to a 3.0 version caused significant improvement to the cameras performance, particularly in terms of framerate; it also appeared to barely be disrupted at all by the projector light. After calibrating the camera to better match the scene, it could follow the markers with enough precision to be accepted as the final solution to the matter of tracking.

5.4.3 Virtual and physical space

Whether to imitate water flow or strive to emulate it defined the development of the virtual environment early on, and the benefits of each where considered. Initially, a simpler solution was entertained where the behavior of the water would just be simulated, requiring a set of predetermined outcomes that would activate upon certain conditions. Though, only mimicking the behavior was assumed to limit the interaction and produce a rigid interaction pattern, inhibiting the users' ability to explore the interface. Allowing for more degrees of freedom of the tools at hand was deemed critical because of an otherwise strictly contained environment. As this realization became evident, the idea to simply imitate water flowing was abandoned.

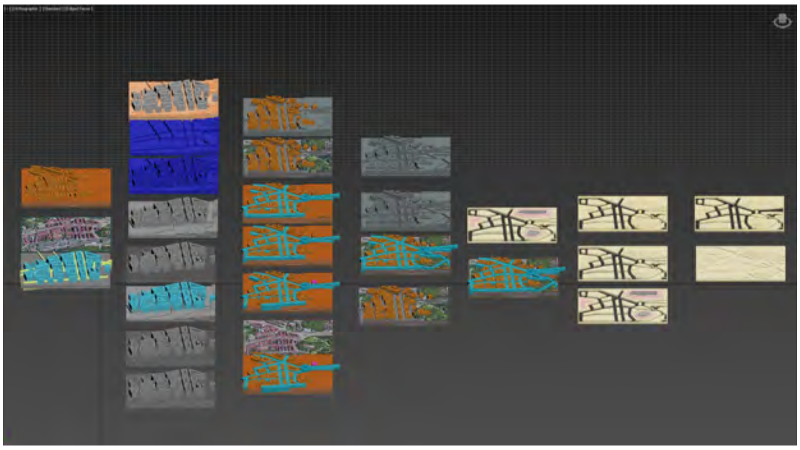


Figure 5.9: 3D models of the virtual terrain used for the projection mapping scene. Developed in iterations from left to right, with the leftmost models being the ones used for the final prototype.

The 3D model of the virtual environment emerged in gradual increments whenever the rest of the design required an updated model. At first, a simple model was used for the sole purpose of understanding how the 3D and physics engine in Touchdesigner would behave. virtual actors were introduced, different 3D geometries, and mapped to tracked physical objects to understand how a user could be considered to affect the seen using a physical token. This model was used until there was a solid tracking solution in place, reacTIVision, which meant individual virtual actors could be introduced and connected to fiducial markers in physical space. With more actors, a more advanced environment model could be developed to enable the design of the behavior of the particle system.

As the resolution of the model increased, so did its resemblance to the real world area of Linnéstaden. In order for the interaction and narrative of the experience to integrate well, a familiarity factor was designed for. Adjustments were made to the size of the streets and spatial relation between terrain objects in order for there to be enough space in which the user could interact and place tokens. A considerable amount of time was dedicated to finding the right height difference between ground, terrain, and dam levels of the model. The dams were the source location of the particle system, meaning a certain amount had to collect there. The design of the space also had to direct particles to critical locations, such as the center street "Linnégatan". To make sure this occured, the model was given a slight concave bend, and terrain was modified to lead particles in certain directions. These adjustments came from realizations during usability testing and informal evaluation sessions.

Figure 5.10: 3D models of the physical actors with fiducial markers placed atop. Printed to become tangible tokens for interaction.

Recognition was considered for the physical actors as well, the barrier and the well. As these were the physical tokens through which the user would manipulate the scene, establishing a mental model by matching the appearance to their properties was necessary to form the interaction. Different form factors were modeled and printed, to see how they could fit the scene. A significant limitation concerned the fiducial markers, which needed to be placed on top of each token and of a certain size in order to be seen by the camera. Finding the correct form became a question of priority rather than feasibility, ranked in order of camera tracking ability, small enough to fit the streets, and large enough to have recognizable features.

Figure 5.11: Picture of the finalized prototype, with the physical model made of plywood, and 3D-printed wells and barriers.

5.4.4 Formative evaluation

When the behavior of the projection mapping interface was stable, usability tests were conducted to receive feedback in order to inform the design process further. As functionalities were implemented gradually, Wizard of Oz solutions were used when necessary. In the first fully coordinated usability test, the wells were activated manually by one of the researchers as they were placed at various positions by the users. Interacting with the interface, the users had two types of objects to manipulate, 1) barriers that would block the water, forcing it to change routes, and 2) wells that were pockets in the 3D model collecting particles, differing from the final design in that they were static. The locations where it was possible to place the wells were marked with a colored cap that would disappear, on a command from the wizard, whenever the user placed a physical well-token at the location.

A scenario was created to put the interaction in context, giving the users meaning to their actions. The users were told that they were public investigators for the city's department for water protection, and that it was their task to prevent the negative effects evoked from flooding. Therefore, they were to try and minimize the negative effects from the flooding of a dam in a simulated scenario (for a full description, see appendix A). A time limit was set to 5 minutes, to put pressure on the users, and to invoke a sense of seriousness causing engagement. After having provided sufficient instructions and presented the scenario, the simulation was initiated and the users left to begin the task.

Two coordinated usability tests were conducted, with each test containing two users working in pairs. Each test was followed by an evaluation session lasting approximately 20 minutes in which semi-structured interviews regarding the interface and interactions were held. The purpose to this was to stimulate an free association and open discussion rather than obtaining the answers to a fixed set of questions. There was a set of questions that were asked to all participants (see appendix A), but the format allowed for deviations from these.

By observing the behavior of the participants during their interactions, some things were noticed that demonstrated the need to clarify and extend the instructions given before the scenario. For example, it turned out that the participants did not actively reflect on the connection between the camera tracking of the physical objects and how this affected the positions of the virtual objects. While thinking of where to place the objects, participants often held them in their hands, or had their hands over them, obscuring the line of sight of the camera. Loosing track of the position of the markers led the system to behave unpredictably, an valuable insight that led to quality improvements for the tracking. In addition, it was noted that instructions preceding the scenario should include a brief description about the tracking system, and how occlusion might affect the system and result in unexpected behavior.

All participants stated that they enjoyed the interaction and appreciated the cooperative aspect of the interface. All took the task at hand seriously, and throughout the scenario discussed how they should approach the problems as they unfolded. As new situations appeared throughout the scenario, new plans emerged to solve the problems, creating a dynamic and natural interaction.

Below some of the key takeaways from the usability test are listed, carefully considered in later improvements of the interface.

- Graphical representations of the virtual object are important to show that the physical object is mapped. Whenever tracking is lost, this will be apparent since the virtual object will not overlap the physical token.
- Scenario was too easy to complete. Both user groups managed to solve the scenario without much effort. The level of difficulty can be affected through the manipulation of particle radius, number of particles generated, etc.
- The capacity of the wells was unclear and sometimes the participants were confused if the wells were active or not.
- It would be nice to see the inclination of the streets, especially since many participants will have preconceptions about what it looks like in Linnéstaden.
- Some participants felt immersed in the scenario and focused much on the task, while other was seeing the interaction more like a game, and felt that the aim was to "beat the system".

5.5 Prototyping: Virtual Reality

In parallel with the development of the Projection Mapping prototype, the development of a VR prototype was initiated. The rationale for creating two separate prototypes using various technologies was to enable for comparative evaluations between the two technologies. This type of comparison is interesting since the technologies provide the users with considerably different physical prerequisites for interaction as well as varying aggregated user experiences. Currently, Virtual Gothenburg Lab puts significant focus on a digital twin in their work towards creating testbeds for information visualizations and research projects. The digital twin will often be experienced by users in VR, something that will bring with it the limitations that are inherent to VR technology. The authors sought to evaluate the technologies, in contrast to the implemented interfaces, in an evaluation workshop addressing both, as to outline their potentials in terms of user experience, interaction goals and context of use.

5.5.1 Creating the scene

A 3D model of a central and well-known city district called "Linnéstaden" was obtained from the city's urban planning department. Exporting the file to the game engine Unreal Engine [23] enabled the implementation of functionality and interactivity, as well as the option to enter the model using VR.

The 3D model of Linnéstaden was in scale 1:1 and covered an area of approximately 800x200 meters (see figure 5.12). The model contained one large road (Linnégatan), from which several smaller roads emanated. Alongside Linnégatan were houses, blocking the sight of the observer and the vacuum surrounding the model. However, moving out on one of the remaining streets would soon lead the user to the end of the model, where the virtual world ended. In order to illustrate the boundaries of the model and at the same time occlude the empty space lying beyond them, a semi-opaque bright wall was placed around the whole model, indicating its end.

Figure 5.12: The 3D model obtained from the city's department for urban planning.

In the VR prototype, more focus was put on the authenticity and immersiveness of the experience than that of the functionality. VR possesses the ability to generate immersive experiences due to its high resemblance to the real world and the first-hand perspective of the user. Since the model was experienced in 1:1 full scale, all experienced perspectives regarding size of objects and distances were identical to those of reality, something that further enforces immersion. The interaction is often what distinguishes VR from the real world, since it is commonly accommodated by hand controls that utilizes arbitrary button functionalities to manipulate objects in the environment. As this type of interaction is not natural, it has to be learned, and it is challenging to implement intuitive signifiers in the interface that communicates the buttons' intended usage.

Coming from the development of the projection mapping prototype, the initial effort to create a flooding scenario in VR included the use of a particle system to generate a material with similar properties as water, as was done in the PM interface. However, it was discovered early that it would be impossible for the CPU to generate enough particles to simulate a flooding scenario in a scene of scale 1:1. Hence, in order to obtain the perception of heavy rainfall as reason to the flooding, a local rain was placed slightly above the VR player, that followed the player wherever it went in the VR scene, creating a visual illusion of heavy rainfall in the whole scene (see figure 5.13).

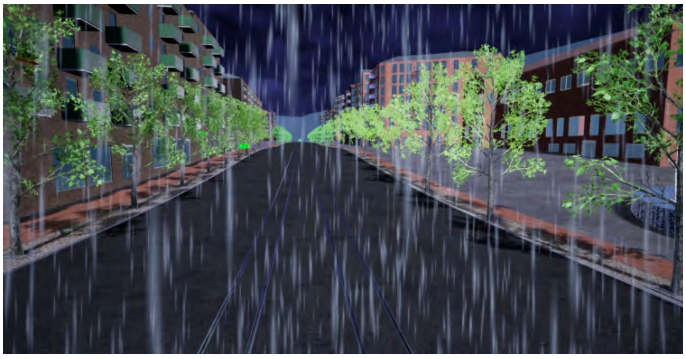


Figure 5.13: The sensation of heavy rainfall was created by placing a local rainfall above the players head.

As for the rising water on ground level, a box geometry was created and placed on top of the streets, and given a dynamically moving material that looked similar to water, but with the physical attributes of the geometry upon which it was placed.

The 3D-model from the urban planning department was obtained without materials and textures, making it difficult to recognize as the well-known Linnéstaden. Materials and textures were added to all surfaces, as well as foliage of different sorts such as trees, bushes etc. Pictures from Linnéstaden were observed closely as to match the real appearance of the district in the 3D scene.

5.5.2 Building functionality

Unreal engine offers a wide range of functionalities that can be used when creating games and animated video content. Thanks to its visual scripting system called blueprint scripting, Unreal Engine offers an easier alternative to its underlying programming language C++. Since the authors did not have experience from coding in C++, the blueprints turned out to be an essential asset to the implementations made, lowering the threshold for creating complex behavior relatively quickly. The visual scripting looks similar to that of Touchdesigner, in the sense that connections of nodes can create complex and large networks with functionalities and connectivities between objects in a scene. However, in Unreal Engine, blueprints are distributed over the actors of a scene rather than adjacent in a layered but central network. This distribution of code relies more on working communication paths between interdependent actors, and added to the complexity when trying to learn the blueprint scripting system.

5.5.3 Interacting in the scene

Since the model was in scale 1:1, navigating the scene by foot would require the same areas as the city district itself, posing a practical impossibility to the interface design. Therefore, it was decided that the navigation would instead be teleportation based, with the users navigating by holding a button on the hand controller to aim for the position to which they wished to teleport, and release the button to activate the teleportation, transferring their body to that position. Teleportation is a conventional way of navigating in VR, and an already implemented function in Unreal Engine, making it a reliable and well-tested solution to apply to the interface. The teleportation in Unreal Engine includes a function for signifying physical constraints, something that communicates the outer bounds that have been set to the users' navigation. Whenever a user attempted to enter an area that was preset to unavailable, the hand controllers would vibrate and the teleportation would only transport the user to the closest approved point before the outer bound was reached.

Placed in the scene were four objects that could be clicked, that would result in their "activation". A long pointer object was placed in the virtual hand actor in the scene, with which the user could aim at objects to interact with. The objects were the same as the ones used as flooding preventers in the PM interface, namely wells and barriers. Three wells and one barrier were placed at different locations in the scene, hovering slightly above ground, with a green glow indicating their interactivity (see figure 5.14). When a user clicked a well, it changed material from glowing green to a rusty brown, and it was then lowered down into the ground. When inactive, the barrier was not visible, but instead a small cube in glowing green was visible, working as an icon for the barrier indicating its presence. When the cube was pressed, the barrier was faded into the scene around the icon through a scalar increase in the opacity of the barrier material. When activated, the barrier covered the street upon which is was placed entirely, which was why the authors chose to have it invisible when inactive, as to not occlude the scene behind it.

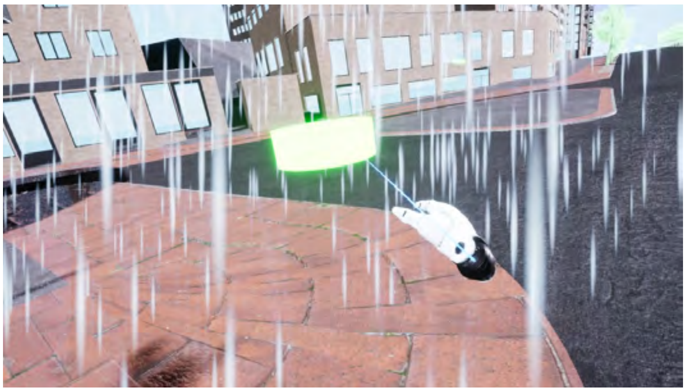


Figure 5.14: Picture of an inactive object, with a green glow indicating its interactivity. The virtual hand can be seen, with the pointer object attached to it.

5.5.4 Simulation scenarios

The flooding of the city model was a simulation that required several iterations, as it proved challenging to create realistic visuals of raising water, especially if the water would originate from a higher point in the 3D scene as to match user expectations of flowing water and street topography. The virtual material that generated the behavior and looks of water required a virtual object (static mesh) to be placed upon, as to exist in the VR scene. In several trials, flat tiles were placed in the scene, slightly above street level, rotated differently as to match the slope of the street. However, as these tiles were to act dynamically upon the activation of a flooding scenario, it became tedious to - in code - address the large amount of tiles required to cover all streets in the scene. Instead, a copy of the 3D model was created, within which all buildings were excluded, leaving only the streets. These were then given a height and the water material was applied, resulting in a large water object able to create the illusion of a layer of water laying on top of all streets in the VR scene. The originating of the water from a higher altitude was created with rotation of the water tile, so that an incremental heightening of the whole material would results in the streets higher up in the level being flooded earlier.

The solution with the large water tile resulted in a simulation where the water levels did not look similarly realistic from all parts in the scene. The rotation of the tile meant that water levels at higher positions in the 3D model had to be made very high as for the low-level positions to obtain any water in the simulation. Therefore, it was decided to create a fixed space in the scene from where the user would observe the simulation, as to assure a higher realism, with the rotation and height of the water tile adjusted to look good from that viewpoint. A cylindrical hole was cut out of the water tile, creating a space inside of the water in which the user could stand and observe raising water in all directions, while simultaneously being protected from the flooding. Fitting the size of the hole a platform was created with a control panel on, with buttons that would initiate the scenarios. In addition, the control panel displayed information about were the user was located, as well as information about the simulation that would play out (see figure 5.15). There were three buttons present beneath the information text, one of which was green and two reds, with the green button indicating its active state. Once the green button had been pressed, the first flooding scenario would be initiated. With every button press, the simulation state would change with a delay, as well as the colors of the buttons so that only one button would be active at any given time, making it impossible to click the wrong button. This design choice narrowed down the users' ability to choose freely among possible interactions and was made with the evaluation workshop in mind, and the limited amount of time that each user would have with the VR interface if all workshop participants were to try it.

Throughout the VR session, the user observed and experienced three different scenarios from the platform. The first scenario was a catastrophic one, in which water levels rose to cover all streets in the scene, with stormy waves that depicted the severity of the situation, and fallen trees floating down the street. The second scenario was initiated when the user drained the water, something that was required for the users to explore the scene, and to activate objects to prevent flooding. Once

Figure 5.15: The control panel, from where the participants received information and instructions regarding the flooding simulations.

all objects had been activated, the user returned to the platform and initiated the third and last scenario. This time, water levels raised, but stopped at a lower level than that of the first scenario, depicting a decrease in severity. In addition, the water material was changed with the initiation of the third scenario, to a material with smaller waves and lesser opacity, creating a calmer tone to the water, to show the positive effects from the activation of objects.

5.5.5 Auditory feedback

Throughout the development of the VR interface, short informal usability tests were conducted to inform the process in a fast and iterative manner. One early test revealed that a lack of feedback resulted in difficulties with understanding the state of the scene, mainly when objects were interacted with. This feedback resulted in the implementation of auditory files that were activated upon interactions. For example, a clicking sound was played whenever an object was hovered over with the pointer, and when clicked, an activation sound was played.

Throughout the whole VR scene, an audio file of heavy rain fall was played in the headset of the user, as to enforce the immersion and sense of presence. Also, thunder sounds were played at some points in the interaction, to enforce the sensation of bad weather.

5.6 Evaluation Workshop

Following the prototyping phase, a workshop was conducted with the aim of generating data that would lead to the answering of the research questions. The aim of the project revolves around the process of involving citizens and outlining the prerequisites for successfully engaging the users rather than creating a single interface that meets these criterion. Therefore the interfaces have, throughout the process, been regarded as instrumental tools for demonstrating the potentials with the technologies with which they were built, rather than being artifacts with intrinsic qualities.

Conducting a workshop that revolves around the technologies rather than the interfaces displayed requires a plan for how to generate non-ambiguous data. Having a discussion about underlying technologies with users that might not be technically experienced might be confusing to some, as it adds a layer of abstraction to the dialogue. For example, it was decided that semantic differential scales was a good way of collecting data that would point directly to differences and similarities between the technologies. However, it would be wrong to assume that the participants' experiences with the interfaces would not affect their attitudes towards the technologies, and it is probable that only collecting quantitative data would result in a situation where it is impossible to distinguish between attitudes towards the interface and the technologies. Therefore, the combination of qualitative data from a focus group session and the quantitative data from the differential scales were thought to complement each other, and together provide nuanced data.

5.6.1 Participants

In total, five evaluation workshop were conducted with three participants in each workshop. The participants can be divided into two categories: professionals/experts (N=6) in the field of Urban Planning, and laymen/citizens (N=8). Two workshops consisted of professionals and the remaining three with laymen. In the two expert groups, all participant were either part of- or well-informed of Virtual Gothenburg Lab, and thus had some insights to the aims of the project. However, none of the participants had seen nor been informed with the details of present project, the evaluation workshop or the interfaces previously. Due to a pandemic outbreak, the citizen groups could not be sampled to be representative of its population, but instead consisted of colleagues and acquaintances. The constellation of participants in the respective groups were however considered, and assembled as to represent various life- and work experiences, to contribute to different perspectives in the focus group sessions.

5.6.2 Workshop phases

The workshop consisted of four phases with the first being introduction and background to the project as to provide a context to the participants. The aim and purpose of Virtual Gothenburg Lab was introduced, but subjects that would later be discussed was only sparsely touched upon, such as the intended target group, learning outcome etc. The aim with this being that the participants would enter the focus group session unbiased to any information unrelated to their personal experiences and interactions with the interfaces and technologies.

Following the introduction, the participants were to try the interfaces, and interact with the scenarios. As to counteract against the potential effects from interacting with one interface before the other, half of the groups interacted with the PM interface first and the other half in VR.

The scenario presented to the PM interface was the same as the one used in the usability test (see appendix A). The VR scenario included instructions on how to interact with objects using the hand controllers. The participants were instructed to

Figure 5.16: Collective efforts to stop the flooding in Linnéstaden in the projection mapping scenario. Showing two participants giving instructions and calling attention to events in the scene, while the third is implementing the measures.

Figure 5.17: One participant currently in the VR scene and two following the perspective through a screen, helping from the "outside".

stand at the platform in the VR scene during all water simulations, and to follow the instructions presented on the control panel. The participants were further encouraged to explore the scene, and that there were no time constraints and they thus did not have to feel any time pressure. While one participant interacted in the scene, the remaining followed the events from a computer screen, and were encouraged to cooperate in locating interactive objects in the scene. For the full instructions, see appendix B. As the observing participants saw where the object were hid in the VR scene, their experience when in the virtual scene would consequently differ from the first participant. Therefore, they were told that they did not have to complete the whole scenario if they did not want to, but should experience the first flooding simulation from the platform, as this simulation was important to understand what a first-hand perspective of a flooding simulation felt like. Although given the chance to not complete the whole scenario, most participants did, and they could focus more on the exploration of the scene than on finding the objects.

After both interfaces had been tried, the participants filled in a form containing 1) a consent form for the collection and usage of their data, and 2) the semantic differential scales. Each semantic differential scale included two words with opposite semantic meaning at the ends of a ten-point scale with boxes (see appendix C). The participants marked on the scale how well the words described the technologies. For each word pair, two scales were filled in, for VR and PM. The participants were instructed to consider the technologies with which the interfaces had been constructed rather than the interfaces themselves.

Following the quantitative data gathering, the focus group session began and lasted for approximately 40 minutes. During this time, several topics were discussed and explored. While guiding the discussion towards topics relevant to the project, the natural conversational flow between the participants was endorsed and closely observed. The participants were told to consider the Urban Planning context in the discussion, and (after first having discussed what target groups they believed the technologies appealed to more) to have the citizen in mind, and reason about the needs and requirements that follows with this heterogeneous target group. Other topics that were explored in depth were factors for engagement and accessibility, potentials for learning, user perspectives, and flexibility to name a few.

In total, the workshops lasted for approximately 1,5 hours. The collected data came from 1) the semantic differential scale forms and 2) interview notes from the focus groups, and both will be presented in the Results section.

5.7 Processing the data

Following the evaluation workshop, the raw data was compiled and analyzed. The qualitative data was categorized into the discussed topics and then reported summatively in text, whereas graphs were generated to represent the quantitative data. The two data sets were compared and scanned for patterns or any discrepancies. For example, if a conformity of opinion was expressed in the focus group sessions but could not be detected in the quantitative data, discussions ensued and explanations were sought for in the remaining categories of data.

5.7.1 Interpreting the results

The aim to contribute to the research field with guidelines that were theoretically grounded in cognitive frameworks and sociotechnical theory preceded the analytical phase. Here, the results were synthesized with theoretical frameworks that hold explanation models able to explain the obtained results.

5.7.2 The emergence of the guidelines

The purpose of the guidelines was to provide a set of compressed and concrete recommendations to follow when planning for a project within urban planning that is involved with information visualization. The guidelines were empirically derived from the processed data from the evaluation workshops. For the sake of transparency, a brief description regarding their connection to the data and analysis was included, that referred back to the origin of the parts preceding each guideline.

6

Results

The evaluation workshops generated quantitative data from the semantic differential scales as well as qualitative data from the focus groups. It is important to point out that the quantitative data was not presented numerically to the participants, but has been transformed for easier processing in the data analysis. The numerical values presented in this chapter does thus not refer to any amounts of the concepts represented by the words in the scales, but rather a position on the scale, where 0 represents the leftmost position and 10 the rightmost. These positions depict how well the words at the ends of the scales describe the participants' experiences with, and thoughts about, the respective technology.

Both the quantitative and qualitative data regard the underlying technologies used to create the respective interfaces rather than the interfaces themselves. It is difficult to conclude that the participants' experiences with the interfaces did not bias their assessments of the the respective technologies in the semantic differential scales. Hence, in present chapter, the quantitative data will not compose reliable material for conclusions on its own, but will instead be interpreted in the light of the qualitative data, that holds the nuance that is lacking in the former. The qualitative data consists of notes from the focus group sessions, and is thus a mix of direct quotations as well as interpretations, by the researchers, of opinions and content. Thus, the data consists of interpretations of meanings, and sometimes added contextual information that the authors deemed relevant to include, as for statements to be correctly interpreted. Adding contextual information as well as interpreting the words of participants always come with the risk of misinterpretations, but is nevertheless necessary in order to understand the intended meaning. The context in which statements are made cannot be neglected if one wishes to get the full picture, and sometimes it can provide more information than the words that make up the sentences. Thus, present chapter will present 1) the quantitative data obtained from the semantic differential scales, mainly in the form of visualizations 2) themes discussed during the focus group sessions, complemented with the authors' comments regarding possible interpretations, and contextual information to provide nuanced data to the reader to make their own interpretations. The qualitative data will also be used instrumentally to make sense of the quantitative data, as the former will be able to explain qualitative differences between the technologies, and shed light on differences left unexplained by the latter. 3) Finally, data gathered from the workshop will be analysed and discussed in the light of the findings from the initial user study conducted in the earlier stages of the project.

6.1 Quantitative data

This section will present the quantitative data collected from the semantic differential scales in graph visualizations. In three graphs, data will be presented for 1) all participants, 2) the expert groups, and 3) the citizen groups. Discrepancies in the data between the technologies will be discussed in the next section, wherein they will be reported in detail together with the qualitative data, for a richer depiction of the context to which they belong.

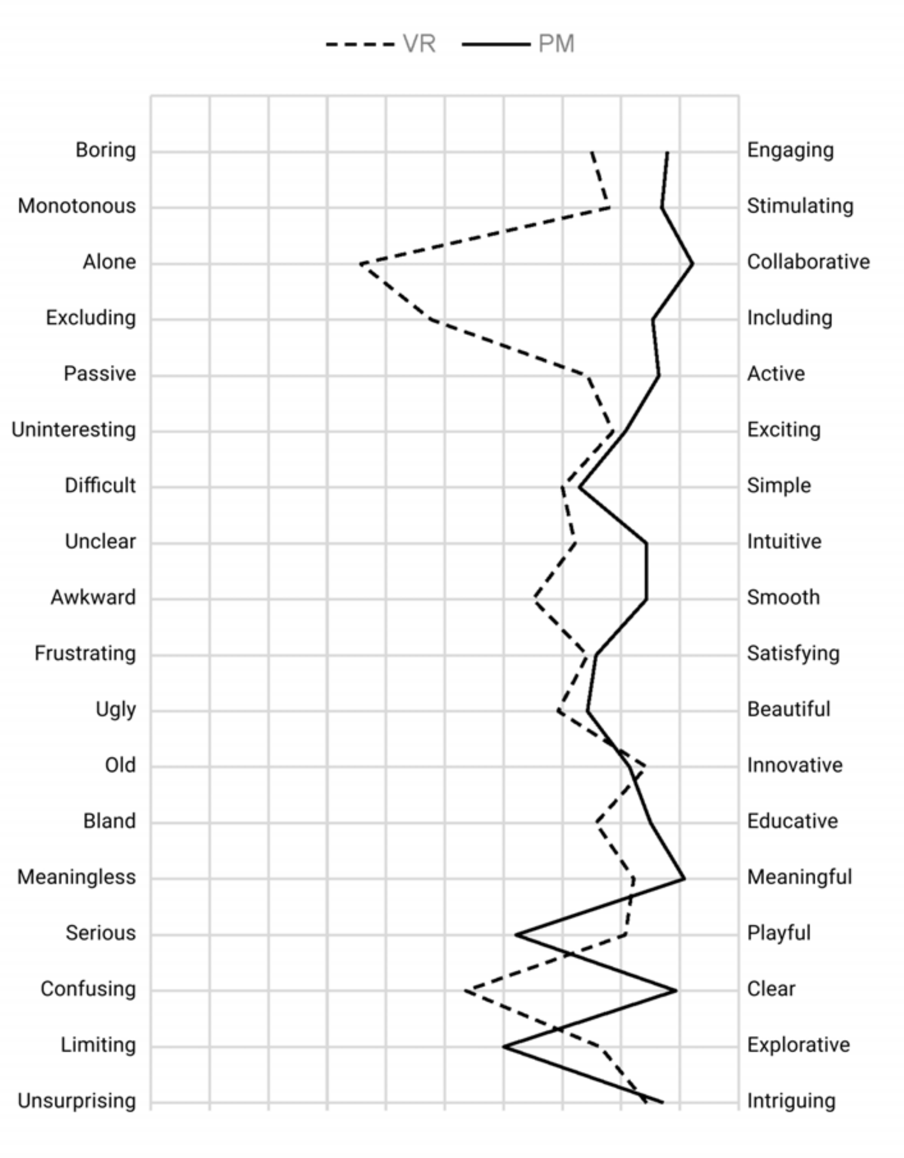


Figure 6.1: Graph depicting the mean scale values from the semantic differential scales, for all participants.

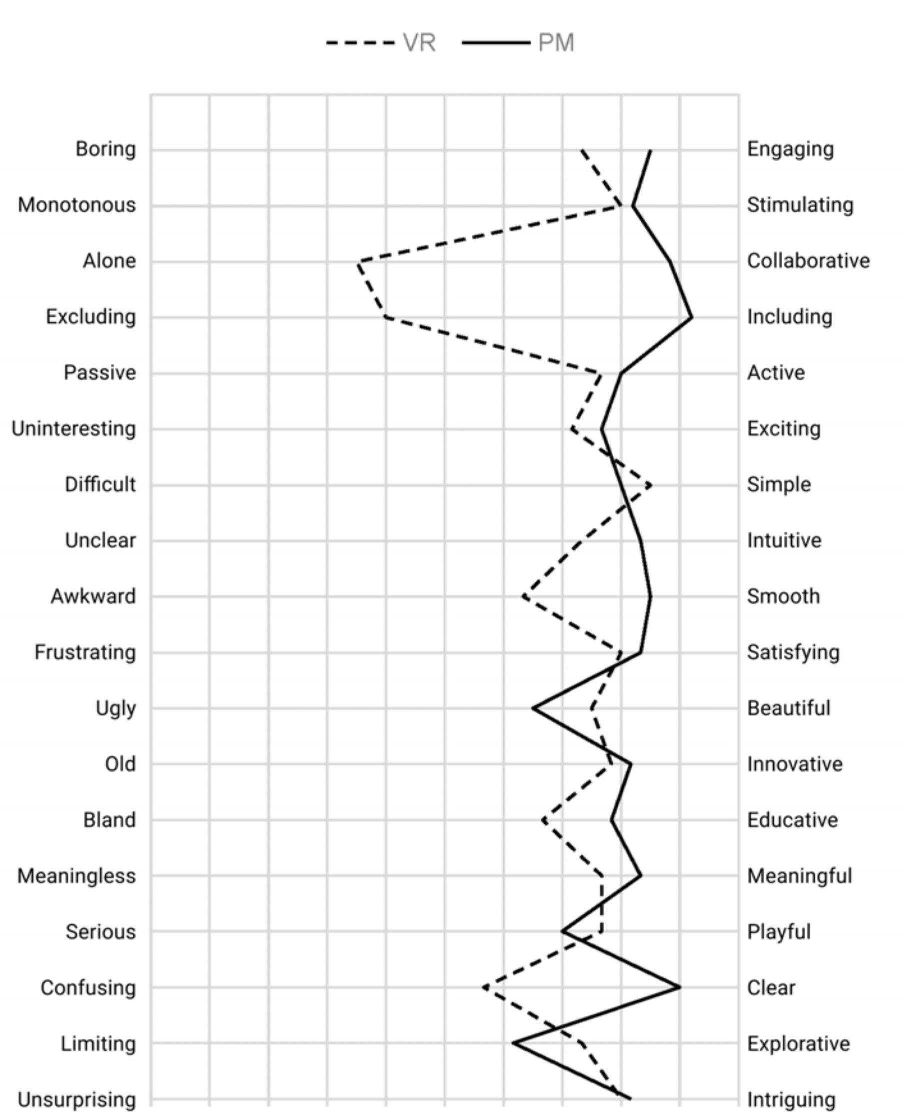


Figure 6.2: Graph depicting the mean scale values on the semantic differential scales, for the experts groups.

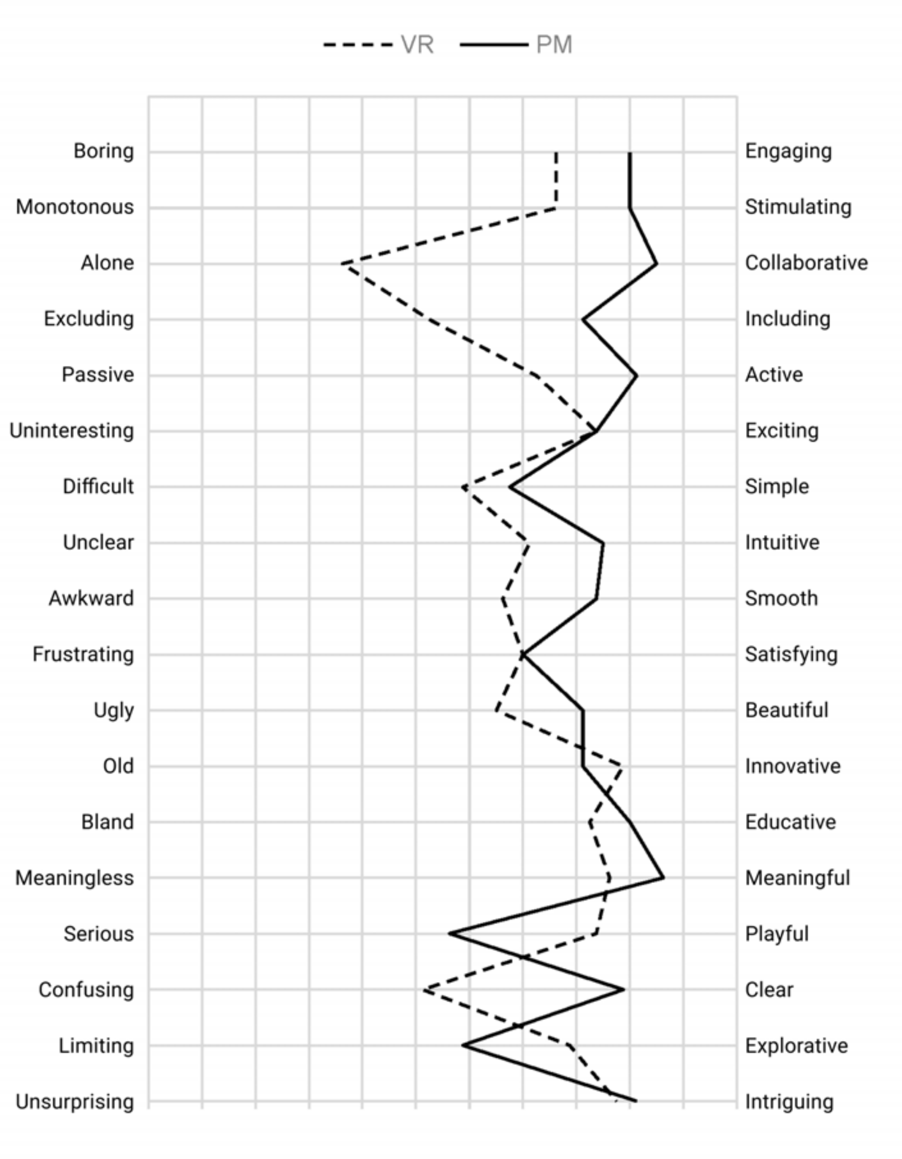


Figure 6.3: Graph depicting the mean scale values on the semantic differential scales, for the citizen groups.

6.2 Qualitative Data

Throughout the focus group sessions, the discussions came to revolve around eight main topics. They were Engagement, Availability, Target Groups, Potentials for Learning, User Perspectives, Collaboration, Fidelity, and Type of Experience. Some of the topics were planned in advance whereas other grew organically into the conversations, and were kept if they provided valuable input to the context.

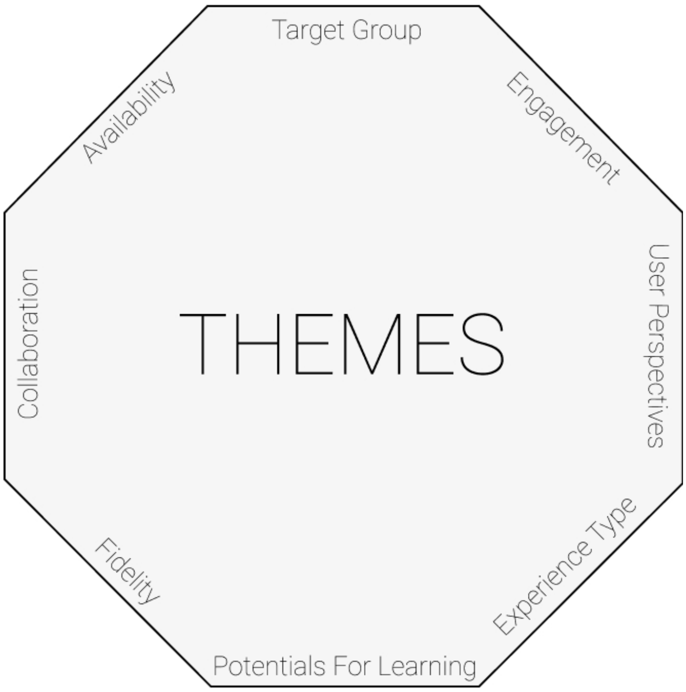


Figure 6.4: Illustration of the main themes discussed during the focus group sessions.

6.2.1 Engagement

Key insights: *Both VR and PM possess the power to engage, but in different ways. Problem solving, immersion, sensation of power to affect the interface are all factors for engagement.*

To begin with, all participants stated that both technologies offered good prerequisites for engagement, and they enjoyed interacting in the interfaces, although for different reasons. The participants mentioned several factors that would make them engaged, such as problem solving, immersion, sensation of power to affect the interface etc. Hence, although not sharing all these factors, the technologies both engaged in their own ways. Some participants stated that their confusion with the interaction buttons on the hand controllers was affecting their experience in VR negatively. For the participants having problems with the interaction in VR, it was expressed that their focus on the environment and their task was neglected, and harder to survey. Despite this, the general opinion regarding the potential for engagement for VR and PM was positive, and recognized also by the participants with interaction problems. This opinion was supported by the quantitative data from the first semantic differential scale *Boring - Engaging*, in which VR received a mean scale value of 7.5 and PM a mean value of 8.79. The differences in what factors are engaging for respective technology will be discussed under the remaining themes.

> “*Both were engaging, but in different ways*”

6.2.2 Availability

Key insights: *High threshold for interaction excludes people from the technology. PM is good for public exhibitions, due to its low interaction threshold. VR can be downloaded from home, but it still might not reach out to many citizens.*

When discussing what VR and PM offers in terms of availability, the threshold for interaction was a recurring notion that was brought up. The ease with which interaction occurred in PM was widely recognized as something positive, and many argued that not having to reflect on the interaction enabled one to focus more on the scenario, and the task at hand. The interaction in VR on the other hand was challenging to several, but not all, participants, and this was experienced as impeding to the focus on scenario and task. No participant experienced any interaction-related challenges with PM, and agreed that the interaction felt natural, and self-explanatory. A frequently ventilated opinion related to this was that the threshold to interacting in PM was lower than that of VR, something that will have an effect on peoples propensity, sometimes even ability, to interact. In one workshop group, it was discussed that the higher threshold for interaction in VR would potentially exclude elderly users as well as users with mental disabilities, as the isolation from the headset prevents assistance from the real world to those interacting in the virtual world.

> "*In VR, you have to learn how to interact, but in PM, you just know how to do it, and you can focus on the task*"

Upon the question of which technology was best fit to be displayed in public spaces, participants agreed that this would be PM, due to its low threshold for interaction, and how it can be understood by most people without instructions. Contrarily, many participants discussed the fact that VR provides the opportunity to reach all those who will not attend to the public spaces. In theory, VR scenes can be run by anyone who owns a VR set. This opportunity was deemed valuable by several participants, and regarded as an important factor to be added to the discussion about availability. However, one participant from one of the expert group discussed that, however true in theory, he did not think that people would download such an application from the municipality in practice. He argued that the reason to this was that the application would suddenly compete against other applications available for download, which is a plethora of applications such as advanced high quality games with which informative applications from the city's department for urban planning cannot compete.

> "*Visualizations in VR from the city's department for urban planning can be downloaded in theory, but I would never do it, and I play lots of games, because then they would compete with games - not community information - and on that arena, there are better alternatives*"

6.2.3 Target groups

Key insights: *VR is better suited for young people with experience from video- and computer games and pose interaction difficulties for those without experience.*

It is the more playful technology of the two. PM is experienced as a more serious technology, good for planning, and appealing to an older audience. PM stimulates the thought and is good for discussion.

Throughout the focus group sessions, a wide consensus was expressed regarding who the technologies were best suited for: VR was for older users and PM for younger. To start with, several participants expressed that it was difficult to interact in the VR interface, as previously mentioned. However, there was not a clear correlation between these users age and how they interacted in the scene. Discussing it revealed that rather habits and practice was considered to determine how well one could interact in VR. All groups of participants discussed the correlation between (video- and computer) games and VR, and how it would be easier for people who grew up playing games to accept the rules of a virtual world and incorporate the hand controllers and their button functionalities. However, in a sense, this notion is coherent with the stated correlation between age and technology habits, as video games is a relatively new invention and was simply not available to older users during their childhood.

> "*VR feels unnatural since I didn't grow up with it. PM is more like LEGO, and I know exactly how it works*"

The different experiences regarding interaction in VR and PM respectively can be seen in the quantitative data, on the scale *Awkward - Smooth*, where the average scale value for VR was 6.5 and for PM 8.43.

A few participants experienced simulation sickness in VR and felt ill. They stated that this affected their experience negatively, and that they thought that the illness was due to their lack of experience with such technologies.

Several participants expressed that VR felt novel and fun. Several discussions revolved around the many possibilities that come with VR, and the fact that the technology can support different perspectives as well as novel ways of collaboration through role play implementations in the interface.

> "*Anything is possible in VR*"

While VR was considered more of a fun experience, with close resemblance to video games, PM was described as a more serious technology, well suited to use as a work tool. This opinion can be seen in the quantitative data of the citizen groups, where the difference in scale values between VR and PM on the scale *Serious - Playful* was 2.75, where PM was experienced as less playful (mean = 5.63) than VR (mean = 8,38). However, the difference between the technologies on this scale was not expressed by the experts, were the technologies were deemed almost equally playful (VR mean = 7.67, PM mean = 7.00), with a difference of 0.67. The difference in opinion in the matter between the groups can be understood from the discussions that followed on the context of use. In the citizen group, it was widely acknowledged that the PM interface felt more like a work tool for planning, that arguably would be relevant for city officials. It was agreed that PM naturally invites discussion and stimulate the thoughts of the user. Although participants of the citizen groups expressed their appeal for these factors, and how they enjoyed the problem-solving aspects of PM, they still argued that the technology was better suited for city planners than for citizens. This discrepancy, where citizens seem to enjoy and value some aspects of a technology - but attributes the same to prove that the technology is better suited for other target groups than their own, might be explained by how the participants reasoned about their own role in urban planning.

> “*Do you want me to think about citizen dialogue as it is, or as an actual dialogue?*”

During the focus group sessions, the authors emphasized that the participants should consider the urban planning context and citizen dialogue when reasoning about the value of the technologies. While doing so, many expressed that VR was a better technology for targeting citizens due to its high resolution and scale 1:1, and ability to communicate established plans. The low resolution in PM was experienced as incomplete, and instead communicated an unfinished plan in need of processing. These reflections seem to point to the fact that citizens do not regard themselves as actors in the urban planning process, and feel that citizen dialogue is in fact not even a dialogue, but rather a one-way communication, in which they are passive receivers. This observation is clearly illustrated in the above quote from a participant from the citizen group, who requested clarification about the authors’ intended meaning of the word "dialogue" in the urban planning context.

> “*I was a city planner in PM, not a citizen - that was fun.*”

Conclusively, Most participants experienced more playfulness in VR than in PM, and associated VR with video- and computer games. The citizen groups felt that VR was better suited for citizens whereas PM was a work tool - better for city planning officials. The expert groups did not express this, and thought that PM was a good tool for citizen dialogue, since it stimulated discussions - an opinion shared by the citizen groups. The difference in opinion regarding the compatibility between PM and citizens seem to stem from the image of the role of the citizen in urban planning, rather than its ability to engage and stimulate the user throughout the interaction.

6.2.4 Potentials for learning

Key insights: *VR is perceived as a game, and participants did not associate it to a learning context. PM was thought to possess good prerequisites for learning due to its ability to present data clearly, from an overview perspective.*

When discussing the technologies’ ability to support learning, a recurring distinction was expressed across all groups, namely that VR is a static technology, in the sense that it requires predetermined conditions, whereas PM is dynamic and flexible, in the sense that the interface adapts to the input of the user, and the outcome will differ between each new interaction. VR is not used as an exploratory technology for situations where the goal is to reach novel insights, simply because the technology requires so much effort to be put on the visual details as well as the interactivity of a scene. For VR interactions to be meaningful, developers must have thought about all possible outcomes of a user scenario, and implemented answers to them in the experience. Several participants expressed positive emotions towards interacting with a dynamic interface, where they had the power to affect the outcome themselves. Some acknowledged that although the predictability of one's interactions can surely be counteracted by design, the feeling of a predetermined outcome would still remain.

The above opinions were supported in the quantitative data for the citizens group, in which PM was assessed as more active than VR on the *Passive - Active* scale, with a mean scale values of 9.12 and 7.25 respectively. However, the difference between VR and PM in the experts group for the same scale was only -0.33.

> "*It feels unnatural to think "learning" in VR. Instead I think in terms of game mechanics when a problem is to be dealt with*"

All groups discussed the benefits from the overview perspective in PM for data visualizations. It was argued that more data could be represented in a graspable way in PM. However, several participants brought up that much information was missing in the PM interface that was naturally represented in VR, such as height differences and inclinations.

Overall, participants agreed that PM provided better prerequisites for learning, much due to its active interaction, the overview perspective, and dynamic technology, in which several types of data visualizations could be presented either parallelly or serially. The biggest obstacle for learning in VR was according to several participants the learning threshold and the arbitrary mappings for interactions in the virtual world.

6.2.5 User perspectives

Key insights: *The two perspectives offered by the technologies are good for different interaction contexts. They both have advantages, and it is important to consider the perspective when designing a visualization. PM offers a clearer overview when data is to be presented.*

VR and PM offers two different perspectives: first-person, and overview. The participants acknowledged these differences as important qualities for the experiences that each technology provided. However, they struggled to compare these perspectives against each other, and extract how they affected the experience. There was no absolute consensus regarding what perspective to be preferred. Some participants argued that VR was the easier perspective to embrace, since the first-person view is how we perceive the world normally, whereas others felt that the overview perspective was better for a holistic overlook. One group discussed that many types of visualizations require an overview perspective, and that it is harder to visualize large amounts of data in VR with a first-person view, since the restricted field of vision does not allow for data to be visualized in layers.

Looking in the quantitative data, there is a clear advantage for PM over VR on the *Confusing - Clear* scale, with PM having a mean scale value of 8.93 and VR 5.36. This difference points to the benefit of presenting information from an overview perspective if the aim is to present data that is easy to overlook.

6.2.6 Collaboration

Key insights: *Participants enjoyed collaboration, and perceived that the group dynamics was stimulating and refreshing. The PM interface offered better prerequisites for collaboration than the VR interface. Many experienced that they were exposed in the isolation from the VR headset, and discussed the impact this could have in public spaces.*

In the user study conducted in the early phase of the project, a traction towards collaborative interfaces in urban planning contexts was identified among the participants. This attitude could partly be confirmed from the qualitative data, and further nuanced.

> "*We like to solve problems in groups - the group dynamics makes you think about different things than when you're alone*"

All participants stated that they enjoyed the collaborative aspects of PM. It was agreed that group dynamics and shared interactions was stimulating, and resulted in new thoughts and insights otherwise lost. As can be seen in the quantitative data, VR was experienced as isolated and PM collaborative on the *Alone - Collaborative*, with mean scale values of 3.57 and 9.21 respectively. However, when discussing differences in collaborative aspects between the technologies, it became clear that most participants did think that VR interfaces could be collaborative, but that the interfaces of focus were clearly distinguishable in this capacity. For example, one participant suggested an interface in which a team of users would be assigned with different roles so that they, although not in the scene simultaneously, could cooperate, and assist in the given task. Another group argued that there could be several users in the virtual world simultaneously, each experiencing the scene from different perspectives. Doing so, their different viewpoints could be complementary and divide the responsibility to contribute towards the goal. However, it was discussed, that communicating between users could be negatively affected since being several users in the virtual world would require avatar representations.

Based on the above discussions on collaboration from the focus groups, the authors believe that the quantitative data was affected by the vast differences in the interfaces in their abilities to support collaboration.

Despite collaboration differences, some things can be concluded about the technologies regarding collaborative aspects.

> "*I felt alone and silly in VR*"

Many participants mentioned that they felt exposed when being in the virtual world due to the isolation from the headset. The experience of feeling vulnerable to the co-participants' enforced a sense of self-awareness and this made some participants uncomfortable. The sensation of being exposed was more often expressed by the participants that were unfamiliar to VR but all participants were able to understand and relate to the feeling. It was argued in most groups that this was a strong downside to VR in the context of urban planning and citizen dialogue in public spaces. The case was made that this type of exposure in public spaces would further raise the threshold for interactions with VR.

6.2.7 Fidelity

Key insights: *High fidelity diverts focus from the interaction, task, and experience. Opinions regarding fidelity and its impact on the interaction differed between the expert- and citizen group, as citizens argued in favor of higher fidelity to increase the level of engagement.*

When discussing the effect of resolution (fidelity) for the technologies, it was noticed that the resolution was a notion only reflected upon with VR. It was apparent that many participants had high expectations on the fidelity in VR, and would discuss details and resemblance to reality when asked about the resolution. Although several participants did recognize the environment in the virtual scene, and could navigate based on their previous knowledge of the city district, they would still comment on such things as the lack of street signs, trams, and other things present in the real world.

It was widely agreed that the focus on details in the VR interface overshadowed the experience and interactions in the scene whereas the lack of details in the PM interface enabled for better focus on the task at hand. Although the level of details was a clear distinction between the two interfaces, several participants argued that VR would be meaningless without its resemblance to reality, and that the level of details was crucial for the experience, which was not the case for PM.

In the expert groups, the positive effects on interaction associated with lower fidelities was acknowledged, and it was discussed that these positive effects were easier to obtain with a technology as PM, whereas high expectations on fidelity in VR tended to shift focus from interactions to interface details of the virtual scene. This distinction was however not made in the citizen groups, in which it was instead discussed that the PM interface would have benefited from higher resolution and more details. For example, some groups argued that more details could have an emotionally engaging effect on the experience, an aspect lacking in the PM interface but not in the VR according to some.

6.2.8 Type of experience

Key insights: *VR is emotionally engaging, and evokes sensations of "realness", partly due to its first-hand perspective. The distance to the information presented in PM make good prerequisites for thoughtfulness and reasoning and was therefore regarded as a good planning tool.*

Judging from the above topics, VR and PM have different qualities that distinguishes them from each other on several levels, such as engagement, interaction, and level of detail. However, to describe the type of experience with a technology, a holistic perspective is required, descriptive of the accumulated sensation generated from several features of the technologies, and can partly be extracted from the participants' expressions throughout their interactions. For example, during interactions in PM, most participants expressed low-affective to no emotions, and were focused on solving the problem and discuss possible actions with their co-users. On the contrary, VR seemed to engage participants emotionally in a different way.

Participants in the virtual world often expressed their emotions vividly, and reacted stronger to events in the scene. In the focus groups, many spoke of VR's ability to engage emotionally whereas PM evoked a more serious feeling. Reconnecting to the *Serious - Playful* scale, it seems that the perception of VR as playful and game-like enforces its ability to evoke emotions in the users. Some participants mentioned that the flooding scenarios in the VR interface felt scary, something that no participant expressed with PM.

Conclusively, VR was experienced as the more playful technology, that evoked emotional responses in the users, whereas PM was perceived to be a serious working tool, good for city planning officials, and ideal for exploring multiple alternatives when the outcome is unknown.

7

Analysis

In an effort to provide a more conclusive perspective on the contributions of this thesis, this section will attempt to synthesize the acquired data with relevant theoretic frameworks and reconnect to previous research. The intended outcome of this discussion will compose a rationale and case for the guidelines of interactive technologies for citizen dialogue that will be presented in the forthcoming chapter.

7.1 Embedded, Embodied and Situated

Results from the semantic differential scales suggest that on average, the users found that projection mapping provides more clear and smooth interaction than VR. This opinion was confirmed during the focus groups where feelings like concreteness and comprehensibility were expressed when approaching and interacting with the interface. Several times, this opinion was stated with an expression of implication or common sense assumption, with a general approval within the group. When they were requested to elaborate on this feeling, it was motivated by notions such as a low threshold to engage and real time consequences of their actions.

A prerequisite for a human agent (user) to experience to be situated, or embedded, in an environment concerns how they through their body can interact and affect the world around them [19]. This predisposition allows for flexible and responsive cognitive processing of a situation, and is therefore an ideal state of mind. In the real world, the agent is continuously conditioned by the relation to its social, cultural and physical environment. Dourish [22] argues that embodiment is to be thought of as a participative status rather than physical relations; implying that anything that occurs in and depends on the real world, such as a conversation or an apple, are both examples of embodiment. In more strict physical terms, the world and the objects within it provide affordances [33] to the agent, which communicate particular properties that enable actions to be taken. Over time, structures and textures of objects become internalized properties for interaction, and embedded in the context. Thus, humans require a body to access some of the information stored in their environment, that they need in order to make proper assessments and become situated. In order for a virtual reality to compete at same level of perceptual resolution as the real world, every property of every element of the scene has to be created through a conscious intention to emulate the features of a real world environment down to a grain. Failing to recreate these conditions means the agent will have to fall back on rules of a pretend world, instead relying on mental models to mediate affordances through mental representations [33].

In two far from perfect interfaces created to evaluate the capabilities of two very different technologies, the interface that utilizes preconditions of the real world has an advantage in terms of affordances. The interaction context of the projection mapping interface is familiar to the user because they hold embodied and multimodal knowledge about it. The virtual environment may also be familiar if one has experienced it before, but arguably primarily through a then established mental model. Barring the haptic feedback from the hand controller, the environment itself does not provide opportunities for physical perceptive input, impeding the process of situating the user. As one of the expert users from the workshops put it *"Abstraction is a problem for virtual reality"*, referring to for example the teleportation function that allows the user to reposition in the scene. To begin with, establishing a mutual understanding of the concept of teleporting to a user which has not previously experienced it is not granted. This method of movement requires a preceding instruction, a process of trial and error once entering the scene, and the occasional confusion of forgetting which button on the controller that causes the action. Dourish [22] acknowledges that when the idea of something, such as teleportation, is the only aspect of the mental representation known to us, then there is no embodied construct. During the preceding user study, a preference toward physical models were also expressed with the same arguments of concreteness, clarity and intuitive interaction. If these types of visualization technologies are used in a public space to convey information or initiate dialogue, a technology which utilizes affordances to encourage interaction seem to be an essential feature.

When it comes to citizen dialogue, aspects not strictly tied to the interaction have to be considered. During the initial user study, desired functionalities like novelty and information exploration in terms of higher resolution data, were expressed. During the workshop many comments took a positive stance on VR's ability to realize a fictive scenario, or project a vision of the future and make it available to the user. As a virtual environment does not have to conform to any real world boundaries, the potential is unlimited. Comments on emotional reactions, discomfort and holding ones breath, during the rising water levels surrounding the user, witness about a dimension of embodiment that arguably can not be accessed through a real world interface. While the sensory perception of these events will not fulfill in VR, the impact of the simulated feeling will still be experienced to some degree. Oosterwijk and Barrett [61] mean that the perception of emotion is a construct of exteroceptive and interoceptive information integrated with knowledge representations. Translated to this scenario, an external stimulation of the visual perception of rising water levels coinciding with a changing heart rate and the knowledge of the potential outcome of being unsheltered from a moving body of water, could together elicit the feeling of panic or fear. Even in the absense of direct sensory input/physical stimulation, the body simulates experiential states [61], such as when one thinks about something painful, suggesting virtual conceptualizations can situate the user and trigger an emotional reaction. The ability to mediate a vision about a future event that may or may not occur depending on urban planning and development, such as extreme downfall causing a flooding, coupled with an emotional response could communicate a sense of urgency or gravity to the user, citizen, urban planning official, or policymaker.

7.2 Fidelity

It has been established within the field of Interaction Design that fidelity matters and decides to a high degree what the user will focus on in an interface. It is therefore important to consider when deciding on a technology for a given context. Depending on the type of technology, users will naturally focus on different things, and it is up to the designer to decide what aspects of the interface should be the focal point. This can be accomplished both through interface details, but, as found in the presented data, the type of technology will also have an effect. It was clear that the participants had higher expectations on the fidelity in VR than in PM, due to its high resemblance to reality. Contrarily, PM makes no claims by appearance regarding functionality or interaction possibilities, as there are no cemented implicit rules or cultural conventions to how information should be presented from an overview perspective. Some participants expressed their associations between the PM interface and 2D maps, something that might shine light on the matter. Approaching an interface with similar expectations as for a less sophisticated technology in terms of interaction possibilities and functionalities will naturally result in a positive experience. To the contrary, unmet expectations will result in frustration and confusion that might affect the user experience as a whole.

7.3 Space

Urban planning is a spatial endeavour, which require any medium for dialogue, such as interactive experiences, to consider how crucial information is presented to accommodate for this aspect. The two technologies and their respective interfaces work on different spatial premises. One operates in the hybrid space of a combined physical-digital world and the other in a completely digital space. Dourish [22] argues that while while VR systems exploit the user's acquaintance with features of the real world like gravity and three dimensions, it also removes them from the physical world. The digital world and the physical world co-exist unconditionally, exclusive of each other. In contrast, hybrid spaces do not project the user into a separate dimension, but instead combines the physical and digital, making them aspects of one another [22]. Maintaining the presence of the user in the physical world secures the preconditions for multimodal interaction, between users and interface. For multiuser interfaces that are governed by social interaction, eye contact, facial expressions, gestures, orientation and other embodied behaviors occur in shared physical spaces and therefore depend on its existence in order to make sense. In other words, if the interaction with the interface is embodied, meaning and effectiveness of multimodal communication is situated, and thus rely heavily on the nature of the context. In the VR interface, the user's are distributed in two different spatial domains, one in the digital and the rest in the physical, which embeds them in different interaction contexts and disrupts their communication.

A key aspect to interaction also regards intention and how it pertains to space. When the user picks up a barrier and places it in the PM interface, they do so with intent. The object, through which the intention is mediated, is connected through a metaphorical or abstract coupling that the user knows will transform their action to implementation. All user's present in the PM interface have a synchronized understanding of this action as they share the space within which it occurs, the temporal and causal relation between action and effect, and the embodied knowledge about how it is produced. In the VR interface, the intentionality primarily regards "activation" of objects rather than a function it can perform, leaving it up to the user to connect the action to its delayed effect of flooding prevention. Comments on the appreciation of real time consequences of actions came up regularly during the focus groups when discussing the PM interface, while comments on confusion regarding results of actions were brought forward for VR. This issue can most likely be attributed to design flaws of the VR interface rather than the technology itself, but its also a testament to the difficulties of working in scale 1:1 and the expectations it puts on performance. No current hardware could produce a particle system with generated physics substantial enough to cover Linnéstaden at scale 1:1. Even the mass equivalent to that of the PM interface (600 000 particles) would translate to giant floating particles that have no resemblance to water. Thus, any VR solution would have to involve a simulated effect as opposed to an actualized effect, being the PM case.

Physical space affords interaction constraints by nature. Humans are accustomed to these constraints because they govern the place in which the body, including the mind, is manifested. A designed object can exploit these constraints and take advantage of the innate human cognitive processing of its features in order to create intuitive interaction. A reality where these constraints are adhered to makes sense, and one that does not will confuse. Embodied processing aside, the advantage to creating novel modes of interaction, such as teleportation, is that it opens new design dimensions. During the workshop, the users previously unfamiliar with VR accepted the rule of the new world where one moves by teleportation, but they also posed questions like *"Am I able to walk through walls?"* or *"Can I stand in the water?"*, unable to process fundamental elements of interaction bodily, and instead constructing a new mental representation of the interaction constraints of their current reality. Since the rules of the real world are broken by instructions before entering the VR scene, its possible that a user might be unwilling to make their own inferences about what is possible.

7.4 Collaboration and Non-verbal Communication

Results from the evaluation workshops revealed significant differences in how participants perceived VR and PM in terms of their collaborative aspects. Although acknowledging that it is indeed possible to implement collaboration to VR, and that the interfaces of foci differed significantly, it was widely agreed that PM was a technology with inherent qualities supportive of collaboration. Contrarily, there were some negative aspects associated with VR if experienced simultaneously by several users - especially in public spaces and the context of urban planning.

According to many, a major deficiency of VR regarding its potential to support collaborative experiences was the exposure to surrounding people in combination with the isolation from the headset. This, it was argued, resulted in a self-awareness that inhibited immersion in the virtual world. Participants stated that, while in the virtual world, they felt uncomfortable when failing to perform the correct actions or not understanding the task. Regarding PM, the shared responsibility for the performance in the interface was comforting, and they did not feel self-aware in the same way, nor did they feel exposed.

During the workshops, participants were allowed and encouraged to collaborate, however, collaboration looked very different in the two interfaces. As the participants could see each other around the PM interface, their interplay included several aspects of natural communication such as verbal communication and gestures. Throughout the scenario in the PM interface, participants were forced to constantly come up with new strategies to reduce the negative consequences from the flooding and plan ahead for how the water would respond to their actions. The planning was manifested in both verbal- and non-verbal communication, and included indexical- as well as action gestures for communicating ideas and thoughts. If one were to listen in on the planning conversations without contextual knowledge or visuals of the non-verbal expressions, the conversations would make no sense, as most references to objects of the interface as well as potential actions were referenced in gestures.

Following the theory of socially distributed cognition, the participants as well as the interface they interact with can be viewed as one sociotechnical system, including human actors as well as technological artifacts, each with its own properties in terms of knowledge and abilities. Whenever a participant performs an action gesture, the represented action is simulated by the system, autonomously and distributed among the human actors, thereby informing the system of the potential effects the action might result in. As each communicated thought is integrated into the system through the distributed simulation of represented actions, a conformity of mind can appear among its actors, establishing the commonly mentioned sensation of shared responsibility. The idea here is that the participants do not regard themselves as individual actors collaborating with other individuals, but rather as a part of a uniform system, that strives to reach a common goal. To their aid while collaborating, the participants have a powerful tool to assist mutual understanding between the systems' actors, namely the neural mirroring that occurs subconsciously [71]. Whenever an action is communicated, either through verbal- or non-verbal communication, the neural patterns corresponding to that action are activated, thus generating a similar motorcortical response as the real action would [39]. Enforcing verbal expressions with semantically redundant gestures, or communicating new information with semantically incongruent gestures, either way enrichens the intended message, thus making it easier to conceive to the receiving part [43].

Analyzing the VR interface through the same framework, an explanation to the self-awareness and discomfort experienced while in the virtual world might be that there are two sociotechnical systems instead of one, with one of them only housing one human actor - the virtual visitor. These systems do not have a shared perception, nor a non-verbal communication tool for how to attend to the same things, creating a conceptual gap to overcome with their communication. They have to rely on their verbal communication being precise enough to convey meaningful messages. Additionally, the two systems are unequal regarding their perceptions of the corresponding system, since the "physical-world" system can attend to the all forms of communication conveyed from the "virtual-world" system whereas communication conveyed in the opposite direction is only attended to auditorily. The latter obstacle became apparent in several workshops, when observing participants performed pointing gestures in combination with incomplete verbal instructions to assist the virtual visitor, to solve the task. For example, when a participant was searching for the inactive flooding preventers in the virtual scene, the observing participants would exclaim *"there it is"* and point to the objects on the computer screen from which they were experiencing the virtual scene, obviously to no help for the virtual visitor. Often, they would not recognize their mistake before it was pointed out by a fellow participant.

On the other hand, pointing gestures proved to be easily conveyed from the virtual system to the physical, as the pointer actor attached to the virtual hand was visually prevailing, and could communicate locations in the virtual world with large precision. The virtual pointer was an interface detail that effectively made nonverbal communication easier, however, the same effect for action gestures could not be reached with interface details. Action gestures often consist of fine-tuned, subtle gestures involving simultaneous movements in space of different limbs, often arms, hands, and fingers [47]. In VR, the user's hands and fingers are tied to the hand controller through a grip, locking some vital gesture-producing limbs to an static position. Additionally, the possible actions in the VR scene are arbitrarily coupled to buttons on the controller that do not carry signifiers to the actions they will produce. Consequently, any gestures representing interactions with the hand controllers are perceptually disconnected from the affected object in the VR scene except from pointing gestures. Also, they do not relate to the action that would produce the same impact on the object if in the physical world, thus further increasing the conceptual gap between the two systems. The arbitrary mapping between buttons and their effects can however be internalized through experience, and must thereby not produce excessive mental load. However, the internalization of buttons will be dependent on visual stimuli from the virtual scene, as the same actions will not produce the same effect in the physical world - thus tying the internalized interactions to the embeddedness of one's body to the virtual scene. Hence, when observing a participant interacting with the hand controller in the virtual world from the physical, mirror neurons will not necessarily fire, since the observed action means something else in the physical world, and the visual percept of the observer is not congruent with the virtual scene.

In some focus groups, it was discussed that being simultaneously in the virtual world would facilitate cooperation between participants. Shared presence in the virtual world would arguably establish a mutual understanding for what the other participants experience, something that would facilitate communication. Pointing gestures would be easier understood, since shared attention is crucial for indexed object to be detected by an observer. However, being simultaneously in the virtual world would still not facilitate more complex forms of non-verbal communication, since these are often dependent on subtle finger movements which, first of all, are not existent in the virtual world and secondly, are impossible to perform if tied to a grip around a hand controller.

8

Guidelines

The aim of the guidelines presented in this chapter is to appeal to a broad selection of projects involved with urban planning in general, but specifically to those targeting citizens as user group. They should be considered as the processed and condensed outcome from present thesis, and have been empirically derived from the data generated from the user study and workshop, including a mix of quantitative and qualitative data.

1. Consider the goal of the visualization.

Different visualization technologies have different qualities that will support various goals. No technology possess the ability to satisfy all aspects of information visualization and context will thus matter. It is essential to consider the goal of a visualization to successfully select the technology best suited for the purpose.

The empirical data suggests that the visualization technologies of focus can not be located on a scale of good to bad in general. They both possess qualities that makes them favourable in certain contexts of use, as well as qualities that can be negative, if they are incongruent with the visualization purpose.

2. Naturally mapped interaction lowers threshold to engage.

Users do not want to feel incompetent. Arbitrary mapping results in higher thresholds for interaction that will deter some people from participating.

Many participants experienced difficulties when interacting with the hand controllers in VR and stated that this removed focus from the experience as a whole. Discussions regarding target groups revealed a shared opinion that interactions in VR are easier for users with technological experience or experience from similar interaction modes such as computer- or video games. For non-experienced users, the arbitrary interaction with hand controllers might pose a barrier that they must overcome before focus can shift entirely to the experience itself.

If an interface takes advantage of the embodied and multimodal knowledge that the user hold, interaction will benefit, as it will not require higher cognitive processing. The user will then be able to focus more attentional and cognitive resources on other aspects of the experience, such as the task, or learning outcomes.

3. Intuitive interaction supports information encoding.

> *Users that struggle with interacting in an interface will focus less on things as task, environment, learning outcomes etc. Intuitive interaction will provide the best prerequisites for encoding as well as retention of information.*

The less a user has to focus on interacting in an interface, the more mental resources will be available for processing information relevant to any any other objective, such as learning outcomes. In the focus group sessions, many participants experienced an induced workload from the arbitrary interactions in VR. Those who had no experience from interacting in VR stated that the process of interacting in the interface was more of a conscious process. Projection mapping on the other hand was easy and the interactions perceived as unconscious.

4. Users hold preferences that they are unaware of.

> *Regular users are not used to reason about the underlying factors for their engagement, and extracting these requires a thorough investigation and patience.*

There was a discrepancy between the citizen and the expert group regarding their opinions about fidelity. In the citizen group, the notion that higher fidelity would seize focus from interaction was not reflected upon. Instead it was argued that higher fidelity was always desirable. In the expert group however, there was a general consensus that higher fidelity impedes interaction and shifts focus towards interface details and look and feel.

Designing for users requires an understanding for their needs and an awareness of this sort of discrepancies. Being unaware of the factors that separates an expert from a user will create a gap between the groups that might pose a problem in encounters that aim to stimulate dialogue and create mutual understanding between the two.

5. Higher fidelity favor attention to detail over bigger picture.

> *Data can be harder to detect and attend to in an interface with high fidelity. High resolution details will draw attention and attract users' interest, sacrificing the holistic perspective on the interface.*

The empirical data tangents a general consensus in the field of interaction design, saying that the higher fidelity to higher expectations will be put on the performance of an artefact. Also, with higher fidelity, the user will focus on appearance rather than the interaction.

6. Immersion and presence in an interface can cause emotional engagement.

> *The sensation of being immersed in a realistic interface is powerful and will engage users emotionally. This engagement will enforce the sensation of gravity in relation to given tasks.*

It was widely agreed that the VR interface offered an immersive experience due to its resemblance to reality in terms of scale and perspective. These factors led the participants to engage in the virtual context and environment, since they were located inside it. The engagement triggered emotional responses in the VR interface when the participants for example experienced flooding scenarios.

7. Interfaces that support multimodal interaction ensure efficient communication.

The advantage of multimodal, interpersonal communication in multi-user scenarios should be harnessed to the benefit of the interaction. The socio-cognitive stimulation that follows enables users to act in the environment they have evolved for.

It was both observed during interaction and extracted from the focus groups that collaboration benefited from multimodal interaction and non-verbal expressions. In the PM interface, gestures were heavily relied on to enforce meaning to verbal expressions, as well as convey action propositions in planning contexts. Non-verbal communication occurred also in the VR interface, but to less benefit, as the communicating actors were located in different spaces and could not see each other - something that points to the fact that much of natural communication happens subconsciously, and is lost whenever the normal mediating mode is disrupted.

8. Context is as important as content.

Context will situate the users to enable them to engage with an interface more naturally. It is a precondition for optimal cognitive processing and will ensure users focus on an intended task, and avoid confusion over non-relevant details of the interaction.

Without an interaction context, participants will not attribute meaning to the interaction, and thus not engage with the experience. This insight became apparent in early tests of usability and was further observed during the evaluation workshop. Some participants mentioned that the context in the VR interface felt forced, and that their experience was perceived as more disrupted than the coherent PM scenario.

9. Actions expect consequences.

Users expect a temporal connection between actions and effect. Clearly temporally mapped feedback is essential for users to grasp the environment in which they interact.

Discussions regarding the visibility of states of interface elements revealed that unclear temporal mapping was a factor for confusion. Participants expressed appreciation for, and enjoyed real-time feedback regarding several interface details, and attributed several situations of confusion to temporal delay between action and perceivable effect.

10. Design requires interpretation.

Design decisions are a result of the interpretation and synthesis of a designers knowledge, experience and data. Data without interpretation is just results, and will not make for an engaging or novel experience.

The generated empirical data shows that the technologies were perceived differently in terms of experiences and the reactions they evoke. Naturally, the technology that a designer chooses to use for visualizing information will heavily affect how the end user receives and perceives that information. As such, the designer concerned with interfacing issues of urban planning with a citizen has a responsibility to consider how their role affects the users perception of that issue, and adjust to ensure the experience is correctly addressed.

The information to be conveyed through a design needs to be transformed by the designer to a fitting format for the given technology and interaction context. Context is subjective, and the end result relies on the designer's ideas on how the information is best conveyed, whether the aim is to evoke emotions or to present data as objectively as possible. This transformation is a necessary phase, as there would not be a designed interaction without a conscious intent to create an experience dedicated to a certain cause.

11. Constraint help form the mental models of users.

If the interaction constraints of the interface are not clear, confusion will follow.

In physical space, affordances offer reliable clues to how one can interact with the environment. The real world is predictable by the embodiment of its constituents, and thus provides us with constraints that form our mental models and guide interaction. In virtual space, the environment does not provide clear affordances, but the user instead rely on preconceptions and rules regarding the specific interface. If the constraints are unclear, interaction will suffer, and mental allocations will be needed to interpret the environment.

12. Collaboration stimulate thinking and discussion.

Cognition emerges in the interplay between collaborating actors of a sociotechnical system. Also, such systems provide a safety net for mistakes as the responsibility for the interaction outcome is distributed among its actors.

A theme of discussion in the focus groups was the exposure to observers that was disturbing, especially when interacting in VR, a technology that isolates the user from the physical environment. Participants argued that the social context that was facilitated by the PM interface led to a feeling of shared responsibility, something that released tensions from being exposed to observers. This notion can be explained in terms of socially distributed cognition, stating that propositions of actions are simulated mentally among the human actors of a sociotechnical system, leading to conformity and mutual understanding as well as a sensation of belonging to a bigger system. In the isolation of VR, participants do not regard themselves as parts of a sociotechnical system but instead as an individual actor, separated by a dimensional barrier from the social system located in the physical space.

13. Public spaces are an exposed setting.

Exposure to observers can be intimidating to some and comprise a threshold for interaction. Being isolated in a sociotechnichal system adds to the sensation of alienation. Users are comfortable when embedded in the same context of interaction.

The focus group sessions revealed that many participants considered the public spaces to constitute exposed settings. The isolation from a VR headset for example, was associated with discomfort, when interaction takes place in new social contexts. A possible explanation to the discomfort is that participants feel more relaxed when they share a sociotechnical system with other human actors, and thus also share interaction responsibilities.

14. Non-verbal communication supports collaboration.

Different information is conveyed non-verbally that words cannot compensate for. In addition to provide new information, non-verbal communication can enforce information conveyed verbally.

Theoretical frameworks through which the results were interpreted states that non-verbal communication not only enforces spoken utterances, but provides new information that might be essential for a receiver to conceive the intended message. Connotations are often communicated non-verbally, and can be hard to replace when two actors interact in different sociotechnical systems.

15. Overview perspective is good for complex/layered data visualization.

Feeling in control, with a clear perspective of all available information will ensure a deeper understanding for the presented data.

The results from the focus group session revealed that participants favored an overview perspective before a first-hand perspective when complex or layered data is presented. It was argued that perceptual excess occluded relevant information and led to a shift in focus.

9

Discussion

The aim of present thesis was to contribute to the field of Urban Planning with an evaluation method capable of determining the efficiency and values of visualization technologies in relation to each other, irrespective of their different interaction opportunities. Introducing this chapter will be a response to the thesis' research questions, followed by a discussion surrounding the implications and limitations of results of the study.

The research questions for the thesis were:

1. *How can projects within urban planning be improved in terms of citizen interest, by adhering to a compound evaluation process of user experience and cognition?*

Projects within urban planning and development inherently compose wicked problems, where there is never a single correct answer to an issue, much like there is never a single solution to design problems. In urban planning, the target groups are as diverse as they come, mainly between groups but also within. When a project looks to involve its citizens, the target group itself cannot be expected to be fully captured from user studies alone. In order to sustain uncertainty on such a level, a project should expect that their point of origin is just somewhere in the design space and from there operate iteratively to uncover knowledge incrementally. As brought up in the thesis background, participatory design does not guarantee a successful outcome. During the evaluation workshops a number of differences between the citizen group and the expert group were discovered, which attests to this fact. As design evaluation is mostly concerned with understanding the user, problems arise when looking to the user for all the answers. Filtering the data through a lense of cognitive science enabled this thesis to work and question the results from a frame of information processing, which is central to the interaction between human and artefact. The inclusion of cognitive processing as feature of evaluation was expected to provide the angle of interaction unable to be captured by the self-reported nature of user experience data. To exemplify this reasoning, a technology could be recognized as the source of engagement by either 1) being reported qualitatively by the users themselves, and 2) being derived by the researchers from analyzing the evaluation data in light of relevant cognitive frameworks. By making sense of the data on a meta level, not simply compiling it and decide how it should guide process, a large scale urban development project might avoid some unnecessary pitfalls. The discussion of interface fidelity separated the user group and the citizen group because the users could simply not imagine why high resolution details could be an issue to the cause, pointing to this very effect. Typical for design research is to keep a prototype at the lowest possible fidelity level necessary to accomplish its aim, because it provides the frame of reference from which the user will make their assumptions. From a cognitive standpoint, the more detailed a prototype is, the less room there is for interpretation, exploration and problem-solving on a basis of free association. Depending on the objective of an interaction, this may or may not be a problem. Regardless, it has to be designed for.

While the guidelines produced as the result of this thesis have yet to be tested and verified, being generated from several levels of refinement (data - interpretation - analysis), where perspectives on the findings are provided, ought to make them more applicable and reliable. User experience evaluation provides the necessary foundation of target data in order to ensure a data-driven process, and cognitive frameworks provide the theoretical underpinning that establishes the connection between the findings of a single study and known factors of human information processing.

2. *Which are the key factors for creating engaging, valuable and meaningful interactive visualization interfaces to support citizen dialogue in the context of urban planning?*

To a large extent, the answer to this question is the guidelines in chapter eight. In contrast to how the first research question takes the structural perspective of the thesis, this question formulates the practical approach. The guidelines are a produce of a deliberate research strategy to target meaningful factors that need to be considered when designing for citizens to engage in dialogue. While the intent is to provide a strategy for practice, the guidelines are adapted to a level of abstraction that ought to make them general enough to be applied to any design of interactive and visual experiences to support dialogue. As such, the key factors can be considered a frame of relevance that lets the designer address the right aspects, irrespective of the nature of the technology at hand.

3. *How do the visualization technologies of focus relate to each other, what attributes set them apart, when do they prosper, and under what circumstances do they underperform.*

This question has served more of an internal purpose to the thesis by giving structure to how to succeed in comparing interactive technologies with few shared attributes. As the interfaces created to represent each technology rely on the reliability of the comparison in order to be legitimate, the act of formalizing this relation to a research question facilitated a vital part of the work.

Chapter five, six and seven are all testaments to the differences between virtual reality and projection mapping. Execution provides contrast by highlighting the interaction aspects and literal differences of each, describing how and why each component is crucial to the technology in order for the user experience to execute properly. Here, the construction of the prototype creates the foundation for further analysis by identifying all the elements that define it. The results chapter contributes with the actual user data of what the *perceived* differences are between them, equally significant when designing for interactive experiences. Threshold for interaction and experience types, among others, were result-specific topics brought up and acknowledged to be aspects where the technologies diverged. Analysis concludes the comparison by discussing the underlying cognitive processing that each technology affords, and why the collected results reflect these facts. Contributions from this stage include, for instance, elaborations on embodiment and situatedness and why their significance impacts each technology. Combined, these chapters comprise a holistic perspective on how projection mapping and virtual reality relate to each other, as well as their optimal preconditions for interaction.

By reviewing existing literature and work on urban planning and interactive technologies, an understanding of priority, feasibility, and research value could be mapped out to set an initial boundary for the scope of the project. This compilation suggested that evaluating user experience alone would not provide a complete account and response to the research objectives, and that stretching the user experience domain beyond could produce a more faceted perspective of the advantages of each technology. Analyzing the needs of the user group before any forming of concepts or design took place ensured the research approach to be data-driven and user centered. The user study conducted prior to the design stage informed the process by providing a set of themes generated by values and needs of the target group. Systematic interpretation of the themes developed an understanding of what type of visualization technologies had the most potential to meet those demands. These aspects were central to establishing a solid foundation for the study.

The data collected from the workshops suggest a difference in opinion between the citizen group and the expert group on some aspects. These results point to one of the most valuable outcomes of the project - *the difference between these groups is likely where dialogue fails.* The group responsible for executing an urban development project is rarely, if ever, subject to the study itself, meaning opinions and knowledge they possess is not structured in the form of data that can be collected and compared, but rather directed to developing methods for analyzing the intended target group. For instance, the expert group mentioned the complex of problems regarding interface resolution where they comply with what previous research claims on the matter. The citizen group on the other hand considered higher resolution to be nothing but an advantage. While the expert group has good reason for their opinion, the discrepancy between the groups has to be acknowledged and accounted for in order to supply the right preconditions for dialogue. If citizens appreciate to attend to interface details and planning officials want to initiate dialogue, the question should concern how interaction can be designed around that premise rather than focus on from which opinion the correct solution will follow. By analyzing through this type of pseudo between-groups research design, critical factors where attitudes diverge can be identified and brought forward.

Having a clear intent as to why to involve the target group, and how that can be expected to affect the process, provides meaning to the data and consequently useful instructions to design for. Participatory design for the sake of participation, without a conscious objective of how or why that data will be a useful contribution, can not be expected to provide value. Design requires interpretation. Without preconceptions of possible useful implementations of collected data the result might not be novel or engaging. The initial user study in this thesis contributed by identifying key themes of engagement, appreciation and interest, through which interpretation would come to constitute the requirements for selecting visualization technologies.

Using a mixed method approach with more than one type of method for data collection improves reliability of the results and allows for more nuanced data. Different types of data provide context and perspective, while reducing the risk of misinterpretation. On their own, the quantitative results from the semantic differential scales could only point to attitude tendencies, but when coupled with the qualitative data the conclusions could be considered more decisive. Not attempting to approach the evaluation in a summative fashion, when data of that nature would likely not result in any meaningful insight, meant efforts could be directed toward optimizing the method after the type of data likely to be acquired.

9.1 Limitations

The main limitations to this study concern the design of the interfaces as representations of the to be evaluated technologies, as well as the fact that it only encompasses two technologies from a range of options. The interfaces were mid-fidelity prototypes that would have benefited from more time dedicated to development to ensure a better connection between representation (prototype) and technology, and relieve the users of the pressure to suppress their experience of the interface and think about the potential of the technology. Moreover, expanding the range of technologies involved in the study may have affected the outlook on the guidelines. While striving to be as general and applicable as possible, the guidelines are to a large extent a result of the data. However, the analysis chapter attempts to synthesize the data with theoretical frameworks, abstracting it to a more general level of insight, improving the guidelines.

9.1.1 Evaluation of technologies

Critique can be directed toward the way the interaction patterns of each interface was constructed. The PM interface had actualized effects of the users' actions, where the system responded to their input in real time by altering the visuals. Whereas the actions in the VR interface simulated a delayed effect, disconnected temporally and with only narrative support to imply an action-consequence relation. Besides this, the projection mapping interface required a significantly longer amount of time to develop and implement to a level of fidelity that made it possible to evaluate. The result being a more limited amount of time left to develop the VR interface, leaving an overall lower fidelity prototype with less functionality and interaction than desired. This fidelity discrepancy constitutes motive for questioning the level of comparability between the interfaces. However, as pointed out continuously to the users, they should be considered to be prototypes and the focus of the evaluation should concern the underlying technologies.

9.1.2 Alternative technologies

A wider range of visualization technologies included in the study could have enabled a richer analysis and an ability to capture and represent the user themes to a larger extent. Potential insights from a more comprehensive study could provide new perspectives for the guidelines where other technologies might be more inclined toward a certain theme. For instance, a mobile technology would have a close connection to the Availability theme, potentially causing users to reason differently around the other technologies and their attitudes towards them.

9.1.3 Data validity

Due to the Corona pandemic outbreak of 2020, the citizen user groups had to be limited to colleagues and acquaintances rather than a diverse set of individuals accessible through public spaces. It is possible that the data might have looked differently in such an instance, and if a similar study were to be replicated that would be the recommended approach.

The quantitative data is limited in its capacity to generate conclusions because of a low sample size. However, when collecting qualitative data, the number of participants must always be weighted against the amount of data extracted from each participant. Thus, although working with fewer participants in present study, large amounts of qualitative data was generated, well able to cover the topics of interest. That being said, the semantic differential scales should be considered more of a tool through which the qualitative data can be analyzed, and make suggestions of possible tendencies.

9.1.4 Research design

There was an explicit formative approach to this study because of the difficulties with comparing technologies with a diverse set of attributes. While the advantage of such an approach is that evaluation of parts otherwise hard to compare is made possible, the downside is that the results can not be considered to be conclusive. Should a future version of this study strive for a generalizable result, it is recommended to attempt to equalize the conditions between technologies and establish a controlled test environment. In order to make statistically significant claims, clear measurables such as time on task and tests of recall, possibly combined with biometric measurements, could be investigated.

9.2 Ethical Issues

This study aimed to inform a large-scale urban development project with guidelines for why, when and how to approach interactive technologies to support citizen dialogue. Doing so, it had the potential to influence its outcome that by extension might provide the data basis upon which impactful decisions are made. Creating unbiased, objective interactive visualisations is difficult since a designer is required to interpret results and realize design concepts in order for there to be something to evaluate, thus influencing how they are perceived. Therefore, it is important to investigate how visualization technologies are received by the user, and evaluate them from a cognitive perspective. The results from this study outline the strengths and weaknesses of two visualization technologies, which together with a theoretical analysis will make up guidelines that enable Virtual Gothenburg Lab to utilize these technologies more efficiently.

A dilemma regarding the user studies is that data was collected from a sample that does not fully represent the population of Gothenburg, mainly due to the Corona pandemic. Limiting the sample for availability reasons poses a democratic dilemma, as some groups will not be represented in the data, and will thus be excluded from any measures taken towards a better understanding of visualization technologies. In a more ideal situation, measures would be taken to counteract the risk of receiving a non-representative sample.

Despite this, it should be added that the groundedness in cognitive theories through which the data was analyzed counteracts against the limited representation of citizens. The theoretical frameworks used to interpret the data makes greater claims of generalizability, and it is conceivable that the guidelines originated in the theoretical analysis would transfer to wider samples of the citizen population. This proves why it is important to be theoretically grounded in research, and how it can be a question of democracy.

9.3 Future Research

For future work on this topic it would be interesting to see an evaluation including a larger number of interactive technologies, possibly spanning wider than having visualization as the primary feature. Moreover, such a study could benefit from collecting complementary data that is more objectively oriented, such as biometric data for measures of engagement, cognitive load, eye gaze, or skin conductance. Doing so might provide additional insights to the results of present study regarding implicit data, that might not have been obtained since it is hard to articulate or derive from introspection only, as was required from the interview format used in this study.

Scenario-based vocal interaction could provide a strong narrative that would enrichen the user experience, regardless of the underlying technology. The scenario provided to the PM interface was appreciated by the participants, and lacking in VR. This study revealed the importance of providing the user with a solid context, in which the interactions and tasks makes sense and are perceived as meaningful. The authors want to call on future work to investigate further the effect of adding strong narratives to the interactions, and especially look further into whether the effect differs between technologies. The strong effect from narratives, and scenario-based interaction has been well-documented in Interaction Design studies, but the potential difference in effect between technologies remain an unexplored field. Having the comparative approach exercised in present thesis presents a way to combine different methods to attain information on the complex matter of how interfaces of different technologies are perceived. For visualization technologies to be used in the contexts within which they will perform at their best, more studies are recommended to follow this approach.

10

Conclusion

Present thesis aimed to contribute to the fields of Urban Planning and Interaction Design with insights regarding the potentials and shortfalls of different visualization technologies in the context of citizen dialogue. Comparing technologies with fundamentally different qualities in terms of interaction possibilities and modes of visualization is important for them to be appropriately applied in their right contexts and thus reach their full potential.

Two interfaces were developed with the aim of visualizing the same information using the tools and methods significant for the respective technology. A data-driven design process ensured that user interests were of prominent foci in the development phase. An evaluation workshop was conducted with the purpose of exploring what qualities are associated with the technologies, and compare them in regard to their ability to support both user interests, as well as cognitive mechanisms important for human interactions.

Analysing the results from the workshop, findings were interpreted through cognitive frameworks including embodied, situated, and socially distributed cognition, as well as multimodal interaction. Qualitative differences between virtual, physical, and hybrid spaces was discussed, as well as their implications within experience design and information visualization. Insights from the processed data together with conclusions from the analysis were then condensed into 15 guidelines to consider in future endeavors within the field of urban development and citizen dialogue in particular.

The guidelines are defined in such a way that they should be widely applicable to similar projects, when exploring technologies to consider for visualization interfaces, or when considering design details during development phases. They span various aspects of information visualizations such as how to think about the fidelity of an interface, the role of contextualizing information, the importance of designing for communication and collaboration, the issues of using public spaces as places of exhibition, and more.

Adhering to these guidelines will by no means guarantee a successful design, but will provide tools and directions to explore when discovering a certain problem space in this field. The broad character of the guidelines is an approach that is well suited to deal with uncertainty and wicked problems, common in both design and urban development. Designers working through narrow frameworks will phase challenges whenever a context demands novel solutions to unexplored situations, and a framework that does not take uncertainties into account will thus prevent the designer from deviating from their path. The presented guidelines are not prescriptive in that sense, but instead aims to establish a common terminology to the field regarding visualization technologies, and will hopefully be applied and further developed in future studies in the field of Urban Planning and Interaction Design.

References

[1] Martha Alibali, Rebecca Boncoddo, and Autumn Hostetter. “Gesture in reasoning: an embodied perspective”. In: *The Routledge Handbook of Embodied Cognition.* Ed. by Lawrence Shapiro. London & New York: Taylor & Francis Group, 2014, pp. 150–159.

[2] Devon Allcoat and Adrian von Mühlenen. “Learning in virtual reality: Effects on performance, emotion and engagement”. In: *Research in Learning Technology* 26 (2018), pp. 1–13.

[3] Jens Allwood. “Språk, kommunikation och kognition”. In: *Kognitionsvetenskap.* Ed. by Jens Allwood and Mikael Jensen. Lund: Studentlitteratur, 2012, pp. 395–408.

[4] Luis Alonso et al. “CityScope: A Data-Driven Interactive Simulation Tool for Urban Design. Use Case Volpe”. In: *Springer Proceedings in Complexity.* Springer, 2018, pp. 253–261.

[5] Visual Arena. *Virtual Gothenburg Lab.* 2020. URL: `https://visualarena.lindholmen.se/projekt-forskning/virtual-gothenburg-lab` (visited on 05/28/2020).

[6] *Augmented Urbans.* 2020. URL: `https://www.augmentedurbans.eu/` (visited on 05/31/2020).

[7] Eran Ben-Joseph et al. “Urban simulation and the luminous planning table: Bridging the gap between the digital and the tangible”. In: *Journal of Planning Education and Research* 21.2 (2001), pp. 196–203.

[8] D Benyon. “Envisionment”. In: *Designing Interactive Systems: A comprehensive guide to HCI, UX and interaction design.* Ed. by D Benyon. Pearson Education Canada, 2014, pp. 177–197.

[9] Monica Billger, Liane Thuvander, and Beata Stahre Wästberg. “In search of visualization challenges: The development and implementation of visualization tools for supporting dialogue in urban planning processes”. In: *Environment and Planning B: Urban Analytics and City Science* 44.6 (2017), pp. 1012–1035.

[10] Kathryn Bock and Susan M Garnsey. “Language processing”. In: *A companion to cognitive science:* ed. by William Bechtel and George Graham. 4th ed. Oxford: Blackwell, 2006. Chap. 14, pp. 226–234.

[11] Michelle A. Borkin et al. “Beyond Memorability: Visualization Recognition and Recall”. In: *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), pp. 519–528.

[12] Michelle A. Borkin et al. “What makes a visualization memorable”. In: *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), pp. 2306–2315.

[13] Andrea Botero and Joanna Saad-Sulonen. “Co-designing for new city-citizen interaction possibilities: weaving prototypes and interventions in the design and development of Urban Mediator”. In: *Proceedings of the Tenth Conference on Participatory Design, PDC 2008*. January. 2008, pp. 269–272.

[14] Chalmers. *VirtualCity@Chalmers*. 2018. URL: `https://virtualcity.chalmers.se/about/` (visited on 04/17/2020).

[15] A Cooper et al. *The Essentials of Interaction Design*. Vol. 13. 6. 2014, pp. 593–596.

[16] N Dahlbäck, J Rambusch, and T Susi. “Distribuerad kognition”. In: *Kognitionsvetenskap*. Ed. by J Allwood and M Jensen. Lund: Studentlitteratur, 2012, pp. 487–495.

[17] Peter Dalsgaard. “Participatory Design in Large-Scale Public Projects: Challenges and Opportunities”. en. In: *Design Issues* 28.3 (2012), pp. 34–47.

[18] Peter Dalsgaard and Eva Eriksson. “Large-scale participation: a case study of a participatory approach to developing a new public library”. en. In: *the SIGCHI Conference*. ACM Press, 2013, p. 399.

[19] Michael Dawson. “Embedded and situated cognition”. In: *The Routledge Handbook of Embodied Cognition*. Ed. by Lawrence Shapiro. London & New York: Taylor & Francis Group, 2014, pp. 59–67.

[20] Thomas Delman et al. “Utopiatyping- New tools & strategies for citizen participationin urban development”. en. In: *ACSP/AESOP Joint Conference Bridging the Divide: Celebrating the City*. Chicago, 2008.

[21] Derivative. *Touchdesigner*. 2020. URL: `https://derivative.ca/`.

[22] Paul Dourish. *Where the action is*. Vol. 31. 2. 2004.

[23] Epic Games. *Unreal Engine*. 2020. URL: `https://www.unrealengine.com`.

[24] Eva Eriksson and Josef Wideström. “Staging the Interaction – Explorative Interventions for Engaging Citizens in the Development of Public Knowledge Institutions”. en. In: *Design Research Society’s Conference 2014*. November 2019. 2014, p. 14.

[25] Eva Eriksson and Josef Wideström. “The Virtual Culture House – shaping the identity of a public knowledge institution”. en. In: *European Academy of Design Conference Proceedings 2015*. Sheffield Hallam University, 2015.

[26] Jodi Forlizzi et al. “Quality control: A panel on the critique and criticism of design research”. In: *Conference on Human Factors in Computing Systems - Proceedings*. 2011, pp. 823–826.

[27] Bo Fu, Natalya F. Noy, and Margaret Anne Storey. “Eye tracking the user experience - An evaluation of ontology visualization techniques”. In: *Semantic Web* 8.1 (2017), pp. 23–41.

[28] S Gallgher. *How the body shapes the mind*. Oxford: Oxford University Press, 2005.

[29] Peter Gärdenfors. “The Role of Intersubjectivity in Animal and Human Cooperation”. In: *Biological Theory* 3.1 (2008), pp. 51–62.

[30] William Gaver. “What should we expect from research through design?” In: *Conference on Human Factors in Computing Systems - Proceedings* November (2012), pp. 937–946.

[31] *Gävle: Making places for biodiversity awareness - Augmented Urbans - Medium.* 2020. URL: `https://medium.com/augmented-urbans/g%7B%5C%22%7Ba%7D%7Dvle-making-places-for-biodiversity-awareness-5a7c4b090d75` (visited on 04/17/2020).

[32] Jan Gehl. *Life, space, buildings — in that order.* Island Press, 2010.

[33] James Gibson. *An ecological approach.* Psychology Press, 1979.

[34] Eric Gordon and Jessica Baldwin-Philippi. *Playful Civic Learning: Enabling Reflection and Lateral Trust in Game-based Public Participation.* Tech. rep. 2014, pp. 759–786.

[35] A Gulz, M Jensen, and J Rambusch. “Lärande och kognition”. In: *Kognitionsvetenskap.* Ed. by J Allwood and M Jensen. Lund: Studentlitteratur, 2012, pp. 219–228.

[36] Rex Hartson and Pardha S. Pyla. *The UX Book - Process and Guidelines for Ensuring a Quality User Experience.* en. Waltham, USA: Elsevier, 2012, pp. 1–937.

[37] Jayne Al-hindawe. “Considerations when constructing a semantic differential scale”. In: *La Trobe working papers in linguistics* 9.7 (1996), pp. 1–9.

[38] James Hollan, Edwin Hutchins, and David Kirsh. “Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research”. In: *ACM Transactions on Computer-Human Interaction* 7.2 (2000), pp. 174–196.

[39] Autumn B. Hostetter and Martha W. Alibali. *Visible embodiment: Gestures as simulated action.* 2008.

[40] Jamil Hussain et al. “A multimodal deep log-based user ex perience (UX) platform for UX evaluation”. In: *Sensors (Switzerland)* 18.5 (2018).

[41] Judith E. Innes and David E. Booher. “Reframing public participation: Strategies for the 21st century”. In: *Planning Theory and Practice* 5.4 (2004), pp. 419–436.

[42] *Investment in Digital Twins will revolutionise urban development.* 2020. URL: `https://www.chalmers.se/en/departments/ace/news/Pages/Digital-Twins-will-revolutionise-urban-development.aspx` (visited on 04/17/2020).

[43] Mikael Jensen. *Interpersonell kommunikation.* Lund: Studentlitteratur, 2015.

[44] Jeff Johnson and Austin Henderson. “Conceptual models: begin by designing what to design”. In: *Interactions* 9.1 (2002), pp. 25–32.

[45] Jeffrey A Juarez and Kyle D Brown. “Extracting or enpowering: A critique of participatory methods for marginalized populations”. In: *Landscape Journal* (2008).

[46] M Kaltenbrunner and R Bencina. “reacTIVision: A Computer-Vision Framework for Table-Based Tangible Interaction”. In: *Proceedings of the first international conference on "Tangible and Embedded Interaction (TEI07).* Baton Rouge, Lousiana. URL: `http://reactivision.sourceforge.net/`.

[47] Adam Kendon. *Gesture: Visible actions as utterances.* Cambridge: Cambridge University Press, 2004.

[48] K Kobayashi et al. “Collaborative Simulation Interface for Planning Disaster Measures”. In: *CHI 2006*. ACM, 2006.

[49] S. F. Kuliga et al. “Virtual reality as an empirical research tool - Exploring user experience in a real building and a corresponding virtual model”. In: *Computers, Environment and Urban Systems* 54 (2015), pp. 363–375.

[50] Heidi Lam et al. “Empirical studies in information visualization: Seven scenarios”. In: *IEEE Transactions on Visualization and Computer Graphics* 18.9 (2012), pp. 1520–1536.

[51] Thomas Leduc, Francis Miguet, and Vincent Tourre. “Motion Perspectives Integration in the Qualification of the Urban Spaces: Towards a 2D- and a 3D-Enrichment of the S-Partition Method”. In: *Urban Design and Representation: A Multidisciplinary and Multisensory Approach*. Cham, Switzerland: Springer International Publishing, 2017. Chap. 6, pp. 69–82.

[52] Youn Kyung Lim, Erik Stolterman, and Josh Tenenberg. “The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas”. In: *ACM Transactions on Computer-Human Interaction* 15.2 (2008).

[53] J Lindblom, T Susi, and A Tysk. “Social kognition”. In: *Kognitionsvetenskap*. Ed. by J Allwood and M Jensen. Lund: Studentlitteratur, 2012, pp. 383–392.

[54] Jonas Löwgren and Erik Stolterman. *Design av informationsteknik: Materialet utan egenskaper*. Lund: Studentlitteratur, 2004.

[55] Bruce Martin and Martin Hanington. *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions*. en. Vol. 49. 10. Beverly, MA: Rockport Publishers, 2012.

[56] E McCrady. *Bibliothèque Nationale: A Building Hostile to Preservation and Access*. 1998. URL: `http://cool.conservation-us.org/byorg/abbey/an/an22/an22-4/an22-405.html`.

[57] Helena Mitasova et al. “Scientific visualization of landscapes and landforms”. In: *Geomorphology* 137.1 (2012), pp. 122–137.

[58] Donald A. Norman. “Signifiers, not affordances”. In: *ACM Interactions* 15.6 (2008), pp. 18–19.

[59] Donald A. Norman. *The design of everyday things*. New York: Basic Books, 2013.

[60] NRF. *Virtual Singapore*. 2020. URL: `https://www.nrf.gov.sg/programmes/virtual-singapore` (visited on 05/31/2020).

[61] Suzanne Oosterwijk and Lisa Feldman Barrett. “Embodiment in the construction of emotion experience and emotion understanding”. In: *The Routledge Handbook of Embodied Cognition*. Ed. by Lawrence Shapiro. London & New York: Taylor & Francis Group, 2014, pp. 250–260.

[62] Anne Marie Oostveen and Peter Van Den Besselaar. “From small scale to large scale user participation: A case study of participatory design in e-government systems”. In: *PDC 2004 - Proceedings of the Eight Participatory Design Conference 2004 - Artful Integration: Interweaving Media, Materials and Practices*. 2004, pp. 173–182.

[63] Sharon Oviatt et al., eds. *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1*. Vol. 1. ACM, 2017.

[64] Anna Petrasova et al. “Tangible modeling with open source GIS”. In: *Tangible Modeling with Open Source GIS* (2015), pp. 1–135.

[65] Barbara E.A. Piga and Rossella Salerno. *Urban design and representation: A multidisciplinary and multisensory approach.* Cham, Switzerland: Springer International Publishing, 2017, pp. 1–223.

[66] Ben Piper, Carlo Ratti, and Hiroshi Ishii. “Illuminating Clay: A Tangible Interface with potential GRASS applications”. In: *Open source GIS - GRASS users conference 2002.* 2002.

[67] M. E. Portman, A. Natapov, and D. Fisher-Gewirtzman. “To go where no man has gone before: Virtual reality in architecture, landscape architecture and environmental planning”. In: *Computers, Environment and Urban Systems* 54 (2015), pp. 376–384.

[68] J Preece, H Sharp, and Y Rogers. *Interaction design: Beyond human-computer interaction.* New York: John Wiley and Sons, 2015.

[69] Carlo Ratti et al. “Tangible User Interfaces (TUIs): A Novel Paradigm for GIS”. In: *Transactions in GIS* 8.4 (2004), pp. 407–421.

[70] Horst W.J. Rittel and Melvin M. Webber. “Dilemmas in a General Theory of Planning Published by : Springer”. In: *Policy Sciences* 4.2 (1973), pp. 155–169.

[71] G Rizzolatti. “The mirror neuron system and its function in humans”. In: *Anatomical Embryology* 210 (2005), pp. 419–421.

[72] Mary Beth Rosson and John M Carroll. “Scenario-Based Design”. In: *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* Ed. by J Jacko and A Sears. Lawrence Erlbaum Associates, 2002. Chap. 53, pp. 1032–1050.

[73] Bahador Saket, Alex Endert, and John Stasko. “Beyond usability and performance: A review of user experience-focused evaluations in Visualization”. In: *ACM International Conference Proceeding Series* 24-October (2016), pp. 133–142.

[74] L Sanders. “Exploring co-creation on a large scale. In Designing for, with and from User Experience”. In: *StudioLab Press* (2010).

[75] D Schön. *The reflective practitioner: How professionals think in action.* New York: Basic Books, 1983.

[76] Orit Shaer and Eva Hornecker. “Tangible User Interfaces: Past, present, and future directions”. In: *Foundations and Trends in Human-Computer Interaction* 3.1-2 (2009), pp. 1–137.

[77] Michael Tomasello. “Cognitive linguistics”. In: *A companion to cognitive science.* Ed. by William Bechtel and George Graham. 4th ed. Oxford: Blackwell, 2006. Chap. 37, pp. 477–487.

[78] Brygg Ullmer and Hiroshi Ishii. “The metaDESK: Models and Prototypes for Tangible User Interfaces”. In: *Proceedings of UIST ’97.* ACM, 1997.

[79] Jos P. Van Leeuwen et al. “Effectiveness of virtual reality in participatory urban planning”. In: *ACM International Conference Proceeding Series* (2018), pp. 128–136.

[80] Colin Ware. *Information Visualization.* Elsevier Inc., 2012.

[81] Robert Wilson and Lucia Foglia. “Embodied Cognition”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N Zalta. Metaphysics Research Lab, Stanford University, 2017.

A

Usability test - Dam breakings in Slottsskogen

A.1 Scenario script

We are currently overlooking Linnéstaden *show the park and the streets*. It has been raining heavily for several days and people are worried that the dams in the park are going to break and flood the area. As a public investigator for the city's water protection, it is your task to prevent the negative effects evoked from floodings. Therefore, we are going to simulate a scenario in which you should try to minimize the negative effects from the flooding of a dam. To your assist, you will get to use two different tools:

- 3 wells, that will drain water to its best capacity.
- 3 walls, which will help you steer the water and keep it away from important roads and places

The city has put up a goal for how to prioritize in situations of extreme flooding scenarios at Linnéstaden, because this city district composes a crucial area for communications that are critical for the infrastructure of the whole city.

Dag Hammarskjöldsleden *point to the road* (large road with incoming traffic to Gothenburg from the south) may not flood. If so, all in- and outgoing traffic from the south will be heavily affected, with large negative outcomes as an effect. Linnégatan (large street) should be protected to the greatestqulg's largest hospital) and the central of Gothenburg, through at least one of the two larger streets in Linnéstaden: Linnégatan or Övre Husargatan, as for ambulances to pass.

A.2 Evaluation questions

- How well did the interface/interaction match your expectations?
    - Something that surprised you?
    - Something that was particularly hard/easy?
    - Did the objects behave the way you expected?
- Did you have a plan for limiting the spread of the water?

    - How did you decide on your actions?
        - What did you think?
- How much did you think about the scenario?
    - Did you experience stress?
    - What did you think of the water's behavior?
        - Its speed?
- Did you recognize Linnéstaden?
    - Did you consider the geographical location when interacting with the interface?
- What did you think of the bird's eye perspective?
    - Pro's and con's?
- Did you collaborate in the interface?
    - How/Why?

B

Evaluation Workshop - instructions for VR

B.1 Scenario script

In this interface, you will get to explore Linnéstaden from a first-hand perspective, and experience the effects from a real flooding. On the hand controller, there are two buttons that will get to use *show teleport- and click button*.

There is a platform in the VR environment from which a flooding scenario will simulated. You should stay on the platform as long as the streets are flooded. There will be information regarding your task on a screen by the platformm. Follow the instructions on the screen and click the buttons to initiate a new scenario.

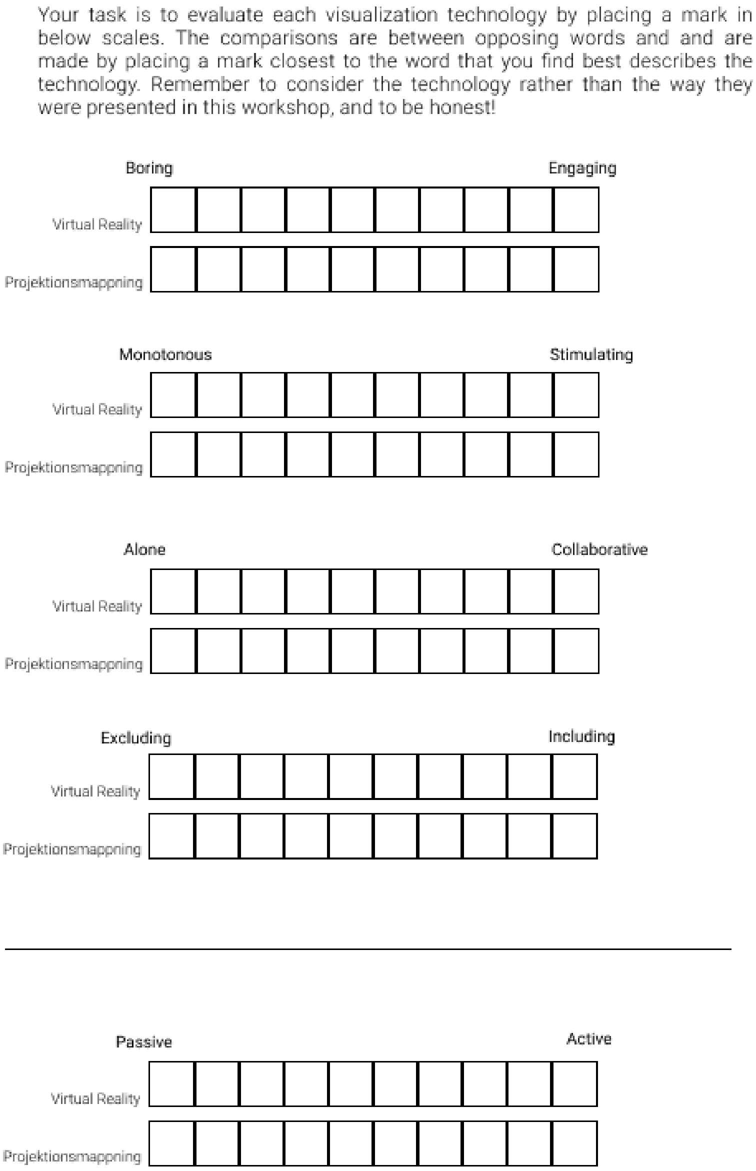
Objects that can be interacted with in the environment are glowing green - you can click these.

You mission will be to activate all objects around the city that can prevent a flooding of Linnéstaden. In total, there are four inactive objects placed at various locations, and they are activated by being pressed.

There is no time limit, so there is no need to feel under time pressure. You can ask if anything is unclear while in the virtual world, we will be right next to you.

C

Semantic Differential Scales



Your task is to evaluate each visualization technology by placing a mark in below scales. The comparisons are between opposing words and and are made by placing a mark closest to the word that you find best describes the technology. Remember to consider the technology rather than the way they were presented in this workshop, and to be honest!

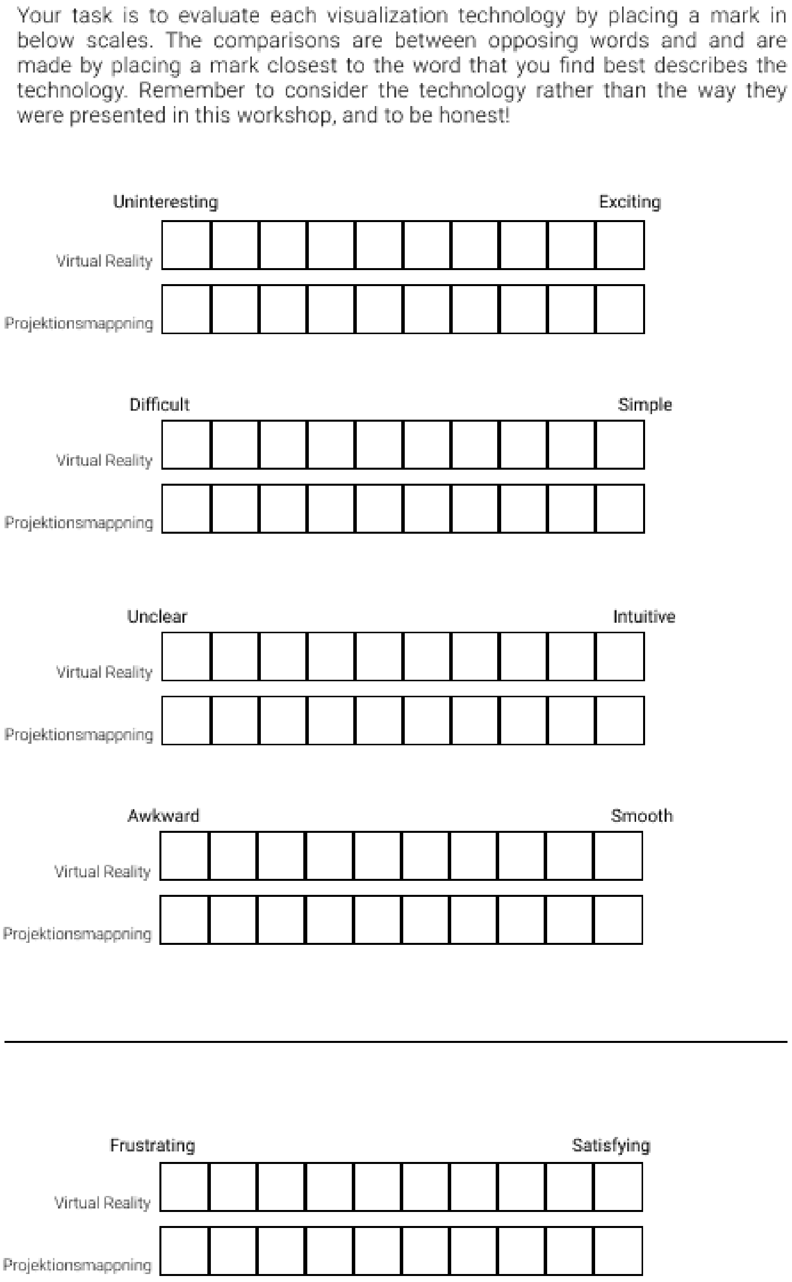
Boring ☐☐☐☐☐☐☐☐☐☐ Engaging
Virtual Reality ☐☐☐☐☐☐☐☐☐☐
Projektionsmappning ☐☐☐☐☐☐☐☐☐☐

Monotonous ☐☐☐☐☐☐☐☐☐☐ Stimulating
Virtual Reality ☐☐☐☐☐☐☐☐☐☐
Projektionsmappning ☐☐☐☐☐☐☐☐☐☐

Alone ☐☐☐☐☐☐☐☐☐☐ Collaborative
Virtual Reality ☐☐☐☐☐☐☐☐☐☐
Projektionsmappning ☐☐☐☐☐☐☐☐☐☐

Excluding ☐☐☐☐☐☐☐☐☐☐ Including
Virtual Reality ☐☐☐☐☐☐☐☐☐☐
Projektionsmappning ☐☐☐☐☐☐☐☐☐☐

Passive ☐☐☐☐☐☐☐☐☐☐ Active
Virtual Reality ☐☐☐☐☐☐☐☐☐☐
Projektionsmappning ☐☐☐☐☐☐☐☐☐☐



Your task is to evaluate each visualization technology by placing a mark in below scales. The comparisons are between opposing words and and are made by placing a mark closest to the word that you find best describes the technology. Remember to consider the technology rather than the way they were presented in this workshop, and to be honest!

Uninteresting **Exciting**

Virtual Reality ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Projektionsmappning ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Difficult **Simple**

Virtual Reality ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Projektionsmappning ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Unclear **Intuitive**

Virtual Reality ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Projektionsmappning ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Awkward **Smooth**

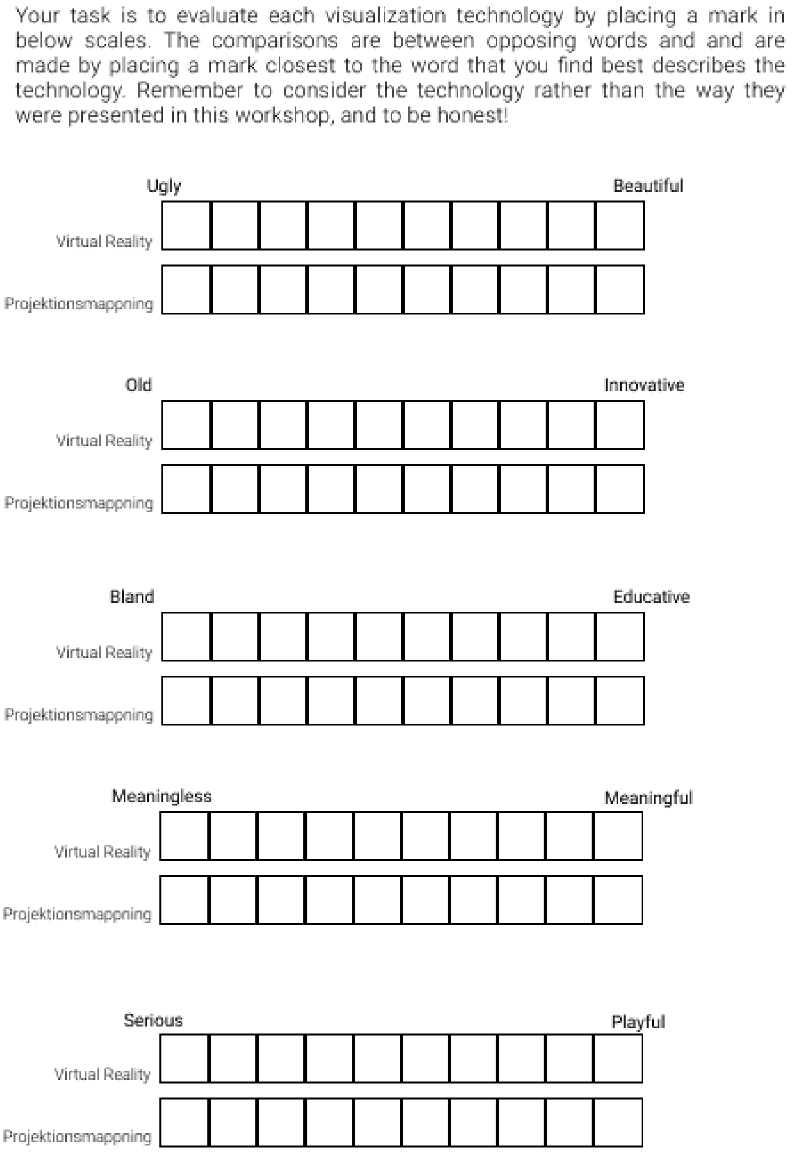
Virtual Reality ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Projektionsmappning ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Frustrating **Satisfying**

Virtual Reality ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Projektionsmappning ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐



Your task is to evaluate each visualization technology by placing a mark in below scales. The comparisons are between opposing words and and are made by placing a mark closest to the word that you find best describes the technology. Remember to consider the technology rather than the way they were presented in this workshop, and to be honest!

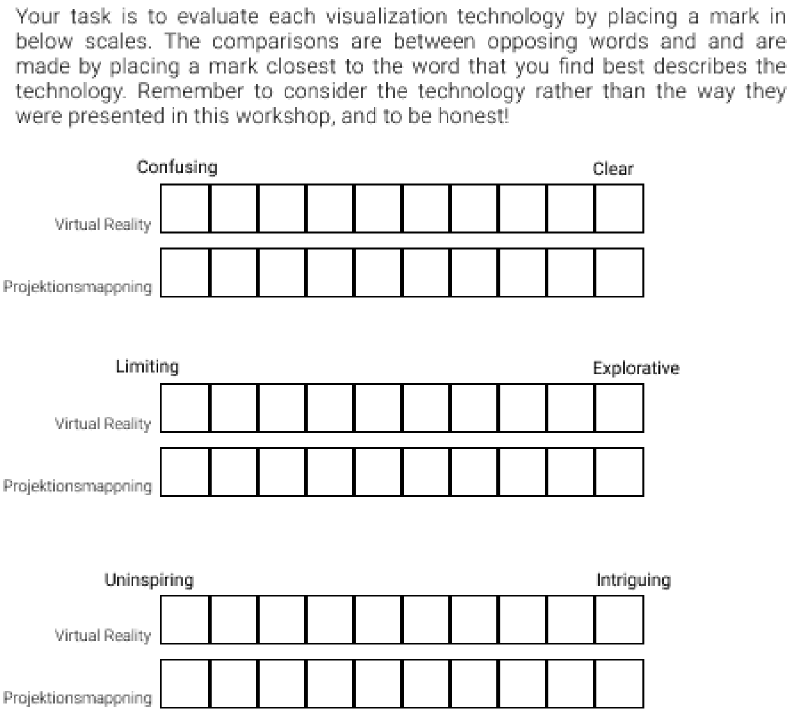
	Ugly									Beautiful
Virtual Reality	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Projektionsmappning	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Old									Innovative
Virtual Reality	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Projektionsmappning	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Bland									Educative
Virtual Reality	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Projektionsmappning	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Meaningless									Meaningful
Virtual Reality	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Projektionsmappning	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Serious									Playful
Virtual Reality	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Projektionsmappning	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐



Your task is to evaluate each visualization technology by placing a mark in below scales. The comparisons are between opposing words and and are made by placing a mark closest to the word that you find best describes the technology. Remember to consider the technology rather than the way they were presented in this workshop, and to be honest!

	Confusing									Clear
Virtual Reality	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Projektionsmappning	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Limiting									Explorative
Virtual Reality	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Projektionsmappning	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Uninspiring									Intriguing
Virtual Reality	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Projektionsmappning	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐