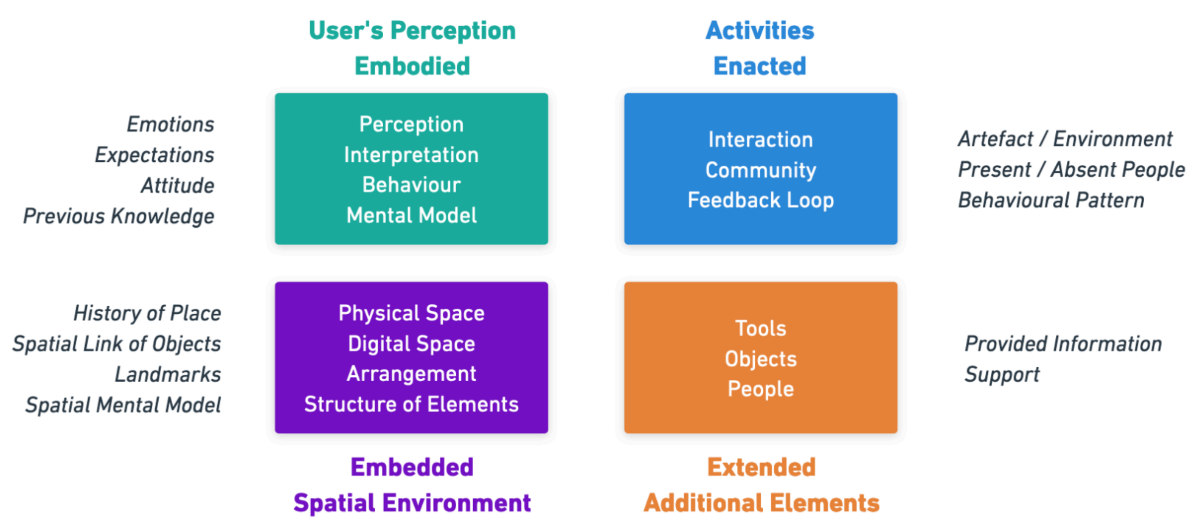
University of Applied Sciences and Arts Northwestern Switzerland
School of Engineering

Spatial Computing Evaluation Framework

A Holistic Approach to evaluate the User Experience in Mixed/ Augmented Reality



Windisch, August 2024

Master's Thesis - MSc in Computer Science

Katja Pott
Student

Prof. Dr. Doris Agotai
Advisor

Abstract

Spatial computing is an emerging technology which combines the physical and digital space and enables interaction with spatial meaning. The extension of reality through virtual worlds creates new opportunities, but also uncertainties regarding UX and UX evaluation. This is reflected in the current state of the art, where human factors in spatial computing tend to be overlooked and the focus is often on technical considerations. Consequently, there is a lack of knowledge in HCI about which aspects should be considered in the UX and evaluation of spatial computing applications. However, contextual factors such as the emotions of a place, the surrounding community, and the activities in the augmented space are crucial as the user experience in spatial computing is fundamentally shaped by context.

To support researchers and experts in HCI, the Spatial Computing Evaluation Framework has been developed. It provides support and guidance for the UX evaluation of spatial computing applications, without limiting exploration. The framework aims to evaluate the situation holistically, highlighting the connection between the physical, digital and social context.

The framework was developed through an extensive literature review followed by a user study using the Apple Vision Pro. In order to emphasise the interconnectedness of the physical, digital and social environments in spatial computing, the theory of 4E cognition was used to build the framework. The specific guidance the framework provides is a novel contribution and can be particularly supportive in the early stages of the field to rapidly build up knowledge and create awareness.

Contents

List of Figures

List of Tables

Acknowledgements

Special thanks go to

- Prof. Dr. Carmen Zahn for the insightful discussions about 4E Cognition
- Flavia Brogle for her wisdom in developing for the Apple Vision Pro
- Yves Simmen for his valuable feedback to my publications and support in creating the application for the Apple Vision Pro
- Prof. Dr. Anton Fedosov for his professional advices in scientific methodologies
- Aileen Bamford and Ladina Miller for their ongoing support and reflection
- Mia Braunwalder for her ad-hoc expertise in visualisations
- Stefan Graser and Reto Senn for their feedback and insightful discussions about the Evaluation Framework

and especially to

Prof. Dr. Doris Agotai for her continuos support and insightful discussions for this master's thesis.

1 Introduction

Spatial computing extends reality through virtual worlds, changing the situation fundamentally by introducing digital content. This changes the context and how individuals interpret the situation [10]. Although the user experience is fundamentally shaped by context, it is often overlooked in spatial computing and only the technical aspects of space are typically considered [33]. This approach fails to recognise the significant impact of context on the user experience in spatial computing, and consequently its enormous potential [34].

There is a shortage in human-computer interaction to understand interaction & context in spatial computing and its implications [34], [35]. The field of human factors is underdeveloped and overlooked, with research focusing solely on the technical components of spatial computing [34]. However, the socio-technical aspect is crucial, as reality is being augmented by virtual worlds that are experienced through interaction and are situated in the physical and social environment [33], [36].

The extension of reality through virtual worlds presents both opportunities and uncertainties in terms of user experience and UX evaluation. Figure 1 shows the uncertainty of UX evaluation on the RV continuum [1]. At either end, UX evaluation methods are well established, as the technologies have longer development histories and large investments by the gaming or military industries. However, in the middle, where spatial computing is located, the uncertainty is high.

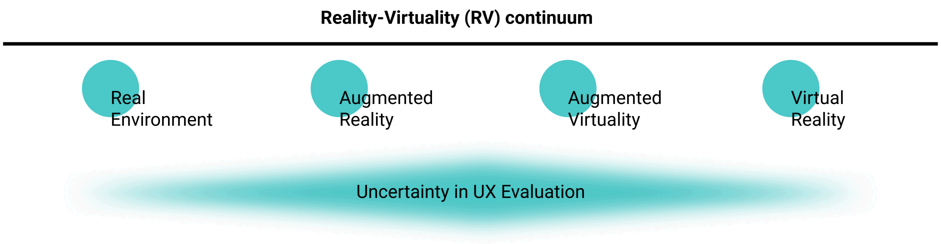


Figure 1: Reality - Virtuality Continuum visualising the amount of uncertainty in UX Evaluation [1]

In the context of Spatial Computing, new aspects are relevant that have yet to be established in HCI, and there is currently no coherent methodology for evaluation [37], [35], [38], [39], [40]. To address this knowledge and methodological gap, the *Spatial Computing Evaluation Framework* has been constructed to support the UX evaluation of spatial computing applications. Its specific guidance is a novel contribution and can be particularly helpful in the early stages of the field.

This thesis is a first step in addressing the gap and adding to the body of knowledge in this still small area of research. Hopefully, it will stimulate relevant discussions for further improvement and support of designers & developers in the creation of outstanding user experiences with spatial computing applications.

2 Motivation

Spatial Computing offers new user experiences that are different from traditional mobile device interactions. However, these experiences also create uncertainties about how to evaluate the user experience. Many previously irrelevant aspects need to be considered, highlighting the need to enrich the current state of the art in UX and UX evaluation. Currently, it is unclear which elements should be analysed to holistically evaluate the user experience in spatial computing and a more comprehensive analysis is required as it is embedded in the environment and includes spatial aspects. In addition, the context and situation in which the experience takes place are fundamental factors that shape the experience & perception and need to be carefully considered. Furthermore, this emphasises that spatial computing is susceptible to various external influences that cannot be controlled but must still be taken into account during development.

In current evaluations, only the most obvious aspects are often analysed, using very broad instruments such as the System Usability Scale (SUS) or the NASA-TLX for the measurement of cognitive load. However, it is not clear whether these instruments are suitable for spatial computing and they also fail to take into account the many aspects that this new technology brings with it. There is a lack of attention to a number of issues which are relevant to spatial computing, and this is probably mostly unconscious. The influence of many of these is unknown and difficult to assess, which poses further challenges for evaluation.

3 Methods

This section provides an overview of the methods used in this thesis and how the final result is derived, based on the literature review and analysis conducted, as well as the modelling approach and evaluation technique.

3.1 Approach

This paper implements an open-ended interpretive methodology to analyse literature, selecting specific examples that may have been published with an alternative focus. The literature is analysed and interpreted in the context of spatial computing.

The term *situatedness* is used to categorise the different findings within the literature. This term allows for a holistic perspective and includes many topics relevant to spatial computing. Most literature focuses solely on spatial aspects and digital content related to it. However, an HCI perspective acknowledges that other factors in the situation significantly influence the user experience of spatial computing applications. Therefore, our analysis of situatedness emphasises the importance of considering a holistic view of the situation in which the user is located.

3.2 Literature Review

Several categorizations of spatial computing applications are existing. In this project the categorization of mobile AR, Head-Mounted Displays and Spatial Augmented Reality is used as the user experience of the three is different. For more technical projects the categorization of Visual See Through and Optical See Through may be used.

In each of these areas, an appropriate example application is selected to investigate the application's characteristics in relation to spatial computing. The characteristics are identified using literature from the fields of augmented reality, information visualisation, human-computer interaction, and/or spatial computing. The example applications are subject of investigation and will be analysed in detail to identify potential aspects relevant to the user experience. Further aspects and literature could be identified by snowballing. To ensure a holistic evaluation of the scene, the characteristics are categorised based on the definition of situatedness. The characteristics of an example are then evaluated for relevance and further elaborated. For each identified characteristic, potential evaluation options are researched and evaluated. Finally, the characteristics are presented in a schema for each scenario in order to find potential similarities between all three scenarios.

3.3 Selection of Examples for Investigation

Each application should allow to investigate points for consideration in the user experience. Other points that are relevant for the selection are the design, the area of use, the place of development and the target user. The aim is to have a diverse set of applications with different areas of use. Another important point is that the applications should have been developed in Switzerland in order to promote and recognise technological innovation in the region.

3.4 Modelling

An overall model to illustrate the influence of context in spatial computing seems appropriate as some similarities of the three example applications could be identified. For this purpose, some existing models found in the researched literature are analysed based on the term situatedness. Furthermore, the schemas created for the three examples are analysed for elements to condense the information into a model in combination with 4E cognition theory. To evaluate the model, a user study will be conducted based on a proof of concept implementation.

3.5 User Study

To evaluate the constructed model, a user study is conducted. The aim is to understand whether users interact with spatial computing in the way that is assumed and visualised in the theoretical model. To narrow down the potential findings, the research question is posed to understand which elements influence the orientation of a user in a step-by-step guidance process with a 3D model and natural language instructions. The user study is conducted with the Apple Vision Pro where a part of the Arigo application is reproduced considering the findings in the literature review. The study is conducted with 9 participants representing three different cohorts.

The study is conducted using the think-aloud method and statements are transcribed and coded with MAXQDA with inductive and deductive coding.

3.6 Evaluation Framework

Based on the constructed model and the results of the user study, an evaluation framework is constructed to evaluate the user experience in spatial computing. The model is divided into the four different dimensions from 4E cognition. This is to combine the theory of 4E cognition with spatial computing for a holistic view on the situation. Each dimension of 4E cognition contains the relevant points found in the literature review and are further explained as questions for reflection. The questions are intended to help designers and developers to interpret the identified points in order to derive relevant insights.

The evaluation framework is completed for one case of the conducted user study to evaluate the usability. Furthermore, in order to evaluate the relevance and practicability, a short usability test is carried out with a mobile AR application and reviews with experts, in which one conducts a real usability test with the framework.

4 Related Work

Spatial Computing is an emerging term in both academia and industry. The term is further explained in this section as well as the term *Situatedness*. The origin and use of the term situatedness will also be discussed. Finally, for a better understanding of the approach used in this thesis, the field of Human-Computer Interaction is briefly introduced.

4.1 Spatial Computing

Spatial computing brings the physical and digital closer together, bridging the two and allowing users to interact in the augmented world [41]. Devices are aware of their spatial context and can model their environment [42]. Figure 2 shows an example of a spatial computing application using a MetaQuest3 [2]. It is a music application where the different albums are shown as digital 3D tapes. Based on the collection of tapes the user can choose one and place it in the recorder to listen to it.

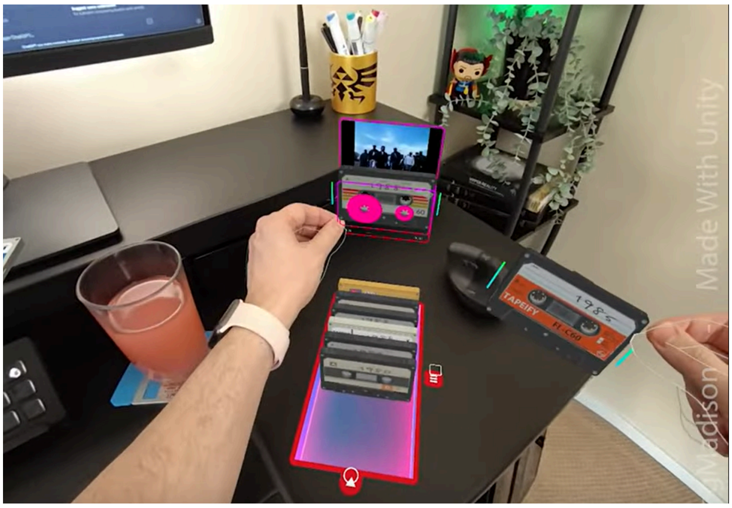


Figure 2: Example of a spatial computing application where a user can switch music by choosing digital tapes and place them in the recorder [2]

Spatial Computing offers new user experiences that are different from traditional interactions with mobile devices. The context and situation in which the interaction takes place are fundamental factors that shape the experience and perception. Recent advances in hardware such as computer vision, sensors and devices make this technology possible [41]. With these capabilities, digital content can be integrated into the real world, opening up new opportunities for many sectors, such as productivity gains in manufacturing or safety in healthcare [41]. The devices are aware of the environment in which they are placed and enable human actions with spatial meaning [42]. Users can interact with the physical and digital world that is aligned with the real spatial world [42].

Spatial computing encompasses a wide range of different technologies. These can be divided into the categories of (1) Mobile Augmented Reality, use of mobile devices to augment the physical space with digital content that can be viewed through the devices. (2) Head-Mounted Displays, headsets which give the possibility to interact with the physical space and the digital content embedded within. (3) Spatial Augmented Reality, projection of digital content onto physical objects, taking into account their shape and allowing to interact with the projection.

4.2 Situatedness

Individuals move within their environment and engage in activities, yet they do not do so independently. Each action is intertwined with a situation, encompassing numerous components and social interactions. All of these components can be viewed as context for a given task. Thus, we recognise context not solely from a viewpoint of spatial organisation and its corresponding objects, but also in consideration of the social dynamics present within the space and within human activities. Context is situational, thus it is essential to consider what is relevant in each circumstance. This is referred to as the *practical objectivity of situations* [43]. Humans create the situation and make sense of their surroundings through their actions and social interactions [44].

The concept of *situatedness* is defined as the notion that the human mind is influenced by its natural, social and cultural environment [45]. According to this theory, the actions of individuals should be interpreted in relation to their specific environmental and contextual factors [45]. The origin of *situatedness* can be traced back to the field of psychology, where its first impact can be observed in 1890 [45]. Many psychologists of the 19th and 20th centuries, including Kurt Lewin and proponents of Gestalt theory, contributed to the definition of situatedness [45]. The concept of situatedness describes the close coupling of the agent and environment, and is now used in various domains including philosophy, education, robotics, and computer science [45].

Bressa et. al use the theory of situatedness for a holistic perspective on augmented reality applications visualising data and considering its context, namely situated visualisations [44]. The study analyses different examples using the five perspectives of situatedness: space, time, place, activity, and community.

Space - Space is the most prominent aspect in research for augmented reality. Space encompasses the physical environment and its digital content. It refers to the spatial representation of information, including its properties and characteristics in three-dimensional space.

Place - Place is concerned with location, beyond a purely spatial perspective. It is the meaning of a space in relation to the people who inhabit it, the activities that take place within it and its historical context. The cultural or social meaning of a place is what forms meaningful relationships between individuals and their environment.

Activity - Activities should not be viewed in isolation. They are components that contribute to broader goals and can be carried out in different spaces, over different periods of time and in collaboration with others. The relationship between a person and the goals he or she wishes to achieve is complex and influenced by the social and cultural context, which is defined as *activity theory* [46]. Individuals do not act *with* technology, but rather *through* it [47]. Understanding the context of activities people perform is essential as this determines which visualisations are appropriate.

Community - Community refers to individuals in relation to a particular place. The focus is on how a specific community utilizes technology or a physical location to achieve their objectives. It is closely tied to the notion of place, which derives its meaning over time from the activities of the community.

Time - The concept of time concerns the recording of data for a visualization and its display, as well as the relationship between these two moments. Additionally, time can be considered beyond pure objectivity or subjectivity and instead be viewed as socially constructed, defined as *Social Time* [48]. Social time can be found in the agreement of the 24-hour day or time zones.

Further, it also encompasses temporal relevance which considers social norms, activities and conventions regarding to time.

4.3 Human-Computer Interaction

Human-Computer Interaction (HCI) is a broad field with research coming from different disciplines. It combines knowledge and tools not only from computer science, but also from social sciences and natural sciences to understand and reason about the interaction between humans and technology [3]. This multidisciplinary approach allows it to analyse a situation from different perspectives, resulting in a holistic view.

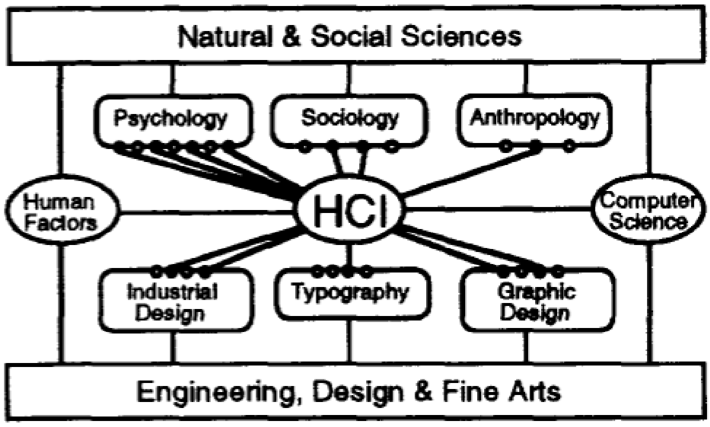


Figure 3: *Human-computer interaction, like human factors and computer science, is a multidisciplinary field that draws from both scientific and design disciplines* [3]

Figure 3 shows the different sciences that influence the field of HCI [3]. The different influences also provide the opportunity to use the tools from their field of research and to apply them to the technology. As useful as this multidisciplinary approach is, it also presents some obstacles. In the literature of the different fields, the naming of the tools used can lead to confusion without specific knowledge. An example of this is when conducting a user study where users are asked to verbalise their thoughts in order to gain deeper insights into their thinking process. In psychology this approach is called *protocol analysis*, whereas in computer science it is called *thinking aloud*. Even though it is the same technique, the insights from the analysis can be different depending on the knowledge of the people analysing the results, whether they have a background in psychology or computer science. This thesis takes advantage of this multidisciplinary field of HCI and shows that not only tools from computer science can be used to evaluate technology. It shows that in order to analyse a technology, several tools from the social or natural sciences can be useful.

5 Subject 1: Head-Mounted Displays

In this section, the example application for head-mounted displays and its relevant characteristics are introduced on the basis of the term situatedness. The characteristics are discussed in terms of their theory and relevance to the example. Potential evaluation techniques are presented for each aspect, followed by an overall analysis of the scenario, its characteristics and potential evaluation techniques.

5.1 Example Application - Arigo

Head-mounted displays (HMD) are suitable for scenarios where both hands are required to perform actions. In this way, information can be extended into the physical space without restricting the dexterity of the user. HMDs can be useful for guiding users through sensitive processes, providing step-by-step guidance while remaining aware of the environment [49]. Such an application represents Arigo, a non-overlayed AR guidance using a 3D model in radiopharmacy (Figure 4) [4].

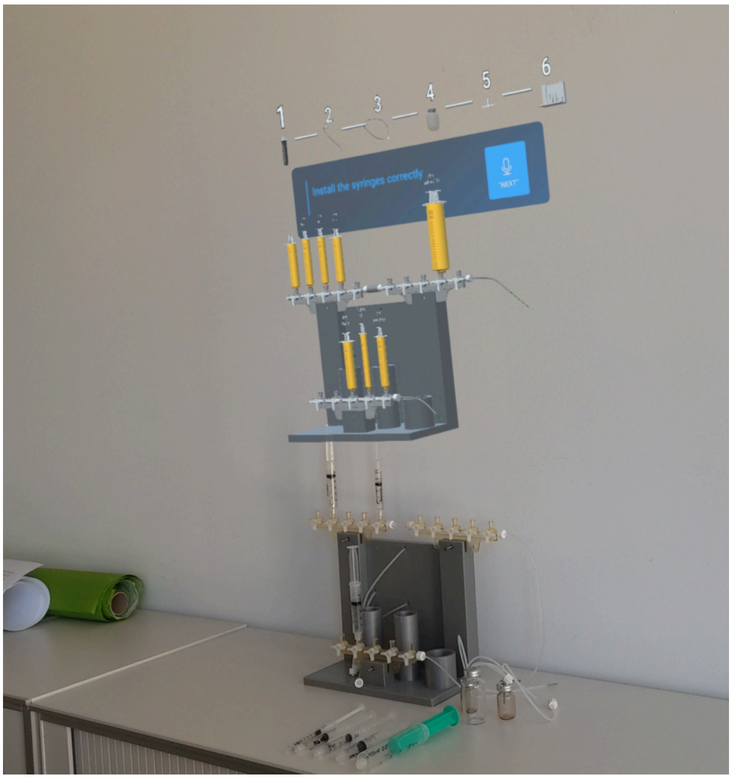


Figure 4: Non-Overlayed AR User Guidance with 3D-Model [4]

Radiopharmaceutical laboratory employees are required to produce mixtures of different chemicals. This process is complex and time-sensitive, and is particularly prone to errors. Typically, paper-based instructions are utilised to guarantee the correct procedure for creating radio pharmaceuticals. However, Simmen et. al's research project aims to substitute paper instructions with Augmented Reality (AR) technology. This involves using a non-overlayed 3D model via a Microsoft Hololens to support the procedural task of mixing chemicals.

In order to produce a mixture, the workers have to find the right items in the laboratory and prepare them for the mixing process. Moreover, they must ensure the precise amount of various fluids is prepared and loaded into the syringes.

Once the components have been prepared, they must be placed correctly onto the machine. It is essential that the fluids are mixed in a specific order to ensure correct mixing. This task requires positioning the syringes correctly on the machine and opening the appropriate vials to mix the fluids accordingly. The augmented reality model offers guidance for every step regarding the placement of each syringe in its designated position, which vial to open, and at what time.

For the scenario presented, several points are relevant to consider for the user experience. These points are categorised based on the notion of situatedness. Figure 5 gives an overview of the characteristics that are relevant and have been selected for the analysis of the given scenario.

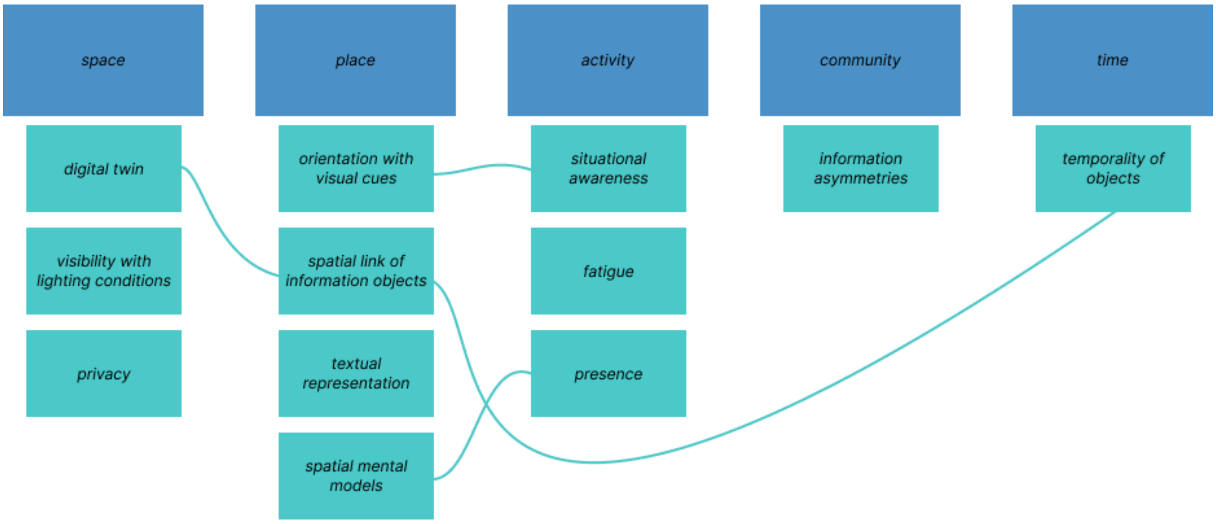


Figure 5: Overview of relevant and chosen characteristics for the scenario of Arigo, categorized by the term situatedness

5.2 Space

This subsection introduces the relevant characteristics to the topic of space. Space encompasses the physical environment and its digital content. It refers to the spatial representation of information, including its properties and characteristics in three-dimensional space.

5.2.1 Digital Twin

A digital twin is a contextualised digital representation of a physical object. It allows the behaviour and appearance of an object to be replicated digitally. The number of properties of the real object to be included in the digital representation can be determined by the developer of the model [50]. This also defines which aspects are important in the digital representation and the context of the situation, and how close a digital twin is to its real counterpart and behaviour [50]. The complexity of a digital twin in its development is also defined by the selected properties [50].

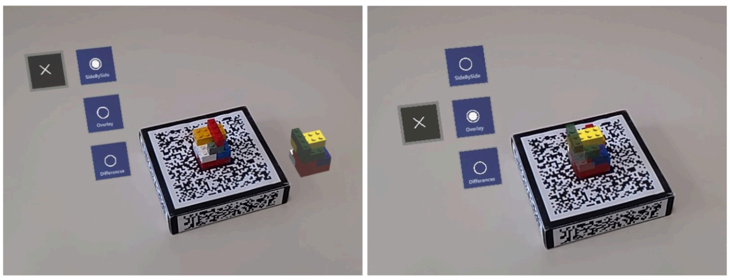


Figure 6: digital twin visualisation: side-by-side or overlayed [5]

To visualise a digital twin in AR several possibilities are possible. Letter et. al present different visualisation methods for version control with tangible objects [5]. The digital representation can be shown next to the physical object, following the principle of juxtaposition, presenting the *side-by-side* mode (Figure 6, left). They place the digital twin next to the physical object to allow comparison of the two objects. Further, they mimic the object's rotation of the real world in the digital twin to ensure correct comparison from the different angles by moving and rotating the physical object. The second implementation follows the principle of superposition, overlaying the physical object with the digital twin (Figure 6, right). In this, also the movement and rotation of the physical object is represented in the digital twin to ensure correct overlay. The overlay in Letter et. al's work uses a semi-transparent overlay visualisation.

Tönnis et al.distinguish the spatial dimension of a model, which can be visualised through a 2D or 3D digital twin. A practical application of this is implemented in the case of Arigo, where the machine's 2D and 3D visualisations are compared (Figure 7).

Figure 7: digital twin visualisation: 2D or 3D representation [4]

It is also an open question as to how realistic a digital twin has to be in order to perform certain actions. Studies in virtual reality show that visual realism of the environment does not always yield the best results as for example in spatial learning [51], [52]. It shows that only the elements that the user perceives as relevant need to be rendered realistically, while the rest of the scene is abstract [51]. This mix of renderings gives the best results for spatial learning of a path in a virtual environment [51]. For digital twins, for example, the results show that the size of a digital twin overlay can vary by up to 5% before users notice a difference in size [53]. Therefore, variance in visual realism could also have an impact on the perception of a digital twin.

From a developer's perspective, the choice of visualisation for a digital twin falls under the category of *space*. However, the placement of the digital twin into the physical world and its interpretation by the user belongs to the category of *place*.

The digital twin in the example of Arigo has a high visual realism compared to its real counterpart. They also follow the principle of juxtaposition, presenting the digital twin spatially above the real machine. As they compared the effectiveness of a 2D and 3D digital twin, both dimensions were developed. However, if the high visual realism is fully necessary is an open question.

5.2.2 Visibility with lighting conditions

Spaces have different lighting conditions which affect the perception of digital content in AR, especially in optical see-through (OST) AR. In OST, virtual content is superimposed on the physical space, making the digital overlay appear transparent. The light of OST is additive to the natural lighting conditions of the environment [53]. The perception of digital content in OST AR is therefore strongly influenced by the natural lighting conditions of a space [53].

Kahl et. al. conducted a study to observe the perception of OST AR in three different lighting conditions: dark, medium and bright [53]. In brighter light, the digital overlay appears more transparent, which also allows for greater variation in the size of the digital content compared to its physical counterpart, without interfering with the identification of the object [53]. However, in brighter conditions, both the presence and the performance of the interaction decreases [53]. They assume that darker lighting conditions were preferred by the user because the digital content appeared more real and easier to perceive due to the intense illustration in the dark environment and the low visual distraction from the physical environment and one's own hands [53].

Digital objects in AR do not automatically have the *consistency* that we perceive from objects in the real world. Visualisation with respect to the lighting and background conditions of the real world are important [54]. Other studies have also shown that light has a strong influence on the readability of text in augmented reality [55].

As the application of Arigo is used in a laboratory, it can be assumed that the lighting is rather strong and bright, which could decrease presence and performance of the user. It is also unclear how the lighting conditions influence the readability of the text annotations used in the application, as the step indicator in the 3D-model is presented with white colour.

5.2.3 Privacy

HMDs are often used in public or semi-public environments, including factories, manufacturing plants and maintenance facilities. Although the digital content is only visible to the person wearing the HMD, the interaction with this content is visible through the movement of the body, particularly the hands. This is a problem in situations where sensitive information, such as numeric or text data, is entered through a digital medium. [56]. The act of observing a user while they input private information is commonly known as *shoulder-surfing* [56]. In HMD applications, user authentication is a common scenario. Users wearing HMDs are particularly at risk of shoulder-surfing as their peripheral awareness is reduced while wearing the device [56]. Several studies indicate that authentication methods that are resistant to shoulder-surfing include selecting an object [56], selecting an object through eye-gaze [57], or answering semantic security questions [58], each after scanning a personalised QR code.

The current implementation of Arigo is not susceptible to shoulder surfing. This is because no authentication is required and no personal information is stored. However, in future scenarios, the recording of the mixing of the substance and the provision of personalised instructions could be implemented. To provide personalisation and protect users from misuse, authentication would be required.

5.3 Place

This subsection introduces the relevant characteristics to the topic of place. Place is concerned with location, beyond a purely spatial perspective. It is the meaning of a space in relation to the people who inhabit it, the activities that take place within it and its historical context.

5.3.1 Orientation with Visual Cues

As the field of view is limited by the HMD, the user needs support in orientation and navigation in the scene [59], [54]. The users attention can be guided but distraction and confusion should be avoided and the attention should only be brought to important objects [6]. However, the awareness of the surrounding should not be diminished as the creation of a tunnel vision should be avoided [54]. Useful information on how to guide a user's attention with visual cues can be found in 360° videos as in [6]. These cues can be diegetic, part of the scene, or non-diegetic, external to the scene in which the user is immersed. Further, the cues can be explicit, communicated directly to the user, or implicit, indirectly guide the attention, and limiting vs.

non-limiting the interaction of the user. Using these three distinction, a design space with three axes can be formed to define the visual cues as in Figure 8.

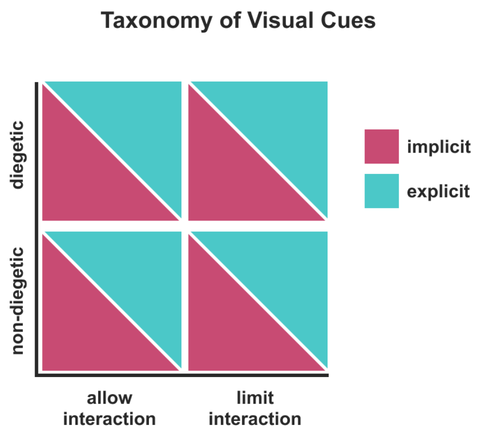


Figure 8: Taxonomy of visual cues given by Speicher et. al [6]

As the digital content is visualised separately from the main immersion scene, Arigo's visual cues are non-diegetic. If the guidance would be overlaid directly on the physical machine, it would be diegetic. The cues are explicit because they communicate directly to the user, either through written instructions displayed at the top, or through the syringe visualisations of the digital twin. If instructions were presented using changing colour luminance or similar elements to indicate where to place a syringe, the cues would be implicit. Furthermore, the engagement of the user with the digital or physical world is not restricted by the cues.

5.3.2 Spatial Mental Models

In order to navigate our environment, we generate cognitive maps of our surroundings [7]. These maps are personalized and based on our subjective perception of space. There are two types of maps: allocentric, which use a reference frame that is external to the user, and egocentric, which use a reference frame that is relative to the user [60]. The distinction between the two reference frames can be illustrated by drawing a hiking map and then using it for navigation. The initial stage of map creation is an allocentric task [60]. The user draws a map that is object-centred or world-centred Figure 9. After starting a walk, the user has to transform the allocentric map into reality and identify his initial position on the map for orientation. The user then has to construct a representation of the route based on his current position, which is a body-centred representation. The cognitive map, which is allocentric in nature, has to be transformed into an egocentric model that allows the user to execute it, which requires cognitive effort [60]. Cognitive maps comprise not just physical reference points, but also digital content in AR can facilitate their development [61]. However, spatial cues in the digital environment need to be recognised and understood [61].

Figure 9: Cognitive Map with allocentric view [7]

Some argue that the term cognitive maps is inadequate for the complexity of knowledge people have about their environment [62]. The metric relation of objects or landmarks are often distorted in our knowledge [62]. Therefore, the term *map* is unsuitable and *spatial mental model* may be used instead to account for this distortion [62]. In addition, spatial mental models use landmarks or objects to establish spatial relationships between each other for orientation in situations where perspective is not directly stored in memory [62]. This facilitates perspective taking and navigation from varying starting points [62]. Spatial references from objects are crucial for providing natural language guidance [63]. When describing procedures, we commonly use references to other objects, such as “to the left” or “on top,” [62]. In addition, it is important to note that spatial mental models are different from cognitive maps in that they take into account the unique experiences of an individual in a particular environment [62].

In the example of Arigo, employees have extensive knowledge of the laboratory environment, incorporating different kind of knowledge for orientation and wayfinding. Consequently, they form spatial mental models that incorporate spatial relationships between objects and landmarks, which they memorise. These mental models allow them to easily create allocentric or egocentric cognitive maps. They also form spatial mental models of the machines they use on a daily basis. This makes it possible to give instructions on where to place the syringes and which valves to open in relation to the machine or another object on the machine, for example: “Place the syringe in the second opening on top and set the valve to a central position.”.

5.3.3 Spatial Link of Information Objects

There are different possibilities of spatial connection of information objects in AR [54]. They can be *direct physical information objects*, this category describes physical information objects placed in the real world. Information objects from the real world which the user perceives in a mediated way through AR are *indirect physical information objects.* Information objects with a spatial relation to the real-world but presented only digitally are *spatial virtual information objects* such as in the case of digital twins. Digital information objects which are referenced to a point or object are *spatially referenced information objects* such as labels. Lastly, digital information objects which have no spatial relation and are of general nature are *detached virtual information objects.*

Figure 10 shows the different categories based on the example application of Arigo. The physical machine is a *direct physical information object*, placed in the real-world. The digital twin, visible through AR, is a *spatial virtual information object.* The labels describing the content of the syringes in the digital twin are *spatially referenced information objects.* Lastly, the step indicator on top of the digital twin is a *detached virtual information object.*

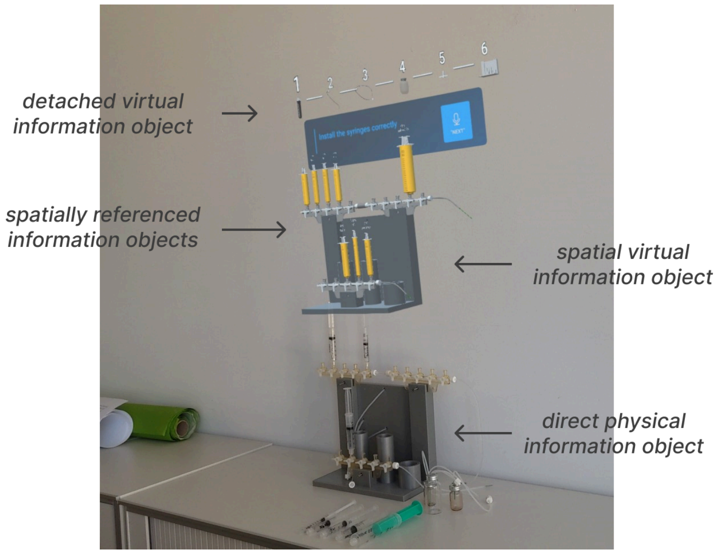


Figure 10: Information objects in the example of Arigo [4]

5.3.4 Textual representation

The *spatially referenced information objects* and the *spatially referenced virtual information objects* are usually displayed as text. There are a number of ways in which these information objects can be integrated into AR, depending on their characteristics [64]. The *complexity* of the physical environment has to be taken into account for readability, as well as the *movement* of the user or the environment and the *temporal properties* of the annotation [64]. Furthermore, the *semantic relevance* of an annotation to an object and the *complexity* of the annotations need to be analysed [64]. Finally, the *interaction* with an annotation must be considered and how this interaction could be facilitated, taking into account the cognitive load of the interaction as well as the overall understanding and interpretation of the annotation [64].

In the example of Arigo, the complexity of the physical space is rather low, given the context of the laboratory. Since all the necessary objects are prepared in advance and placed next to the machine, the user's movement is also limited. As the procedural task of mixing radio pharmaceuticals is guided step by step, the written instructions for each step disappear once the step has been performed and confirmed by the user. As the users are experts and use the scientific terms for the objects, the annotations are displayed on the digital object with the scientific term. In the case of the example, it is not possible to interact with the annotations as they are of a purely informative nature.

5.4 Activity

This subsection introduces the relevant characteristics to the topic of activity. Activities should not be viewed in isolation. They are components that contribute to broader goals and can be carried out in different spaces, over different periods of time and in collaboration with others. The relationship between a person and the goals he or she wishes to achieve is complex and influenced by the social and cultural context.

5.4.1 Presence

Augmented Reality (AR) offers an experience that is completely different from two-dimensional scenarios, placing digital content in the physical space surrounding of the user and relating it to the real world. The psychological state of being consciously engaged in the augmented world is known as *presence* [65]. This phenomenon differs from immersion, which refers to the extent of the capacity of technical device to provide an illusion of reality [65]. The formation of presence, or being present in a scene, stems from the perception of stimuli from the environment leading to presence through cognitive processing those stimuli [65]. The mind creates a mental representation of the surrounding environment through the stimuli, meaning that what the user experiences as reality, or mixed reality in the case of AR, is an outcome of internal processes and not the direct input received via the senses. [65].

Presence is also influenced by the environmental affordances, as proposed by Schubert et. al [65]. Through the established mental model of the environment, the user evaluates the actions that can be performed in the given space. The environment is conceptualized by the user in terms of actions. The perception of the possible actions can be divided into two categories: projectable and non-projectable properties of action. Projectable properties define the actions that can be taken within a space, as perceived from the environment. Non-projectable properties consist of patterns of actions that are based on the user's implicit memory. The mental model of the perceived actions by the environment must be translated into bodily actions that involve the manipulation of objects, orientation, and interaction. The existence of presence is demonstrated by this conversion of cognitive perception into bodily action. However, this conversion must be learned. Users must perceive stimuli from both the physical and digital environments to establish a mental model, suppress unnecessary information, and create a spatial-functional model. These processes require cognitive resources which easily collapse under pressure and require awareness of the construction for it. Witmer and Singer define presence as the combination of *involvement* and *immersion*: *"Involvement is a psychological state experienced as a consequence of focusing one' s energy and attention on a coherent set of stimuli or meaningfully related activities and events. [Psychological] immersion is a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences."* [66].

In the example of Arigo's step-by-step instructions, the user has to understand the instructions given by the digital twin, which is presented through AR. This requires a lot of cognitive resources for orientation and interaction, as the perceived instructions have to be translated into bodily actions in the right place and in the right order. However, despite the high cognitive effort required, the successful execution of the whole procedure shows that a sense of presence could have developed if a successful translation from perceiving to bodily actions had occurred. The stimuli provided by the digital content must also be processed through cognitive processes into a spatial mental model, which is necessary for the development of presence.

5.4.2 Situational Awareness

Taking into account the context in which the user is located and the situation in which the activity is taking place is crucial in order to understand the need for situational awareness. Applications such as Pokemon Go have shown that users are easily distracted by AR, leading to many accidents. Especially in dangerous environments, situational awareness of the surroundings is important to prevent accidents. AR can be a source of distraction, as shown in a study of drivers wearing an HMD [49]. Due to reduced attention caused by the use of AR, important signs and other elements of the road were missed [49]. This is a phenomenon often seen in HMD applications, called information tunnelling, attentional fixation or cognitive tunnelling [67]. In this case, attention is drawn to one channel of information, mostly derived from the digital overlay, for longer than is ideal, neglecting another channel of information [67]. Not only does the information tunnel reduce attention in peripheral vision, but it is also shown that objects in foveal vision are missed as users focus too intensely on other elements presented in the HMD [67]. Cluttered displays exacerbate the problem of inattentional blindness as well as when objects are not presented within the scene of the environment [67]. Objects should be anchored in the real world, blending digital and physical information, to avoid a dominant perception of digital content [68]. Furthermore, visual cues can draw attention to a particular object, leading to the neglect of other objects in the scene [67].

However, there are also studies where situational awareness could be enhanced by the provision of information about the environment that the user would not naturally be able to see [49].

In the example scenario given by Arigo, the need for situational awareness is high, as the users are in a laboratory where radio pharmaceuticals are being produced. However, the application does not have any features that would increase the level of situational awareness. The information presented by digital means is minimalist, which could also have the effect of increasing situational awareness, as it could be perceived as not being distracting. Nevertheless, the digital information does not blend in with the real environment, as it is presented on a digital twin that is displayed on top of the physical object. By focusing too much on the digital overlay of the HMD, this could lead to an attention tunnel.

5.4.3 Fatigue

When using AR, the user's sensory and cognitive effort is higher due to the additional information provided. This can lead to increased cognitive load and fatigue [54]. A potential cause of fatigue is the switch between different focal distances and contexts of digital content overlaid in AR and the real world [59], [69]. There is also a difference in the increase in fatigue between the two HMD technologies, video see-through (VST) and optical see-through (OST) [70]. OST devices are prone to the problem of virtual images not being projected correctly, leading to perceptual errors that cause fatigue [70]. Headaches, eye strain or difficulty focusing after wearing an HMD are also associated with fatigue, often referred to as oculomotor problems [71].

Arigo's example scenario reported an average completion time of around four minutes [4]. Wearing a HMD for this short duration is unlikely to cause fatigue. Nonetheless, the augmented reality (AR) visualisations may increase cognitive load and could lead to potential fatigue. However, no results have been reported or measured concerning this. It would perhaps be of interest to see the contrast in the fatigue levels between the two models, displayed in both 2D and 3D.

5.5 Community

This subsection introduces the relevant characteristics to the topic of community. Community refers to individuals in relation to a particular place. The focus is on how a specific community utilizes technology or a physical location to achieve their objectives. It is closely tied to the notion of place, which derives its meaning over time from the activities of the community.

5.5.1 Information Asymmetries

As AR superimposes information on the real world, the local spatial dimension needs to be considered in terms of the norms and values of a community inhabiting it [72]. It has the power to change the meaning of a space by adding information with mental, physical or communicative effects [72]. It alters social interactions in public or private spaces, creating new power relations between users and non-user [72] as in the example of Pokemon Go (Figure 11).

Figure 11: Example of Pokemon Go which creates information asymmetries with the use of AR in public spaces [8]

Users have digital overlays of the real world that differentiate them experientially from non-users, but still intervene in the experience of non-users [72]. Users' experiences can also be differentiated through personalised content. All forms lead to information asymmetries between people and the construction of experiential bubbles [72]. AR's additional digital content to the real environment can introduce new spatial meanings to a place, potentially leading to the creation of cognitive distance from experiencing space by focusing only on the superimposed digital content [72]. As a result, people are no longer engaged in a shared experience of a place, changing the social norms of that place with implications for social life [72].

The introduction of AR in a laboratory with Arigo is relatively new, as there are few use cases in such a regulated and sterile environment. It has the potential to impact social interactions among employees in the laboratory. The wearing of an HMD can create a barrier to social interaction, introducing new communication obstacles. Additionally, employees using an HMD receive more guidance than those following paper-based instructions. The application has also the potential to provide customised content for each employee, supporting their specific requirements and highlighting certain steps for additional assistance. It is crucial to investigate the situational awareness of users as focusing solely on digital content could reduce their awareness and create hazards in the laboratory.

5.6 Time

This subsection introduces the relevant characteristics to the topic of time. The concept of time concerns the recording of data for a visualization and its display, as well as the relationship between these two moments. Additionally, time can be considered beyond pure objectivity or subjectivity and instead be viewed as socially constructed, defined as Social Time. Further, it also encompasses temporal relevance which considers social norms, activities and conventions regarding to time.

5.6.1 Temporality of objects

When analysing the temporal aspect of objects in AR it is important to distinguish the several elements and categorize them according to their spatial link [54]. For example, direct physical information objects are placed permanently in the real world and only disappear when the user moves away from them. Nonetheless, the temporal relevance of the object may differ from its presence. Temporal relevance can be combined with the temporal appearance for digital objects, which is not possible for physical objects. Procedural tasks are characterized by discrete steps that must be performed in a specific order [54]. Therefore, the information objects are subject to discrete changes as the information objects, or at least some of them, will change with each step of the task that is performed [54].

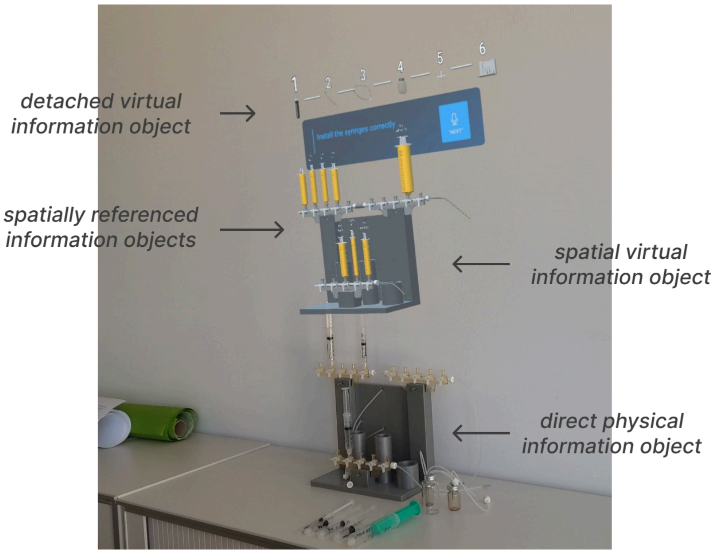


Figure 12: Information objects in the example of Arigo [4]

An analysis of Arigo's information objects with their spatial link in relation to their temporality shows that there is a need for further distinction. For example, the spatial virtual information object does not change completely with each discrete time step. However, some elements in the information object do change because some syringes have to be placed on the machine. Similarly, the detached virtual information object, shown as a step indicator at the top of the digital content, does not disappear throughout the process. However, within the information object the content changes at each discrete time step.

5.7 Potential Evaluation of the User Experience

The example of Arigo provides the opportunity to analyse the user experience from a number of different points of view. As it is used in the lab, situational awareness is important. It is important to be aware of potentially dangerous situations and to avoid tunnel vision of the digital content. Therefore, it should be analysed whether the digital content requires too much focus from the user and how the situational awareness is. As the employees will wear the HMD for a longer period of time for the creation of several mixtures, the physical effort for the perception of the digital content through AR should be taken into account. Wearing an HMD for long periods of time can cause fatigue and reduce concentration. The symptoms of simulation sickness, such as fatigue and tired eyes, should therefore be assessed.

As the user moves in the space, the legibility of the text annotated instructions should be analysed. Lighting conditions are not always the same, as employees also work at night. This also needs to be considered when evaluating the user experience.

5.7.1 Classification of Visual Realism of Digital Twins

Digital objects in augmented reality (AR) can have different levels of complexity in visual realism, depending on the technical complexity and intended use [73]. The level of visual realism has to fit to the intended task to reach the desired effect of the digital object as this has an influence on immersion and presence [73]. Schmied-Kowarzik et al. [73] have developed a *visual realism classification scale* (VRCS) in order to classify digital objects, such as digital twins, on the basis of their visual realism. Based on several studies, they account for several factors which influence perceived visual realism of digital objects and analyse the objects based on shadow, light, texture, and form complexity of 3D models. Each category was scored on a scale of 0-2 and the sum of all categories determined the visual realism of the object, ranging from 0 (not resembling a real object) to 10 (highly realistic).

- Soft shadows are considered more realistic than hard shadows, leading to a rating system for the shadows of digital objects which includes (0) no shadow, (1) hard shadow, or (2) soft shadow.
- Regarding light, the classification consist of (0) absent light, (1) general environmental light, or (2) directional light.
- The quality of texture used for rendering plays a crucial role in defining the digital object's appearance. It is defined in the VRCS as (0) no intentional colour, (1) intentional colour, (2) colour that shows a pattern and an additional attribute (1) if the texture appears 3-dimensional.
- When the number of polygons used for rendering is reduced, it is possible to change the form of the object. The VRCS defines the form as either (0) unintentionally shaped, (1) intentionally shaped, or (2) includes detailed shape adjustment.

Different levels of visual realism have different effects depending on the chosen scenario. As for example in exposition therapy to treat acrophobia (fear of heights) in virtual reality, people without phobias perceived a difference in presence depending on the visual realism where as people with phobias show lower difference in presence in the varying visual realism scenarios, indicating that emotions have an influence on perceived visual realism and presence [74].

Some researchers argue that high visual realism can distract users and reduce performance [75]. However, recent research shows that the desired level of visual realism depends on the task and no general recommendation can be made about visual realism [75]. The results suggest that when digital objects are used for shape learning, high levels of visual realism are appropriate [75]. However, when digital objects are used for learning processes and abstract knowledge, lower levels of visual realism seem appropriate [75]. To further evaluate if the right level of visual realism is chosen, more metrics from attention, cognitive load or eye gaze may provide useful insights [75]. When the level of visual realism of digital objects is excessive, attention to detail is very high [75]. This could be shown by analysing gaze and attention. In addition, the cognitive load will be increased if the level of visual realism of the digital object is too low or too high [75].

5.7.2 Effectiveness of Guidance Methods

Mental models can be assessed using five techniques, namely problem solving, verbal report, drawing, categorisation and conceptual pattern representation [76]. To evaluate a spatial mental model, the technique of drawing in combination with verbal report seems to be appropriate [77]. This allows the user to draw the spatial references they use for orientation and to verbalise the relationship between the different objects. It can then be assessed whether the orientation instructions given correspond to the users' spatial mental models. The drawn spatial mental model can also be used to evaluate whether the spatial links of the information objects are appropriate, as users visualise the spatial reference of the objects they have in mind.

As mental models differ significantly across user groups, it is essential to conduct careful evaluations of the target user group to ensure the intended application is designed with the appropriate mental model [78]. This, in turn, ensures that the application aligns with end-user mental models, making interactions less cognitively demanding [78]. The cognitive load of the instructions can be assessed with the questionnaire developed by Leppink et al. as in Table 1 [28]. It measures the intrinsic load (the complexity of the task and the prior knowledge of the user), the extraneous load (instructional features that are not suitable for learning) and the germane load (instructional features that are beneficial for learning). With a ten-item survey, it can be assessed whether the instructions match the knowledge level of the target user group and which elements are advantageous or disadvantageous for learning [28]. The distinction in the cognitive load could also indicate whether the temporal aspects of the step-by-step instructions are being used in an appropriate way.

	Question	**Load**
1	The topic/topics covered in the activity was/were very complex.	intrinsic
2	The activity covered formulas that I perceived as very complex.	intrinsic
3	The activity covered concepts and definitions that I perceived as very complex.	intrinsic
4	The instructions and/or explanations during the activity were very unclear.	extrinsic
5	The instructions and/or explanations were, in terms of learning, very ineffective.	extrinsic
6	The instructions and/or explanations were full of unclear language.	extrinsic
7	The activity really enhanced my understanding of the topic(s) covered.	germane
8	The activity really enhanced my knowledge and understanding of statistics.	germane
9	The activity really enhanced my understanding of the formulas covered.	germane
10	The topic/topics covered in the activity was/were very complex.	germane

Table 1: Questionnaire items to measure cognitive load by Leppink et al [28]

5.7.3 Analysing the Semiotics of Space

By incorporating digital content into real-world environments, we have an impact on the meaning of physical space, subsequently altering human behaviour and attitudes [79]. This influence can be examined analysing the semiotics of space [79]. Semiotics examines the interaction of humans with their environment by analysing their perception of sign and objects in space, the need for interpretation by the interpretant, leading to actions, forming the chain of interpretation as defined by Määttanen (Figure 13) [9], [10].

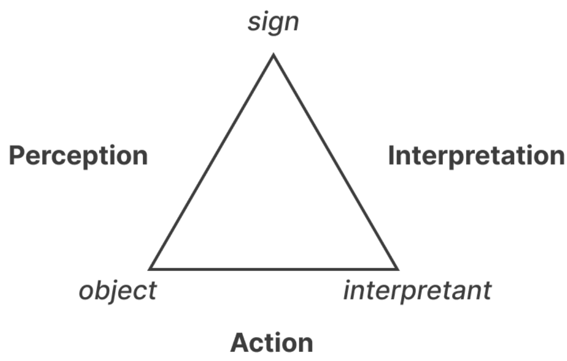


Figure 13: Semiotic triangle defined by Määttanen [9], [10].

Semiotic analysis uses qualitative techniques commonly used in the social sciences [80]. By conducting interviews, it is possible to categorise and label the relevant features that contribute to the transformation of the meaning of space [80]. Ishizawa et al. propose a categorisation into seven domains to analyse the semiotics of space: Living Things, Objects, Landscape, Information, Institution, Occurrences, and Transmedia Storytelling [80].

- *Living things* primarily concern humans with whom we form relationships and build empathy.
- *Objects* refer to physical or digital entities that exist in either the real or virtual space. Both should respect physical laws and provide a sense of reality. People show emotions towards objects, especially in social interactions. In addition, the authentic value of an object must be considered, which is the reason for its existence.
- *Landscape* analyses the empathic and aesthetic value of a user in relation to a space, in particular the realism of the space.
- *Information* is typically superimposed in augmented reality (AR). The authenticity of the information presented is crucial for building trust. Therefore, the authenticity of the text in relation to its placement must be analysed, as well as its persuasiveness.
- *Institutional mechanisms*, such as the currency system or social norms, shape our daily lives. AR can alter or highlight human attitudes and behaviours through social influence, considering the ideological and authentic value of the form of influence in AR.
- *Occurrences* can evoke strong emotional experiences that users remember. These emotions can be generated by ideological messages or by increasing the positive value of a space.
- *Transmedia Storytelling* can increase the sense of objects being present in the real scene by providing information through multiple means, evoking emotions and enhancing their positive value.

By using this method it is possible to analyse the visible, behavioural, interactive and factual aspects of a given space, thus influencing the interpretation of the environment and social relations. Further, it is also possible to analyse visual cues used in AR. The perception of a visual cue (sign) and its interpretation can be examined by interviewing the user and analyse the followed action of the interpretation. Combined with spatial mental model analysis, it provides more detail if the visual cues are linked correctly, which also influences interpretation and the resulting action.

5.7.4 Readability of Annotations

To evaluate the readability of text annotations in AR most researchers define their own questionnaires for subjective feedback [81]. Objective data is gathered through performance measurement such as idendification of the correct letter or words displayed [81] such as [82], [83], [84]. As there is no standardized instrument the results can't be compared among the studies, as well as it does not analyse why text is not readable. For this an in-depth analysis is necessary based on the introduced characteristics for annotations. The following questions should be posed in order to evaluate the user experience for text annotations:

1. How complex is the environment where the text annotation is placed?
2. How and when does the user move in the environment?
3. When is the text annotation needed?
4. What is the semantic relevance of the text annotation to an object?
5. How complex is the text annotation itself?

Further information could be derived from a cognitive walkthrough conducted by an expert to analyse the above points, as well as to analyse whether the content is visible and the text readable given the lighting conditions of the environment.

5.7.5 Level of Situational Awareness

Many Situational Awareness (SA) questionnaires come from the military and focus on combat situations, such as the Quantitative Analysis of Situational Awareness (QUASA) [85] or Situational Awareness Rating Technique (SART) [86]. In addition, many SA questionnaires are used in general aviation, where SA is extremely important and operators need to be aware of many elements in their environment, notice changes immediately and react to them very quickly [81]. Many of these questionnaires are not suitable for use in the private sector due to their specific application and questions. However, due to the lack of alternatives, many researchers use these questionnaires and modify the questions according to their scenario [81]. Further information for SA could also deliver questionnaires in the field of *presence*, as many of these also account for attention, engagement and immersion. More information can be found in the chapter of *Presence*.

In addition to questionnaires, there are several other ways of analysing people's SA [87]. The *Freeze Probe Technique* freezes the scenario in a certain situation and analyses the SA by asking specific questions. However, this method is very intrusive and unnatural. A less intrusive but similar technique is the *Real-Time Probe Technique*, which also asks questions about the SA but does not freeze the scenario. Questionnaires can be used in *Self-Rating Techniques* or *Observer-Rating Techniques*. Eyetracking can also be a useful tool for SA, as it provides information about the user's gaze and where their attention was. Further, performance data as time-to-completion or reaction-time to a specific event is often measured to evaluate SA [81], [88].

5.7.6 Evaluation of Presence including Situational Awareness

Presence is a complex construct that requires a number of different elements. To account for different variables, questionnaires must consider multiple elements. A recent example of a presence questionnaire is presented by Georgiu and Kyza, the *Augmented Reality Immersion* (ARI) questionnaire [89]. It considers a user's engagement, engrossment and immersion in AR. Engagement includes items related to interest and usability, engrossment relates to emotional attachment and focus of attention, and immersion relates to presence and flow. It is a potentially useful tool that covers a number of aspects and can provide useful cues for situational awareness, by including usability and attention. The *Mixed Reality Experience Questionnaire* (MREQ) has also been developed specifically for mixed reality experiences [90]. It assesses how well the digital world and the physical world blend together, but it does not include items such as attention or usability.

The *Igroup Presence Questionnaire* (IPQ) is a widely used tool in virtual reality (VR) that assesses spatial presence, involvement and experienced realism in a virtual environment [65].

Although it was developed for VR, many items could be used in AR or need slight adaptation to adhere to AR. The IPQ is based on the work of Witmer and Singer, who developed the *Presence Questionnaire for Virtual Environments* (PQ) [66]. Although designed for VR, the PQ could provide interesting insights for AR applications as it also takes into account involvement, interface awareness and reality. It also provides an assessment of a user's immersive tendencies that is unrelated to the specific scenario. This information could also provide interesting information and correlations to the results of the general PQ.

5.7.7 Level of Fatigue after Task Execution

The level of fatigue is commonly measured through Simulator Sickness Questionnaire (SSQ) mostly used in VR [71]. The subscale *Oculomotor* includes items measuring fatigue, eye strain and difficulties in focus. However, the SSQ was developed in 1993, therefore many researchers think its application is no longer up-to-date as technology has changed completely since then. Never-the-less, it is still used heavily as it measures relevant factors of motion sickness introduced by technology.

More detailed information on the symptoms of AR exposure delivers the Virtual Reality Symptom Questionnaire (VRSQ) as in Figure 14 [11]. Even though it was developed for the application in virtual reality, many symptoms can also be found in the use of AR. The more detailed results on what symptoms occur after wearing an HMD could provide interesting insights and stimulate discussion on how to improve digital content to reduce symptoms. For example, if many participants report eye fatigue, eyestrain and difficulty focusing, it is likely that too many switches between focal distances are required to complete the task. This may indicate the need for a better visualisation that requires fewer switches and a more natural integration of the digital content into the physical space.

Subject Code…………………………….. Correction: CL Specs None
Date………………………………………

	None	slight	moderate	severe	None	slight	moderate	severe	None	slight	moderate	severe
GENERAL BODY SYMPTOMS												
General discomfort	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Fatigue	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Boredom	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Drowsiness	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Headache	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Dizziness	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Difficulty Concentrating	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Nausea	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
EYE RELATED SYMPTOMS												
Tired eyes	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Sore/aching eyes	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Eyestrain	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Blurred vision	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6
Difficulty focusing	0	1 2	3 4	5 6	0	1 2	3 4	5 6	0	1 2	3 4	5 6

Other symptoms/feelings			

Figure 14: Virtual Reality Symptom Questionnaire (VRSQ) [11].

5.8 Conclusions

The analysis shows that, depending on the user's needs and the context in which they are placed, AR can be a powerful technology to support users in different ways. Perception and spatial learning can be influenced by the visual realism of a digital twin, as well as the way it is presented in 2D or 3D. Furthermore, a higher level of visual realism does not always lead to better results. The display and presentation of the digital twin is also dependent on the lighting conditions in the environment. This, as well as the background of the digital content being presented, should be considered to ensure readability and appropriate perception. The varying levels of visual realism can also affect users' situational awareness, making them more prone to shoulder surfing as they are less aware of observers around them. Furthermore, reduced situational awareness can lead to a higher risk of accidents and distraction, known as information tunnelling. It is therefore important to blend the digital content with the virtual environment to reduce the dominant perception of only the digital content. A higher level of blending could also lead to a higher sense of presence, which is defined as *being consciously engaged in the augmented world*. To achieve this state, the spatial connection of the digital content should also be considered in order to anchor the elements appropriately in the real environment and not to disturb the spatial mental model of the users. Visual cues can also disrupt presence if they are not properly integrated. It is important to consider how they are integrated, either as part of the scene or outside of immersion, and whether they are communicated explicitly or implicitly. As the example of Arigo is a procedural task, the user's work steps are discrete in nature. Therefore, the information provided is typically discrete as well, reducing the complexity of when information is relevant to the user in the development of the application.

While AR with HMDs offers a lot of possibilities, it can also lead to problems. The augmentation of the real world can result in increased cognitive load, leading to fatigue. In addition, prolonged use of an HMD could lead to perceptual errors, resulting in other symptoms known as oculomotor problems. People wearing HMDs may also change the meaning of a place by creating experiential bubbles, or change the social interaction of places.

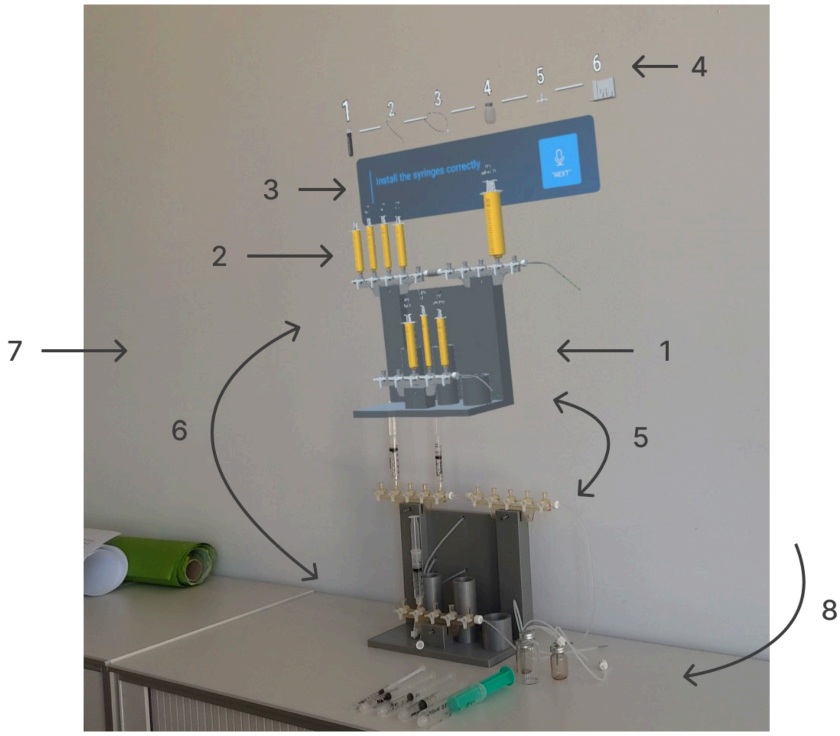


Figure 15: Example of Arigo and relevant characteristics for scenario analysed in this thesis

Figure 15 shows which elements are evaluated independently in this analysis based on the scenario of Arigo.

1. The digital twin can be analysed in terms of its visual realism and suitability for the task at hand. The complexity of the digital twin must be taken into account, as well as the user's attention to detail and the resulting cognitive load.

2. Visual cues can be analysed to determine if the spatial link is appropriate, and the user's interpretation can be assessed.
3. The spatial mental model provides information on whether the spatial reference in the instructions is appropriate and how much cognitive effort is required to interpret the instructions.
4. In order to analyse the readability of AR text, it is necessary to take into account the complexity of the text, its relevance and the consideration of the environment.
5. Analysis of fatigue after wearing an HMD can provide insights depending on the user's symptoms.
6. Presence provides more information about a user's engagement, involvement and immersion, as well as the blending of digital and physical worlds.
7. Semiotics of space can reveal information asymmetries introduced by AR in spaces, as well as changes in the meaning of space for users, resulting in different behaviour and interaction.
8. Situational awareness analyses the user's attention and whether awareness of the environment is still present.

Many elements influence the UX in AR, as the analysis of the chosen scenario for the HMD application shows. The evaluation of these factors can be carried out independently or in combination in order to obtain a holistic view of the UX, and sometimes a combination of the two may be required. There are cases where a combination of methods may be necessary because the results are interrelated and only a combination of approaches can effectively analyse the root causes of certain findings. An example of such an analysis can be seen in the evaluation of visual cues. In order to accurately evaluate visual cues, it is necessary to combine the analysis of the spatial mental model and semiotics of space. These two approaches are interrelated and affect one another. It can be analysed how a visual cue is perceived, what the interpretation is by the followed action. If combined with the analysis of the spatial mental model it can be analysed if the spatial cues is linked incorrectly, therefore leading to different perception and interpretation.

6 Subject 2: Mobile Augmented Reality

In this section, the example application for mobile augmented reality and its relevant characteristics are introduced on the basis of the term situatedness. The characteristics are discussed in terms of their theory and relevance to the scenario. Potential evaluation techniques are presented for each aspect, followed by an overall analysis of the scenario, its characteristics and potential evaluation techniques.

6.1 Example Application - feey

Mobile Augmented Reality can be a useful way to visualise products in the physical space and help users make decisions through simulation [91]. Elements can be analysed based on aspects such as size, appearance and integration into the real world. The online shop *feey* provides such an example for buying plants as illustrated in Figure 16.

Figure 16: Mobile AR application of feey to support users purchasing plants in collaboration with an expert

The AR application of feey is used in consultation session with an expert to find the right plant for the user. It is a web application which allows to have a video call with the expert and at the same time show the environment of the user. In collaboration, the expert and the user can discuss the settings of the room as well as the needs of the user regarding plants. Both participants can see the same scene on their mobile device to foster communication. Further, they can select various plants and place them in the room to analyse their appearance. The expert can support the user in the selection of plants regarding their suitability of the different environmental properties. They can discuss together through the visualisation on the mobile device the preferences and suitability of the plant, simplifying communication and decision through simulation.

For the scenario presented, several points are relevant to consider for the user experience. These points are categorised based on the notion of situatedness. Figure 17 gives an overview of the characteristics that are relevant and have been selected for the analysis of the given scenario.

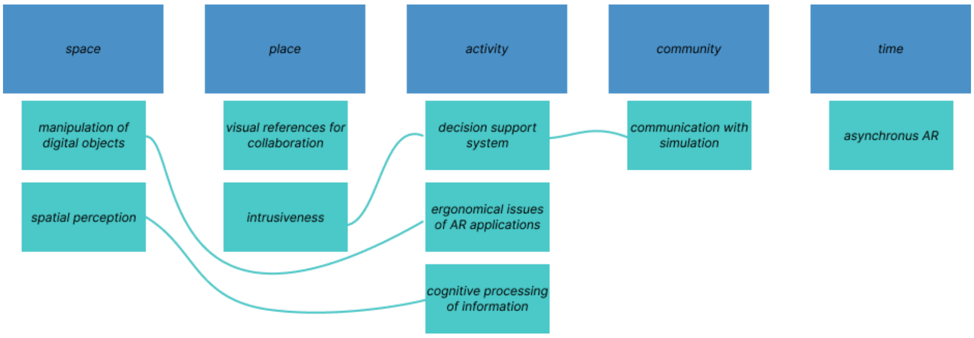


Figure 17: Overview of relevant and chosen characteristics for the scenario of feey, categorized by the term situatedness

6.2 Space

This subsection introduces the relevant characteristics to the topic of space. Space encompasses the physical environment and its digital content. It refers to the spatial representation of information, including its properties and characteristics in three-dimensional space.

6.2.1 Manipulation of digital objects

Digital objects in AR can be manipulated in 6 degrees of freedom (6-DoF) [92]. The 6 degrees can be divided into translation and rotation. Translation moves an object along the x-, y- or z-axis of space, while rotation turns an object around its own axis for all three axes. These are also referred to as yaw (turning the object around its own z-axis), pitch (rotation around the x-axis) and roll (rotation around the y-axis). Further interactions can be made by selecting and scaling objects. All these interactions pose some problems, especially in mobile AR, as the interaction is usually done through touch gestures on the screen, where the fingers cover a large part of the screen [93]. Furthermore, the need to control multiple DoFs can make the interaction uncomfortable and complex [94].

Several solutions to this problem have been proposed. Cohe et al. [95] suggests touch sliders on the screen for translational movements, the inertial sensors of mobile devices for rotational movements and scaling by pinching two fingers. Mossel et al. [92] propose a solution for 3-DoF by selecting and holding an object by touch on the mobile device and moving the object along the x and y axes. To move the object on the z-axis, the mobile device must be tilted so that it is over the object (Figure 18).

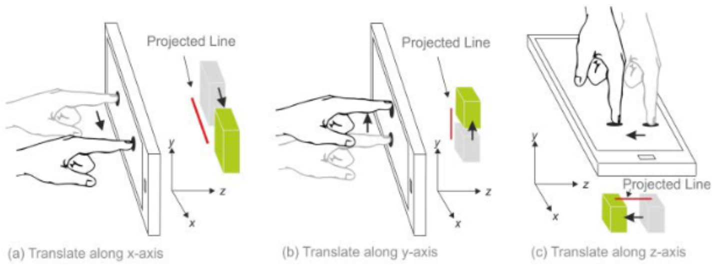


Figure 18: Select and hold object, moving finger up and down for translation on x and y-axis, tilt device for translation on z-axis.

Monteiro et al. [96] link the manipulation of digital objects to tangible objects. The scaling of objects can be tied to a physical slider that controls the size of the digital object (Figure 19, top). Rotation around its own z-axis can be controlled by turning a coffee cup, while translation movement can be bound to a plate that acts as a steering wheel, familiar from steering a car (Figure 19, bottom).

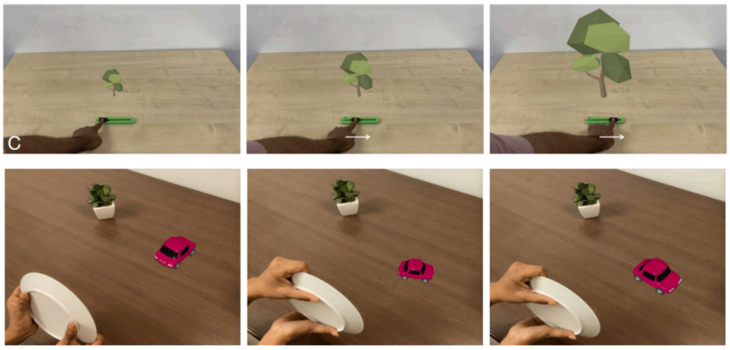


Figure 19: (top) scaling of object bound to a physical slider, (bottom) translation of object bound to a plate acting as a steering wheel

In the example of feey, only 3-DoF need to be controlled. Movement along the x and z axes, as the plant should not float in the air or under the ground, and rotation along its own y axis, as it should stand vertically on the ground without tilting. There is also no need for scaling, as the virtual plant should be the same size as it is in the real world. The interaction in the application is implemented with a select and hold approach, where the plant can be selected and moved around the room by holding it down.

6.2.2 Spatial Perception of 3D Objects

Spatial perception in AR is the accurate perception of size, position, distance and orientation of objects in the real or virtual environment [59]. Colour, luminance as well as the relative difference to the background can affect depth perception [97]. When an object's colour appears warmer, it is perceived as closer than when it is colder [14]. This effect is particularly pronounced when the object is 2-dimensional, whereas the effect is smaller for 3-dimensional objects in AR [14]. The fidelity of the object also influences depth perception, with low fidelity being more influenced by colour and luminance [14].

Perception in AR is well studied, especially in the area of head-mounted displays [97]. An analysis shows that results are not transferable between HMD AR, Mobile AR and VR [97]. Depth perception is influenced by different variables, while each technology has its own characteristics that influence depth perception [97]. Furthermore, the results of depth perception of 2D and 3D objects vary, indicating that the properties of the two different renderings have an influence on the perception [97].

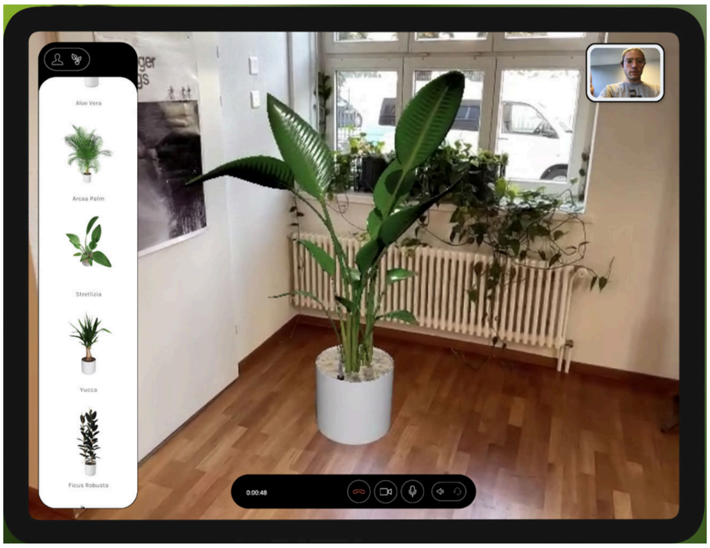


Figure 20: Placement of plant in room as high-fidelity 3D model

Spatial perception is important in the example of feey as users place plants int their environment to analyse their look for purchase decision.The plants presented as 3-dimensional objects, which are mainly green, with varying shades of green as seen in Figure 20. As green often appears cooler, this can lead to distortions in depth perception. Furthermore, the background for the plant placement cannot be influenced, as it depends on the environment from the user. It can therefore be assumed that this could also distort the depth perception of the 3-dimensional plant as relative brightness and colour depends on the background. Lastly, also the colour of the pot could influence perception.

6.3 Place

This subsection introduces the relevant characteristics to the topic of place. Place is concerned with location, beyond a purely spatial perspective. It is the meaning of a space in relation to the people who inhabit it, the activities that take place within it and its historical context.

6.3.1 Visual References for Collaboration

In collaborative tasks, groups usually try to establish *common ground* or *grounding* [98]. It is the shared understanding of a task or an environment by means of verbal or non-verbal acknowledgements [98]. To build a common understanding of the task within the group, successful grounding is important [99]. In collaborative AR, grounding can be achieved through references to objects or items in the space that are visible to all participants. Those references can either be general (e.g. “north”), definite (named entities), detailed (e.g. “the blue ball”) or deictic (e.g. “that one”) [98]. Collaborative AR is particularly useful for deictic references as they simplify the production of communication with the principle of least collaborative effort [98]. However, to be able to use deictic references, the objects in AR need to be visually referenced [100]. This can be achieved by manual highlighting of objects by the user [101], shared eye gaze [102] or head gaze [103], indicating finger pointing [102], indicating with annotation [104], or by moving/placing objects in areas with spatial meaning [98]. This highlights the referenced object for both users, allowing deictic reference [98]. Without knowing which object is being referred to, the sentences with deictic reference make no sense [105].

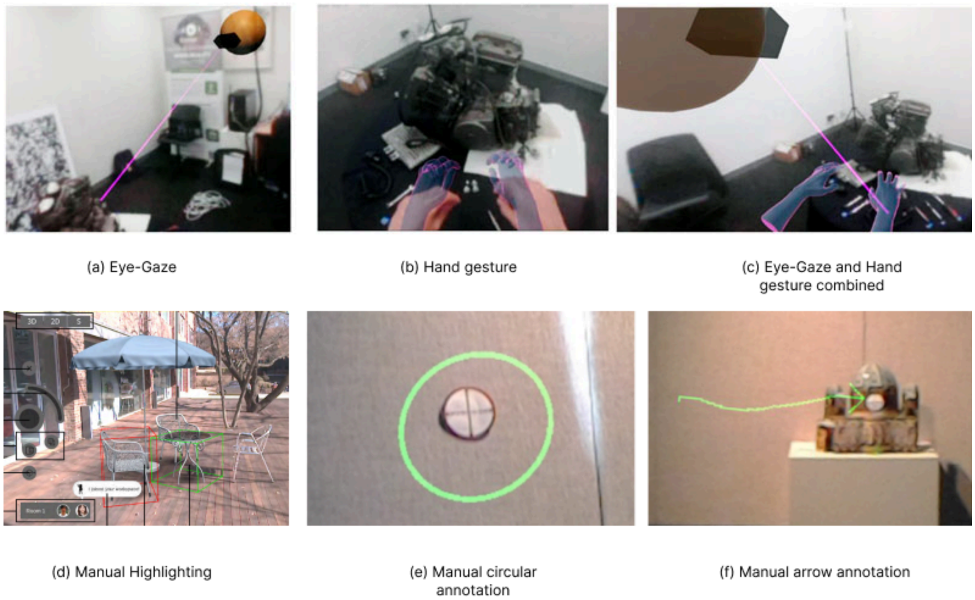


Figure 21: Options for creating visual reference in AR

Figure 21 shows different options for visual reference: (a) the user's eye-gaze is shared [102] (b) the user's hand gesture is shared [102] (c) the eye-gaze from (a) and the hand gesture from (b) are combined [102] (d) objects can be manually highlighted with a coloured square around the object [101] (e) manual annotation of an object with a circle by hand interaction [104] and (f) manual annotation of an object with an arrow by hand interaction [104].

Eye-gaze can be used not only to improve communication, but also to convey implicit information such as what the user is thinking [106] . Such as in the example of support in coding from an expert [106]. Implicit information is conveyed by the shared gaze of the students, which tells the expert which part of the code the user is analysing. They also implemented sharing the user's cursor, which conveys explicit information as the user can move the cursor deliberately to communicate. Several studies show that such tools facilitate simpler communication in collaborative environments and also reduce the number of misunderstandings [106], [98]. However, depending on the task, eye-gaze is not always the best option for creating a visual reference, as shown in the example of repairing an engine using AR [102]. Users had to repair an engine with the help of an expert who either used eye gaze, hand gestures or a combination of both as shown in Figure 21 (a), (b) and (c) [102]. The results show that with superimposed hand gestures, participants were able to complete the tasks faster and provided more information for the assembly of the parts. With eye gaze, they reported missing some information because they did not know how to manipulate the objects. These studies show that visual cues need to match the task and the knowledge level of the target user group.

The Feey example offers the possibility of successful grounding because it places virtual objects in the user's real environment, visible to both expert and user. In addition, deictic references can be used by the participants, but the elements are only highlighted when they are held and moved around in the environment. As the end user's video camera is streamed to the expert, manual highlighting of objects could also be done by the user by moving the camera to a particular object. However, the application could also potentially use other means to make communicating easier and more efficient. Eye gaze could be a useful way of analysing where the end user is looking, which could be an indication of potential spatial thoughts in relation to the plant. It could also be useful to use annotations to draw in the environment to discuss elements in the room with deictic references rather than explicit references, which can be prone to error.

6.3.2 Intrusiveness

Augmented reality can raise feelings of intrusiveness and privacy concerns for several reasons. Firstly, the user has to give the camera access to the application in order to use AR functionality, which can lead to a reduced sense of control and increased intrusiveness [107]. Secondly, the camera films the user's surroundings and superimposes digital objects on the environment in real time, which could feel intrusive and lead to negative emotions as it is perceived as intruding on the user's physical world [107]. These concerns would also reduce spatial presence, as the use of AR is perceived as a risk of vulnerability [108]. The reduction in spatial presence further reduces the decision comfort of AR mobile shopping applications [108].

Studies show that people perform a cost-benefit analysis when using AR applications, analysing whether the benefits of using AR outweigh concerns such as privacy and intrusiveness [109]. In most cases, the benefits seem to outweigh the concerns, especially for applications that place objects in the user's physical environment to assist with shopping [109]. Smink et al. [107] compared an AR and a non-AR application for furniture shopping. The AR application was perceived as less intrusive, more enjoyable and more immersive than the non-AR experience. The authors argue that with AR the purpose of the application, to sell products, is more hidden than with non-AR applications, leading to more enjoyment, which reduces intrusiveness.

Feey's example scenario aims to help people buy plants for their home using AR and video chat with an expert. As the user has to film their personal home and show it to a stranger, this could lead to feelings of intrusiveness. However, since the user actively chooses to have this expert discussion with an AR app, it can be assumed that the attitude towards AR is positive, which means that the benefits of using AR are perceived as higher than the costs. Therefore, it may also not reduce the spatial presence, which supports the decision comfort for the shopping assistance. Furthermore, as shown in the study by Smink et al. [107], AR applications for furniture shopping assistance are perceived as less intrusive than traditional non-AR shopping applications. Therefore, it can be assumed that intrusiveness and privacy concerns are low and do not affect spatial presence.

6.4 Activity

This subsection introduces the relevant characteristics to the topic of activity. Activities should not be viewed in isolation. They are components that contribute to broader goals and can be carried out in different spaces, over different periods of time and in collaboration with others. The relationship between a person and the goals he or she wishes to achieve is complex and influenced by the social and cultural context.

6.4.1 Decision Support System

Decision making is a complex process that requires cognitive effort to process information, especially when large amounts of data are present, which can lead to poor decisions [110]. To reduce the cognitive load, augmented reality can be a good solution. It allows the user to load an *external memory* rather than using their own memory capacity [91], [111]. The visualisation should be spatially situated in order to improve the sense and decision making [91]. Through the relationship of the visualisation and the environment in which it is placed, the user can gain meaning through the combination and relationship between the two [112]. The information presented can then be analysed based on contextual data, leading to better decision making and also reduced cognitive load [91], [110], [111]. The digital information is embedded in the environment and can be moved naturally within the physical environment, leading to an increase in decision comfort [113]. As a DSS in combination with AR is not tied to a screen, it falls into the category of pervasive computing.

Decision Support Systems (DSS) can have different drivers as a basis [112]. For example, Marques et al. define five drivers [112]: model-driven (based on simulation models), communication-driven (supporting collaboration), data-driven (facilitating access to data), document-driven (managing information) and knowledge-driven (providing expertise). The information provided by the DSS can reduce uncertainty, either actively, passively or interactively [114]. Active and passive are distinguished by whether the user has to actively search for new information or whether the information is passively received, for example in reviews [114].

The Feey example provides a DSS with multiple drivers. As it shows a simulation of the plant in the real environment, it is a model-based DSS. In addition, it is a communication and knowledge-based DSS because of the collaboration with experts, which allows for professional expertise. The visualisation of the plant is 3-dimensional and can be moved around the room on the digital devices, allowing different placements and views to be compared. The visualisation allows the user to understand how the plant looks in the room and how it fits into the space. In addition, by communicating with the expert, the user can get more information to help him or her decide which plant to buy.

6.4.2 Cognitive processing of Information

People seek, represent, process and retrieve information in different ways, which can be reflected in *cognitive styles* [115]. A popular distinction in cognitive styles is the *Visualiser-Verbaliser* (VV), which differentiates the processing and representation of information either visually or verbally [115]. People with a preference for the visualiser style use mental imagery and spatial representations to process information [116]. In comparison, the verbaliser style prefers words and linguistic representation of information [116]. The visualiser type can be further divided into spatial and object visualisers [116]. Spatial visualisers construct mental images to represent spatial relationships between objects, whereas object visualisers construct vivid and detailed mental images of objects [116].

Research shows that the performance and satisfaction of applications, as well as the effectiveness of object representation, is influenced by the VV cognitive style preference [115], [108]. Verbalizers benefit more from AR because they can rely less on their mental imagery than visualizers [108]. Object visualisers also benefit more from AR than spatial visualisers [111].

For the Feey example, verbalisers can be expected to benefit more from placing AR plants in their environment. Visualisers, especially spatial visualisers, are likely to rely more on their own mental imagery. However, with the combination of the video call functionality with an expert, further benefits are possible for this group.

6.4.3 Ergonomics issues of AR applications

Interaction with AR is a broad topic, especially considering the different technologies that allow for different interaction possibilities. In mobile AR, interaction is usually done by manipulating the digital objects on the screen of the mobile device [12]. However, there are several other possibilities such as the use of different sensors like accelerometers or gaze tracking. All these possibilities have different effects on the body and its posture [12]. It is therefore necessary to consider human factors and ergonomics (HFE) when developing an AR application [12]. Poor posture over a long period of time can lead to pain and affect performance and satisfaction as illustrated in Figure 22 [12].

Figure 22: Bad body posture through the use of AR which can lead to pain [12]

When interacting with mobile AR for long periods of time, users usually sit down to stabilise the mobile device [117]. Otherwise, posture support should be provided [12]. Further issues can regarding ergonomics could be fatigue due to extended use of the application, pain due to heavy devices, or difficulty manipulating the digital objects through the screen [31].

In the case of the AR application of Feey, the expected interaction time is longer. The user has an appointment with an expert to discuss the suitability of different plants for the given space. A smartphone or tablet is used to facilitate the video call. As the room needs to be scanned at the beginning and the different plants need to be placed in the room using AR later on, the user needs to hold the device in front of their body for a longer period of time. Also, to see the plants from different angles, the user has to walk around the room to analyse the appearance. All of these factors can lead to poor posture, causing pain and fatigue. It should be considered how the user could be supported regarding ergonomics in the use of the application.

6.5 Community

This subsection introduces the relevant characteristics to the topic of community. Community refers to individuals in relation to a particular place. The focus is on how a specific community utilizes technology or a physical location to achieve their objectives. It is closely tied to the notion of place, which derives its meaning over time from the activities of the community.

6.5.1 Communication with Simulation

Augmented Reality makes it possible to place virtual 3D objects in a real environment and to share this experience digitally with other people [118]. This enables multi-party collaboration and communication of potential design interventions in the environment [118], [119]. Users can use AR to illustrate their design ideas and show their impact [119]. It allows them to visually share their point of view (POV), which is essential for social decision-making [120], [119]. This enables effective communication between different parties, as other participants can gain contextual information about the design intervention and understand the circumstances [120]. Shared understanding with the support of in-situ visualisations improves the chances of better decision making [118]. Also in an expert-novice planning process, situated visualisation can improve decision making [120]. The expert gains a richer experience of the location, enabling informed recommendations [120]. Meanwhile, the user can share their environment and provide relevant contextual information to the expert [120].

Studies of collaborative environments in virtual reality also show that participants spend more time talking to each other in immersive collaborative scenarios than on traditional desktop devices, and that conversations are more social [99].

The Feey AR application allows users to share their environment with an expert, making it easier to gather contextual information and understand environmental features to provide valuable recommendations. By placing the virtual plant in the user's environment, the design implications of the plant can be visualised. The in-situ visualisation allows the expert and user to explore the options and implications of different plants, facilitating effective communication. The visualisation, together with the expert's discussion of different options and understanding of the context, enables the user to make better decisions.

6.6 Time

This subsection introduces the relevant characteristics to the topic of time. The concept of time concerns the recording of data for a visualization and its display, as well as the relationship between these two moments. Additionally, time can be considered beyond pure objectivity or subjectivity and instead be viewed as socially constructed, defined as Social Time. Further, it also encompasses temporal relevance which considers social norms, activities and conventions regarding to time.

6.6.1 Asynchronous AR

Collaboration in AR can be either synchronous or asynchronous. Most applications that focus on collaboration use synchronous communication [13]. However, few studies show that AR is also useful for asynchronous collaboration, such as in expert-novice scenarios, particularly when spatial information is relevant [13]. Communicating spatial information via paper is difficult and can lead to problems and performance degradation [13]. Cho et al. [13] present a use case for assembly training in which an expert uses an AR headset to record the steps of an assembly process. Later, the novice can view the recording in AR, with the expert and his steps superimposed on the real environment. In addition, the novice can add annotations in the augmented environment where he or she has questions about the steps that the expert has performed. The expert can also review these annotations in AR to answer the open questions. Not only does the application show the steps performed through a virtual replica of the object, but the expert is also visualised as an avatar, increasing co-presence. Figure 23 shows the avatar of the expert and the objects used for assembly in AR, as well as the annotations made by the novice in AR.

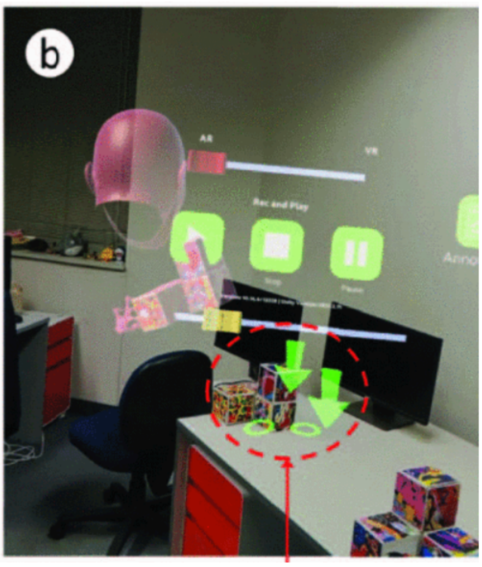


Figure 23: Avatar visualisation of expert and objects through AR, annotations in AR from novice user highlighted with red circle in image [13]

Another application for asynchronous collaboration in AR is presented for interior designers [121]. Typically, they only have access to the client's space when they are physically present. During this time, they can use AR to design the environment and present it to the client. However, they cannot make changes or experience the space after they leave. Asynchronous AR allows clients to scan their environment with lidar sensors in an iPad and the software creates a replica of the environment. The designer can place objects in the virtual replica and visualise the space through AR in their office, simulating in-situ AR. Once the design is complete, clients can experience the design through AR on site. This provides an immersive and non-immersive experience of designing a space, which is proving to be very effective [121].

Feey's example scenario uses only synchronous AR for collaboration between the expert and the novice user. However, asynchronous AR collaboration may have great potential for this use case. The user could scan their home with their smartphone, allowing the expert to experience the space and get a sense of the environment prior to the consultation. This would also allow the expert to place some plants in the simulated environment and analyse their appearance in advance. The consultation could then take place, with the novice and expert working together synchronously to discuss the different options. The asynchronous AR would also allow the novice to re-experience the simulation after the counselling session, where several options could be analysed again for a better decision.

6.7 Potential Evaluation of the User Experience

This section presents options for the evaluation of the identified characteristics that are relevant to the example. They can either be assessed in isolation, or they need to be analysed in combination with other aspects, as they are closely linked and interdependent.

6.7.1 Effectiveness of a Decision Support System

There are several ways to evaluate the effectiveness of a Decision Support System (DSS). As it should support the user in making decisions and reduce uncertainty, either about product performance or product fit, the cognitive load of making the decision should be lower than without the DSS by providing externally relevant information [110]. This can be analysed by separating the intrinsic, extrinsic and germane load [110]. Intrinsic load is caused by the task itself, extrinsic load is caused by the additional information presented and germane load is caused by the creation of mental models and the processing of information [110]. This can be assessed with the cognitive load questionnaire by Leppink et al. [28]. As the online shop example of feey has two drivers, model and communication, we cannot distinguish which driver influences the different cognitive load categories.

Another widely used method is presented by Häubl et al. [122]. In their approach the evaluate *amount of search for product information*, *consideration set size*, *consideration set quality*, and *decision quality* for a DSS. The amount of search for product information is measured as the number of products for which detailed information is obtained, which is an indication of the amount of search effort for alternatives. Consideration set size is the number of products the user considers for purchase and is an indication of the uncertainty of the purchase decision. Consideration set quality is the proportion of products in the consideration set that are suitable for given attributes. Decision quality can be measured objectively and subjectively. The objective measures analyse whether the final product chosen is the best product given the desired attributes of a product. Further, it could also be assessed whether users want to switch to another purchase alternative when given the opportunity. Switching is an indication of poor decision quality. The subjective measure assesses the degree of confidence in a purchase decision, which can be assessed using a rating scale by asking participants: "How confident are you that the product you just purchased is really the best choice for you?"

Both of these methods could also be combined with semi-structured interviews in order to obtain more information about the types of product uncertainty and whether the methods presented could support the users in the reduction of their uncertainty [110]. For example, as in Häubl et al.'s [122] approach, the assessment of decision quality could be further divided into performance and fit uncertainty, assessing whether the uncertainties arise from concerns that the user does not have enough information to make a good decision or whether they are unsure whether the product will meet their needs.

Another combination is to distinguish the cognitive style of the test users, as this influences satisfaction and effectiveness [115], [108]. To assess whether someone processes information visually or verbally, several questionnaires have been developed [115]. As these are time-consuming and tiring, Petsas et al. developed a game to assess the preferred style [115]. The cognitive style information could provide more information about which group the AR app is more appealing to and effective for. It can also be used to assess whether the sample is unbalanced in terms of cognitive style.

6.7.2 Visual reference Cues for Collaboration and Communication

The effect of simulation and visual reference cues in AR on communication and collaboration could be assessed with the Social Presence Questionnaire (SoPQ) [29] and the MEC Spatial Presence Questionnaire (SpPQ) [30]. The Social Presence Questionnaire assesses both attention and social presence with subscales of co-presence, attention allocation and perceived message understanding (Table 2) [29]. It evaluates whether attention is more focused on the partner in collaborative settings or on other items such as the simulation presented in AR. It can also be used to assess whether communication was easily facilitated, even using deictic references, using the Least Effort Principle for collaboration. The effectiveness of the visual reference cues used in AR is also indicated by this subscale.

The Spatial Simulation Model subscale of the Spatial Presence Questionnaire (Table 3) [30] can be used to assess the effect of simulation in AR. It analyses whether the user would be able to draw a mental image in the presence of the simulation and how well it is perceived. It uses a five-point Likert scale, and three versions of the scale can be constructed, depending on the desired length and depth of evaluation, as indicated in the last column of Table 3.

	Question	Subscale
1	I noticed (my partner).	CP
2	(My partner) noticed me.	CP
3	(My partner's) presence was obvious to me.	CP
4	My presence was obvious to (my partner).	CP
5	(My partner) caught my attention.	CP
6	I caught (my partner's) attention.	CP
7	I was easily distracted from (my partner) when other things were going on.	AA
8	(My partner) was easily distracted from me when other things were going on.	AA
9	I remained focused on (my partner) throughout our interaction.	AA
10	(My partner) remained focused on me throughout our interaction.	AA
11	(My partner) did not receive my full attention.	AA
12	I did not receive (my partner's) full attention.	AA
13	My thoughts were clear to (my partner).	PMU
14	(My partner's) thoughts were clear to me.	PMU
15	It was easy to understand (my partner).	PMU
16	(My partner) found it easy to understand me.	PMU
17	Understanding (my partner) was difficult.	PMU
18	(My partner) had difficulty understanding me.	PMU

Table 2: Social Presence Questionnaire (SoPQ), subscales Co-Presence (CP), Attentional Allocation (AA), and Perceived Message Understanding (PMU) [29]

	Question	Version
1	In my mind's eye, I was able to clearly see the arrangement of the objects presented/described	short
2	Even now, I could still draw a plan of the spatial environment in the presentation	short
3	I was able to make a good estimate of how far apart things were from each other.	medium
4	Even now, I could still find my way around the spatial environment in the presentation.	medium
5	I was able to imagine the arrangement of the spaces presented in the [medium] very well.	long
6	I had a precise idea of the spatial surroundings presented in the [medium].	long
7	I was able to make a good estimate of the size of the presented space.	long
8	Even now, I still have a concrete mental image of the spatial environment.	long

Table 3: MEC Spatial presence Questionnaire, subscale Spatial Situation Model [30]

To assess the effectiveness of visual cues, the Perceived Message Understanding subscale of the Social Presence Questionnaire can be used. Further details could be evaluated by measuring the amount of deictic references in comparison to the amount of direct references in the conversation. A high proportion of deictic references would indicate a simple production of the communication, whereas a high proportion of direct references, while visual reference cues are present, would indicate potential improvements in cue selection. Further information could be obtained by means of semi-structured interviews. These would ask whether communication could be produced easily and effectively, and whether the visual reference cues were helpful.

6.7.3 Spatial Perception of Objects

For the assessment of spatial perception of objects in AR, *paired comparison tests* or *perceptual matching tests* are most commonly used [97]. In a paired comparison test, two objects are placed side by side [14]. The user has to decide which of the objects appears more to the respective condition, in this case which object appears to be closer [14]. This combination is repeated for a given set of objects. The results can be presented in a matrix where the preferred object is shown as in Figure 24 [14]. The sum of a row is the score of how often that object was preferred, indicating which object was most preferred [14].

								Score
	x	7	11	13	12	11	8	62
	8	x	6	7	10	8	10	49
	4	9	x	9	8	7	6	43
	2	8	6	x	11	7	4	38
	3	5	7	4	x	6	6	31
	4	7	8	8	9	x	9	45
	7	5	9	11	9	6	x	47

Figure 24: Paired Comparison Test Matrix of depth perception for different coloured tea pots by [14].

In perceptual matching tests, users are also shown two objects, but instead of comparing depth, they have to place the objects in AR so that they appear to be the same depth [97]. This is usually done with a reference object placed in the environment and the user has to match one object [97].

6.7.4 Manipulation of digital Objects

The manipulation of digital objects for mobile AR is often evaluated in terms of the overall usability of the application [31]. The System Usability Scale (SUS) and the NASA TLX for cognitive load are often used for this purpose [31]. However, these do not take into account the specifics of manipulating digital objects in AR [31]. Santos et al. therefore developed the Handheld Augmented Reality Usability Scale (HARUS) [31]. It specifically evaluates manipulability and comprehensibility for mobile AR. Manipulability assesses the ease with which the application can be manipulated during use, identifying ergonomic issues such as *Hand interactions are difficult to perform*. Comprehensibility assesses the ease with which the content presented in AR can be understood, taking into account perceptual issues such as *The virtual objects are not well registered*. During the development of the questionnaire, further explanations are given for each question, providing further insight (Table 4, column Relevance).

	Question	Relevance
	Manipulability	
1	I think that interacting with this application requires a lot of body muscle effort.	HAR is often used while moving around the real environment.
2	I felt that using the application was comfortable for my arms and hands.	HAR strains the hands and arms the most.
3	HAR strains the hands and arms the most.	HAR has grip and pose issues.
4	I found it easy to input information through the application.	HAR introduces novel interaction metaphors.
5	I felt that my arm or hand became tired after using the application.	HAR strains the hands and arms the most.
6	I think the application is easy to control.	HAR introduces novel interaction metaphors.
7	HAR introduces novel interaction metaphors.	HAR has grip and pose issues.
8	I think the operation of this application is simple and uncomplicated.	HAR introduces novel interaction metaphors.
	Comprehensibility	
1	I think that interacting with this application requires a lot of mental effort.	HAR is susceptible to presenting too much information on a small screen.
2	I thought the amount of information displayed on screen was appropriate.	HAR introduces novel visualization metaphors.
3	I thought that the information displayed on screen was difficult to read.	HAR has legibility issues due to ambient light, glare, etc.
4	I felt that the information display was responding fast enough.	HAR has latency issues due to the limited processing power and network connection.
5	I thought that the information displayed on screen was confusing.	HAR introduces novel visualization metaphors.
6	I thought the words and symbols on screen were easy to read.	HAR has legibility issues due to ambient light, glare, etc.
7	I felt that the display was flickering too much.	HAR is susceptible to tracking and registration errors due to many factors, such as dynamics of lighting.
8	I thought that the information displayed on screen was consistent.	I thought that the information displayed on screen was consistent.

Table 4: Handheld Augmented Reality Usability Scale [31]

6.7.5 Analysing Intrusiveness

Several methods can be used to assess the perceived intrusiveness of an AR application. Smink et al. [109] use a seven-point Likert scale to indicate whether the use of the application was interfering, invasive, intrusive, unpleasant, and disturbing, thus adapting the approach used by Li et al. [123]. As general privacy concerns influence feelings of intrusiveness, further data can be analysed by assessing awareness of privacy practices. A five-point Likert scale was used to assess awareness in AR shopping apps [108] an adapted approach by Malhotra et al. [124]:

1. *Companies using this online try-on tool should disclose the way the personal information and images are collected, processed, and used.*

2. *A good consumer online privacy policy to accompany this online try-on tool should have a clear and conspicuous disclosure.*
3. *It is very important to me that I am aware and knowledgeable about how my personal information and images will be used.*

The general attitude towards AR can also influence intrusiveness and privacy concerns. Therefore, the *Augmented Reality Applications Attitude Scale (ARAAS)* could be helpful to assess general attitudes [125]. It assesses satisfaction, anxiety and willingness to use AR applications with a 15-item questionnaire [125].

6.7.6 Evaluation of Ergonomics

To evaluate the ergonomics of an AR application, the *Rapid Upper Limb Assessment* (RULA) can be used [12]. It evaluates the user's upper body posture based on the position of the upper and lower arm, wrist, neck, trunk and leg. Each component of the analysis is given a score from 1 to 4, with extra points for heavy weights or twisting of body parts as in Figure 25 [15]. The final score indicates whether the posture is acceptable for the application or requires further investigation, indicating possible changes, either soon or immediately. In the case of the AR application of Feey, the final score is 4, indicating that research is needed to improve the ergonomics of the application.

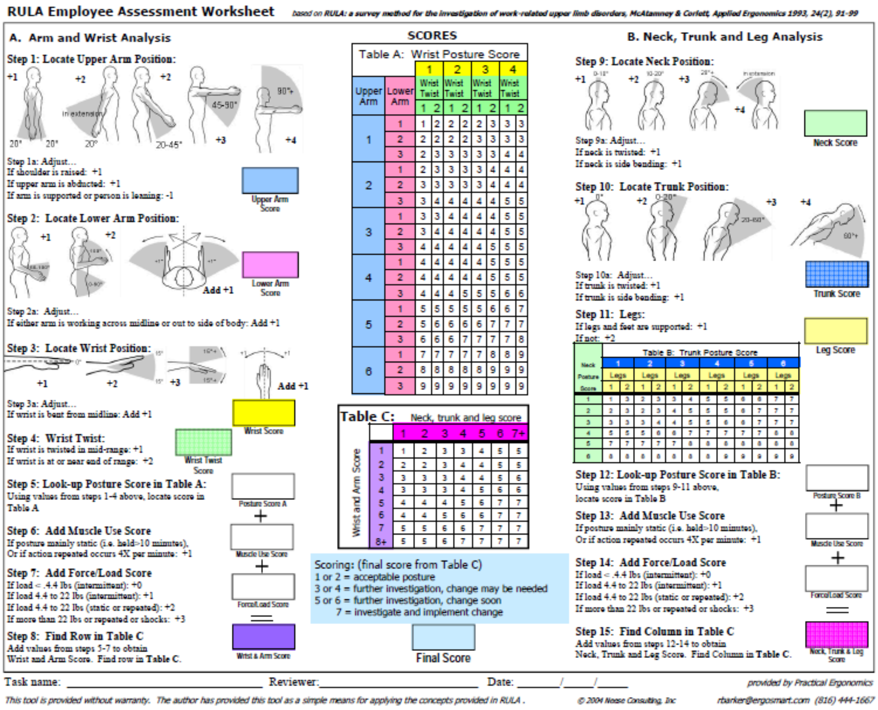


Figure 25: Rapid Upper Limb Assessment (RULA) [15]

6.7.7 Perceived Usefulness and Competence through the use of AR

To evaluate the overall use case scenario of feey the *Intrinsic Motivation Inventory* (IMI) could be useful to analyse the perceived usefulness and competence after using the application [126]. This could deliver insights as how useful the application is seen overall for the use case of purchasing a plant and if the feelings of competence for making a purchase decision could be increased. This could also be compared to a purchase decision made only with the traditional desktop application or an application supporting asynchronous AR support of an expert.

The IMI uses a seven-point likert scale to evaluate Interest/Enjoyment, Perceived Competence, Effort/Importance, Pressure/Tension, Perceived Choice, Value/Usefulness, and Relatedness. However, for the application of feey only the subscales of Perceived Competence and Usefulness may be applicable, which are shown below:

Perceived Competence
1. I think I am pretty good at this activity.
2. I think I did pretty well at this activity, compared to other students.
3. After working at this activity for awhile, I felt pretty competent.
4. I am satisfied with my performance at this task.
5. I was pretty skilled at this activity.
6. This was an activity that I couldn't do very well. (reversed)

Value/Usefulness
1. I believe this activity could be of some value to me.
2. I think that doing this activity is useful for
3. I think this is important to do because it can
4. I would be willing to do this again because it has some value to me.
5. I think doing this activity could help me to
6. I believe doing this activity could be beneficial to me.
7. I think this is an important activity.

6.8 Conclusion

AR has great potential in collaborative decision-making scenarios, as the analysis of the Feey example shows. Contextual sense-making is enabled by the simulation embedded in the environment. This can be helpful in supporting decision making, especially when combined with expert collaboration for further support. However, AR simulation is subject to spatial perception and there is a risk of misperception of objects in the environment, which is influenced by colour, fidelity or relative difference to the background.

However, simulation with AR has many positive effects, for example on communication, as in-situ visualisations make it easier and more efficient to create utterances with deictic references. This, in turn, requires visual references to enable deictic references. Another possibility to improve collaborating and communicating with AR is offered by asynchronous AR. This provides new opportunities for improving client advice from experts. However, this also carries the risk of increasing feelings of intrusiveness, as users' private surroundings are analysed. Every use of AR is subject to a cost-benefit analysis by the user, which assesses the risks and benefits of the use. The individual user's information processing style also influences this analysis. People process information differently, and a possible distinction can be made between visualisers and verbalisers, with verbalisers benefiting more from AR simulation. In addition, previous experience with AR is another factor influencing the cost-benefit analysis, as well as overall privacy awareness.

The potential for ergonomic issues arising from the use or manipulation of digital objects in AR is another factor to consider when designing AR experiences. Mobile AR is prone to poor posture, which can lead to pain and fatigue over time. In addition, manipulating digital objects through the screen of the mobile device can be tiring and require mental effort.

Figure 26: Example of feey and relevant characteristics for scenario analysed in this thesis

Figure 26 shows which elements can be evaluated based on the scenario of feey.

1. The *Handheld Augmented Reality Usability Scale* (HARUS) could be used to assess whether the manipulation of objects through mobile AR can be facilitated with ease. It also assesses whether the content presented through AR is easy to understand, taking into account perceptual issues.
2. To analyse the spatial perception of objects in AR, a *paired comparison test* or a *perceptual matching test* could be performed, depending on whether two objects are to be compared or one is to be manipulated in terms of depth perception.
3. For effective AR collaboration, visual references are needed. Their suitability can be assessed with the *Spatial Presence Questionnaire*, which analyses the effect of simulation in AR. The *Social Presence Questionnaire* also seems to be suitable, especially in combination with an analysis of the communication. It analyses the co-presence of collaborators, the allocation of attention and the perceived understanding of messages.
4. Augmented Reality in collaborative scenarios where private space needs to be shared can be perceived as intrusive. Whether the application is perceived as such could be assessed using a Likert scale to indicate whether the user found the application intrusive, invasive, intrusive, unpleasant and disturbing. Further analysis could be carried out by assessing the general level of concern about privacy and the awareness of privacy practices.
5. Simulation can be helpful for decision support, which can reduce uncertainty and cognitive load. In order to assess whether the application could achieve these goals, analysis of intrinsic, extrinsic and germane load could be useful. Further data could be obtained by analysing *search effort for product information*, *consideration set size*, *consideration set quality* and *decision quality* through the simulation. Further insights could be gained through semi-structured interviews to see if decision uncertainty could be reduced.
6. Mobile AR is prone to ergonomic problems, especially when a device has to be held for a long time. The *Rapid Upper Limb Assessment* (RULA) can assess a user's posture very quickly and provide results when further investigation is needed to improve ergonomics.
7. Cognitive processing style analysis could provide results on whether the simulation is perceived as engaging and effective by the intended user group. The information could also be useful in assessing whether the test sample is unbalanced, thus influencing other results.
8. Communication in collaborative scenarios follows the principle of *least effort*. Deictic cues can be used to support this principle, but require visual cues. To analyse whether communication could be facilitated easily, the amount of deictic versus direct references could be

compared. Also the subscale *Perceived Message Understanding* of the Social Presence Questionnaire could be useful to analyse communication through AR.

9. The overall usefulness and the perceived competence through the possibility of AR could be assessed with the *Intrinsic Motivation Inventory* after the use of the application. This could also be used in an A/B test of synchronous and asynchronous AR.

7 Subject 3: Spatial Augmented Reality

In this section, the example application for spatial augmented reality and its relevant characteristics are introduced on the basis of the term situatedness. The characteristics are discussed in terms of their theory and relevance to the scenario. Potential evaluation techniques are presented for each aspect, followed by an overall analysis of the scenario, its characteristics and potential evaluation techniques.

7.1 Example Application - interactive Books

Exhibitions often pursue hedonic qualities in technology which promotes discussions about the content in society, in ways other presentation forms couldn't [127]. Projection Mapping can promote those qualities as it breaks out of the traditional screen and shows information in a different form which can promote visual storytelling [128]. Such an installation are the interactive books from i-art [129]. They use projectors and sensors placed on the ceiling to show information on a physical book and allow for interaction with the digital content as in Figure 27.

Figure 27: Interactive Books from i-art

The books are usually placed in museums to provide an interactive experience that is novel to many visitors. The combination of physical integration opens up new possibilities and allows multi-user scenarios with different interaction possibilities than on traditional screens. The books are placed on special stands that track the user's touch gestures on the pages. The installation is connected to a beamer installed in the ceiling that projects the digital content onto the physical books. As the stand is fixed in one place, the beamer and its projection do not need to be dynamic. The user can explore new content by turning the pages of the book or interacting with the projection on the current page. Turning the page is a familiar way of interacting and does not require much explanation. However, combined with the digital projection on the books, it is a new way of visualising information for most people.

For the scenario presented, several points are relevant to consider for the user experience. These points are categorised based on the notion of situatedness. Figure 28 gives an overview of the characteristics that are relevant and have been selected for the analysis of the given scenario

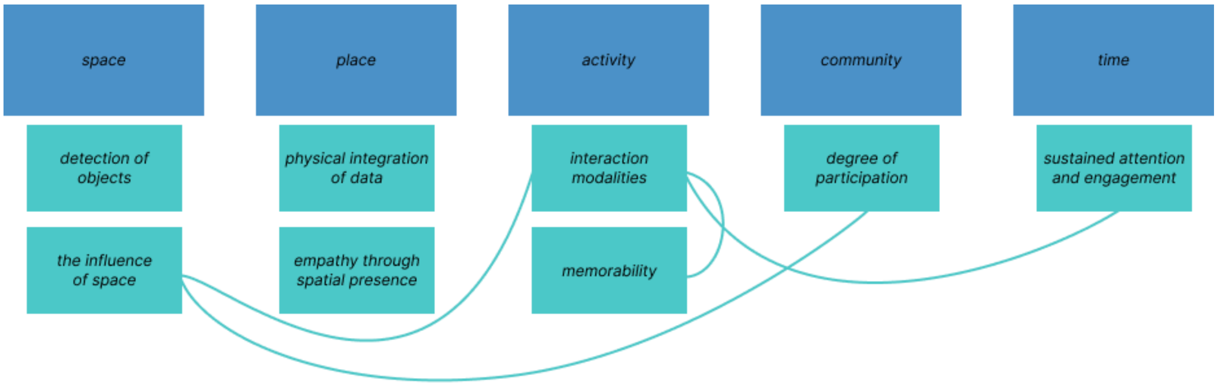


Figure 28: Overview of relevant and chosen characteristics for the scenario of the interactive books, categorized by the term situatedness

7.2 Space

This subsection introduces the relevant characteristics to the topic of space. Space encompasses the physical environment and its digital content. It refers to the spatial representation of information, including its properties and characteristics in three-dimensional space.

7.2.1 Detection of Objects

In Spatial Augmented Reality, physical objects are often used to project onto, which requires markers to be placed on the objects to be recognised by the system [16]. There are two categories of object recognition: marker-based and marker-less. Marker-based uses visual markers placed on the objects, which must be large enough to be detected [16]. There is also a requirement that the surface must be flat in order for the entire marker to be detected, which means that shaped objects usually do not work with marker-based solutions [16]. In addition, they are visible on the object, which can reduce immersion [16]. Infrared ink can also be used, making the marker invisible to the user, but the functionality is limited [16]. Markerless object recognition does not require visible markers, as recognition is based on shape, depth and edge detection [16]. Although the marker-less approach does not require visible markers on the objects, which is beneficial for immersion, objects that are symmetrical or periodically shaped cannot be successfully estimated by the system [16].

Tone et al. present a solution for active markers in dynamic projection mapping by implementing optical fibres in 3D printed objects [16]. The detection, even in scenarios with heavy occlusion, is very stable for several difficult objects, such as a rabbit that is uneven and curved, which is usually not supported by conventional marker techniques. Not only was the pose estimation stable, also the processing time by their algorithm was very fast as well were the fiber markers inside the objects not visibile to the human eye.

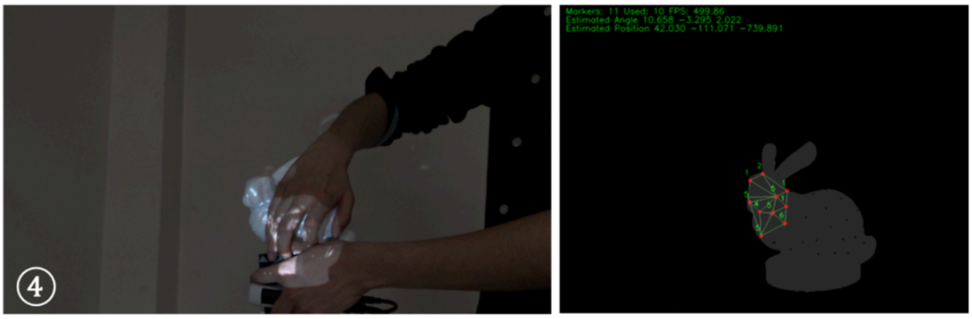


Figure 29: Projection Mapping and Token detection of occluded marker as 3D-printed bunny [16]

The interactive books are markerless objects whose shape and position are detected by the system. As they are permanently placed in the same position and the projector is also permanently installed in the ceiling, the books only need to be detected in one position. Touch is detected by sensors placed underneath the book, allowing for interactivity. In future scenarios, the pages of the book will have touch foils that can track users' touch interactions and allow for further interactions such as drawing or multi-touch.

7.2.2 The Influence of Space

Social interactions in public spaces are influenced by the physical layout of the environment [23]. Depending on the space, it can encourage or discourage social interaction [23]. The physical environment influences the subjective experience of people, while the technology present in the public space is an active part of that experience [23]. It is important to understand the relationship between the physical environment and the technology placed in the public space in order to facilitate social interactions and positive experiences [23]. Fischer et al. introduce the term *Shared Encounters*, ephemeral technology-enabled interactions between people in public spaces where co-presence is experienced [17]. The technology needs to be integrated into the physical environment. Fischer et al. divide the environment into different spaces around the technology in order to consider how the environment influences the technology and vice versa (Figure 30) [17]:

- ***Display Spaces*** *are the areas from which a display can be seen.*
- ***Interaction Space*** *is the space used to carry out a form of communication with the installation. It belongs to a single person, but can overlap with other person's interaction space.*
- ***Potential Interaction Spaces*** *(PIS) are spaces where the interaction between system and performer can potentially occur.*
- ***Gap Spaces*** *are spaces that create distance, either between human and system or among humans.*
- ***Social Interaction Spaces*** *(SIS) are those areas where people congregate, being attracted by the system, and have a Shared Encounter.*
- ***Comfort Spaces*** *provide a sense of physical and psychological ease. Protective features like walls, pillars, trees, etc. draw people subconsciously towards them.*
- ***Activation Spaces*** *are spaces where some displays can be seen from, often triggering curiosity, but interaction is not possible.*

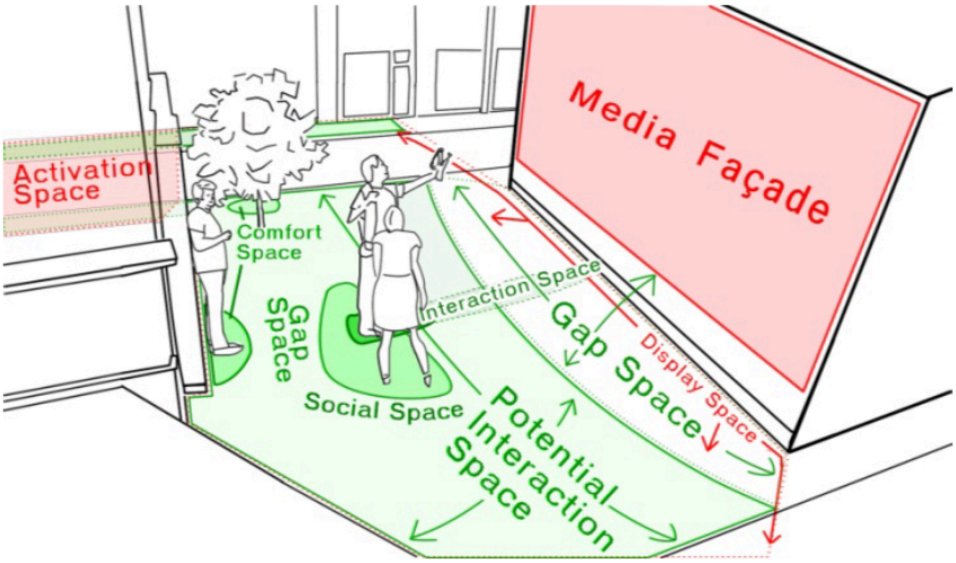


Figure 30: Visualisation of the different spaces identified by Fischer et al. for urban technology installations [17]

Distinguishing between different spaces in an environment highlights that the experience of technology installed in public places is not simply shaped by a given setting [17]. It is created by the relationship between the installation, the space in which it is placed, the architectural design and layout of the space, and the other people in the space [17]. This creates a complex model of influence that needs to be taken into account in public space installations for successful social interaction [17].

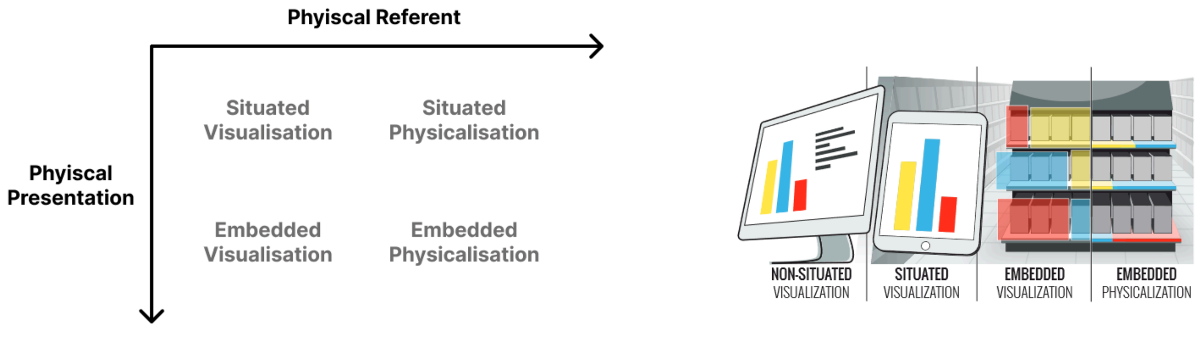
The interactive books are placed horizontally in public spaces, reducing the distance between the installation and the active users. However, the display space is still relatively large, as the books are installed elevated, allowing for comfortable interactions while standing. The interaction space is rather small and is located on one side of the installation, directly in front of it. The potential interaction space is almost the same as the interaction space, as it is only possible to interact directly when standing in front of the installation. The social interaction space is however slightly larger as the installation shows an animated projection even when it is not being actively used, allowing social interactions from a distance by observing the content. The comfort space is given by the physical layout of the environment, which changes in the installations of the interactive books. The activation space could also potentially be larger, depending on the physical layout of the space, as the books are elevated and can be seen from distance.

7.3 Place

This subsection introduces the relevant characteristics to the topic of place. Place is concerned with location, beyond a purely spatial perspective. It is the meaning of a space in relation to the people who inhabit it, the activities that take place within it and its historical context.

7.3.1 Physical Integration of Data

Spatial Augmented Reality makes it possible to situate data for visualisation, in contrast to traditional screen-based solutions, but the spatial relationship has to be taken into account in the development and design [130]. Placing data in the environment offers the potential for the physical object to act as an information-carrying medium as opposed to simply being an object [130]. Willett et al. present a model with two axes for positioning a visualisation depending on the *physical referent* and *physical presentation* of the visualisation [18]. The physical referent is the physical object or space to which the data relates, whereas the physical presentation is the object that makes the data visible.



(a) (b)

Figure 31: (a) Design Space of Physical Integration of Visualisations based on Willett et al. [18], (b) Examples of different characterisations [18]

The relationship between the two axes can be seen as the spatial proximity of the data referring to an object and where it is represented. Figure 31, (a) shows the space with the two axes of physical referent and physical presentation. In the design space the different characterisations, shown in Figure 31 (b), can be positioned. A low intensity in physical reference and presentation results in a *situated visualisation*, such as a visualisation on a phone close to the relevant object or space. A higher degree of physical reference leads to a *situated physicalisation*, where a physical object encodes the data. *Embedded visualisations* have a high degree of physical representation but a low degree of physical reference, usually represented as overlays, projections or other techniques that present the data close to the physical object. Finally, *embedded physicalisation* present data on physical objects that are closely related to the physical referent. It may also be the case that the physical reference is not explicit, but rather a semantic relationship, because the object collecting the data does not present the data in a meaningful way. The physical integration of data into the environment allows people to remain aware of the context and may perceive the data differently, providing additional insights that would not be possible with traditional media.

The interactive books can be placed in the middle of the *physical reference* axis of the design space in Figure 31 (a), as the information is neither purely visual nor purely physical. With the use of a physical book for the visualisation of information on it, the example shows the continuum of the axis. On the *physical representation* axis, the books are situated rather than embedded, and the relationship is also more semantic rather than actual.

7.3.2 Empathy through Spatial Presence

Spatial Augmented Reality can enhance spatial and social presence, which helps to create empathy [131]. It is therefore a suitable solution for education and awareness-raising in society on difficult issues, which would not be possible by other means [127]. Empathy involves seeing the world from a different perspective and responding emotionally to that different perspective [132]. SAR is able to transform an environment and thus to show it from a different point of view. It is shown that a higher level of spatial presence can be achieved with projection mapping, which influences attitudes and possible changes in behaviour [131].

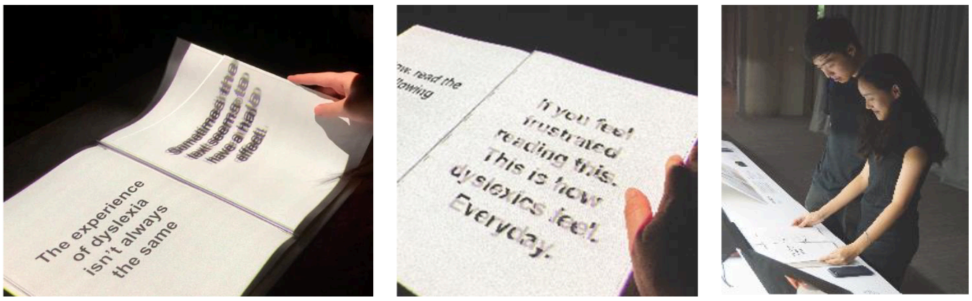


Figure 32: *Dyslexperience* installation to create empathy and awareness for dyslexia [19]

One such example is the exhibition *Dyslexperience* (Figure 32), which raises awareness of dyslexia through the experience of SAR [19]. It uses a physical book on which the content is projected in a special font that can be used to simulate dyslexia in order to create empathy and understanding. Another experiment uses projection mapping to support smoking cessation [131]. They compared traditional information about the negative effects of smoking with a projection of the consequences on the participants' bodies. The results showed that projection increased spatial presence, leading to a higher persuasion effect.

Interactive books are often used as part of exhibitions. One such example is 'Arabian Journeys', which tells the story of nature and culture in Arabia. These books are used to engage and immerse the user in the history of Arabia, to create an understanding of it and to educate them about its rich culture and heritage.

7.4 Activity

This subsection introduces the relevant characteristics to the topic of activity. Activities should not be viewed in isolation. They are components that contribute to broader goals and can be carried out in different spaces, over different periods of time and in collaboration with others. The relationship between a person and the goals he or she wishes to achieve is complex and influenced by the social and cultural context.

7.4.1 Interaction Modalities

Besancon et al. [133] distinguish several interaction modalities: 1) tactile and pen-based interaction, 2) tangible and haptic interaction, 3) mid-air gestural interaction, and 4) hybrid interaction. *Tactile and pen-based interaction* enables direct control via pen or touch-based interaction that is well known and well established. *Tangible and haptic interaction* is based on the use of physical objects for the synchronous manipulation of digital objects. In this way, it aims to make use of the natural abilities of human beings to interact with digital content. *Mid-air gestural interaction* also makes use of natural human interaction in the real world, either through wearable technologies or optical tracking. Optical tracking allows for interaction without any additional device, either through gaze, gesture or facial expression, captured by tracking the body. *Hybrid interaction* capabilities are usually developed to overcome the limitations of the above techniques by combining them, such as mid-air gesture interaction with tactile interaction.

For an opera in China, a novel way of mid-air interaction was developed [20]. To enable interaction with the background projection, it uses sensors hidden in the actors' costumes. This allows the actor to manipulate the digital content projected in the background without being visible to the audience. For example, the actor can light a lamp and move it with hand gestures as in Figure 33.

Figure 33: Interaction with background projection in opera with bluetooth [20]

In terms of tangible interaction, Walsh et al. present ephemeral tangible user interfaces for the rapid creation of controls based on the current context and affordances in the environment [134]. This allows the user to create disposable, temporary controls that are based on the objects in their environment. For example, in the cooking scenario, it is likely that a timer will be needed, but the designer cannot predict what tools will be available in the user's environment. For example, a kitchen utensil could be used as a timer that starts and shows remaining time as it is turned.

In the example of i-art's interactive books, tactile and pen-based interaction is used to interact with the content projected in the book. Tactile and haptic interaction is also used as users can turn the pages of the book. Mid-air gestures are not yet used, but would have great potential.

7.4.2 Memorability

Spatial Augmented Reality often uses physical objects for interaction. Since early childhood, humans have been interacting with and learning from physical objects [135]. Tangible User Interfaces (TUIs) can build upon these abilities, which can be beneficial for spatial memory, especially in the case of spatial data [135]. Additionally, TUIs not only improve spatial memory but also enhance creativity in spatial design tasks by positively affecting spatial cognition [136]. The manipulation of TUIs allows for the formation of new spatial information and relationships that are not observable in 2-dimensional interactions [136]. TUIs are also beneficial for recalling spatial relationships between objects and information associated with the object [135], [137].

The interactive books allow physical interaction through an object with the digital projection, therefore they are defined as TUIs. As they are mainly used in exhibitions to present information to the public, they are especially suited to convey information to enhance memorability of the information. The books can be further accompanied with a physical pen to draw digital content into the book. The pen is applied in the exhibition of *Picasso*, where users can draw the outlines of a projected painting with the pen. This allows for reproduction of the characteristics of a painting from picasso, which would also allow for the possibility to print the own drawing to take it home.

7.5 Community

This subsection introduces the relevant characteristics to the topic of community. Community refers to individuals in relation to a particular place. The focus is on how a specific community utilizes technology or a physical location to achieve their objectives. It is closely tied to the notion of place, which derives its meaning over time from the activities of the community.

7.5.1 Degree of Participation

Spatial Augmented Reality has the ability to involve multiple people to varying degrees. As the projection is usually on a larger scale, it is not only visible to the person actively interacting with the technology [138], [139]. Bystanders or observers can also see the projection as well as the interaction [138]. The passive presence of some people also creates an opportunity for social learning by observing how others interact with their environment or specifically with a technology [23]. SAR should not only be designed with a focus on an enjoyable user experience for active participation, it should also be an invitation to participation for bystanders [23]. By observing others, people also create expectations about the activities of the active users [23]. These explicit or implicit social interactions can be seen as influences that shape the experience [23].

Not only can SAR be used actively or passively, there are also times when a system is not in use. This should not be forgotten and should be seen as inviting the user to move from passively observing to becoming an active participant [23]. Potential factors to motivate people to engage could be curiosity, appeal of the content or technology, challenge, or collaboration [138].

Figure 34: Interactive books in the *Arabian Journey* exhibition allowing for active participation by several people

As i-art's interactive books are quite large, this allows for active participation by several people, as in Figure 34. In particular, the interaction through touch allows for interaction in a group and also creates a situation in which social learning can take place. The mere presence of the books also enables social learning for bystanders and passive users, depending on the spatial layout of the environment. They also consider the state of not being used by a person in the installations where the interactive books are used. In this state, they show an animated projection that invites the user to actively participate by creating curiosity.

7.6 Time

This subsection introduces the relevant characteristics to the topic of time. The concept of time concerns the recording of data for a visualization and its display, as well as the relationship between these two moments. Additionally, time can be considered beyond pure objectivity or subjectivity and instead be viewed as socially constructed, defined as Social Time. Further, it also encompasses temporal relevance which considers social norms, activities and conventions regarding to time.

7.6.1 Sustained Attention and Engagement

Maintaining users' attention and engagement is important for extended interaction with an object or installation [21]. Hong et al. have created a framework in which the experience of a museum visit is divided into a pre-visit phase, a during-visit phase and a post-visit phase [21]. In the during-visit phase, attention, engagement and immersion play a crucial role in determining whether and for how long a person interacts with an installation. In a first step, the installation must successfully attract the user's attention. The user is more likely to explore the installation if his/her interests and expectations are met. Not only are interests and expectations important in attracting attention, but also the attractiveness of the installation can be crucial. Once the user starts interacting with an installation, they need to understand very quickly how to interact with it and what it is trying to communicate. Otherwise, the user will stop interacting with it. To increase the time spent interacting with an exhibit, the level of immersion must be high. The level of difficulty of the content also has an influence on the engagement, as it should neither be too difficult nor too easy.

Figure 35: Interactive books at the Landesmuseum Zürich, showing the digital content with common signifiers for interaction

It is likely that the interactive books attract the attention of users in a museum as the projection on a physical book is very uncommon and novel for most people. Therefore, it could be assumed that the attractiveness of the installation is high but no estimation can be made about the interest and expectation of the users. Once the user starts to interact with the books as it successfully attracted attention, it has to be communicated on how to interact with the books. As most people know the interaction from physical books, the page turing interaction should not pose a problem. However, it has to be conveyed that the books are interactive and touch-gestures are possible to derive further information about the content presented. They use common signifiers for interaction which are mostly known from the use of mobile devices as in Figure 35. To sustain engagement with the books, the immersion should be high which also depends on the content and context.

7.7 Potential Evaluation of the User Experience

This section presents options for the evaluation of the identified characteristics that are relevant to the example. They can either be assessed in isolation, or they need to be analysed in combination with other aspects, as they are closely linked and interdependent.

7.7.1 Consider the different degrees of Participation

As SAR provides the opportunity for different levels of participation, these should be considered in the evaluation of the user experience. From this point of view, it is also understandable that not only the number of active interactions should be taken into account in the evaluation for the analysis of the success of the installation, as bystanders do not actively interact, but still have an experience of the installation [23].

Wouters et al. conducted contextual enquiries to analyse the installation in a public space [23]. They observed the situation over several days and conducted further analysis with active and passive participants through semi-structured interviews. This would provide further information about the motivation for participation, the social learning of passive participants and the social interactions of active participants. In addition, they analysed quantitative data collected from the system, including number of active participants, length of active participation, and configuration of people who interacted with the installation.

7.7.2 Physical Integration

The different levels of physical integration of the data make it possible to perceive it differently than on traditional screens. However, it is not clear how the different levels of integration change the perception of the users [18]. In order to analyse the perception, but also the different interpretation of the data, the think-aloud method, or also called *protocol analysis*, can be useful [136]. This could be used to analyse the underlying cognitive processes and perceptions of different integrations, either by comparing two different integrations or by evaluating a given integration of data. It would be important to have a clear definition of the objectives of the visualisation and of what it is intended to convey. The results of the evaluation would allow discussion of whether the intensities in the physical reference or presentation are appropriate and whether the reference is explicit or more semantic with less spatial proximity.

7.7.3 Empathy

To evaluate sympathy and empathy induced through media the *Ad Response Sympathy* (ARS) and *Ad Response Empathy* (ARE) questionnaire [32] can be used and adapted as in [140]. Both instruments were developed the evaluate sympathy and empathy in advertisement videos, however they argue that it can be used for several types of media [140]. Both have five questions which are evaluated with a seven-point likert scale as in Table 5.

	ARS	**ARE**
1	Based on what was happening in the commercial, I understood what the characters were feeling.	While watching the ad, I experienced feeling as if the events were really happening to me.
2	Based on what was happening in the commercial, I understood what was bothering the characters.	While watching the ad, I felt as though I were one of the characters.
3	While watching the ad, I tried to understand the events as they occurred.	While watching the ad, I felt as though the events in the ad were happening to me.
4	While watching the ad, I tried to understand the characters' motivation.	While watching the commercial, I experienced many of the same feelings that the characters portrayed.
5	I was able to recognize the problems that the characters in the ad had.	While watching the commercial, I felt as if the characters' feelings were my own.

Table 5: *Ad Response Sympathy* (ARS) and *Ad Response Empathy* (ARE) questionnaire [32]

7.7.4 Memorability

Memories can be semantic, general facts and knowledge about a particular event, or they can be episodic, factual memories related to the personal experience of [141]. The episodic memories are important in predicting whether a visit is perceived as positive and whether there is a desired for future visits [141]. The most memorable events are those that are distinctive, stand out and are social in nature [141].

Mai et al. showed that emotions are crucial when it comes to memorability [142]. They evaluated the memorability of TV commercials on the basis of the emotions they evoke. The results show that positive emotions and comprehensibility play a crucial role in memorability, with advertisements that evoke the most positive emotions and are easy to understand being the most memorable.

An appropriate questionnaire to assess emotions and comprehensibility is the *User Experience Questionnaire* (UEQ) [143]. It has six subscales to assess 1) attractiveness, the overall impression of the product, 2) efficiency, whether the product is efficient and easy to use, 3) perspicuity, whether the product is easy to understand and learn, 4) dependability, whether the user is in control of the interaction, 5) stimulation, whether it is interesting and exciting, and 6) novelty, whether the product is creative and innovative.

7.7.5 Interaction Modalities

The interaction with the technology should not require a high level of mental and physical effort. In addition, it should be easy to use, quick to learn, and appropriate for the given task and the given technology. In order to evaluate the usefulness and the usability of the interaction method, the first subscale of the questionnaire *Modular Evaluation of key Components of User Experience* (meCUE) can be used [144]. This questionnaire uses a seven-point Likert scale for the evaluation of a product. The questionnaire has further modules for perception of non-instrumental qualities, user emotions, consequences of use and overall evaluation.

7.7.6 Object Detection

From a user perspective, SAR needs to be stable and fast in detecting objects. The projection mapping should not have any offset and it should be able to work even in the case of occlusion. Depending on the scenario, it should also support dynamic use of the object, such as translation or movement. The *User Experience Questionnaire* (UEQ) is one way of assessing how object detection and projection mapping are perceived, as it assesses efficiency, perspicuity and dependability [143].

Most studies on new marker detection methods do not evaluate the user experience and focus mainly on projecting accuracy and rendering speed. However, the presence of relatively large markers on objects could affect the usability as well as the accuracy of detecting the object and the degrees of freedom allowed to interact with the object.

7.7.7 Analysing the Space in Installations

In order to analyse the different spaces that an installation has depending on the physical environment, a contextual enquiry can be carried out, such as by Fischer et al. [17]. Observations could be filmed for further analysis of how people interact with each other and how spaces are distributed around the installations. It is also possible to observe whether shared encounters occur or are even possible given the physical layout around the installation.

For a more quantitative analysis, they propose the methods of coming from advertisement to calculate the so-called *G-value*, which depends on the angle of view, distance, competing nearby stimuli and the number of passers-by [17]. The value gives an estimate of how many people per hour will remember the advertisement on a billboard in a public space.

7.7.8 Sustained Attention and Engagement

One way to assess how engagement and attention are sustained is with the *Augmented Reality Immersion* (ARI) questionnaire [89]. It takes into account the user's engagement, involvement and immersion in AR. Another suitable solution for measuring attention is eye-tracking [145] or even eye-pupil measurement for assessing attention vigilance [146].

Further quantitative data could also be collected with the *User Engagement Self-Report Questionnaire* (VisEngage) [147]. They identified 11 characteristics of engagement that had the highest frequency of occurrence in the literature and that were the most relevant to engagement: Aesthetics, Captivation, Challenge, Control, Discovery, Exploration, Creativity, Attention, Interest, Novelty, and Autotelism.

For qualitative evaluation, contextual enquiries for observation and interviews for deeper understanding are appropriate and common tools [21].

7.8 Conclusion

The analysis of aspects relevant to Spatial Augmented Reality shows that space is used differently in SAR than in MAR or HMD. As projection mapping installations are placed in the environment, the space, community and interaction are different and require further consideration. SAR also allows for social interactions and learning to take place through the installations, provided that the space allows for these interactions. It is also possible with SAR to integrate data into the physical space, transforming a simple object into an information-carrying medium. In this way, additional insights can be gained and data can be perceived in a different way. Tangible interaction has also been shown to increase the memorability of information and evoke emotions. As SAR is a new technology for most people, it is highly likely that people will be drawn to it to explore its functionality. Once the user's attention is captured, it can be maintained through high levels of engagement and immersion.

Unlike other AR technologies, SAR not only allows interaction as an active user, but passive bystanders can also be considered users in SAR. They can observe active users and learn to become active users in the next phase. This is an important difference and must be taken into account in the design of the installation as well as in the design of the space around the installation. As it is only when the space allows for social learning that passive users can see active users in order to learn and be attracted to them. The interaction modalities can also differ from traditional technologies, as users can interact not only with touch gestures, but also with tangible objects, mid-air gestures or with a pen. This makes it a versatile technology that breaks out of the traditional screen and can be used in a variety of contexts to convey information that is not possible with other technologies. For this reason, it can also be particularly useful for evoking feelings of empathy, which can be stronger with greater social and spatial presence. As a result, it can influence attitudes and possible changes in behaviour. To avoid disrupting the immersion provided by SAR, the object detection needs to be facilitated accordingly. Either with appropriate markers that blend into the physical space, or without markers, which presents other difficulties in detection. However, the right technology for object detection must be chosen by considering the space in which the objects are placed.

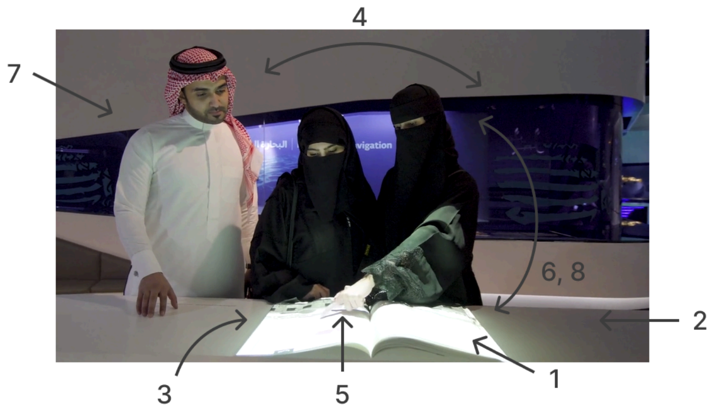


Figure 36: Example of interactive books and relevant characteristics for scenario analysed in this thesis

Figure 36 shows which elements can be evaluated based on the scenario of the interactive books.

1. To enable projection mapping, objects must be detected by the system, either with or without markers. Object recognition can be evaluated from a technical point of view, but also from a user experience point of view, whether it is perceived as efficient, reliable and understandable. This could be assessed using the *User Experience Questionnaire*.
2. Installations in public spaces have several spaces with different functionalities. In order to analyse whether shared encounters are possible and how people interact with each other through the different spaces, contextual enquiries may be suitable. Quantitative data can be derived by calculating the G-value, which gives an estimate of how many people per hour will remember an installation.
3. Spatial Augmented Reality allows data to be integrated into the physical environment with different intensities depending on the physical reference and presentation. The perception of this integration could be assessed using the think-aloud method to understand the underlying cognitive processes and perceptions.
4. Spatial and social presence can support the creation of empathy, suggesting that spatial augmented reality is particularly useful for this application. To assess sympathy and empathy, the *Ad Response Sympathy* and *Ad Response Empathy* questionnaires could be used and adapted.
5. Different modalities can be used to interact with AR. To evaluate their usefulness and usability, the *module evaluation of key components of user experience* can be used.
6. Memorability is strongly influenced by the emotion and tangibility of the interaction. This could be assessed using the *User Experience Questionnaire*, which analyses several factors relevant to memorability.
7. Spatial Augmented Reality installations allow for different levels of participation, from active interaction to passive bystanders or observers. In order to analyse whether the installation allows for social interactions and learning by supporting different forms of interaction, contextual enquiries may be appropriate.
8. Attention and Engagement are important to extend the interaction time with an installation. This could be assessed using the *Augmented Reality Immersion* questionnaire or the *User Engagement Self-Report Questionnaire* (VisEnagage).

8 Discussion of Literature Research

Spatial computing differs from traditional screen-based technologies such as web or mobile applications. The content presented through AR is placed in the user's real environment. It should therefore take into account the space in which the user is placed as well as the user's context. Each technology needs to recognise the space differently because the content is presented and placed differently. The modality of interaction with either the technology or the content is also different and requires its own consideration. However, what becomes visible are the overarching similarities common to all spatial computing technologies: artificially created artefacts that blend into the context, taking it into account, with which people interact in their natural contextual environment.

The aspects relevant to spatial computing have different characteristics depending on the technology, context and scenario. This requires a thorough analysis of the context and its facets, which can be broken down to the definition of situatedness: space, place, activity, community and time.

The use of space differs significantly between technologies. SAR uses space in a broader sense than HMD or MAR. The user is usually not interacting independently, but with other people, either actively or passively. In HMD or MAR, social interaction can also take place, but either through the device to see the same content, or in the real environment, where usually only one person can either see or interact with the digital content. However, the space has an impact on perception in all three areas and shapes the user's perception of the place.

We associate different norms and values with spaces, depending on the context. Not only does the space influence how we perceive a place, but the community inhabiting the space and their interaction with it shapes our experience, and our interaction with AR shapes their experience of the space, either actively or passively.

The analysis shows that spatial computing needs a multidisciplinary approach, as is common in HCI, to realise its full potential. This is because many aspects from different research areas influence the experience of spatial computing. It also shows that in order to evaluate the novel interaction, it is necessary to use evaluation tools from other fields. However, it is questionable whether all tools are applicable to different spatial computing applications. There are still many unanswered questions about user experience and its evaluation in the current state of the art. This is also a reflection of the fact that the field is not as mature as virtual reality or traditional web or mobile development.

The shortcomings of the current state of the art are also reflected in the current applications of spatial computing and how to evaluate them. There are many aspects that are relevant to development, but most development teams have limited time and are unable to conduct such extensive research. Consequently, they look at current applications and evaluation methods of others. Thus they end up using traditional and simple methods. However, what becomes apparent is that methods and considerations do exist, but they are not so well established and therefore require a lot of research. This leads to the question:

How can additional considerations and methods relevant to spatial computing become established within this field to develop better experiences, leading to a more holistic evaluation?

9 HCI Models in Augmented Reality

There are various models in HCI regarding interaction or perception. However, many of these models have shortcomings, either focusing on a few aspects in the situation or neglecting the complexity of the interaction. This chapter introduces the theoretical basis of complexity science and modelling in HCI. Based on the literature research some models are analysed and which aspects they cover of the term *situatedness*. Lastly, the development of a model showing the situational influence in spatial computing is outlined.

9.1 Complexity of Interaction

The interaction of a human with technology is dynamic, the behaviour results from the interaction of these two elements, shaping the experience [33]. However, it is not only the technology and the human that drive the behaviour, the environment and other humans also influence the interaction, making it a complex system of behaviour with multiple elements [33]. The elements influence each other through feedback. As the user has a goal when interacting with the technology, the interaction changes the state of the system, which affects the user's behaviour, making it an endless cycle of changing elements [33]. There is still a lack of understanding of the dynamics of interactions in HCI, as the use of complexity science in combination with HCI is still quite new [33]. Interactions evolve and adapt over time depending on the context [33]. This viewpoint moves away from the older view of understanding an interaction as a linear process of human intentions through a computer and more towards an agency-based approach that emerges from the context [33].

9.2 Modelling of Complex Systems

When analysing a situation, people have mental models that illustrate the causal attributes of different elements. However, many of these mental models are flawed because few take into account the feedback loops of the elements and think in terms of single strand causal chains rather than complex systems with feedback loops [148]. This way of thinking can lead to ignorance of side effects, as it assumes that actions have only one effect, rather than creating a series of effects that change many elements in the system [148]. Not only are feedback loops ignored, time delays of consequences are often also not considered [148]. Causality is often searched with temporal or spatial proximity [148]. In complex systems, however, cause and effect are often distant in time and space [148].

A situation has many elements influencing each other, creating a complex interdependent system. Modelling can be a useful tool for understanding the influences and interactions of the different elements. The aim is to illustrate complex processes in order to show the dynamics of the elements and to lead to an effective policy-making [148]. Successful models include not only technical components but also political, social and environmental impacts [148]. This shows the complex system and its dynamics emerging from the interaction of the different components in the model [148].

Sterman states that a complex system is made up of several elements, and that in order to build a model of a complex world, different dynamics must be taken into account [148]. For example, systems are *constantly changing*. Systems always evolve and change, although the scale of change may vary. Also, the elements in a system are *tightly coupled*. Actors within a model interact strongly with each other and their environment. Also, often forgotten in dynamic systems, actions have *feedback* that must be considered. The decision taken changes the some elements, causing changes in other elements of the system, thus influencing again other elements in the system.

When building a model, boundaries are important. There are many influences to consider and the number of external influences should be as small as possible to understand the system and not rely on external elements [148]. External influences should also not be omitted as it is then expected that they have no influence on the system [148]. The built model needs to be evaluated with different sources of data from observations, theory and experts [148]. In this way it may be possible to illustrate complex processes and understand their dynamics for successful interventions [148].

9.3 Development Process of a Model in HCI

There are two different starting points for the development of a model, either the observation of a phenomenon in the real world or the formulation of a hypothesis about a phenomenon [3]. The first is called the *inductive* method where researchers observe people in the environment without any preconceived idea, and try to describe their behaviour later to build a model based on their observations [3]. The second approach is called the *deductive* method where researchers formulate hypotheses in order to explain phenomena in the real world [3]. They then conduct an experiment in a controlled environment to test their hypotheses. The results are analysed and the hypotheses are adapted to the findings in order to adjust the model originally created [3]. The deductive method strives for reliability, where the same conditions produce the same results, and validity, where the results can be generalised given the conditions in the controlled setting [3].

Both approaches attempt to model phenomena in the real world, creating cycles of theory and observation to iteratively refine the model created [3]. Since HCI is neither a pure social science nor a pure natural science, both approaches have their validity and may need to be combined, the inductive approach with more quantitative results and the deductive approach with more qualitative results. However, both methods are based on the same assumptions:

- *Natural phenomena exist and can be isolated for study.*
- *Observers are unbiased.*
- *Repeated observations under the same conditions will yield the same results.*
- *Conclusions drawn from observations in one setting can be generalized to other settings.* [3]

The combination of the two approaches may be useful in HCI, as it is not a natural phenomenon that is being studied, but the interaction of human beings with an artificially created artefact [3]. Because people can adapt, the problem is not static but rather dynamic, with people adapting to the use of technology and changing their behaviour over time in response to environmental, social and technological influences [3]. Mackay therefore proposes a combined approach to the creation of models in HCI, in which researchers formulate hypotheses about a behaviour and also make observations in the real world that influence the model [3].

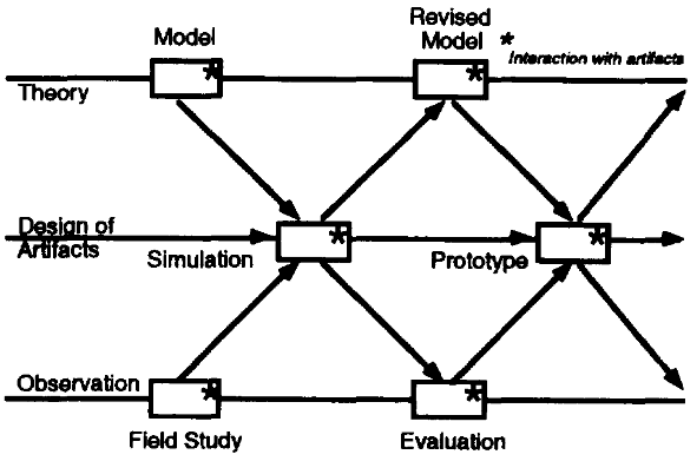


Figure 37: *HCI creates and examines the interaction of people with artificially-created artifacts, moving between theory and empirical observation. The boxes represent the interaction with artificially-created artifacts, rather than independent natural phenomena or design artifacts.*[3]

Figure 37 shows a simplified approach on how to create a model in HCI through a combination of Theory and Observation [3]. However, such an approach is not only common in HCI, it is also being used in several domains of research as the the combination of theory and observation is very useful. Further can the two inputs not be viewed in isolation as when conducting observations, usually hypotheses exist based on theoretical knowledge.

9.4 Models found in Literature

Based in the prior research it becomes visible that elements which influence an experience in spatial computing are interrelated influence from different factors. Several researchers came up with models to show the influence and relationships of factors in technology. Relevant and interesting models from the literature research are being presented hereafter and analysed.

9.4.1 Visitor interactive experience model

Hong et al. follow an inductive approach to form their model of visitors experience in an interactive science museum as in Figure 38 [21].

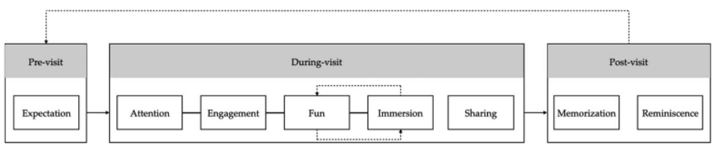


Figure 38: *Visitor interactive experience model, indicating how a visitor interacts within an interactive museum/exhibition context.* [21]

Based on three different stages: pre-visit, during-visit and post-visit, their model illustrates the temporal flow of interaction with an installation. The relevant emotions in each phase are shown, as well as their relationship to each other. For example, in the *during-visit* phase, the emotions for *fun* and *immersion* are related. The model also shows the dependencies of each phase, if the user stops in phase, the whole interaction stops and the journey starts with the next installation.

9.4.2 The Interaction-Attention Continuum

Bakker et al. follow a deductive approach to construct the *interaction-attention continuum* as in the Figure 39 [22]. The pose the hypothesis that *interactions with interactive systems should* *be available at various levels of attention.* [22]. Their model shows the continuum from *focused interaction*, where no other activity can be performed, to *peripheral interaction*, where information can be perceived and interacted with from the periphery, to *implicit interaction*, where no explicit interaction is performed and it relies on the sensor to detect the human and their activity.

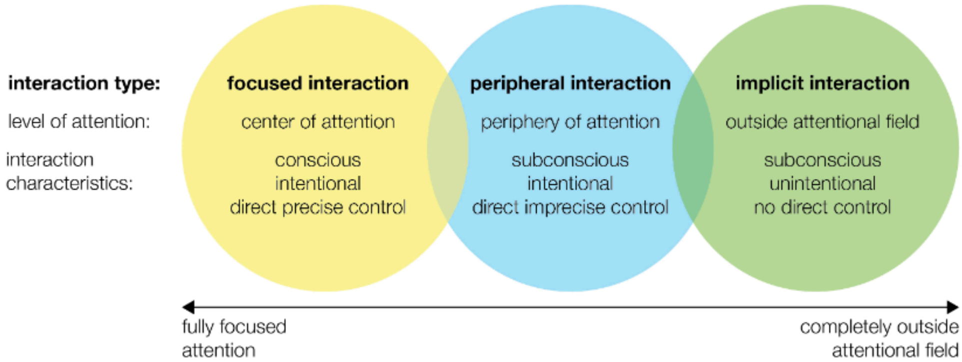


Figure 39: *The interaction-attention continuum.* [22]

9.4.3 Honeypot Model - Engagement with Interactive Systems

Wouters et al. have created a model through an inductive approach to illustrate the so-called honeypot effect, *a social learning influence that causes individuals to be affected by the mere and passive presence or activities of others, regardless of any competition, reward or punishment* [23]. This effect has already been studied in HCI, especially the influence of spatial configuration, but not the drivers and motivations for engaging with an interactive system, especially in public spaces.

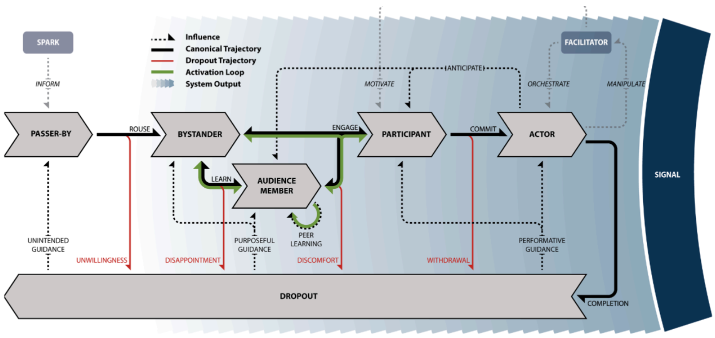


Figure 40: *Honeypot Model containing the user roles, trajectories, influences and triggers that affect how audiences engage with interactive systems.* [23]

Figure 40 presents the honeypot model and user roles, trajectories, influences and triggers for interactive systems. It considers different forms of user roles, from bystander to actor, and the trajectories from this passive to active role.

9.4.4 Value of Public Interactive Displays

There are several factors that influence whether a public interactive display is perceived as useful. Parker et al. used a deductive approach to analyse which factors influence the perceived value of such displays [24]. They conducted a literature review of several public interactive displays and analysed the different aspects. They identified four aspects that contribute to perceived value: People, Location, Community and Time.

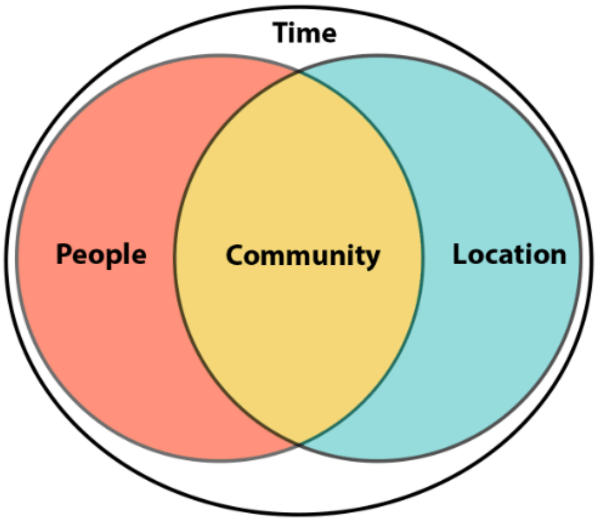


Figure 41: *The key factors found in our analysis: People, location, community, and time.* [24]

The model, as in Figure 41, shows that each of the factors is related to the others and is influenced by the overarching factor of time. The *people* factor describes the individuals or groups interacting with an interactive public display, which is further defined by five aspects: Trust, Privacy, Relevance, Appropriation and Ownership. Location describes the composition of the physical location and Community the collection and type of users found in a particular space.

9.4.5 Perception-Action Cycle

People interact with their environment and make sense of it based on their prior knowledge. This dynamic cycle is captured in Neisser's *Perception-Action Cycle* in Figure 42 [25], produced by a deductive approach. The perception of the environment and its elements is influenced by existing knowledge. The model shows how awareness of a situation is maintained in a dynamic environment, also known as situational awareness. It also illustrates that visual perception is a continuous activity.

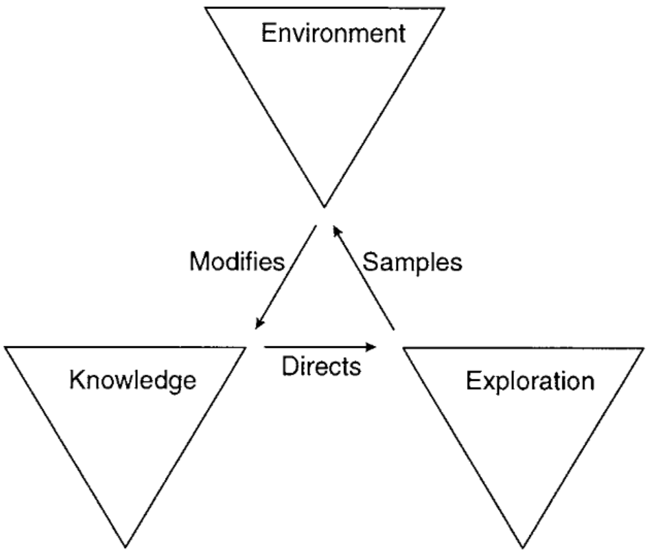


Figure 42: *The perception-action cycle.* [25]

9.4.6 Determinants of Usability

Already in 1991, Bevan et al. stated that usability is not a static quality, but rather a quality of a product for a user that depends on the task being performed in a particular environment [26]. This notion emphasises that technology should be analysed through a socio-technical framework, taking into account the interaction between people, the environment and their goals.

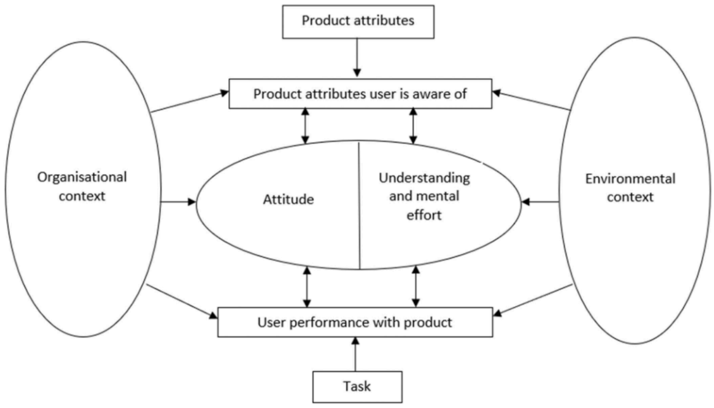


Figure 43: *Multiple determinants of technology usability.* [26]

They created a model that illustrates the determinants of usability through a deductive approach as in Figure 43. The model shows that the environment, task, product and interaction influence the usability of a product.

9.4.7 Semiotics of Space

The semiotic triangle, Figure 44, was developed by Määttanen [10] and is based on Peirce's idea through a deductive approach. They have adapted Peirce's notion for the analysis of a concrete interaction of a living organism with its environment. The model illustrates how spatial meaning is formed through perception, action and interpretation.

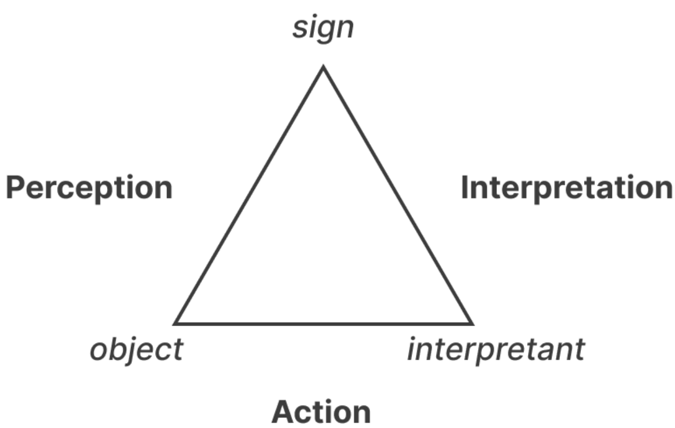


Figure 44: *The semiotic triangle* [10]

Semiotics studies how people interact with their environment by analysing how they perceive signs and objects in space, which require interpretation by the interpreter, leading to actions which form the chain of interpretation.

9.4.8 Conclusion of Models

The presented models all stem from HCI, analysing the interaction of user with technology from different view points or with different aspects. What becomes visible is that based on the term of situatedness, every model only considers some aspects and not the relation of all five factors. Table 6 gives an overview of the presented models and which aspects from the term situatedness they consider.

aspects	**space**	**place**	**activity**	**community**	**time**
Visitor interactive experience model			x		x
The interaction-attention continuum			x		
Honeypot Model			x	x	x
Value of Public Interactive Displays	x			x	x
Perception-Action Cycle	x	x		x	x
Determinants of Usability	x		x	x	
The semiotic triangle	x	x	x		

Table 6: the different models and the aspects considered based on the term situatedness

9.5 Development of a Model

As the analysis of the existing models in the literature shows, most of them have shortcomings and do not consider the situation in its full complexity, which is a simplification of the dynamics in the system. This thesis aims to model the system dynamics of how the situation shapes the experience in spatial computing in order to illustrate the influence on the experience and to highlight the importance of the context. The aim is to support developers and designers in the consideration of environment and situation for the development of outstanding user experiences in spatial computing. The model should help to analyse issues in the user experience and detect the root cause with it as aspects are tightly coupled and interrelated. Identified issues have therefore often many potential sources and dependencies.

9.5.1 Identification of Elements in the System

To develop a successful model which illustrates the system dynamics, the various elements compromising the situation need to be identified. For each scenario which was analysed in the previous part of this thesis, a schema has been developed to illustrate the characteristics and the relationship.

Figure 45 shows the schema for the scenario of Arigo, using a HMD to support users in procedural tasks. In this scenario, the projection is above the real object, relevant for the task. The user wearing the HMD perceives the projection and the real world deepening on the context in which he/she is placed.

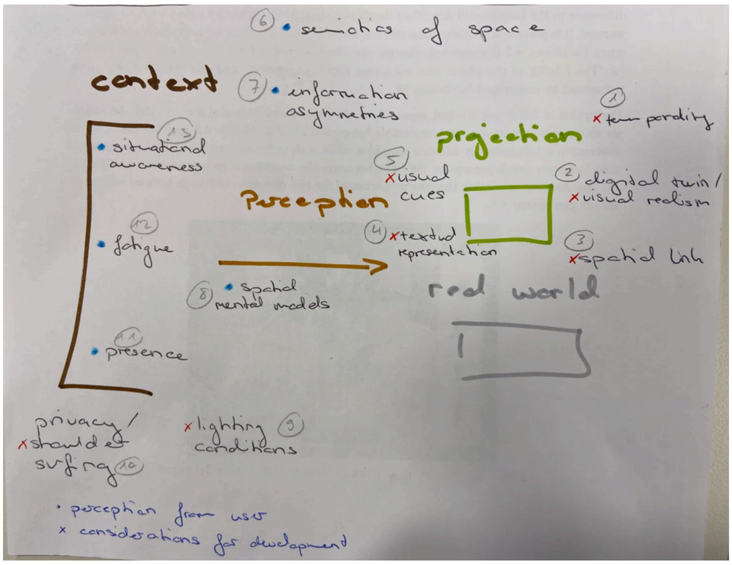


Figure 45: Schema for HMD

Figure 46 shows the schema for the scenario of feey, using MAR to collaborate with an expert to find suitable plants simulating their appearance through MAR. In this scenario, there are two actors present interacting with each other remotely through a video call. The user places the digital plant through the mobile device in the real environment. The perception of the augmented world happens through the mobile device of the user, as well as the remotely participating expert.

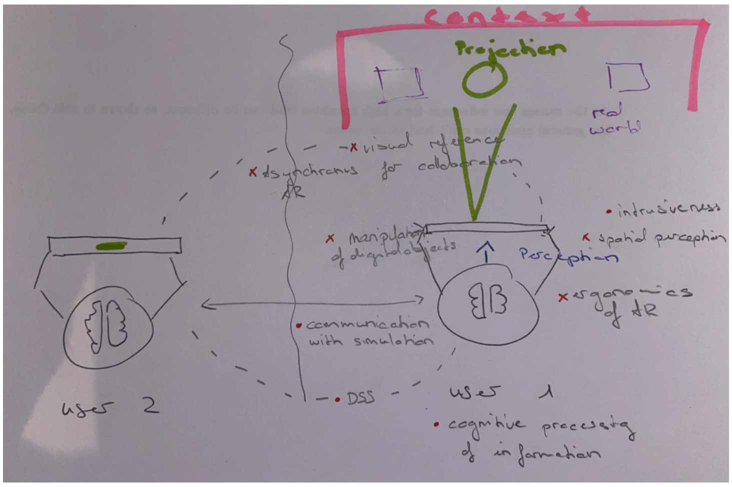


Figure 46: Schema for MAR

Figure 47 shows the schema for the scenario of the interactive books, using SAR to project onto physical books allowing tangible interaction in the augmented world. In this scenario the perception of the physical world in combination with the digital projection influence the interaction of the active but also passive users.

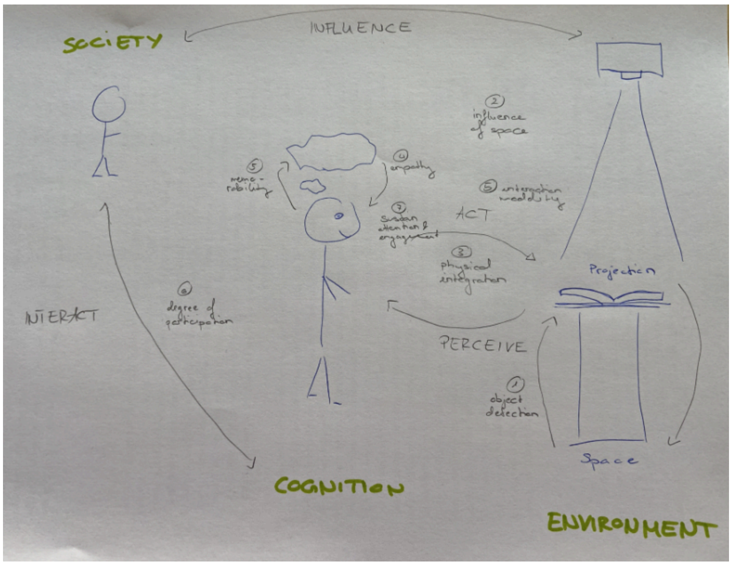


Figure 47: Schema for SAR

If analysed for similarities, several elements can be identified such as:

Place
1. Perception of the user
2. Previous knowledge
3. Emotions
4. Interpretation

Space
1. Physical space
2. Digital content

Community
1. People present in the environment
2. People present in remote environment
3. People absent in the environment

Activity
1. Actions of the user
2. Goals of the user
3. Attitude

Time
1. Temporal flow
2. History

What the user perceives influences how she/he acts. The community, either present or absent, influences the environment, as well does the user interact with the community. These influences and relations were tried to illustrate in Figure 48.

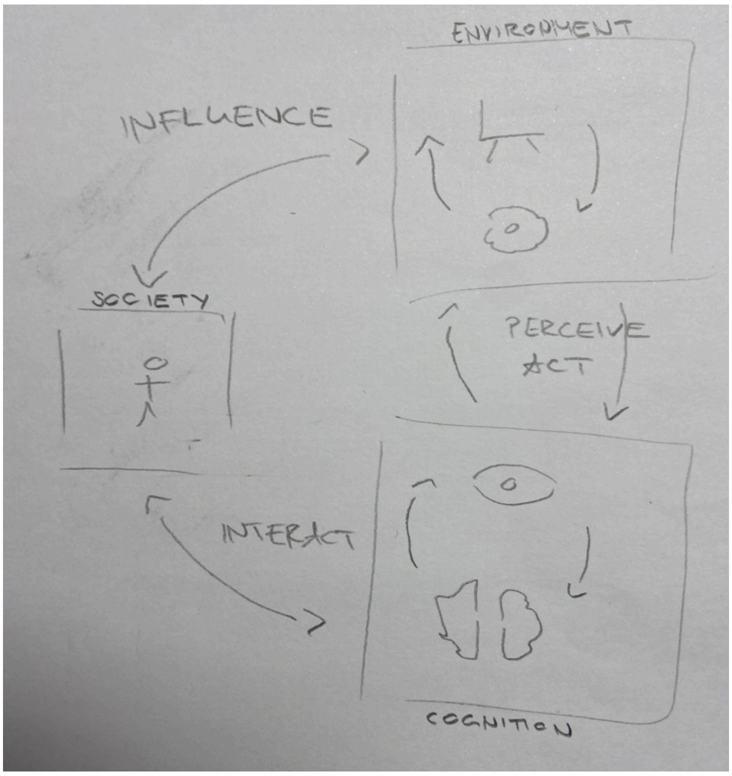


Figure 48: Model to illustrate influence and relation of User, Environment, and Community

Perception and cognition are interrelated, what a user sees needs to be processed through cognition. However, perception and cognition are influenced from the previous knowledge of the user as well as by the environment they are placed in. The environment consists of the physical space as well as the digital projection on the real world, creating an augmented space which is based on perception and cognition. The augmented space only makes sense if the digital projection suits into the physical environment and allows for meaningful interpretation which is followed by bodily actions from the perceived environment. Further is the user influenced by the community, either present in the environment or absent but however shaping the norms and values of the space through its history.

However, several elements are missing in this illustration such as for example the influence of time as well are the overarching goals of the user not included, and the previous knowledge of the user.

9.5.2 Modelling of System Dynamics

Figure 49 presents a first version of a model to illustrate the system dynamics of the identified elements relevant to the user experience in spatial computing. It is divided into two parts: the user and the environment, which is connected by the interaction with the artefact. The user's perception is affected by the overarching goals to be achieved, the user's attitude as well as emotions. This in turn influences the interpretation of the situation, which is also based on the user's previous knowledge. The interpretation and perception of the situation leads to actions that depend on the temporal flow. The model also includes temporal delays, as the interpretation of a situation can be memorised and used as prior knowledge in future scenarios. The user interacts with the environment through the artefact. The physical space is influenced by the history of the place, as well as the digital space, which in combination creates the augmented space with which the user can interact. This is also influenced by the community inhabiting the space, either present or absent.

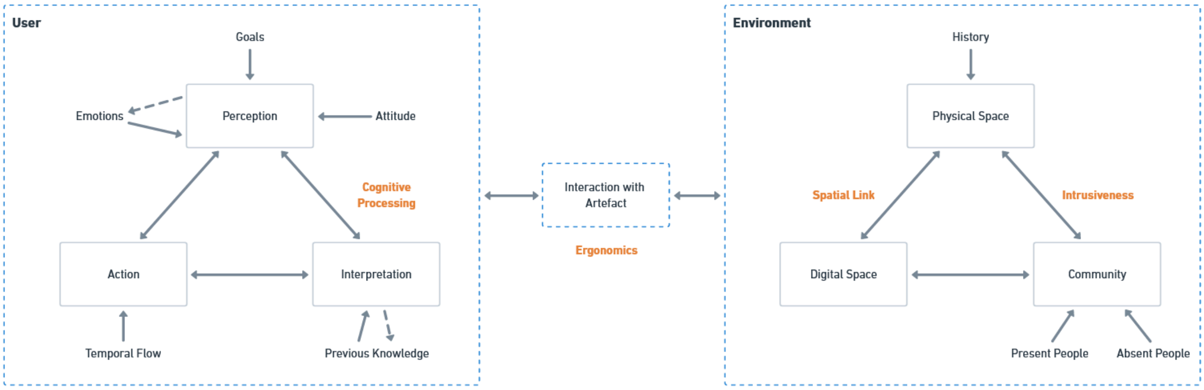


Figure 49: Model to illustrate system dynamics with feedback and time delays, highlighting new aspects relevant to spatial computing in orange

With spatial computing, new aspects are relevant to the user experience, such as the *ergonomics* of the interaction, which is located in the middle of the model in Figure 49. In a traditional mobile or web application, the ergonomics of the interaction could be neglected because the interaction takes place purely on the device. In spatial computing, however, the user is moving around the space and may be holding or wearing the device. The ergonomic issues arising from this need to be considered as it has a huge impact on the user experience as it can lead to pain and fatigue.

The *spatial link* of digital objects to physical space is also important and needs to be considered in spatial computing. This aspect lies in the environment between physical and digital space. It defines whether the digital objects displayed through AR are correctly linked to support the user's mental model and allow for natural language guidance such as "above the item".

Another consideration lies between the physical space and the community with the aspect of *intrusiveness*. Devices are equipped with cameras and sensors to detect the environment in which they operate. As a result, they also scan the private environment of the user or other people present in the space, which can lead to feelings of intrusiveness.

On the user side, the *cognitive processing* of information is now relevant to the development of applications for spatial computing. With AR it is possible to simulate objects in the environment, however not everybody processes information in the same way where as some benefit more from simulations and some rely more on their mental images. The cognitive processing style of the user strongly influences the perceived usefulness of AR simulations.

9.5.3 Difference to 2D-Interactions

The model created in Figure 49 is applicable to spatial computing applications where context needs to be considered and spatial interactions can be performed by the user. It takes into account aspects relevant to the user experience in spatial computing that were previously not very important. In traditional mobile or desktop applications, the context does not have much influence on the user experience. In addition, the interaction only takes place on the screen of the device. The user can be in any environment and have the same interaction with the application. The physical environment does not allow or hinder the user from interacting with a mobile application, but it may in spatial computing. Furthermore, the digital content is always visible on the mobile device and manipulation is limited to touch interactions. In spatial computing, visibility depends on the context and there may be multiple interaction possibilities, resulting in a different user experience.

The left-hand side of the model, which is an illustration of the user's cognition, can also be applied to traditional mobile or desktop applications. However, the mental processes are slightly different in the two cases. In contrast to spatial computing, in mobile applications the environment is not relevant to the perception or interpretation for actions.

9.5.4 Further Iteration of the Model

Based on a review with Prof. Dr. Carmen Zahn as well some inputs from the *World Information Architecture Day* Conference the model was further developed in some points as illustrated in Figure 50.

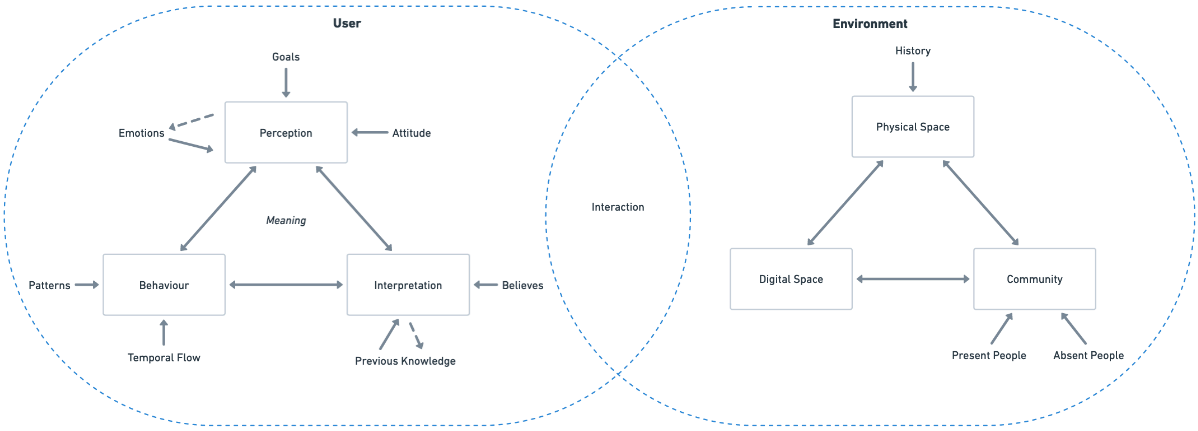


Figure 50: Improved model after review and inputs

As the user is in the physical space and is therefore part of the situation, the two dimensions - user and environment - should not be separated from each other and should overlap. Therefore, the dimensions are drawn as circles overlapping through interaction in the revised version. This visualisation further emphasises that the situation should be analysed holistically as a complex system in which elements are highly intertwined and interact with each other.

Further considerations are added in the dimension of the user. The beliefs of the user will have an influence on the interpretation and the patterns of behaviour will have an influence on the behaviour.

9.5.5 4E Cognition

The mind cannot be considered in isolation to understand how cognition works. It is a process of the brain and how it interacts with the body and the physical and social environment [149]. This forms the theory of *4E cognition*, a relatively new concept that builds on early work in 1993 on the *embodied mind* by Varela et. al [150]. 4E cognition, like embodied mind, states that cognitive processes are dynamically intertwined with the human environment, leading to the four types of cognition that are embodied, embedded, extended and enactive [149]. They can be defined as:

- Embodied: Cognition is not just a mental process, but involves the body and the mind and the interaction between the two.
- Embedded: Cognition is not isolated from context, but is influenced and shaped by it.
- Enacted: Cognition is not passive but an active process in which the user interacts with the environment and other people, resulting in cognition through the actions and behaviour.
- Extended: External elements outside the boundaries of the mind can also be an essential part of cognition, such as objects, tools or other people.

4E Cognition also incorporates complexity theory as the four different types of cognition form a complex system. It also has many similarities to the notion of situatedness, which also has a holistic focus on the situation, where multiple elements in the context shape the situation in which the user is placed.

Information processing and interaction with the environment and other people is different when AR is introduced. This changes all four types of 4E cognition and creates a whole new situation. With AR, we interact with the physical environment and it is through this interaction that we perceive and interpret the world. This highlights the fact that the processing of information in AR should not be seen in the traditional way as with 2D content, but rather as a complex interaction with the situation in which the user is placed. With AR the four types can change as follows:

- Embodied: AR allows the user to interact with the digital content in the physical world, creating a new situation beyond the natural environment. The user is immersed in the scene of the augmented environment through the interaction of the body with the digital space.
- Embedded: The user's context is enriched with digital information that may also allow interaction, creating a new situation in which the user can interact with the augmented world which in turn influences the cognition. The digital content offers new possibilities, either in the digital space on its own or through the enrichment of the physical space.
- Enacted: The introduction of AR changes the way we interact with other people, possibly also disrupting it as there is an information asymmetry or it feels like a barrier to communication when someone is wearing an AR headset. As the behaviour and actions of the user and surrounding people changes, so does the enacted cognition.
- Extended: The digital content of AR can enrich extended cognition by changing the mental load of some tasks, as some thinking can be offloaded. For example, visual instructions overlaid on a machine instead of textual information on paper, or digital simulation instead of mental imagery.

It is important to take into account the change in the situation caused by the introduction of AR. A situation with AR cannot be considered in the same way as a situation without AR, as changes occur on several levels. The interaction between the physical and the digital needs to be analysed in detail and how the interaction with the environment changes. Figure 50 is an illustration of how closely the elements are coupled and interact with each other. The potential difference to 2D interactions is also highlighted by the theory of 4E cognition. As context is not as influential in this area as it is in spatial computing, 4E cognition may not be applicable in 2D-interactions.

The considerations with 4E cognition leads to a revised version of the model, with a more user-centered design approach than the previous version, which was more allocentric. Figure 51 illustrates the coupling of the four levels and the corresponding considerations for each level.

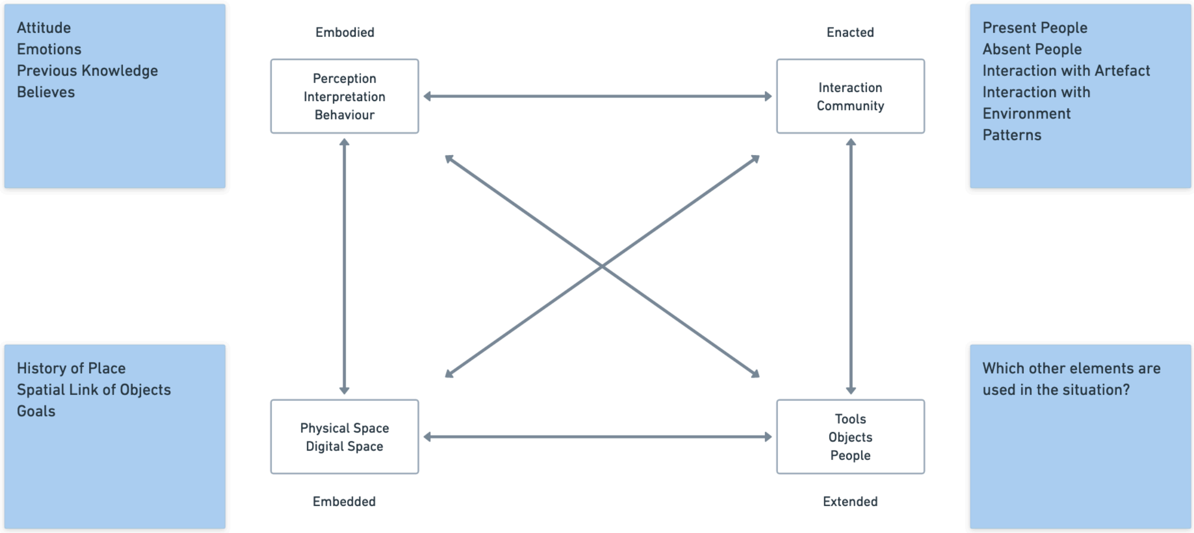


Figure 51: Revised model with 4E Cognition showing the novel combination of 4E Cognition with Spatial Computing

9.5.6 Practical Implications

The constructed model shows relevant aspects to consider in the context of a situation for spatial computing. This theoretical model can be translated into a framework, which can be used in practice to evaluate the user experience and to help identify the causes of issues encountered. As cause and problem are often distant in time or space, a holistic and coherent strategy is needed to evaluate the situation, which can be achieved with the constructed framework. It also recognises that interactions are complex and cyclical rather than linear. Changes in one element of the situation affect other elements, creating a cycle of change. The framework recognises this and after each change in the situation, all elements can be assessed to see if there are any changes due to the change in another element. By analysing all the steps and changes in the situation, it may also be possible to find the cause of a problem.

	Initial Setup	Step 1	Step 2
User			
Perception			
Goals			
Attitude			
Emotions			
Interpretation			
Previous Knowledge			
Action			
Temporal Flow			
Interaction			
Environment			
Physical Space			
History			
Norms & Values			
Digital Space			
Spatial Link			
Community			
Present People			
Absent People			

Figure 52: framework derived from theoretical model to analyse the situation holistically

Figure 52 shows the constructed framework derived from the theoretical model in Figure 51. The first column shows each element of the model in its relevant dimension. The first step is to analyse the initial situation. What are the user's initial goals? What are his/her attitudes and emotions? The perception of the situation can be constructed with this information and each observation can be noted in the framework. After changing one of the factors, for example when the user starts to interact with the technology, the observation may be moved to the second column to note the changes in the factors due to the user's interaction.

10 Qualitative Research

This chapter outlines the theoretical basics of qualitative research to lay the foundation for the user study. It discusses the difference of qualitative and quantitative research as well as explains the features of qualitative research followed by grounded theory and some basic research designs. Further are some sampling strategies outlined and different triangulation possibilities. Qualitative research offers different strategies for data collection which are explained and how the data can be analysed.

10.1 Quantitative vs. Qualitative

There are two main types of research: quantitative and qualitative. Quantitative research is often found in the natural sciences, where a phenomenon is studied under controlled conditions in order to achieve a higher level of validity. Researchers formulate a hypothesis based on literature and define dependent and independent variables in the experiment. The results are statistically analysed and conclusions are drawn. Depending on the statistics, the results allow for a generalisation, which makes it possible to formulate general laws. [151]

Unlike quantitative research, qualitative research takes place in a less controlled environment, analysing people in their natural environment. It also does not require a literature review prior to an experiment, as observation can serve as a basis of knowledge. More recently, however, it has also been suggested that the state of the art should be analysed in advance to find gaps in order to formulate meaningful research questions. Qualitative research is also open to analysing a problem in all its complexity, not reducing it to a few variables. The quality of the results can be defined by whether appropriate methods have been chosen and whether the results are grounded in empirical material. Further is generalisation not achieved by the number of cases studied, but more the quality of the cases analysed which should be based on a solid theoretical background. Numerical generalization does not fit to qualitative research as many studies aim to develop new theories and insights. This approach is more common in the social sciences and often provokes debate as to whether it is scientific or not. [151]

There has been an ongoing debate about the validity of qualitative research and the superiority of quantitative approaches. It is important to understand that the two approaches do not exclude each other and that there is no single superior method. The research approach depends on the research question, for example whether to observe a frequency or a human experience. The two methods can also complement each other, creating a mixed methods study, either as a triangulation or in sequence. In this way, the two approaches can cover each other's shortcomings, so that the frequency of a phenomenon can be understood with a quantitative approach, and the reasons for a phenomenon can be understood with a qualitative approach. [151]

10.2 Features of Qualitative Research

Qualitative research is especially useful in analysing complex situations which encompass many variables which cannot be isolated. The subject under investigation is represented in its entirety embedded in the context. This does not create an artificial situation in a laboratory an tries to analyse the subject in its natural environment which allows to observe practices and interactions of the everyday life. The aim of this approach is to *discover and explore* rather than *test what is known.* [151]

By analysing a phenomenon from the point of view of the participants, this approach can also reveal a variety of perspectives. Each case has different insights, experiences and social backgrounds that can lead to conclusions about the subjective and social meanings associated with it. The analysis of the subjective meaning that individuals attach to their activities and interactions embedded in the environment is known as *symbolic interactionism*. It is based on three premises: (1) Humans act towards things on the basis of the meaning that these things have for them, (2) the meaning of these things are derived out of social interactions, and (3) the meanings are handled through interpretation by the human. This results in a subjective point of view on situations that have different meanings for each individual. [151]

Qualitative research attempts to understand a phenomenon from the subjective point of view of the participants, including the social rules that are relevant to a situation. Individual cases are the subject of in-depth analysis and are the basis for comparisons and generalisations. Each case represents the personal perspective of an individual in a specific social or cultural context. Each case's reality is constructed, either actively produced in conversations or interactions, or latently in social situations. The empirical analysis of these situations is based on the production of texts by the researcher. The perspective of the subject is reconstructed and forms the basis for interpretation.[151]

10.3 Grounded Theory

Qualitative research often aims to develop new theories and insights, which should be "grounded" in empirical material. *Grounded theory* was developed by Glaser and Strauss in 1967 and it is still very prominent and being further developed by researchers. It allows for a very open approach to a field of study, but researchers should still be aware of the theories that are already in place. Although some of the theories should be known, this is not the starting point of the grounded theory research, but rather the end point. [151]

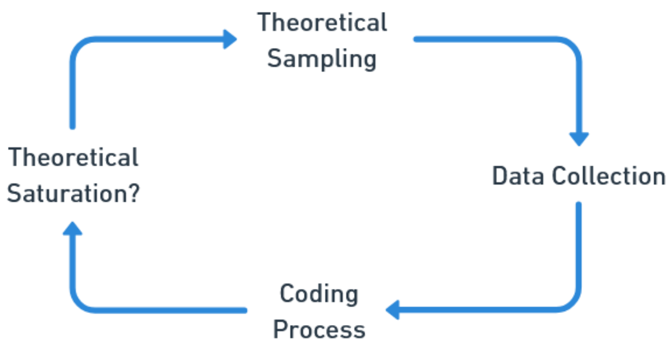


Figure 53: Cyclical Process of Grounded Theory

Grounded theory is an approach in which a variety of materials are collected and analysed in order to have an understanding of a phenomenon. This may include observations and ethnography as well as interviews or documents. The choice of material to be used depends on the cases being analysed and the insights they provide to further develop a theory based on the empirical material. [151]

The process of grounded theory research is based on theoretical sampling, grounded theory coding and theory writing. The focus is on the analysis of data that has been collected in various forms. The aim is to understand the complexity of a situation in its entirety, rather than to break down the complexity of a situation and its context into variables. The process of Sampling, Data Collection, and Data Analysis is cyclical as illustrated in Figure 53. After each iteration the researcher analysis if theoretical saturation is reached which ends the process. [151]

10.4 Basic Designs in Qualitative Research

In qualitative research, there are a few basic designs that can be used in isolation or in combination, depending on the research question.

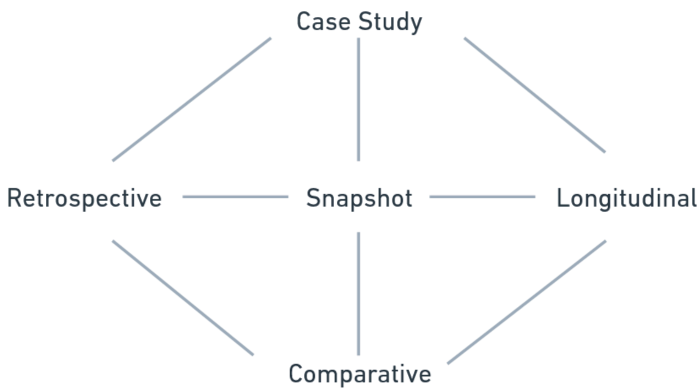


Figure 54: Basic Designs in Qualitative Research forming two axis

Case studies aim to reconstruct and accurately describe a particular case. A case can be an individual, a community, an organisation or an institution. What is important is the relevance of a case to the research question and the consideration of other relationships to the case. A case can be chosen if it is particularly interesting or typical of a general problem. The difficulty is to choose an appropriate case that will allow insights and more general conclusions to be drawn. [151]

In a *comparative study* a number of cases are observed. However, they are not analysed in depth as a whole and in all their complexity as in case studies, but only relevant extracts from the cases are analysed and later compared. The difficulty is in the choice of appropriate dimensions for analysis and in the way in which the complexity of a case is taken into account. [151]

Retrospective studies analyse certain past events and processes from the subjective point of view of the individual to understand their meaning for the individual or collective life history. It is important that the timeframe for analysing the events is clearly defined and which sources should be used in addition to interviews. It is also important to consider current perceptions of the events, which may influence memory and evaluation. [151]

Snapshots are an attempt to describe circumstances at the time the research is being conducted. However, past events and processes have an impact on the current situation, they are not the subject of the research. Snapshots consider the past, however do not try to reconstruct it and focus on the present situation. [151]

Longitudinal studies analyse certain events over a number of time periods. In order to analyse how certain things have changed over time, several periods of data collection need to be defined with a meaningful interval between them. The collection method, such as interviews, needs to be adapted according to the time frame in order to capture the development. This approach offers an opportunity to analyse events or views over time and their evolution. The change can be observed and recorded, as opposed to retrospective studies where past events have to be reconstructed. [151]

10.5 Sampling Strategies

Figure 55: Sampling and Analysis Strategies

Sample selection can be carried out according to the principle of *representativeness*. It is usually based on random sampling, where the potential samples are evenly distributed and a certain number of cases can cover the entire sample to ensure an even representation of characteristics. Another principle is based on *relevance*, where certain characteristics are relevant to the research question and it is therefore decided to include the subject in the sample. In qualitative research, the principle of relevance is usually applied to sampling, although there are different ways of selecting a sample. [151]

The sample can be defined in advance by identifying the relevant dimension to be covered. This approach is similar to quantitative research, but mainly based on the principle of relevance rather than representativeness. A subcategory of sampling beforehand is purposive sampling, where certain cases are selected on the basis of their characteristics. Possibilities include selecting extreme or deviant cases, typical cases, maximum variation in the sample, critical cases, important or sensitive cases, and convenient cases. [151]

Grounded theory often uses step-by-step sampling, where the researcher decides along the way which material or case to analyse next, also called theoretical sampling. With this approach, it is often unclear how long the research will take as samples are taken until theoretical saturation is reached. In this case, new samples would not provide new insights into the research question. [151]

Before deciding which samples to use, it is also important to consider whether sampling width or depth is more important. Width tries to represent the field in its diversity by using many different cases, whereas depth tries to analyse a particular area in the field and understand its underlying structures. [151]

10.6 Triangulation

Complex issues may require multiple methods of investigation. This may be a combination of quantitative and qualitative methods, resulting in a mixed methods study, or a combination of qualitative approaches, using the concept of triangulation. This is the combination of different methods that result in different perspectives on a phenomenon, thus creating a more solid foundation as the scope, depth and consistency can be increased. [151]

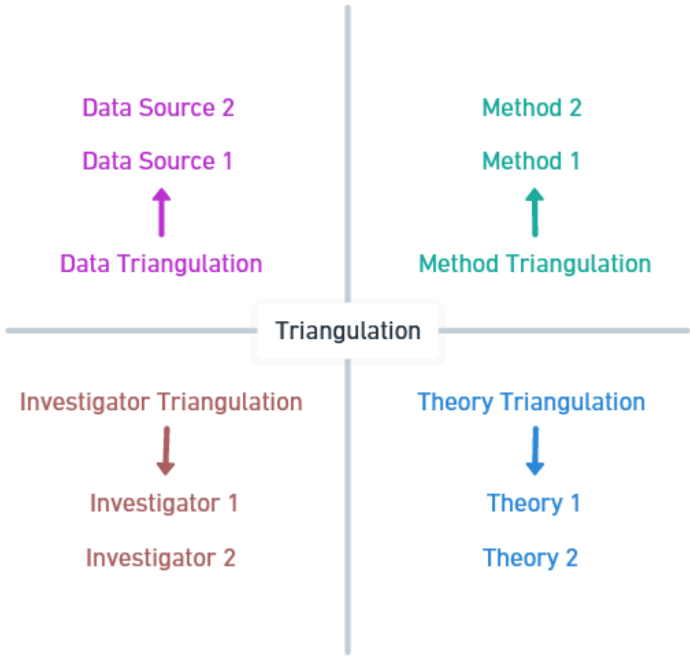


Figure 56: The four types of Triangulation

One type is *data triangulation*, which uses multiple sources of data, distinguished either by time, space or people. Another type is *investigator triangulation*, where different observers or interviewers are used to minimise bias. The results of each researcher are systematically compared for analysis. *Theoretical triangulation* has a different starting point, as the researcher starts with multiple hypotheses to illustrate the different perspectives for research. Finally, *methodological triangulation* proposes a combination of methods to study a phenomenon, either as within-method or between-method. Within-method triangulation seeks to analyse an issue from multiple perspectives using the same method, such as a semi-structured interview combined with narrative sequences. Between-method triangulation combines methodological approaches that differ in their focus and in the data they provide, such as a combination of semi-structured interviews and observation. [151]

10.7 Data Collection

In qualitative research, data is usually collected as verbal data. This allows implicit knowledge to be revealed and further insight to be gained. It can be collected through semi-structured interviews, narrative interviews or focus groups, although there are other forms, but these are the main methods for verbal data. Other data can be gathered through observation, documents or visual data such as film or photography. [151]

10.7.1 Semi-structured Interviews

The *semi-structured interview* is probably the most common type. It is based on a set of prepared questions, mostly open-ended, to guide the interview. However, it allows for some flexibility as the order of the questions does not need to be strictly followed, but rather serves as a guide for the interviewer. Interviews can be further divided into a *focused interview*, a *semi-standardised interview* and a *problem-centred interview*. [151]

The focused interview was introduced to study the effect of mass media on the population. It starts with a stimulus, usually in the form of a film or photograph, and its effect is studied through the interview. It encourages retrospection by presenting specific material or asking specific questions to recall a particular event. [151]

The semi-standardised interview aims to reconstruct the interviewee's subjective theory about a topic. It uses not only open questions but also theory-driven, hypothesis-driven and confrontational questions. It is further complemented by the *structure-laying technique*, where the insights from the interview are graphically represented and later interpreted with the interviewee. [151]

The problem-centred interview uses a combination of open-ended questions and narrative prompts to elicit biographical information about a particular problem. It uses several elements as a conversational strategy between the interviewer and the interviewee: a conversational entry opens the conversation, general and specific prompts lead the interviewee to the specific problem, and ad-hoc questions can be used to elicit additional information. [151]

10.7.2 Narrative Interviews

The narrative interview aims to reconstruct biographies of participants or experiences of certain major events or developments in society. The interviewee is asked to narrate the story of interest, including all relevant events from beginning to end. The interviewer does not interrupt the interviewee during the story and only asks follow-up questions at the end. The interview begins with a narrative question to stimulate the interviewee's thoughts and narrative. A subcategory of the narrative interview is the episodic interview, where only certain episodes are the subject of analysis. [151]

10.7.3 Observations

The term *observation* can be defined very broadly to include the collection of information in various ways, either as a participant or non-participant, and in various forms, in order to understand a particular phenomenon. Most often it is the use of documents or communication to observe interaction and other processes. The observation can be facilitated either as *participant observation*, *ethnography* or the use of *visual data*. It can also be classified as systematic or unsystematic observation of a natural or artificial situation. [151]

In participant observation, the researcher becomes a participant. As an active member of the field, they take part in activities and do not just observe. It also involves informal or formal interviews with other participants and the collection of information about practices in the field in the form of notes or videos/photographs for documentation. [151]

Ethnography is based on longer participation and observation than participant observation and is also more flexible. It can also be combined with group discussions and documents such as diaries, newspapers or letters. It can also be conducted as participant observation, but also allows for non-participation in the field and only observation of the active participants. [151]

Visual data used in the research field can be studied in observations by analysing their content, intentions, composition or use. [151]

10.8 Data Analysis

Qualitative research aims to make statements about implicit and explicit dimensions and to make sense of material by interpreting and classifying verbal or visual data. The analysis of verbal data is usually done through coding. This method uses categories to analyse texts by assigning a label to words, phrases, segments or lines that represent their meaning. There are several methods of coding that can be used depending on the research question and the data to be coded. [151]

Most methods start with an *open coding* of the verbal material to derive codes from the data. Later the codes can be used by other methods such as *theoretical coding*, *axial coding* or *thematic coding*. [151]

Open coding should allow the relevant codes to emerge from the analysed text. The data is segmented and annotated with a code that describes its meaning. The codes can represent the expressions of the interviewees (in vivo codes) or from the literature (constructed codes). Open coding is contrasted with deductive coding, in which pre-defined codes are used to analyse the verbal data and match text segments to the relevant codes. [151]

Axial coding attempts to reveal relationships between codes by using a coding paradigm model (Figure 57). It attempts to clarify the relationships between a phenomenon and its underlying cause and consequences, taking into account the context and the strategies of the participants. [151]

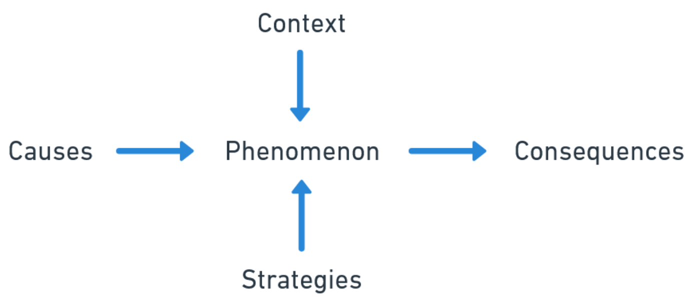


Figure 57: Coding Paradigm Model for Axial Coding

In theoretical coding, codes are grouped into coding families in order to identify underlying structures or principles. Several coding families may be derived from the literature, or families may be constructed from the codes identified. Examples of theoretical families are *Process*, where codes are grouped on the basis of stages, phases, transitions, chains or sequences; *Strategy*, where codes are grouped on the basis of strategies, tactics, techniques or management; or *Culture*, where codes are grouped on the basis of social norms, social values or social beliefs. [151]

Theoretical and axial coding focus on an issue, whereas thematic coding focuses on individuals and analyses the data on a more case-by-case basis than the others. This also allows groups of participants and their contrasting views to be illustrated. Open coding will be applied on a case by case basis and the codes identified for each case will be grouped into thematic areas. [151]

11 User Study to Evaluate the Model

To evaluate the constructed model a user study is conducted. This chapter will outline the goal of it and as well the benefits of combining HCI with Computer Science background. Further are potential applications elaborated which could be developed for the user study. Lastly, the chosen application for implementation is outlined as well as the results presented and discussed from the conducted user study followed by the implications of the results on the constructed model.

11.1 Goal of the User Study

The previous chapters lay the foundation for the construction of a theoretical model that illustrates potential considerations for user experience in spatial computing. To evaluate the constructed model, a user study is conducted. The aim is to understand whether users interact with spatial computing in the way that is assumed and visualised in the theoretical model.

From the theoretical model, a framework is derived that will help developers and designers in practice with the evaluation of the user experience of their spatial computing applications. In the user study, the framework will be evaluated in terms of its suitability for use in a usability test and whether problems and their causes can be identified.

To achieve the objectives of

(a) Evaluation of the theoretical model through analysis of the interaction of users in spatial computing
(b) Evaluate the constructed framework on the basis of the theoretical model for practicality
(c) Evaluation of the codes in the model

the research object must be narrowed down with a specific research question and application. The codes in the model can be evaluated using deductive coding and new codes can be found using inductive coding.

11.2 HCI and Computer Science

There is an ongoing debate in computer science about whether or not human-computer interaction (HCI) should be part of the discipline. Many argue that HCI is more psychology than computer science and should therefore be part of this discipline. However, design also plays a crucial role in HCI, but this fact is often neglected, as many examples show. Rather than arguing whether HCI is psychology or computer science, it should be seen as an interdisciplinary field where skills in one area influence and foster skills in the other as the following points should outline.

Combination of Knowledge leading to new Approaches

Spatial Computing is a growing field in computer science. However, it is usually only considered from a technical point of view, such as performance, resolution or novel technical approaches. Spatial computing transforms the space as a whole. Not only are changes being introduced to the space itself, but it is also changing the way in which we perceive the situation as a whole in the space. The digital content is embedded in a situation where the user interacts not only with the digital content but also with the physically present people and environment.

The combination of spatial computing and 4E Cognition Theory makes it possible to illustrate these changes in the user's situation. 4E Cognition describes four different levels of cognition and is an extension of the theory of embodied cognition that also takes into account the space and the community of the space.

It is only with knowledge and understanding of both fields that this combination of spatial computing, from computer science, and 4E Cognition, from psychology, can be achieved. Computer scientists need to understand that technology affects and transforms our perception of the whole situation in which we find ourselves. How this transformation takes place can be illustrated with 4E cognition. To illustrate the change and understand its implications, knowledge of computer science and psychology is required.

Psychological impact of design

Good design requires an understanding of psychology. When designing an application, it is important to have an understanding of how users perceive elements and how they are likely to be interpreted. For example, elements have affordances that are not universal and vary between user groups. The possible interpretations of the user groups need to be taken into account and how decisions will be made. Only the combination of all three fields - psychology, design and software development - enables the development of outstanding user experiences.

Automatic data collection by the device

The devices usually allow certain data to be collected automatically from the user. In the case of the Vision Pro, head gaze or hand position can be collected. However, in order to collect this data, the application must be specifically designed to enable these features. The window or windows must be placed in an immersive space to enable tracking. Applications are usually placed in the shared space, where several applications can be seen and used at the same time. In the immersive space, the user has to focus on one application as all other applications are hidden in the user interface. In order to take advantage of the possibility to record, the application has to be developed with the idea that it should be placed in an immersive space, either mixed or full, as the change from shared space to immersive space at a late stage of development creates a lot of work. By combining the disciplines of psychology and computer science, the researcher knows the possibilities and limitations of data collection and can strategically enrich the data with these functionalities.

Control of influencing variables in design and development

Conducting an experiment by designing and developing software involves many variables that affect the results. These variables need to be identified, which requires knowledge of design, software development and psychology. For example, an understanding of how the fidelity of a prototype or piece of software affects its perception and the resulting emotions can help in choosing the right level of fidelity to achieve the desired outcome of the experiment. Furthermore, other variables can be designed specifically for the experiment, not only in UX design but also in the software development process. However, this can only be achieved if the different disciplines are combined and synergies are exploited.

Consideration of technical constraints

With knowledge of design and software development, the technical constraints of a technology are known and the design can be pragmatically developed in accordance with these constraints. Because the development effort can be estimated, features that are very difficult to develop are not designed, but more developer-friendly solutions are explored. Even if the evaluation of the software is done by the same person, the focus of the evaluation can be set according to the design. Observing the interaction with the software provides different insights than reading a report from another person. It is possible that some findings will not be reported because the focus is different. It is also possible to distinguish whether a problem is due to the design or technical limitations of the software.

Effective communication

By having knowledge in all three fields of design, psychology and software development communication with various stakeholders can be facilitated differently. The point of view can be better understood and the language used in communication can be adapted accordingly. Different languages and specific terms are used by stakeholders in different areas. Knowing these can help to facilitate effective communication and reduce misunderstandings. Using the same language can not only reduce the number of errors, but it can also help to generate new ideas and solutions.

11.3 Potential Applications

To conduct a user study to evaluate if the model is suitable to analyse issues in the user experience, a proof of concept has to be implemented.

In this part, possibilities for a proof of concept implementation are introduced to conduct a user study. For each possibility, the potential scenario, limitations and implementation steps are estimated. With this information, the suitability of the scenario and technology is evaluated for suitability to evaluate the created model.

11.3.1 Spatial Augmented Reality

The potential scenario of SAR is inspired by the paper: FibAR - Optical Fibres embedded in 3D printed objects for SAR [16]. It uses 3D-printed objects to embed optical fibres in it. The objects are detected through the different blinking patterns of the fibre, which are invisible for the human. With the detection of the distribution of the fibre cables it allows for dynamic projection mapping, as the system can detect the position of the object.

Potential scenario: Users should collaboratively analyse different meal options to understand the impact of food choice on the environment. They can choose several types of meals to allow for visual analytics of the impact. With the selection of a meal, the ingredients are visualised on a 3D-printed plate.

Goal: It has been shown that SAR is suitable for potential attitude change as it can convey information differently to other media forms. It also has the ability to embed information in visualisations, creating tangible interfaces. This allows new information to be revealed, which is not possible when using other forms of visualisations. The aim of the user study is to analyse what influences the attitude and a potential change of it through embedded visualisation.

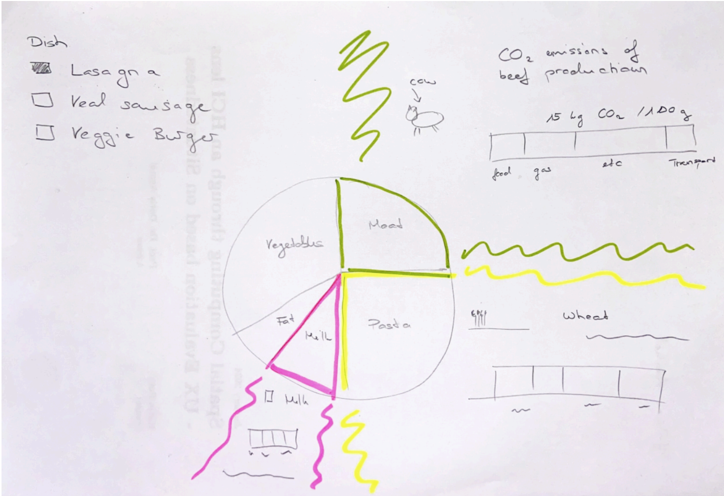


Figure 58: Sketch for scenario of SAR

Potential Limitations: The implementation from the researchers use a camera with 525 fps, where as Lusee has 30 fps and a projector with 1′000 fps which is much lower with Lusee. This leads to a slower detection of the pattern and some seconds of delay when the object is moved. Also, the implementation uses projection mapping from the side where as Lusee projects from the top.

Steps for implementation:

- 3D printing of object with special material to allow for embedding the fibre cables
- Changes in Lusee
  - Recognition of bit pattern of fibre cables
  - Identification of object through pattern
  - Calculation of position through pattern
- Creation of projection mapping

Time estimation: 5 months

Aspects considered:

space	place	activity	community	time
• object detection	• visual references • physical integration • empathy	• situational awareness	• communication	• sustain engagement

Table 7: the different aspects considered based on the term situatedness

Overall evaluation: This scenario seems less suitable for the task of evaluation the created model. It covers only few aspects as well is the development time very long which leads to a short time for conducting the user study and analysis of the model. Further has the lusee setup many limitations which would lead to delays in detection and projection.

11.3.2 Mobile Augmented Reality

The potential scenario of MAR is inspired by the application of feey and some experience in the use of model-viewer to create AR experiences.

Potential scenario: The user should place digital furniture in a given space and analyse its composition whether it fulfils the requirements. The user should analyse the different options available and their suitability. It is possible to place the various elements in the physical room through AR to simulate their appearance.

Goal: It has been shown that simulation can assist users in their decision making by serving as a model-driven decision support system. However, simulation is susceptible to spatial perception, which is influenced by colour and shape. The aim of this user study is to analyse what influences the decision comfort with simulation.

Steps for implementation:
- Create AR elements in blender
- Create Application for AR with WebXR

Time estimation: 4 weeks

Aspects considered:

space	place	activity	community	time
• manipulation of objects • visual realism • spatial perception	• visual references • spatial link	• situational awareness • fatigue • ergonomics • decision support		• sustain engagement

Table 8: the different aspects considered based on the term situatedness

Overall evaluation: Several aspects can be covered in the scenario. Further, this combination has not been studied yet and would contribute to the corpus of knowledge. Through the use of WebXR the implementation can be rather fast and is usable in several environments.

11.3.3 Head-Mounted Displays

The potential scenario of HMD is inspired by Arigo and some students project using Coconut XR for HMD application development.

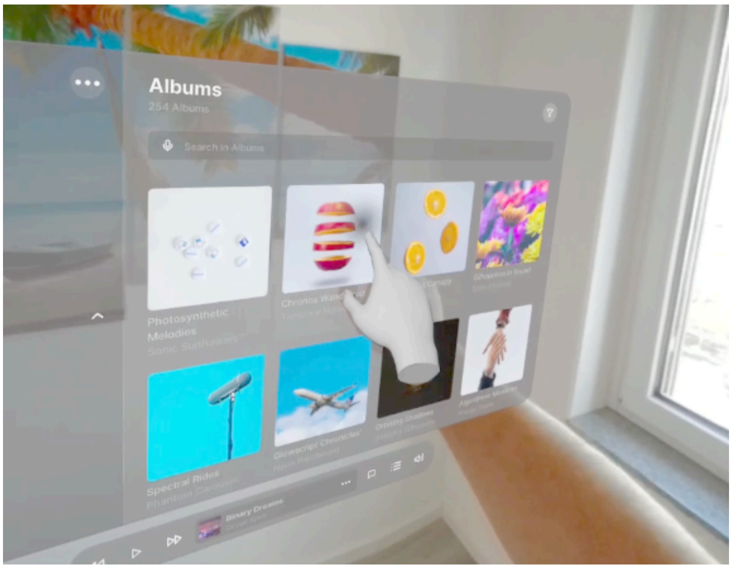


Figure 59: Example Spatial UI from Coconut XR

Potential scenario: The scenario adapted from Arigo to simulate the procedural task with guidance. The user wears the HMD and can see the step-by-step guidance for the procedural task of mixing radio pharmaceuticals. However, the user is a novice who has never performed the complicated task and requires therefore many cognitive resources to perform the steps.

Goal: It has been shown that with an AR model it is possible to give instructions with spatial references in natural language. However, the spatial mental model influences the understanding of these references and can lead to a high cognitive load in understanding and performing the necessary actions. This user study aims to analyse what influences cognitive load when using spatial references in natural language guidance.

Steps for implementation:
- Setup environment for development for potentially MetaQuest
- Develop scene with coconut XR
- Develop 3D elements with unity

Time estimation: 4 weeks

Aspects considered:

space	place	activity	community	time
• spatial perception	• visual cues • textual representation	• situational awareness • fatigue • ergonomics • interaction modality		• sustain engagement

Table 9: the different aspects considered based on the term situatedness

Overall evaluation: Many aspects can be covered and research scenario could be interesting contribution to corpus of knowledge in HCI. The implementation time is rather low and the setup can be used in several environments.

11.3.4 Conclusion

The new Apple Vision Pro had just been launched at the time of deciding which case to choose for the evaluation. The Institute had purchased one and it was available in time for the user study. Therefore, in consultation with the advisor, the case for the HMD was chosen, but instead of using the MetaQuest Pro with Coconut XR, the Apple Vision Pro will be used.

11.4 Implementation

This section will outline the task for the user study as well as the UX/UI Design of the application and the technical implementation of it. Further is the research question outlined for the user study as well as the sampling strategy and method for data collection & analysis.

11.4.1 Task

As the example application is based on the Arigo research project, the task will be the same.

Through the HMD, users see step-by-step instructions for placing syringes and turning valves in a specific order. The instructions are displayed digitally on a 3D model and the user has to perform the necessary steps with the physical elements as in Figure 60. The syringes have to be placed on the physical model and the valves of the physical model have to be turned.

The users have to place 8 syringes in different places on the model and turn the valves of the syringe holder.

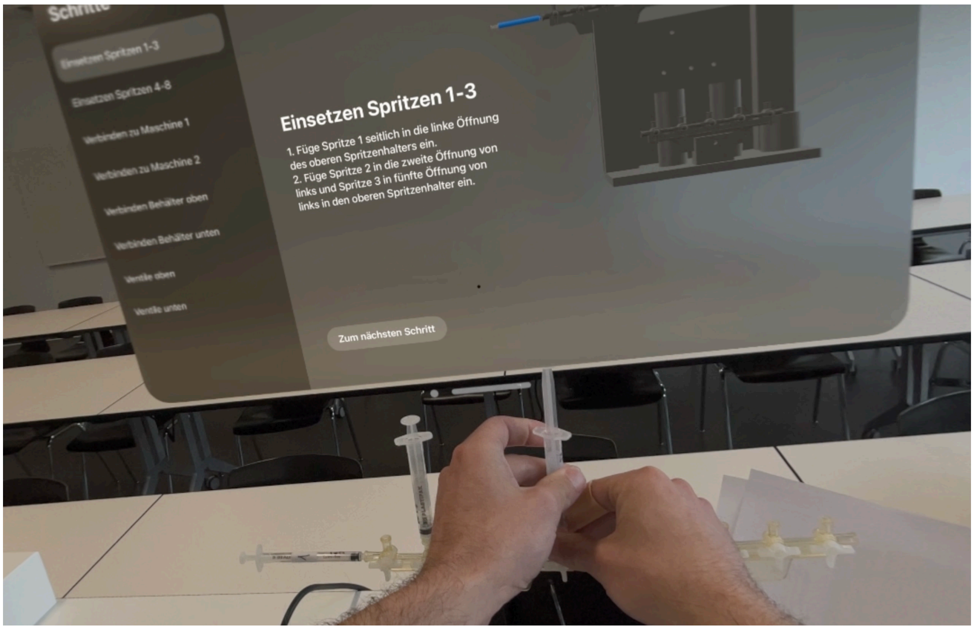


Figure 60: View from the Apple Vision Pro of a user placing the syringes on the physical model with the digital instructions in the back

The instructions are given in small batches, and the user can click on a button to proceed to the next step once the current step has been completed. On the left side of the window in the digital space, the user can see which steps still need to be taken to complete the process, as shown in Figure 61.

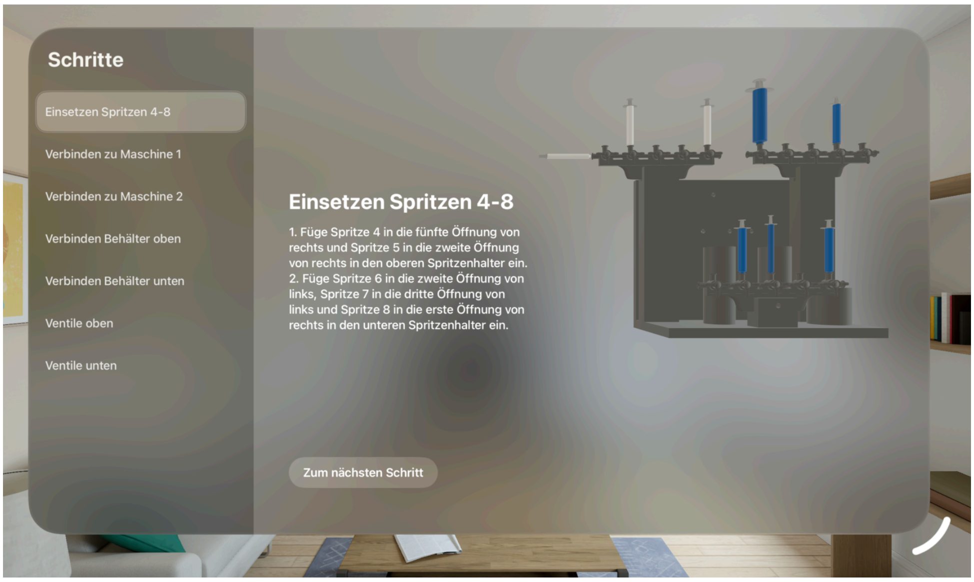


Figure 61: Instructions given through AR with the digital model and textual information in small batches, left upcoming steps are visible to the user

11.4.2 UI / UX Design

As the application is developed with SwiftUI, which is a declarative programming language, the UI design is heavily pre-defined from Apple. The Human-Interface Guidelines of Apple are already implemented in the standard components of SwiftUI. They also suggest to use their standard design in order to have a coherent design language throughout the devices and products. The paradigms and components are already known to many users and therefore simple and intuitive to use.

With VisionOS for the Vision Pro, new design possibilities are introduced as it is especially designed for augmented reality. One example is the glass window style of elements as in Figure 62. On traditional 2D-devices the interfaces used the entire screen. However, in AR the window is placed in the real surrounding of the user and the windows need to fit in accordingly. The glass background effect poses the possibility to subtly integrate the window into the context of the user and ensure that the text is always readable.

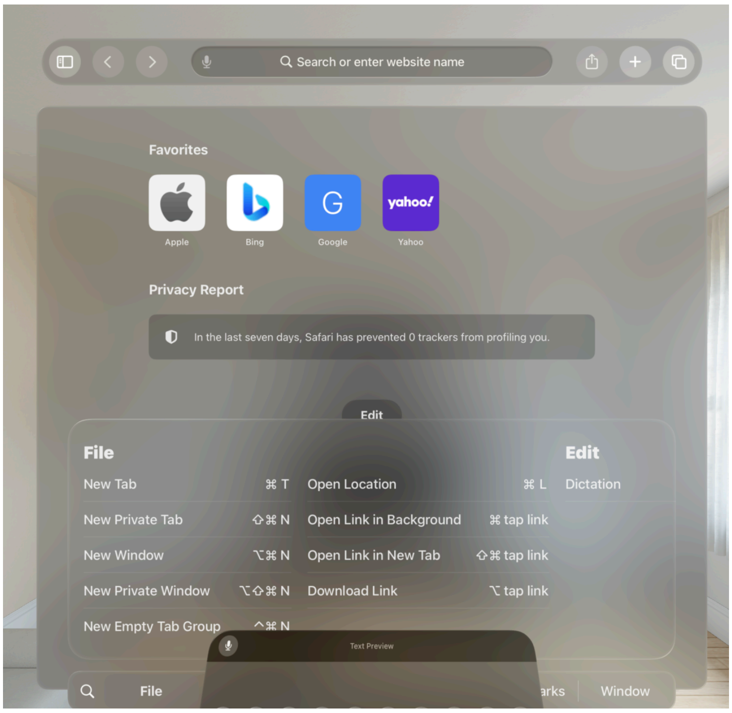


Figure 62: Spatial Window of the Apple Vision Pro using the Glass Background Effect

Not only are the design guidelines from Apple applied, also the considerations outlined in the previous chapters are considered in the application as outlined below and illustrated in Figure 63.

- *Visual cues* are used to indicate which elements are relevant to the current step. In the 3D model, the relevant elements are coloured, as in contrast to the machine, which is grey. The cues are therefore diegetic, explicit and non-limiting the interaction.
- The visual cues, which highlight the information relevant to the current step, allow information to be *easily processed* following the principle of least effort for communication.
- Only textual information that is *temporally relevant* to the current step is displayed in the guidance process. The user can see which steps are ahead in the navigation area, but the detailed text only gives information about the current step.
- The user's *attention* can potentially be held well as they have a high level of interest in the technology and are excited to use it for the first time. Furthermore, the interaction possibilities are easy to understand and learn, the immersion in the augmented space is high, and the difficulty of the task is low due to the step-by-step guidance supported by the visual cues in the 3D model.

- The glass window background and window interaction capabilities in VisionOS allow the user to rearrange the window in the physical space, but the glass background always ensures the *readability* of the text.
- In VisionOS, the window can be resized by the user, with the content such as text and 3D model also automatically resized. This changes the size of the model and can affect the perceived *visual realism* of the digital twin.

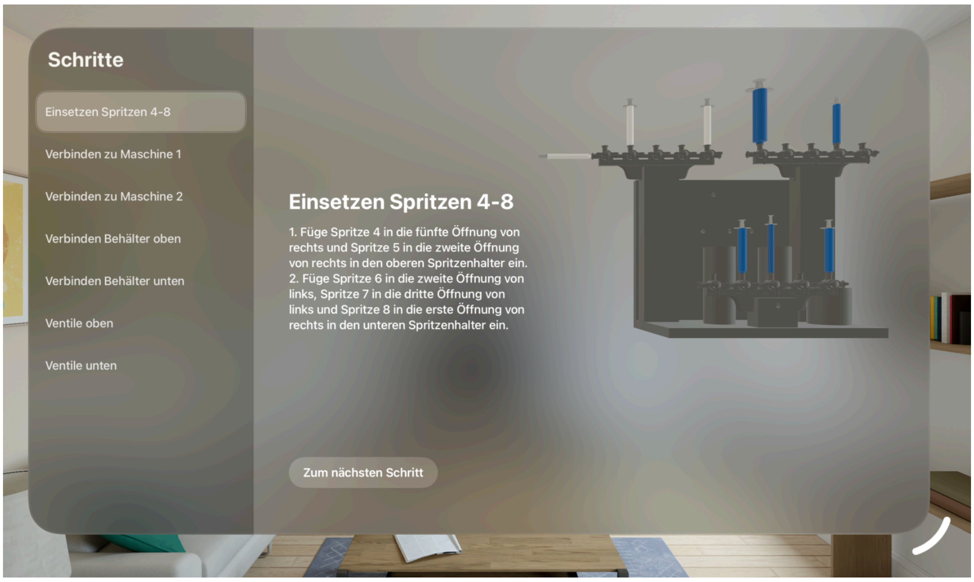


Figure 63: Window showing a step in the guidance process with visual cues

11.4.3 Technical Implementation

The application offers the opportunity to the user to get accustomed with the device and task.

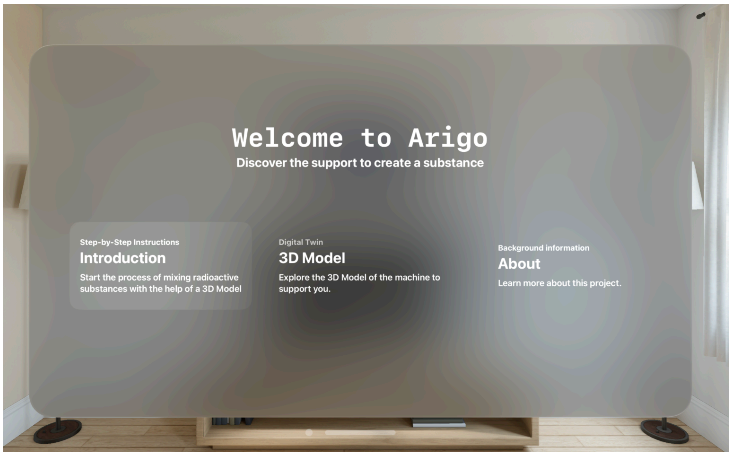


Figure 64: Start screen of the application allowing the user to get accustomed to the device and task

The introduction screen allows the user either start the guidance process, explore the 3D-model of the machine, or learn more about the application as illustrated in Figure 64. The 3D-model of the machine is introduced and the user can explore the model and understand which components the machine has. This may help the user to conduct the process of mixing a substance as he/she knows already the components of the machine and therefore may simplify the understanding of the natural language guidance. This section also has the 3D-model of the machine which is created with Unity and imported into RealityComposerPro, which is needed for importing the 3D-models into the application. On the first screen also an introduction to the step-by-step guidance in shown where the user can also start the process. When starting the process from the introduction screen, the user navigates to a new navigationStack, which is the navigation design of SwiftUI as illustrated in Figure 65.

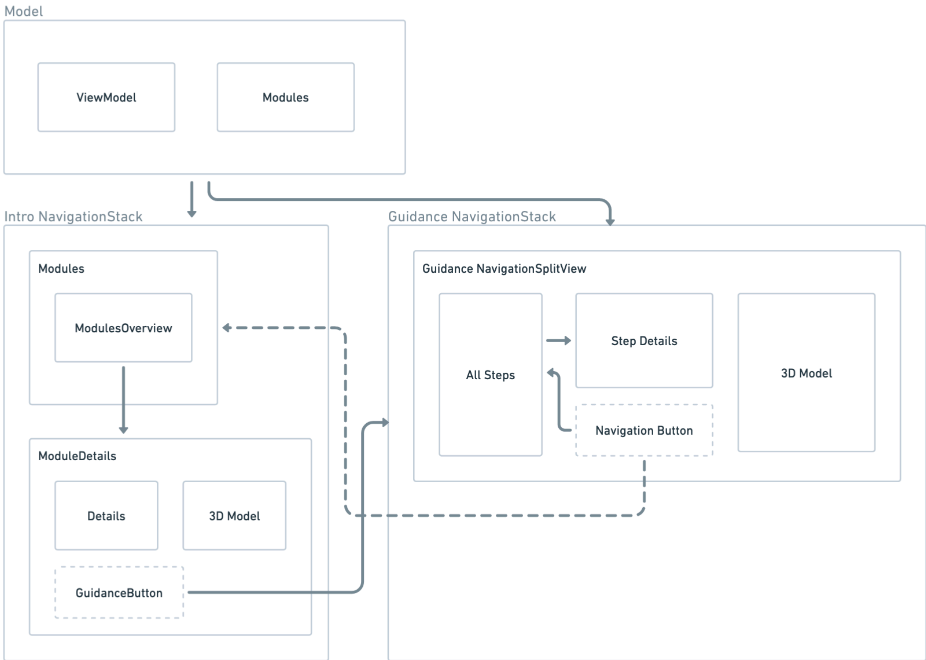


Figure 65: Software Architecture of the Arigo Application for VisionOS

The NavigationStack uses a push/pop paradigm. As the introduction and the step guidance are not interrelated two different stacks are being implemented. At the end of the step guidance the user pops back to the root navigation stack which is the start screen of the application and the entire navigation path gets deleted. This is an efficient way of clearing navigation path and reduces memory usage.

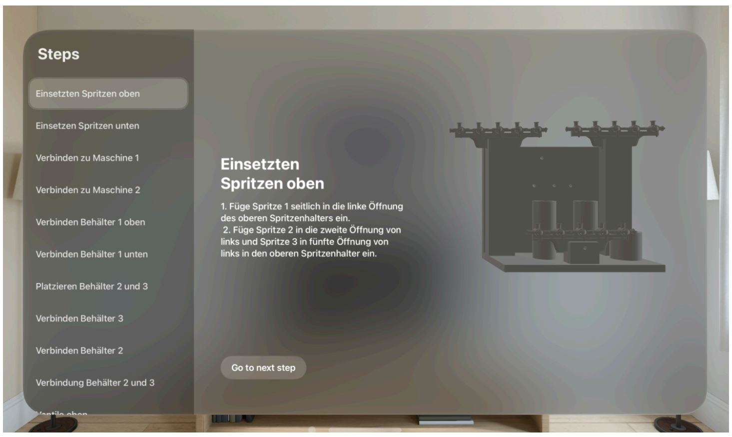


Figure 66: Step-by-step guidance using NavigationSplitview, showing the current and upcoming steps on the left and detailView with the corresponding 3D-model on the right side

The step guidance itself uses a NavigationSplitView in order to indicate the current step to the user on the left side of the window and the guidance text with the 3D-model on the right side of the window as illustrated in Figure 66. The text for the guidance is stored in a enumeration in the Model. This allows to dynamically show the content on the window and iterate through the enumeration. The step description is dynamically pulled from the enumeration and when selected, the details to the current step are shown in the detail view on the left side. The binding to the 3D-model has to be done manually with if-statements. This may not be the most efficient way of binding the models to the relevant step, however with a switch case or dynamic presentation trough the information from the enumeration does not work, assuming that the threads do not have the same processing time and therefore can't load the information when needed on the UI.

In the detail view the user can continue to the next step with a button. This iterates through the enumeration and increments with each click. The increment is also bound to the number of steps shown in the left side, where the steps to be performed are shown. When clicking on the button to go to the next step, the current step disappears in the step indication and the view switches automatically to the next step in the step indication on the left and also in the detail view on the right side of the window. When the user has reached the end of the enumeration the navigation pops back to the root NavigationStack, which is the start screen of the application.

As it is not possible to collect eye tracking information with the Apple Vision Pro, some other solutions must be found to analyse where the user is focusing while interacting with the content. Because the headset has the ability to anchor content to the user's head, it is possible to track the user's gaze. Content can only be anchored to the head in an immersive view. As the user should still be able to see the physical space, the window must be placed in the mixed immersive view. The immersive view must be actively opened and closed by the user, as no other window will be visible in this mode. In the Arigo application for Apple Vision Pro, the start and end of the immersive view is tied to the start and end button of the instruction process. In this way, the user does not notice that the immersive view has started and the head gaze can be tracked. As only visual analysis is possible, an object has to be anchored to the head, which in the case of Arigo should not be distracting. Therefore, a small black dot is placed in front of the headset to visually track the head gaze, as shown in Figure 67.

Figure 67: Tracking of the head gaze with little black dot placed in front of the headset anchored to the head of the user

11.4.4 Research Question

The aim of the study is to understand which elements influence the orientation of a user in a step-by-step guidance process with a 3D model and natural language instructions. Further should be evaluated whether there is a difference in users based on their spatial imagination and binocular vision, as well as on their affinity for visual or verbal information based on their experience.

Through this question, several points of the model can be covered and how they interact with each other can be analysed.

With these questions the constructed model for influences in spatial computing should be analysed in its entirety to discover and explore new influences. In doing so, the constructed model should be evaluated and further considerations should be found that can be added to the model. The situation should be analysed from the participant's point of view to understand the subjective meaning of elements and interactions and how they are interpreted.

The study is designed to analyse the individual cases of each participant to understand the situation in depth. Thus, the study is designed as a case study, analysing snapshots of situations, but considering the impact of the past on the current situation however without further investigation.

11.4.5 Sampling

The study is planned with three different cohorts of three participants each. The first group is experienced in working with 3D models due to their job or education (visual cohort). The second group has a greater affinity with language (verbal cohort) and the third group has a high level of technical knowledge (technical cohort). All participants have a positive attitude towards AR, but mostly limited experience with it.

11.4.6 Data Collection & Analysis

To analyse the elements influence the orientation in spatial computing an ethnography and semi-structured interviews will be conducted.

In preparation to the test the participants will conduct a test to evaluate their spatial imagination and binocular vision. The spatial imagination is evaluated using the *Mental Rotation Test*. The test uses eight figures which are turned in a 0°, 50°, 100° and 150° angle either correctly or inverted as in Figure 68 [152] as originally developed by Vandenberg and Kuse [153]. The participants have to identify the correctly turned figures as fast as possible.

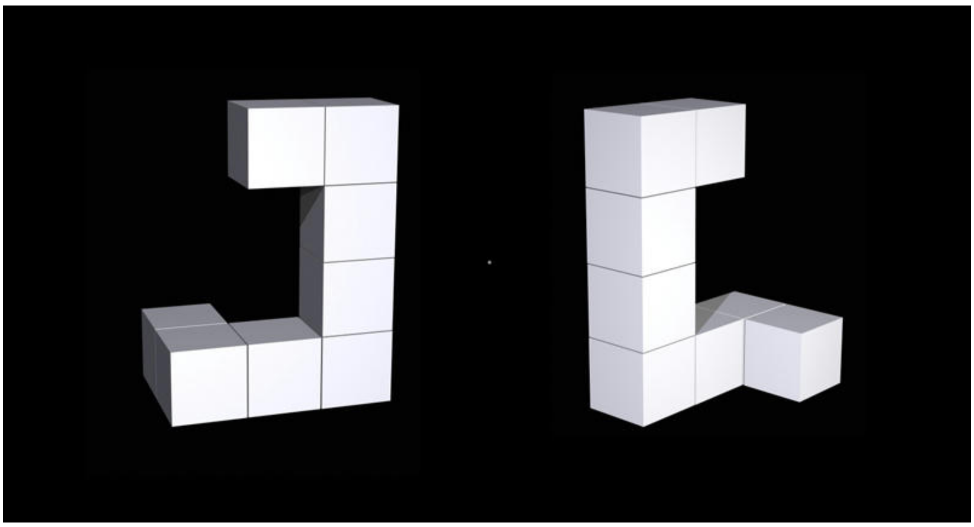


Figure 68: Mental Rotation Test of a Figure turned with a 150° angle

The binocular vision will be evaluated using the *stereoacuity test* as an android application [154] as illustrated in Figure 69. It analyses the binocular vision of participants by showing several images which are only visible with anaglyph glasses reducing the angle with each correctly identified image. As lower the angle is as better is the binocular vision of a participant. The binocular vision can have an influence on the cognitive load of participants as the application uses a 3D-model in AR which is harder to perceive for participants with low binocular vision.

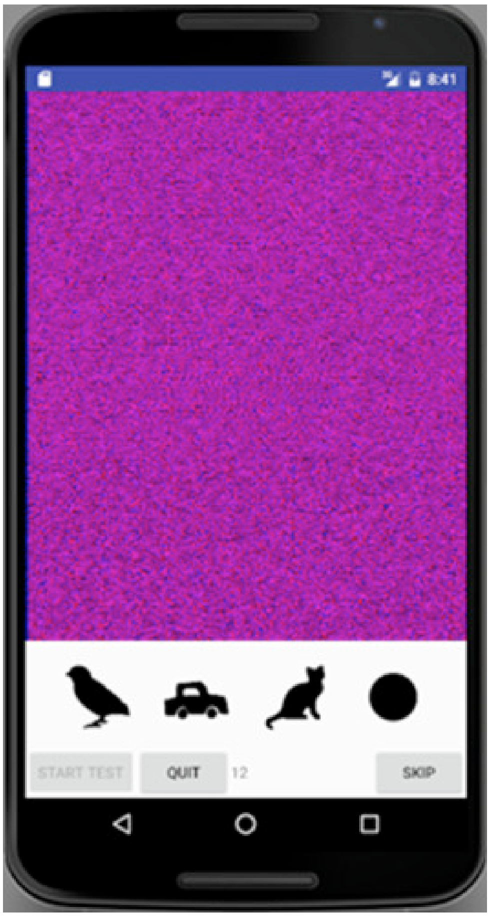


Figure 69: Stereo Acuity Application analysing the binocular vision

In the ethnography, some questions will be asked before and after the participants perform the task combined with an observation during the participants perform the task. The application is further designed to be able to analyse the head gaze of the participants. With this it can be analysed where the focus of attention was.

The semi-structured interview should help to understand how the user has oriented themselves in the augmented space given the instructions either in text or with the 3D-model.

As some aspects relevant for the user experience in spatial computing could already be identified through literature, a combination of open and deductive coding will be used for the analysis of data. With this, it should be evaluated whether the constructed framework with its aspects can be filled out and which other aspects need to be further considered. The codes identified through open coding will be further structured through thematic analysis to reveal general topics of the codes.

Axial Coding will be done to evaluate whether it is possible to find causes for issues by using the constructed framework.

11.4.7 Hypotheses

In preparation to the user study some hypotheses are formed to analyse whether the research question and study setup is matching and the goal of the study can be achieved. Further will the hypothesis after the study be evaluated whether the results were as expected.

Hypotheses:

1. Participants will be able to follow the instructions easily and complete the process without any major problems.
2. Participants will rely too much on the 3D model and will not actively or consciously process the steps performed.
3. The visual cues in the 3D model will help participants to process the information easily.
4. The visual cues will immediately be recognised as the elements which are relevant in the current step.

11.4.8 Investigating Reliance

As the hypothesis is that participants will rely too much on the 3D model, a mismatch between the text of the instructions and the digital model illustrating the steps was implemented in the application.

In the process, the participants have to place eight syringes on the model and connect six tubes to it. They then have to turn the valves so the fluids in the syringes flow in the right direction. Before turning the valves, a number of steps have already been performed, which could have led to a learning effect that the digital model always shows the information in the text, so people rely more and more on the digital model. Therefore, a mismatch has been placed at the end of the process. The instructions tell the user to turn the third valve to the right, however the digital model shows the fourth valve turned to the right (Figure 70).

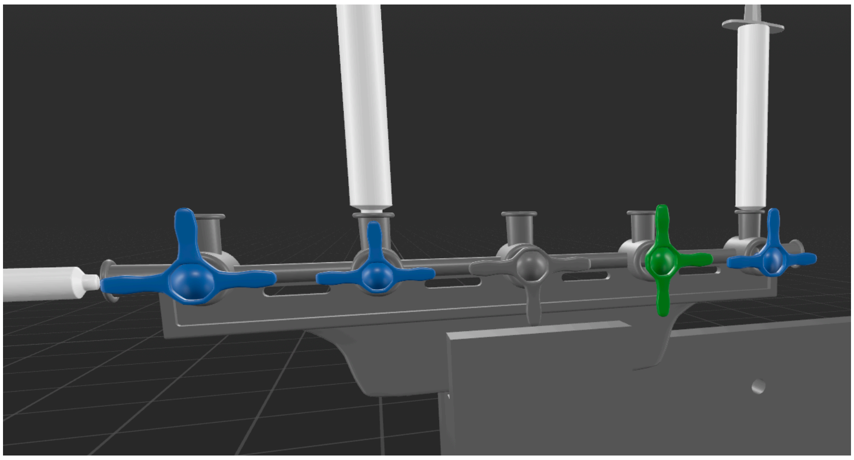


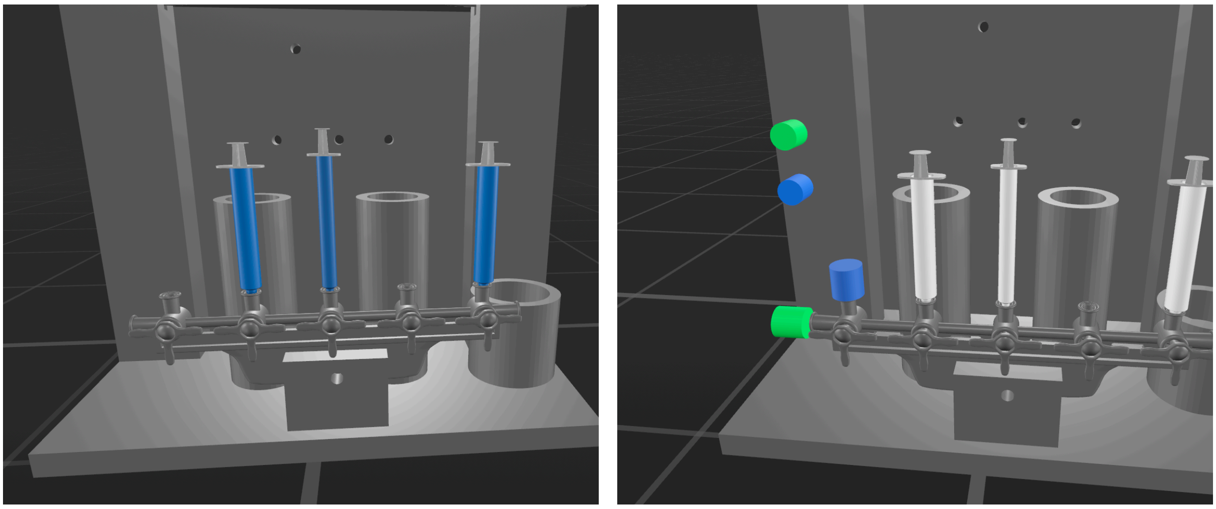
Figure 70: Mismatch of the text and digital model for the valve. In the model valve 4 is turned and highlighted whereas valve 3 is grey and not turned.

With this discrepancy it is possible to find out whether the participants read the text of the instructions, compare it with the digital model or whether they only rely on the digital model.

11.4.9 Investigating Perception of Visual Realism

To illustrate where the tubes should be placed in the digital model, coloured notches are placed in the model with low visual realism. The aim was to find out if the high abstraction of realism was helpful compared to high visual realism.

To analyse the perception of high visual realism, a small discrepancy in the size of syringe seven was placed in the digital model as it has the same size as syringe 6 in the physical model (Figure 71 a). This is to analyse whether the small mismatch is perceived also compared to the low visual realism option for the tubes (Figure 71 b).



a) b)

Figure 71: a) Mismatch in the visual realism of syringe seven as it has the same size as syringe six in physical model and b) Low visual realism visualisation to connect the tubes with the model

11.5 Results

This section present the results of the user study. First the participants are introduced as well as the final code system followed by a case analysis for each participants. Lastly, the general findings of the study are presented as well as the research question answered.

11.5.1 Participants

Table 10 shows the nine participants of the user study, including background information to their profession as well as the measurements for binocular vision and the mental rotation test. The column *cohort* indicates to which cohort the participant belongs to based on their background. The column *focus* indicates where to focus was in the case of the visual and textual mismatch, either they followed the text or the model.

Nr	Background	Measurements	Cohort	Focus
1	BIM Specialist, MAS Digital Construction	Binocular Vision: 66′ MRT: 96%	Visual	Text
2	Web Developer	Binocular Vision: 132′ MRT: 98%	Tech	Text
3	Master Student Virtual and Digital Construction	Binocular Vision: 132′ MRT: 96%	Visual	Model
4	Professor School of Pedagogy	Binocular Vision: 132′ MRT: 74%	Verbal	Model
5	Computer Science Student	Binocular Vision: 198′ MRT: 96%	Tech	Model
6	Teacher Communication School of Economics	Binocular Vision: 330′ MRT: 88%	Verbal	Text
7	Designer Mechanical Engineering	Binocular Vision: 66′ MRT: 98%	Visual	Model
8	Consultant University Development	Binocular Vision: 791′ MRT: 96%	Verbal	Text
9	Full Stack Software Engineer	Binocular Vision: 198′ MRT: 98%	Tech	Text

Table 10: Participants of the User Study with their backgrounds and measurements of Binocular Vision and Mental Rotation Test

11.5.2 Codesystem

During the user study the participants were asked to think aloud. These protocols were transcribed and analysed with an open and deductive coding in MAXQDA. The terms from the created model were used as codes, as well were new codes introduced. Figure 72 shows the resulting codesystem from the analysis of the think-aloud protocols and interviews conducted. New codes were introduced on an ongoing basis and already coded documents were analysed again for the newly introduced codes.

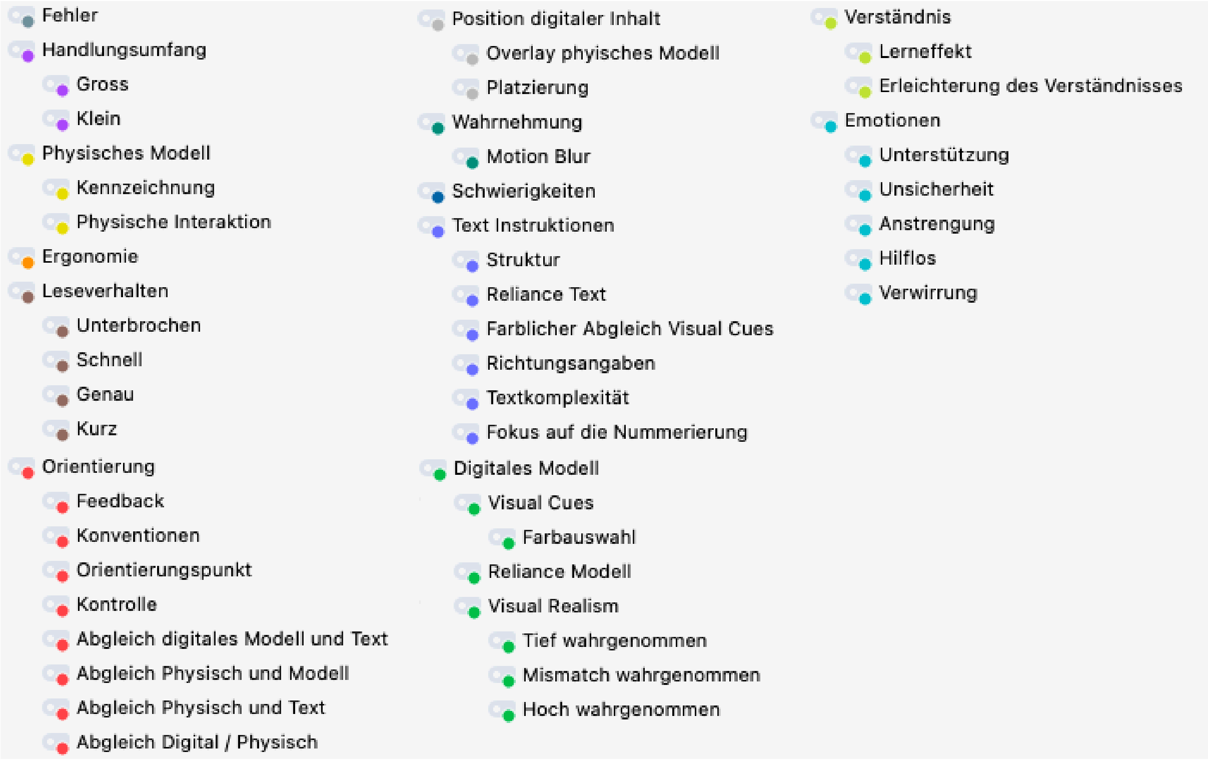


Figure 72: Codesystem from inductive and deductive coding of the user study

11.5.3 Case Analysis

The think-aloud protocol for each participant was coded and analysed in detailed. The case of each participant is briefly outlined below, enriched with background information and the main highlights of the case.

User 1

User 1 is an expert in Building Information Management and is doing an MAS in Digital Construction and therefore belongs to the visual cohort. Binocular vision is very good with an angle of 66′ and mental imagery is also very good with a result of 96% on the MRT. The user also has red-green colour blindness.

The user usually reads the instructions thoroughly. After reading the text, the user compares the instructions with the digital model to better understand the instructions. After comparing the text and the digital model, the user carries out the steps with a small extent of action and verifies the actions through the digital model.

The text is also perceived as rather complex with the change of directional information (from right or left). Therefore, the user sometimes has to read the instructions several times to understand them and compares the text with the digital model. The user also mentioned that the digital model does not provide enough information to rely on it completely. However, in steps where the digital model has enough information to perform the steps, the user does not read the text and relies completely on the digital model.

Due to red-green colour blindness, the user does not see the visual cues in the step to rotate the vials. This leads to confusion as the instructions appear inconsistent. As the user does not see the visual cues, the mismatch between digital model and text for valve three is not seen and the user follows the text instructions.

The user uses the placed syringes and the number of openings on the machine for orientation. In particular, the prominent syringes are mentioned to support orientation in the physical and digital model.

User 2

User 2 is a web developer and student in Computer Science and therefore belongs to the technical cohort. Binocular vision is good with an angle of 132′ and mental imagery is also very good with a result of 98% on the MRT.

The user reads the instructions very thoroughly and memorises the number of the syringe and the opening on the machine. After reading the instructions, the user compares them with the digital model. The user carries out the actions very carefully and follows the instructions very closely. At the end of the action, the user compares the physical model with the digital model for verification.

The user recognises even small mismatches in the digital model that lead to irritation. However, the user still finds the digital model helpful and uses the visual cues for orientation. However, the user does not rely on the digital model because there is not enough information and the text instructions are not detailed enough.

The user recognises the mismatch for the valve and follows the text instructions.

User 3

User 3 is a master student in virtual and digital construction and therefore belongs to the visual cohort. Binocular vision is good with an angle of 132′ and mental imagery is also very good with a result of 96% on the MRT.

In the beginning, the user reads the introduction to the process very carefully and compares the information with the physical model. The user also arranges the syringes according to their number. In the beginning, the user compares the text with the digital model and then carries out the actions on the physical model. However, after a few steps the user reads less and relies more on the digital model, and at the end of the process the user relies completely on the digital model without reading the instructions. This behaviour could indicate that there is a learning effect that the digital model usually shows the same information as the text. The user also mentions that there is a learning effect in the structure of the text, and that the visual cues are supportive in visualising the current actions.

When the user reads the instructions, the extent of action is small, whereas when the user relies on the digital model, the extent of action is large.

As the user relies completely on the digital model and does not read the instructions at the end of the process, the mismatch between the text and the digital model for valve 3 is not recognised and the user follows the instructions from the digital model.

User 4

User 4 is a professor at the School of Pedagogy and therefore belongs to the verbal cohort. Binocular vision is good with an angle of 132′ and mental imagery is also low with a result of 74% on the MRT.

The user reads the instructions very briefly and only the beginning of the sentence. As soon as the user could deduce a possible action from the instructions, a comparison was made with the digital model and the action was performed with the physical model. Therefore, the extent of the action was small. After a few steps, the user relies only on the instructions with the visual cues of the digital model and does not read the text any more. As the user only reads the instructions very briefly, the working practice is somewhat sloppy and an error occurs with syringe 6. The error is noticed at a later stage because of a check between the digital and physical model where the mismatch was observed.

As the user relies completely on the digital model and does not read the instructions at the end of the process, the mismatch between the text and the digital model for valve 3 is not recognised and the user follows the instructions from the digital model.

User 5

User 5 is a Computer Science Student and therefore belongs to the technical cohort. Binocular vision is lower with an angle of 198′ and mental imagery is high with a result of 96% on the MRT.

The user takes some time to get started and reads the first step carefully several times, comparing it with the digital model. However, the user soon starts to read the instructions in an interrupted manner and misses the relevant information at the end of the sentence. This leads to confusion and the user relies on the digital model for guidance. The user understands the visual cues for the syringes, but not for the tubes, and is confused by the missing information on the digital model and the complexity of the text instructions. Similarly, in the second step of connecting the tubes, the user does not understand the visual cues on the digital model. Surprisingly, to turn the valves, the user relies on the digital model and does not read the textual instructions.

The user shows some difficulties in perceiving the digital model and its visual cues, especially in transferring the knowledge from the digital model to the physical model. The user also found the process a bit exhausting and the observed concentration was low.

As the user relies completely on the digital model and does not read the instructions at the end of the process, the mismatch between the text and the digital model for valve 3 is not recognised and the user follows the instructions from the digital model.

User 6

User 6 is a teacher at the School of Economics in Communication and therefore belongs to the verbal cohort. Binocular vision is very low with an angle of 330′ and mental imagery is lower with a result of 88% on the MRT.

The user is very focused on the text and sometimes performs the actions without looking at the digital model. In the beginning, the user reads the introduction very carefully. However, after a few steps, the user reads the instructions interrupted and performs the steps as soon as they can be deduced. The comparison with the digital model is only carried out from time to time, in particular when the textual instructions are too complex or when the user is uncertain. The user is very concentrated and the extent of the action is small. Also in the valve turning step, the user relies on the text and performs each turn step by step in accordance with the text instructions.

The user recognises the mismatch for the valve and follows the text instructions.

User 7

User 7 is a designer in mechanical engineering and therefore belongs to the visual cohort. Binocular vision is very good with an angle of 66′ and mental imagery is also high with a result of 98% on the MRT.

In the beginning, the user reads the instructions very carefully and memorises each number in order to place the syringe in the correct opening. The user soon notices that the digital model shows the same information as the textual instructions and compares the two at each step. The user no longer reads the sentences in their entirety and misses important information about how to perform the actions. The user is therefore confused as there is a perceived mismatch between the text and the digital model. In cases where the visual cues are immediately understood, the user does not read the textual instructions and relies on the digital model, for example in the step of connecting the tubes

Interestingly, for the valve mismatch, the user reads the text instructions aloud but rotates the valves according to the digital model.

User 8

User 8 is a consultant in the development of the university with education in linguistics and therefore belongs to the verbal cohort. Binocular vision is very low with an angle of 791′ and mental imagery is high with a result of 96% on the MRT.

In the beginning, the user reads the text instructions very carefully and compares them with the digital model. The user orders the syringe according to its number for preparation. After a few steps, the user starts to read the text intermittently and performs the action as soon as it can be deduced. As a result, the user misses important information at the end of the sentence and perceives a mismatch between the text and the digital model. If the textual instructions are simple enough to understand, the user relies on the text alone and does not make a comparison with the digital model. The user performs the actions very carefully and remembers the number of the opening for each syringe. The visual cues are immediately recognised and help with orientation, especially in the steps to connect the tubes. When using the digital model for orientation, the user has a strong focus on the colours of the visual cues.

The user recognises the mismatch for the valve and follows the text instructions.

User 9

User 9 is a full stack software engineer and therefore belongs to the technical cohort. Binocular vision is low with an angle of 198′ and mental imagery is high with a result of 98% on the MRT.

In the beginning, the user reads the introduction very carefully and compares the text with the physical model. The user also reads the textual instructions for the first step very carefully and remembers the information about placing the syringe in the correct position. However, the user would count the openings differently and relies on the digital model. After the first step, the user changes the reading behaviour and performs an action as soon as it can be deduced from the instructions. In cases of uncertainty, the user compares the textual instructions with the digital model and reads the text several times. However, if the text is clear, the user performs the action without looking at the digital model. As the textual instructions sometimes contain several steps and the user is strongly focused on the text, syringe 6 is forgotten.

The visual cues are perceived as helpful and the user notices the mismatch in the size of the syringes. The user recognises the mismatch for the valve and follows the text instructions.

11.5.4 Findings

Based on the detailed analysis of each case the following general findings can be deducted:

Visual Cues

The visual cues in the digital model are perceived as very helpful. In the case of uncertainty about the instructions in the text, the users rely on the visual cues in the digital model for their understanding.

The use of colour is also important for understanding which elements are currently relevant and which are previous elements. In the step of connecting the tubes, users often referred to the tube as “the blue tube” or commented “blue to blue” while making the connection to the physical model. The low level of visual realism in the step of connecting the tubes was found to be very helpful and most of the users were able to follow the steps immediately. However, when the visual cues were given with high visual realism already a slight mismatch to the physical object was observed and lead to confusion.

Text Instructions

The textual instructions are perceived as too complex and the structure is not ideal. Users often read the instructions interrupted because there was too much information in one sentence. They therefore carried out the steps without reading the sentence to the end and missed important information placed at the end of the sentence.

Some sections of text have 2-3 steps that are numbered. Users often forgot which sub-steps they had already completed, leading to confusion or even errors. However, in some cases the users also mentioned that there was not enough information in the text, for example in the step to turn the valves. Most users were missing information on which direction to turn the valves, as the only instruction given was to “turn it up”.

Most users focused on the numbering of syringes and openings on the machine. While performing the task, they often repeated verbally the information they had read about the numbering. Directional information was also important and was often repeated while performing the step.

Overall, the process was not perceived as exhausting or demanding by most users. However, reading and understanding the textual instructions was demanding and perceived as overly complex for the actions to be performed.

Information Sources

The textual instructions and the visual cues on the digital model are perceived as two sources of information that need to be compared. However, most of the users mentioned that only one of the current sources would not be sufficient to perform the steps, as both are missing certain information that can only be perceived by combining the two. Having only one source would be preferable, especially if all the information could be visualised in the digital model, which would be supportive for most users.

Consistency

Changes in the structure of the information were often perceived as confusing or sometimes even overlooked, such as the change in direction to place the syringes in the correct opening. The first instructions were given from the left and the second from the right. As the information about the direction is at the end of the sentence and most of the users did not read the instructions in full, the information was missed and confusion arose.

One user also has red-green colour blindness and missed the visual cues in the step to turn the valves. The visual instructions were perceived as inconsistent, leading to confusion and uncertainty.

Personality

It was also observed that the personality of the user had an influence on performance. In particular, the user's attitude had a huge influence, as some users were very detailed and concentrated. They followed the instructions very closely and even felt that the instructions were not detailed enough. Others were less precise and followed the instructions loosely.

The more accurate participants also had a preparation phase before starting the process. They read the introduction and compared the information with the physical model. They also ordered the syringes for preparation to make the process more efficient.

One user also mentioned that the numbering of the openings on the machine differed from the personal mental model. This conflict could only be resolved by comparison with the digital model.

Learning Effect

Most users understood the visual cues in the digital model immediately, especially the visualisation of the syringes. The visual cues in the step of connecting the tubes often took some time. However, in the second step there was a learning effect as the visual cues were understood immediately.

There was also a learning effect as the user relied more and more on the digital model towards the end of the process. Users realised that the text and the digital model showed the same instructions and therefore relied more on the digital model as it required fewer resources to interpret compared to the text.

11.5.5 Research Question

The user study was conducted with the research question:

Which elements influence the orientation of a user in a step-by-step guidance process with a 3D model and natural language?

To answer the research question hypotheses were formed before conducting the user study. Each hypothesis will be introduced and compared to the results below.

1. *Participants will be able to follow the instructions easily and complete the process without any major problems.*

   This hypothesis has to be rejected partly. The process was not perceived as exhausting for most users, however reading the textual instructions was demanding. The structure of the text often lead to confusion and even some errors. However, the digital model was supportive for understanding the textual instructions and simplified the process.

2. *Participants will rely too much on the 3D model and will not actively or consciously process the steps performed.*

   This hypothesis can be confirmed for most users. To analyse this hypothesis the mismatch between the text and digital model in the step to turn the valves was introduced. With this, it can be analysed whether users rely on the text or digital model. As most users turned the valve according to the digital model it seems that the reliance is on the digital model and no comparison was made any more with the text instructions.

Further have most users always compared the text instructions with the digital model and later compared their action with the physical model with the digital model. Very rarely was the action performed only based on the textual instructions without the comparison with the digital model. It can therefore be assumed that the actions are not performed actively based on the understanding of the instructions and only a comparison to the digital model was made.

3. *The visual cues in the 3D model will help participants to process the information easily.*

   This hypothesis can be confirmed. The visual cues were perceived as very supportive during the process and helped users to understand the textual instructions in cases of uncertainty or confusion. The comparison with the digital model was usually made immediately after reading a certain fragment of the sentence for understanding.

4. *The visual cues will immediately be recognised as the elements which are relevant in the current step.*

   This hypothesis can be confirmed partly. Most users understood the visual cues in the step of the syringes immediately. It was clear that the blue syringes are the current ones where as the white syringes are the already placed ones. However in the later steps some users had to learn what the visual cues meant and that they belong to the current step to be performed.

When combining the argumentation for the hypotheses and findings of the case analysis the following elements can be observed to influence orientation:

Digital Model

- Prominent Elements as Landmarks

  In wayfinding people use prominent landmarks for orientation [62]. This behaviour could also be observed in the user study where users used prominent elements as landmarks for orientation. As for example the first syringe which has to be placed horizontally or syringe 4 which is larger than the other ones. When placing syringe 5 or connecting the first tube the larger syringe 4 was often used for orientation and was commented as "placing it next to the big syringe".

- Visual Realism

  The digital model had a high level of visual realism and showed many details of the physical model. The syringes and valves also had a high visual realism by modelling the elements in great detail. Only the connection of the tubes was visualised with low visual realism, as it would have been difficult to see the tubes as they are very thin.

  The high visual realism of the machine itself helped the user to orientate and find out where to place the elements. The combination of the high visual realism machine and low visual realism cues for the tubes yielded very good results. Users immediately recognised the visual cues and knew what actions to perform. However, combining the high visual realism machine with the high visual realism elements can lead to confusion. Because the syringes were modelled with a high level of detail, even a slight mismatch between the physical and digital elements was detected, sometimes leading to confusion. This risk of potential mismatch is reduced with the low visual realism visual cues.

- Mental Separation of Elements

In order to facilitate orientation and to derive the necessary actions from the text and the digital model to the physical model, the users mentally divided the machine into several parts. Some divided it into a left, right and bottom part and referred to it as such verbally, while others divided it into a top and bottom part. Carrying out an action in one of the mental part allowed them to ignore the other parts and focus only on the relevant section. This behaviour was observed when an error occurred. Only when an action was performed in the relevant mental part was it recognised by comparison with the digital model.

Natural Language

- Structure of the Text

  As most of the users did not read the sentences until the end, they missed out on important information, which led to confusion. Also, many users compared the text to the digital model for understanding as soon as an action or meaning could have been deduced from the sentence fragment. The structure of the text is therefore highly relevant when giving instructions. The relevant information has to be placed at the beginning of the sentence in order to increase the possibility that it will be read.

- Provided Information & Mental Models

  The textual instructions were given with indication to the number of the syringe and the number of opening on the syringe holder with directional information such as to the right or left. The user used this information for orientation, counted the openings and ensured they had the correct syringe number. As obvious as it seems that the users used the information provided by the author, it shows the importance of thinking about the mental model of the potential users. If the information provided differs from the mental model, there may be confusion or uncertainty. In one case, the user had a different counting behaviour than the author of the instructions and confusion arose due to this discrepancy. This could only be solved by comparing the text to the digital model.

Combination

- Colours

  The visual cues in the model were recognisable due to the coloured highlighting compared to the grey model. Also the currently relevant elements are coloured differently to the previously relevant elements. In the step to turn the valves the elements were coloured differently depending on the direction in which the valves have to be turned. The users could not make the connection easily with the textual information and highlighting the text in the same colour would have been supportive. In the step to connect the tubes most users associated the colour with the object and referred to it as "the blue tube" or performed the step with the comment "blue to blue". This indicates that the colouring of the elements is used for orientation.

  Some users also mentioned that the model itself could have colours to simplify orientation. Instead of giving directional information the instructions could have been given with colours as for example "place syringe 2 in the second opening on the blue part of the machine".

- Clearly Distinguishable Elements

  During the process, users have to use several items and place them in the correct position on the machine. The syringes were labelled with numbers for identification, whereas the container and the tubes were not labelled. Users could clearly identify which syringe to use as the text referred to the number of the syringe. However, at the beginning of the process some users were confused about syringe number 6 as they were not sure whether it was 6 or 9. However, when they had to perform actions with the tubes, either in the machine or placing them on the machine, they were not sure which item the text referred to. The tubes were indicated in the text with directional information, such as “the tube in the middle of the machine”. With this information, it was not clear to most users which tube to use, and they compared the text to the digital model for clarification, taking more time. In cases where the elements could only be clearly identified from the text, the actions were sometimes carried out without comparing it to the digital model.

- Size of the Instruction

  There were differences in the extent of the instructions given in each step. For example, for placing the syringes in one part, instructions were given for two syringes in one step. For more complex steps, such as connecting the tubes, the extent was smaller or simpler steps, such as turning all the valves in the bottom part up, were given in one sentence. It could be observed in the study that the smaller the scope of the instructions was, the easier it was to carry out the actions. The user did not have to remember certain parts of the sentence and could perform the action immediately after reading it. They also did not need to remember which action they had just performed in order to continue reading the sentence. In the shorter instructions, they could move on immediately to the next step. Many users also criticised the length of the instructions and wished for smaller steps.

- Connection of Text and Model

  The user study showed that users often missed the link from the text instructions to the digital or even physical model. It was suggested that the colour of the visual cues from the digital model should be used in the text instructions, or even that the physical model should be highlighted with the appropriate colour. As the text was placed on one side of the window and the digital model on the other, it was perceived as two separate sources of information that needed to be linked. Some users also mentioned that it would be helpful to have an overlay on top of the physical model and place the instructions directly on top of it. Or, if an overlay is not technically possible, bring the text instructions closer to the digital model to make the connection easier.

- Consistency of the Instructions

  After a few steps with the instruction process, the users showed a learning effect and recognised the structure of the instructions. Understanding the structure made it easier for the users to follow the instructions as less cognitive effort was required. However, when the structure changed, confusion and uncertainty arose. For example, one user has red-green colour blindness and was unable to see the visual cues in the step to turn the valves. This inconsistency in the instructions led to confusion and the user was unsure of what to do. It is likely that another user may have mild red-green colour blindness as the user had severe problems with the identification of the green visual cues. This user was also confused as to why the structure of the instructions had changed and had difficulty following the required steps.

- Cognitive Effort for Interpretation

  Most users reported that it was not difficult to follow the process and perform the necessary steps. However, the most demanding part was reading the instructions and making the connection with the model to understand the required actions. Many had to read certain steps several times to understand the instructions, which required more cognitive effort and was perceived as exhausting.

- Spatial Proximity of Elements

  In the study it could be observed that the majority of the users constantly compared the textual instructions with the digital model and the digital model with the physical model. The result is a triangle of fixations for the understanding of the instructions for the execution of the necessary actions on the physical model. In order to allow for a constant comparison in the triangle of fixations, the elements should have a high degree of spatial proximity. As Apple Vision Pro automatically places the window in the centre of the view in which the eye tracking was performed, many had the window slightly to the right. Although they were given instructions on how to move the window in the digital space, none of the users changed the window, although they mentioned that the placement of the window was not very comfortable for them and that they would have preferred the digital window to be closer to the physical model.

11.6 Implications

The user study has yielded some interesting findings where some seem relevant for most spatial computing applications. For this, the points are analysed and in more general terms summarised to implement them in the constructed model.

The relevant terms in each level are elaborated below:

Embodied

- Mental Model

  The study showed that users have some preconceptions about certain elements, which are shaped by their previous knowledge. These preconceptions can also be referred to as a *mental model*, which is different for each user and depends heavily on a person's previous knowledge and experience. The mental model also has an influence on the interpretation of a situation and therefore also on the behaviour. In order to better understand both, the mental model and the user's prior knowledge should be analysed and is therefore added to the model with the question: What prior knowledge shapes the mental model?

Embedded

- Spatial Arrangement

  In the study, most participants continuously compared the digital content with the physical model. Depending on the placement of the digital content, larger movements of the head were required, which can be tiring and demanding. The digital content was also constructed as a separate window, although some participants mentioned that they would have preferred a stand-alone model that could be placed next to the physical model for faster comparison. This suggests that the spatial arrangement of the digital and physical content is important and should be considered also in combination with the spatial link of elements.

- Landmarks

  To compare the digital and physical spaces, landmarks are used for orientation. The landmarks can be prominent elements in the digital or physical space that are clearly identifiable.

The identified landmarks serve as orientation points that support and allow faster interpretation and connection of the space.

- Structure of Elements

  The way elements are structured, either in digital or physical space, is crucial and has a strong influence on how the situation is perceived and how the interaction is shaped. In the user study, the text was perceived as poorly structured and important information was missed, leading to confusion. Not only the structure but also the consistency of the structure is important, as the users learned the structure after a certain number of steps, and a change in the structure again leads to confusion and uncertainty.

- Spatial Mental Model

  The structure of the elements also shapes the users' spatial mental model. Users create a mental model of the digital and physical space in which they either use the given structure of elements or create their own structure. This is done by dividing the given elements into smaller parts for orientation and interaction. The spatial mental model of the users should be taken into account and not interfere with their understanding of the structure.

Enacted

- Feedback Loop

  The user's interaction influences the perception, which creates a feedback loop, as the perception influences the interaction. This behaviour was also seen in the user study where certain interactions influenced the perception of further steps. This is also an important point when considering 4E cognition, where the situation is seen as a whole and perception is not a linear process that ends with the interaction. It is therefore important to analyse behaviour and perception as a feedback loop that influences each other.

Extended

- Provided Information

  In the user study the user used the information provided for orientation. In the text, certain elements with directional information were mentioned and therefore the user oriented himself or herself by this information. It is therefore important to take into account what information is provided and in what form, and also what information is not provided, as some of the users were also missing certain elements.

- Support

  In the model, present and absent people are considered at the enacted level. However, it should also be considered what support these people can give in a particular situation, as this also influences cognition.

Revised Model

Figure 73 shows the revised model with the findings from the user study implemented.

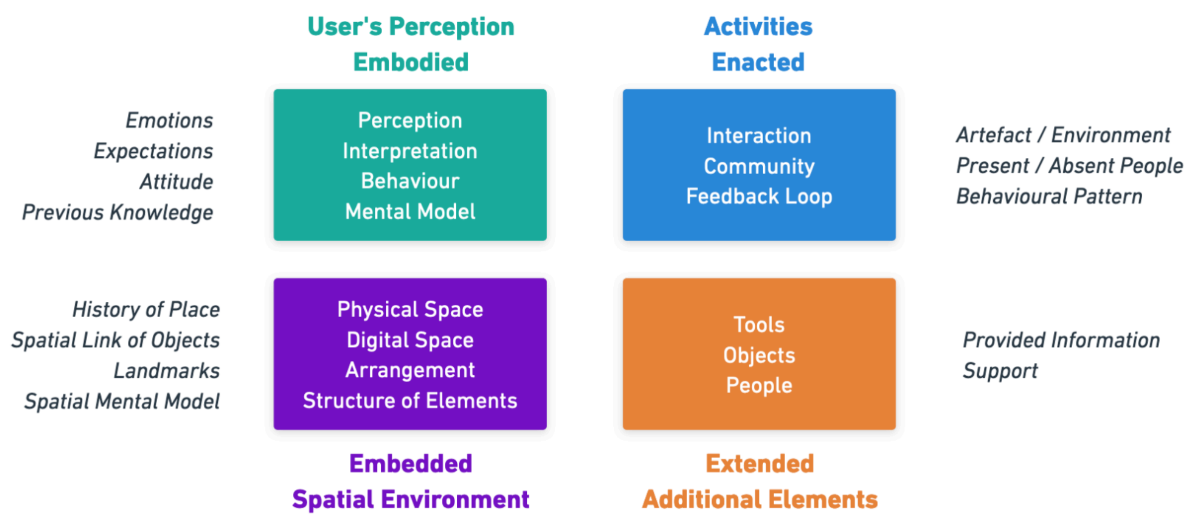


Figure 73: Revised Model with findings from user study considered

11.7 Discussion

The user study revealed some interesting points about how orientation is facilitated in spatial computing. Combined with the hypotheses formulated prior to the study, some behavioural patterns became visible. These findings were generalised and implemented in the constructed model. However, not all findings could be generalised and are only relevant to the specific case of the study. Nevertheless, with the evaluation framework derived from the constructed model, the generalised points should help the designers to uncover these findings.

The study showed that users continuously compared the text with the digital model and also compared the physical model with the digital model. Many steps were carried out only by comparison with the digital model. This may indicate a high degree of reliance on the digital model, leading to a behaviour of simply copying what is visualised, rather than actively processing the actions. In turn, this could have the effect of a lower level of recall of the steps that were performed, as they were not consciously processed. The example of wayfinding with and without digital tools could also be compared for the behaviour of not actively process information. Wayfinding with real-time information on the map about where the user is leads to less recall of the route. This could be even higher with AR wayfinding, where the user does not even have to translate the allocentric map into an egocentric execution, but simply follows the information provided on the smartphone screen through AR.

It would also be interesting to evaluate whether the visual reliance would be even higher if the digital model would hold more information. Many participants mentioned that the text instructions were necessary because the digital model did not provide all the information. It could be hypothesised that with this change, the reliance on the digital model would be even higher and that the participants would read even less. Another change to the study design could be to provide more guidance through colour rather than numbers or directional information. Many participants mentioned that they found the visual cues very helpful and could be used even more to give instructions, also in combination with the physical model.

The study also raises the question of how the results would change if it was carried out in a laboratory rather than a classroom, and if the syringes were not empty. The current set-up was clearly recognisable as a test set-up where there would be no consequences if a mistake were to be made. The users also clearly identified the situation as such and had no constraints in interacting with the elements or the model. In reality, however, the model is placed in a laboratory and the process has to be completed under time pressure. The study only asked users to avoid errors and did not mention time pressure. All these variables could change the results.

Another interesting finding of the study is the perception of the steps with low visual realism. Even with the high level of abstraction, the elements were clearly identified by the participants. It would be interesting to evaluate how interaction and orientation would change if the syringes were also visualised with low visual realism. A potential research question would also be whether low visual realism leads to lower cognitive load.

As well as the situation in which the study took place, the choice of technology also influenced the results. Because the Apple Vision Pro is a new device with a cutting-edge UX, users were excited to test the new device, which influenced their attitudes and emotions. The resolution and interaction capabilities are not comparable to current state-of-the-art technologies such as Meta Quest Pro, which has lower resolution, limited and tiring interaction, higher motion blur and less accurate hand and eye tracking.

12 Evaluation Framework for UX in Spatial Computing

As the literature review has revealed there are currently limited possibilities to evaluate the user experience in spatial computing and no coherent methodology is existing. Extensive research is required to find points for UX consideration and how to evaluate them, which is not practical for UX designers in practice.

This master thesis should be a first step to fill this gap and provide a test instrument to evaluate the UX in spatial computing holistically. To achieve this goal, a literature review was conducted to identify relevant issues based on the term of situatedness. This should reveal relevant aspects in the areas of space, place, community, time and activity. Based on the findings, a model was constructed to show the influences on UX in spatial computing. This model was further iterated and combined with 4E cognition theory to emphasise the user's interconnectedness with the situation and the different elements in the situation.

12.1 Structure of the Evaluation Framework

The constructed model has been translated into an evaluation framework. This should help designers and developers to identify problems in spatial computing applications. The evaluation framework shows the different items and levels from the constructed model, but in a table to be filled in by the user's observed behaviour during the use of the application. As the relevant items are only mentioned with keywords, questions were formulated for each item to help designers and developers to understand the keywords and consider the relevant aspects of the behaviour.The different levels of 4E cognition and the associated keywords are illustrated below (Figure 74 - Figure 77), along with questions formulated to support understanding. The framework is divided into steps of interaction, where for each step the relevant questions of the respective level are to be answered.

With the division of each action and the analysis according to all four dimensions, it is possible to identify issues in the UX. However, the framework only shows the issues. The transformation from uncovering an issue to designing potential improvements still needs to be done by the designer/developer as usual in usability testing.

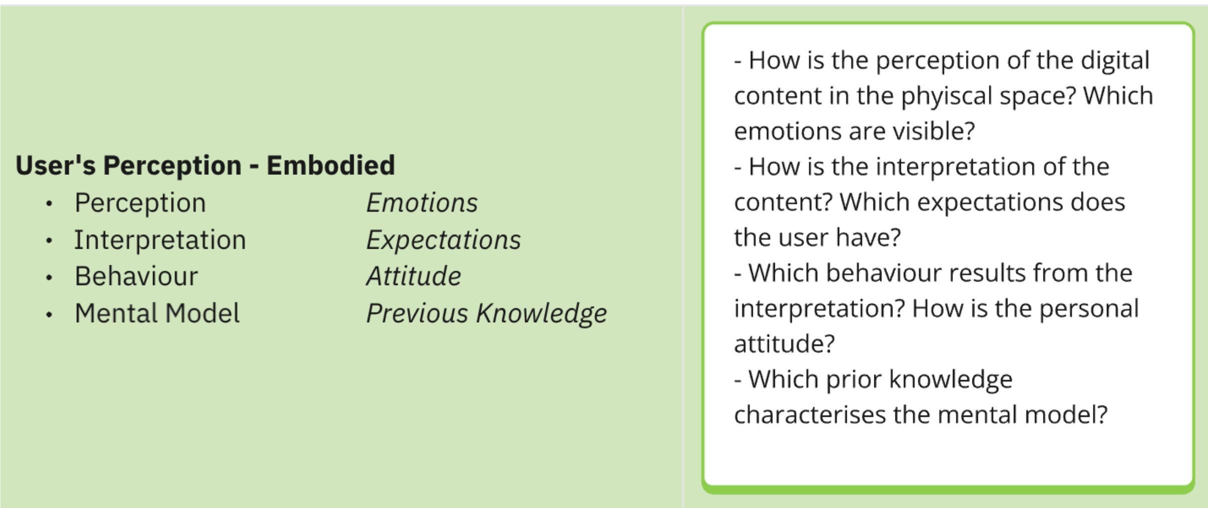


Figure 74: Keywords and associated questions of Embodied from Evaluation Framework

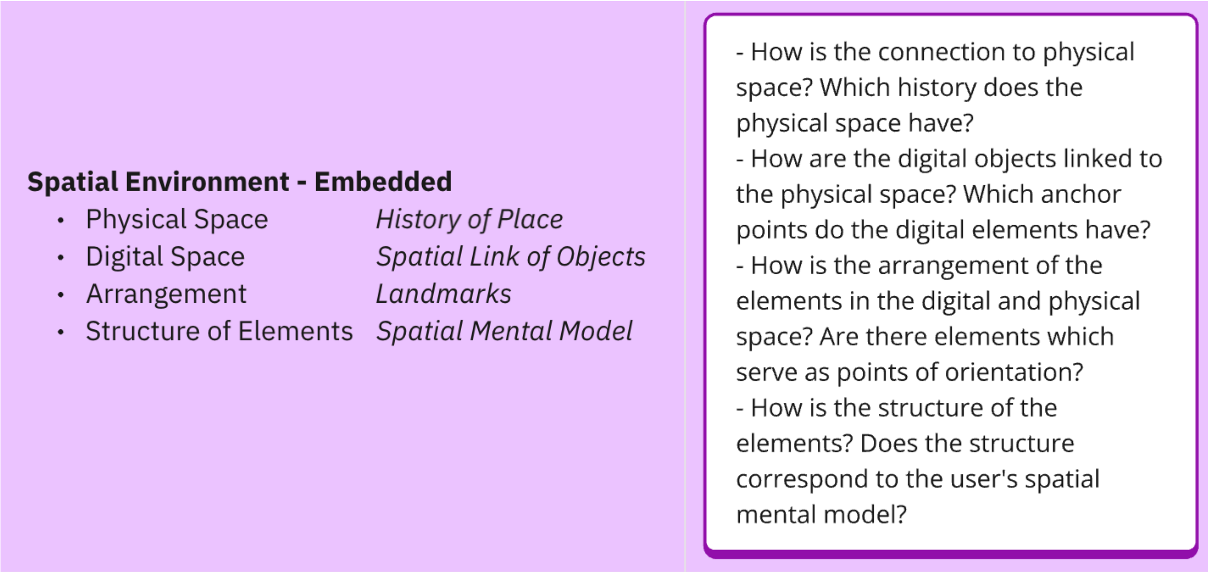


Figure 75: Keywords and associated questions of Embedded from Evaluation Framework

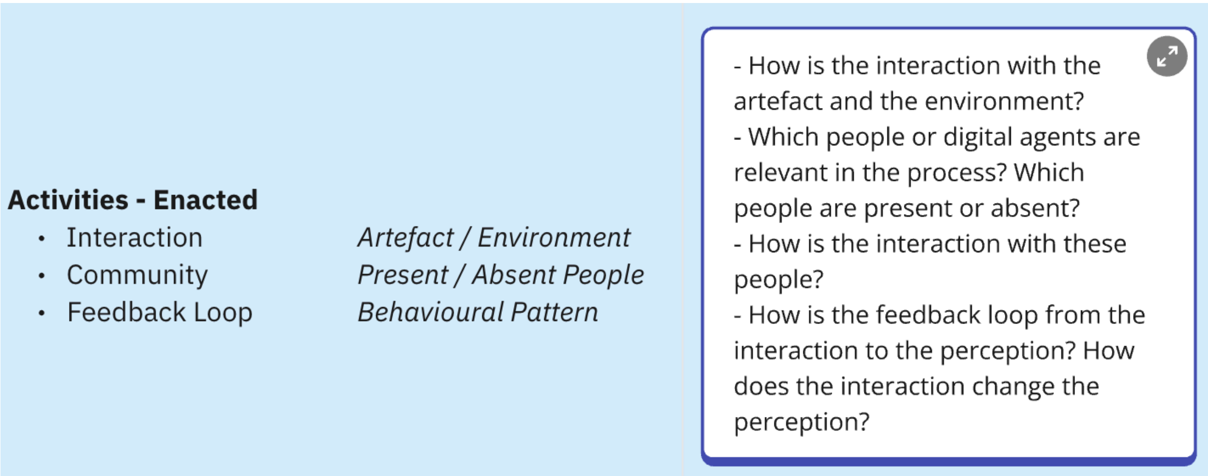


Figure 76: Keywords and associated questions of Enacted from Evaluation Framework

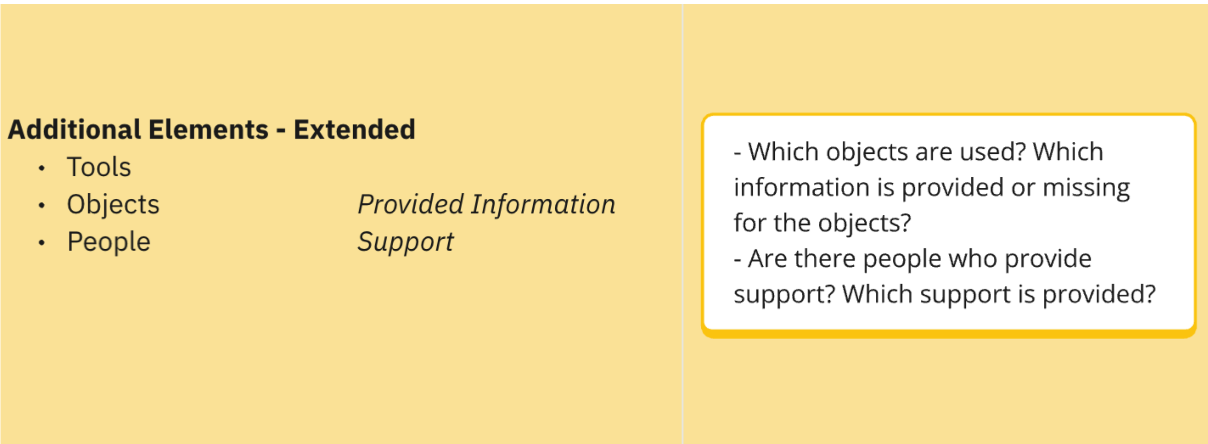


Figure 77: Keywords and associated questions of Extended from Evaluation Framework

12.2 Innovation

Current evaluation methods are usually derived from 2D interactions with mobile phones or desktop applications. However, these methods neglect many aspects relevant to spatial computing, where the user interacts with 3-dimensional space and context is much more important than in 2D applications. Not only is there a gap in evaluation methods, but the state of the art in the design of spatial computing applications lacks general guidelines on what aspects are relevant to the user experience of these applications. A versatile tool is needed to deliver and evaluate the user experience of spatial computing applications, allowing for flexibility depending on the technology and use case.

Current methods do not consider the situation as a whole or focus only on the digital content of the application. In spatial computing, however, the entanglement of physical and digital space needs to be considered, as well as the spatial interaction of the user in both spaces. The spatial component is very important, but is usually not included in the evaluation.

12.2.1 Holistic Evaluation

The created evaluation framework is therefore a first step to fill the gap in the current state of the art and to support designers and developers in their daily work. The evaluation framework tries to consider the situation as a complex system with feedback loops and combines spatial computing with 4E cognition.

4E Cognition and Spatial Computing

In spatial computing we are interacting with the environment and it is through physical interaction that we perceive the world, interpret it and interact with it accordingly. The evaluation framework combines spatial computing with 4E Cognition, which states that cognition is a process of the brain and how it interacts with the body and the physical and social environment [149]. The framework takes into account elements from both the physical and digital space, as well as the user's interaction with both spaces. It also considers the social aspects of the situation and how they influence it. With this combination, the user and the environment are no longer seen as separate, but highly intertwined. The user is situated in the space and influences it as a part of the situation as a whole. This emphasises another aspect of the framework, which no longer considers elements as separate, but analyses a situation holistically.

Interaction as Complex System

Some theories see interaction as a linear process that ends with the action through the user's body. More recent theories see interaction as a complex system with feedback loops. Especially when interacting with 3-dimensional content there is a feedback from the body to the cognition that needs to be considered. The evaluation framework takes this feedback loop into account. It is divided into several steps of actions performed by the user to highlight the changes resulting from previous steps, as a change in one element may result in changes in many other elements of the situation. This not only allows the interaction to be analysed as a complex system, but also gives the opportunity to analyse a problem and find the potential cause in previous steps, as cause and error are often not close in time or space.

Collection of relevant aspects

The literature review showed that many aspects are relevant in the context of spatial computing. However, extensive research is needed to find relevant points for consideration, especially beyond the technical aspects of space. The evaluation framework combines different aspects that are relevant in the situation of spatial computing applications, not only considering the space itself. Considerations based on the notion of situatedness are implemented to provide a holistic way of including place, activity, time and community. This allows us to consider that the user's emotions associated with the space influence the perception as well as the norms and values of that place. Furthermore, it is not seen as the user acting with the technology, but rather acting through the technology to achieve a certain goal. Interaction does not occur in isolation. It is usually surrounded by other interactions, not only with the space, but also with social interactions, either planned or unplanned. The social interactions also shape the place as we perceive it and influence our behaviour.

This holistic view of a situation incorporated in an evaluation possibility is rather uncommon where as instruments usually focus on one single aspect, neglecting the influence of the entire situation. The developed evaluation framework should provide the possibility to analyse a situation holistically and to reveal problems in different areas.

12.2.2 Implications

Spatial computing applications are currently evaluated using a variety of tools. Depending on the time and knowledge of the team, several tools from 2D interaction and VR are combined. However, with limited time or knowledge, traditional tools such as the System Usability Scale or Nasa-TLX are used for evaluation. However, these tools have limited suitability for evaluating spatial computing applications as they focus on only one aspect or usability in a very broad sense.

The developed evaluation framework should be a first step to fill this gap and provide a holistic way to evaluate spatial computing applications. This should allow teams with little knowledge or time constraints to still consider many relevant aspects in the context of their application to identify issues and provide a good user experience.

12.3 Use of the Evaluation Framework for the User Study

The items included in the framework were derived from the literature reviewed and the results of the user study conducted. To evaluate whether the framework is usable one case of the user study is used as an example. The evaluation framework was completed for this case to analyse whether it covered all relevant findings from the study. To do this, each interaction step was analysed in detail and the relevant questions for each level were answered.

The formulated questions are helpful for filling in the framework, but it takes a lot of time to complete the whole process and as more steps are added to the framework, the questions are no longer visible on the screen and it has to be scrolled back and forth to fill it in properly. However besides the minor inconvenience of scrolling it could be filled out easily with the supporting questions and many issues could be revealed with them. However, as the levels are very close and intertwined, it is not always clear to which level a particular point belongs. This is why some points are mentioned in more than one level. Another technical issue is that the cell where the relevant points are noted are not automatically increasing or the lines are not breaking at the end of the cell.

Furthermore, it is questionable whether the same results would be obtained without the prior detailed analysis of coding the entire process to reveal problems through the framework alone. It would therefore be advisable to evaluate another application without previous knowledge by the author.

Changes based on Evaluation

As some technical problems have been discovered, a change of tool is necessary. A similar tool to Whimsical is Miro with different features. Miro offers the ability to use tables, which automatically increase the size of cells as more text is added. It also offers the functionality of *key* *information cards* that can be pinned to the left side of the screen, which is a good way to avoid scrolling to see the relevant questions when filling in a cell, as shown in Figure 78.

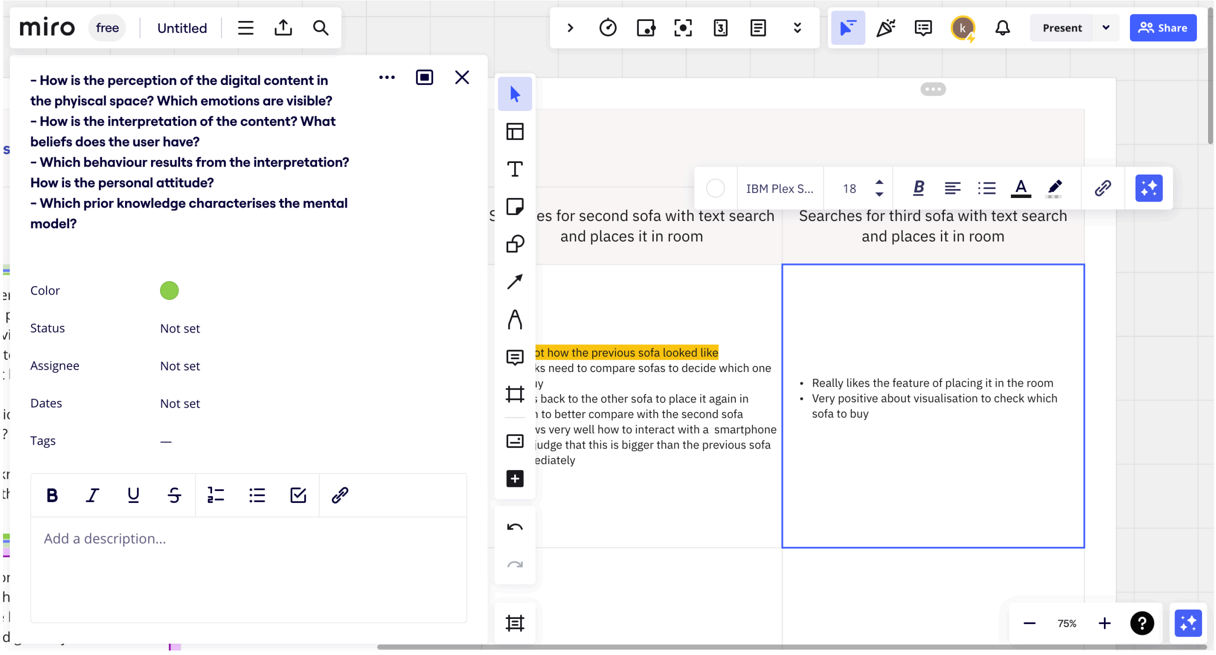


Figure 78: Screenshot from using the Miro board with the relevant questions pinned on the left

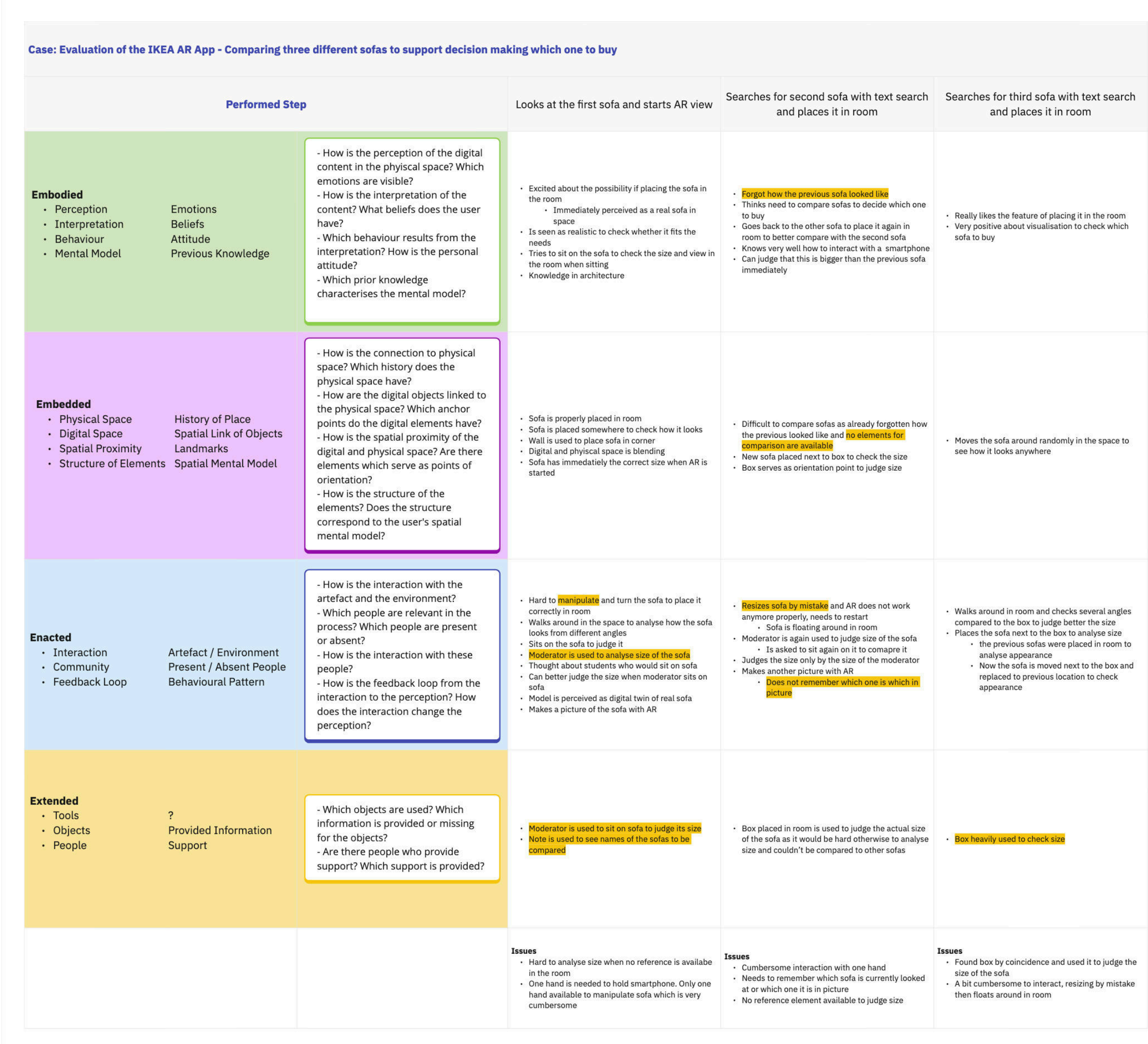
12.4 Use of the Evaluation Framework for the IKEA AR App

As it can be assumed that the evaluation framework is suitable for the cases of the user study, it is unclear how generally applicable it is. It is therefore used with the example of the IKEA AR app, which is close to the feey app that was investigated in the literature review.

Furthermore, it should be evaluated whether the framework can be used with little effort to simulate a real case as in the industry.

To use the IKEA AR app, users should find a suitable sofa for a room and analyse its appearance. They were asked to compare the sofa Linanäs in dark grey, Landskrona and Vimle in AR.

The first user was asked to evaluate the sofas to see if they would be suitable for creating a relaxation area for students at the FHNW. The test lasted 10 minutes and the evaluation framework was filled in immediately after the test.



Case: Evaluation of the IKEA AR App - Comparing three different sofas to support decision making which one to buy

Performed Step		Looks at the first sofa and starts AR view	Searches for second sofa with text search and places it in room	Searches for third sofa with text search and places it in room
Embodied • Perception — Emotions • Interpretation — Beliefs • Behaviour — Attitude • Mental Model — Previous Knowledge	- How is the perception of the digital content in the phyiscal space? Which emotions are visible? - How is the interpretation of the content? What beliefs does the user have? - Which behaviour results from the interpretation? How is the personal attitude? - Which prior knowledge characterises the mental model?	• Excited about the possibility if placing the sofa in the room • Immediately perceived as a real sofa in space • Is seen as realistic to check whether it fits the needs • Tries to sit on the sofa to check the size and view in the room when sitting • Knowledge in architecture	• Forgot how the previous sofa looked like • Thinks need to compare sofas to decide which one to buy • Goes back to the other sofa to place it again in room to better compare with the second sofa • Knows very well how to interact with a smartphone • Can judge that this is bigger than the previous sofa immediately	• Really likes the feature of placing it in the room • Very positive about visualisation to check which sofa to buy
Embedded • Physical Space — History of Place • Digital Space — Spatial Link of Objects • Spatial Proximity — Landmarks • Structure of Elements — Spatial Mental Model	- How is the connection to physical space? Which history does the physical space have? - How are the digital objects linked to the physical space? Which anchor points do the digital elements have? - How is the spatial proximity of the digital and physical space? Are there elements which serve as points of orientation? - How is the structure of the elements? Does the structure correspond to the user's spatial mental model?	• Sofa is properly placed in room • Sofa is placed somewhere to check how it looks • Wall is used to place sofa in corner • Digital and phyiscal space is blending • Sofa has immedatiely the correct size when AR is started	• Difficult to compare sofas as already forgotten how the previous looked like and no elements for comparison are available • New sofa placed next to box to check the size • Box serves as orientation point to judge size	• Moves the sofa around randomly in the space to see how it looks anywhere

Enacted
• Interaction — Artefact / Environment
• Community — Present / Absent People
• Feedback Loop — Behavioural Pattern
	- How is the interaction with the artefact and the environment?
	- Which people are relevant in the process? Which people are present or absent?
	- How is the interaction with these people?
	- How is the feedback loop from the interaction to the perception? How does the interaction change the perception?
		• Hard to manipulate and turn the sofa to place it correctly in room
		• Walks around in the space to analyse how the sofa looks from different angles
		• Sits on the sofa to judge it
		• Moderator is used to analyse size of the sofa
		• Thought about students who would sit on sofa
		• Can better judge the size when moderator sits on sofa
		• Model is perceived as digital twin of real sofa
		• Makes a picture of the sofa with AR
			• Resizes sofa by mistake and AR does not work anymore properly, needs to restart
			  • Sofa is floating around in room
			• Moderator is again used to judge size of the sofa
			  • Is asked to sit again on it to comapre it
			• Judges the size only by the size of the moderator
			• Makes another picture with AR
			  • Does not remember which one is which in picture
				• Walks around in room and checks several angles compared to the box to judge better the size
				• Places the sofa next to the box to analyse size
				  • the previous sofas were placed in room to analyse appearance
				  • Now the sofa is moved next to the box and replaced to previous location to check appearance

Extended • Tools — ? • Objects — Provided Information • People — Support	- Which objects are used? Which information is provided or missing for the objects? - Are there people who provide support? Which support is provided?	• Moderator is used to sit on sofa to judge its size • Note is used to see names of the sofas to be compared	• Box placed in room is used to judge the actual size of the sofa as it would be hard otherwise to analyse size and couldn't be compared to other sofas	• Box heavily used to check size
		Issues • Hard to analyse size when no reference is availabe in the room • One hand is needed to hold smartphone. Only one hand available to manipulate sofa which is very cumbersome	**Issues** • Cumbersome interaction with one hand • Needs to remember which sofa is currently looked at or which one it is in picture • No reference element available to judge size	**Issues** • Found box by coincidence and used it to judge the size of the sofa • A bit cumbersome to interact, resizing by mistake then floats around in room

Figure 79: Filled out evaluation framework for the use of the IKEA AR app to compare three different sofas with the AR functionality

The placement of each sofa in the room was considered as one step in the evaluation framework. For each step, an attempt was made to answer the supporting questions in order to identify problems with the application. The sections highlighted in orange illustrate problems in using the app, which are summarised in the last row of the framework for each step.

The evaluation with the framework shows that the user had some problems with the manipulation of the objects. The interaction with one hand, whereas several fingers are needed to turn the sofa, is cumbersome and leads to errors. For example, the user tried to turn the sofa, but resized it in AR. As a result of the accidental resizing, the AR no longer worked properly and the sofa floated off the floor. It also became apparent that there was no reference object in the empty room to judge the size, so the moderator was used. Later, the user happened to find a grey box and used it as a reference to judge the size of the different sofas. There is also a lack of information about which product is currently being analysed as well as when an image is taken in the AR view for later reference.

Discussion

The use of the framework with the IKEA AR app shows that it reveals several issues. The use of the framework does not require a lot of time and it is up to the user to decide how detailed the evaluation should be in relation to the amount of time that is available. The evaluation of the IKEA AR app was done in three steps, but it could be more detailed.

This short test shows that the framework is usable, but the question of whether it covers enough relevant points in the UX of spatial computing is still open.

12.5 Comparison to Design Guidelines

In some cases, design guidelines for applications also offer the possibility for UX evaluation. In AR there are several guidelines with different recommendations.

Bloksa [155] created a comprehensive list of recommendations for augmented reality applications in his master thesis. However, the guidelines are very broad and only serve as a point of consideration, e.g. *The colour of UI elements can make the perception more natural and visible.*. Furthermore, the guidelines only apply to the UI design of the application and focus solely on the digital content without considering the environment or situation in which the user is placed.

The design guidelines by Kourouthanassis et al. [156] also consider the context of mobile AR applications. However, they only present five guidelines and cover very specific points, as the list below shows:

1. Use the context for providing content
2. Deliver relevant-to-the-task content
3. Inform about content privacy
4. Provide feedback about the infrastructure's behaviour
5. Support user's procedural and semantic memory

As the list of recommendations is very short and the points are specific, the guidelines can't be used in isolation and several guidelines need to be consulted for holistic recommendations.

Vi et al. [39] combine several guidelines for AR and provide a more comprehensive list of recommendations. They also take into account the space in which the user is located, although this is very limited. The focus is still on the digital content and limited consideration is given to the combination and consideration of the physical space.

Conclusion

Design guidelines are helpful in the development and design of applications as they provide several recommendations for the application. This can help developers and designers with limited experience to deliver better solutions and consider issues that would otherwise be overlooked. However, as the brief overview shows, most guidelines for AR focus heavily on the digital content of the application and neglect the environment or situation in which the user is placed.

12.6 Placement of the Evaluation Framework

The Nielsen-Norman Group has created a three dimensional framework to illustrate the various user research methods in the landscape (Figure 80). It is structured with two axes for placement. The X-Axis distinguishes the instruments between Qualitative and Quantitative where as on the Y-Axis they are placed as either Attitudinal or Behavioural. The third dimension illustrates the context of use which is not visualised.

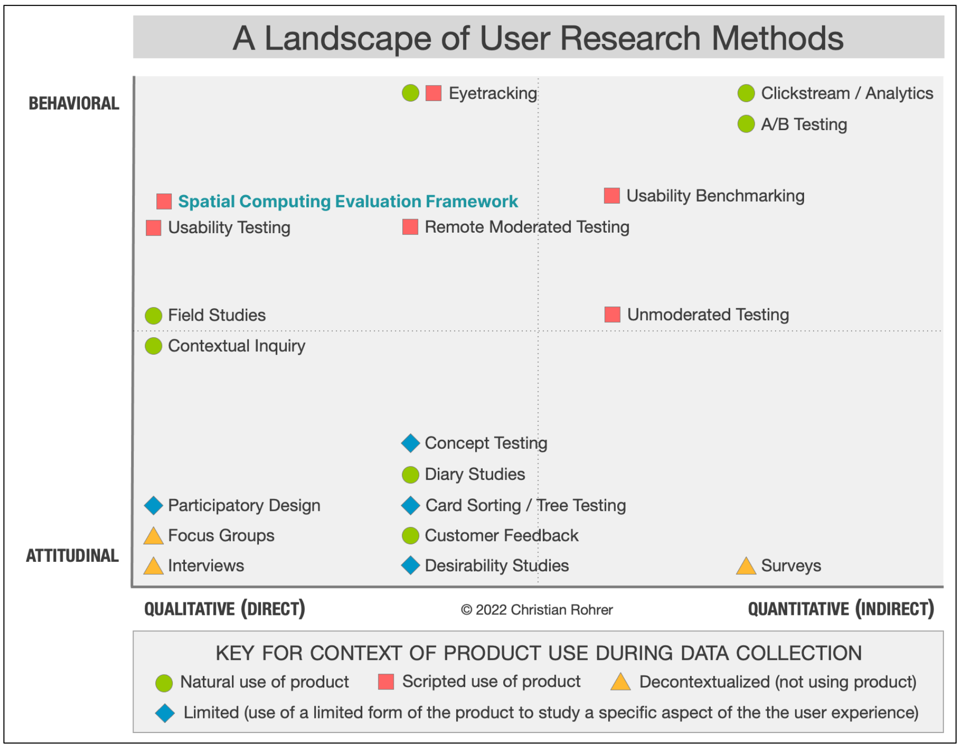


Figure 80: User Research Landscape by Nielsen-Norman Group [27]

The constructed evaluation framework can be placed in the quadrant of qualitative and behavioural instruments. As it is intended to support designers & developers in conducting usability tests and consider relevant aspects in them, the framework can be placed at the same position as the *Usability Testing* itself.

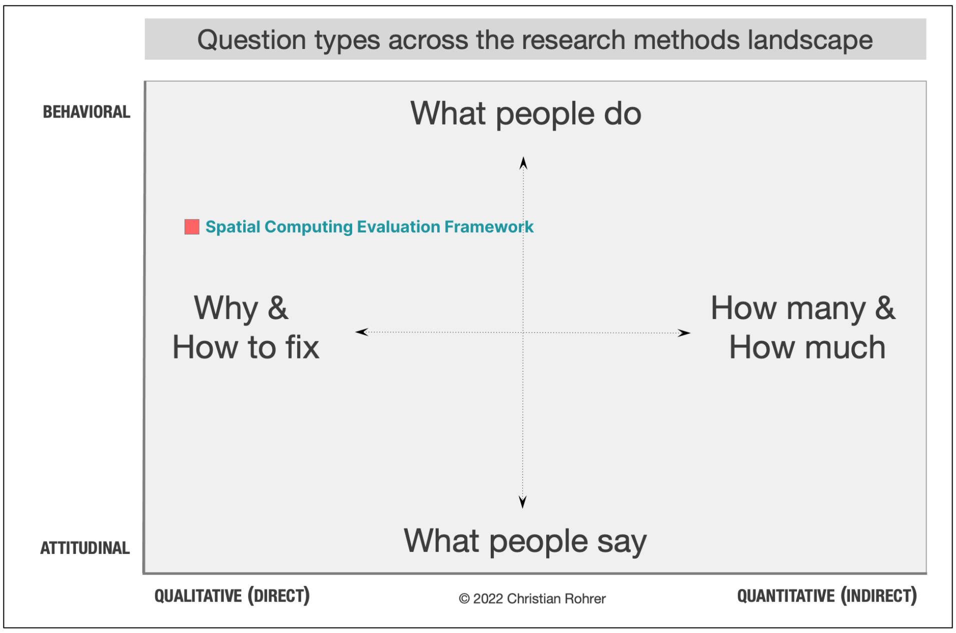


Figure 81: Question types depending on the Research Landscape [27]

Figure 81 shows which question types are usually posed depending on the two axes. As the evaluation framework is located in the upper left quadrant it is concerned with what people actually do instead of what people say. Further is it concerned with the *Why* and reasons for potential usability issues.

The quality of a usability test depends on the moderator and the questions asked. The points which should be observed are defined in advance to the usability test and are defined by the moderator. These observation points vary depending on the application and goal of the usability testing. The same applies to the questions asked during the test, which are also created by the moderator. This also shows that the observation points and formulated questions heavily depend on the knowledge of the moderator in the field of application. As spatial computing is a new area, many professional miss the specific knowledge to analyse the situation holistically.

Here the constructed evaluation framework can provide guidance and orientation in observing relevant points as well as asking contextual questions for the spatial computing application. For traditional usability tests no such guidance instrument is existing as the field of application is too broad. The support and guidance with specific points in a field of application is a novelty and can support especially in the early stage of the field. This can support the community to rapidly build up knowledge and learn many points for consideration.

12.7 Analytical vs. Empirical Evaluation

There are two different approaches for evaluating an application. In psychology this is usually done through empirical evaluation which is structured and combined with validated instruments such a questionnaires. The results can be analysed statistically or deliver also results about the usability. One such instrument is the *AttrakDiff* of Hassenzahl which analyses an application for its hedonic and pragmatic quality [157]. The result is visualised in a quadrant that indicates the attractiveness and value of the application. However, the tool does not provide any information about issues and potential improvements.

On the other hand, there is the analytical evaluation, which is usually performed in design. The application is evaluated through the expertise of the evaluator and potential improvements are explored. This process is usually done iteratively to improve the application with each design iteration. The evaluation is done without any structures as this is seen as limiting the exploration of potential solutions.

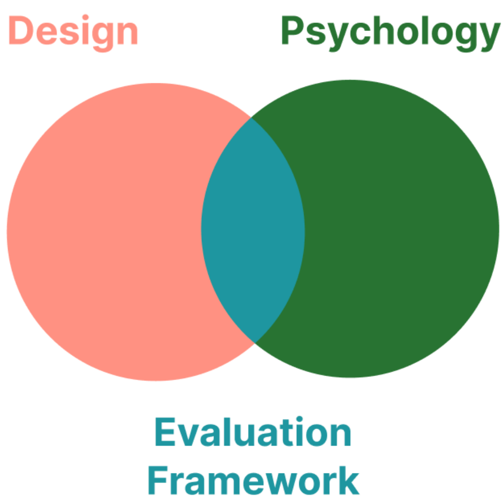


Figure 82: Placement of the Evaluation Framework, combining the analytical approach from Design and empirical approach from Psychology

The constructed evaluation framework lies at the intersection of both approaches, as illustrated in Figure 82. It combines the structured approach by providing guidance and orientation in the evaluation with relevant aspects and potential questions on these aspects without limiting the exploration. By formulating questions that can be asked to evaluate the application as a complex system, where each step of the user is analysed individually, the analysis can provide results where problems exist and the potential root cause can be found by analysing the previous steps of the user.

This combination of approaches allows the reviewer to still rely on expert knowledge to analyse the application, but to be supported by the evaluation framework as to which aspects may be relevant in the situation.

12.8 Examining the Characteristics of Augmented Reality

Azuma defines Augmented Reality by three characteristics to allow for a broad definition, independent of the technology used [158]. These are:

- Combine real and virtual
- Interactive in real-time
- Registered in 3D

The combination of real and virtual is a key aspect of spatial computing. Not only should it be combined as defined by Azuma, but in spatial computing the real and the virtual should be related to each other. The consideration of this characteristic can be found in the Embedded dimension of the framework, which analyses how the physical and digital space are connected and what elements are used to embed the digital elements in the physical space.

In addition, AR should enable real-time interaction according to Azuma. In spatial computing, interaction should not only be in real time, but should also allow interaction with spatial meaning, either in physical or digital space. This consideration can be found in the dimension Enacted in the framework, it analyses the interaction with the artefact, which can be a device, but also with the environment, either physical or digital.

Finally, the digital objects augmented in physical space should be 3-dimensional according to Azuma. In spatial computing, the objects should not only be 3D, but also blend seamlessly into the environment to allow for a high presence in the augmented space. This characteristic can be found either in the Embodied dimension, which analyses the subjective perception of space, its interpretation and the resulting behaviour. Or it can be found in the dimension Embedded, which analyses how the digital objects blend into the physical space.

12.9 Expert Reviews

To review the Spatial Computing Evaluation Framework expert reviews are conducted.

12.9.1 Real-world Usability Test with AR Application

The first review is being conducted by Reto Senn, Head of AR at Bitforge, conducting a real usability test with the Evaluation Framework. The application is already on the market as a product and currently being further developed for a new customer. The company was open to use the Framework during the usability testing to evaluate the application.

The application is a remote support solution called *Yago* for the construction and machinery industry defined as: *The app connects locally stationed experts with technicians in the field and supports them with video telephony, placing hints in AR, file streaming, and the ability to draw freely in the room in AR.*

Figure 83 shows the application in which the user can share the space through the video camera of a mobile device with a remote technician. The remote technician can annotate the screen and give the user instructions through AR.

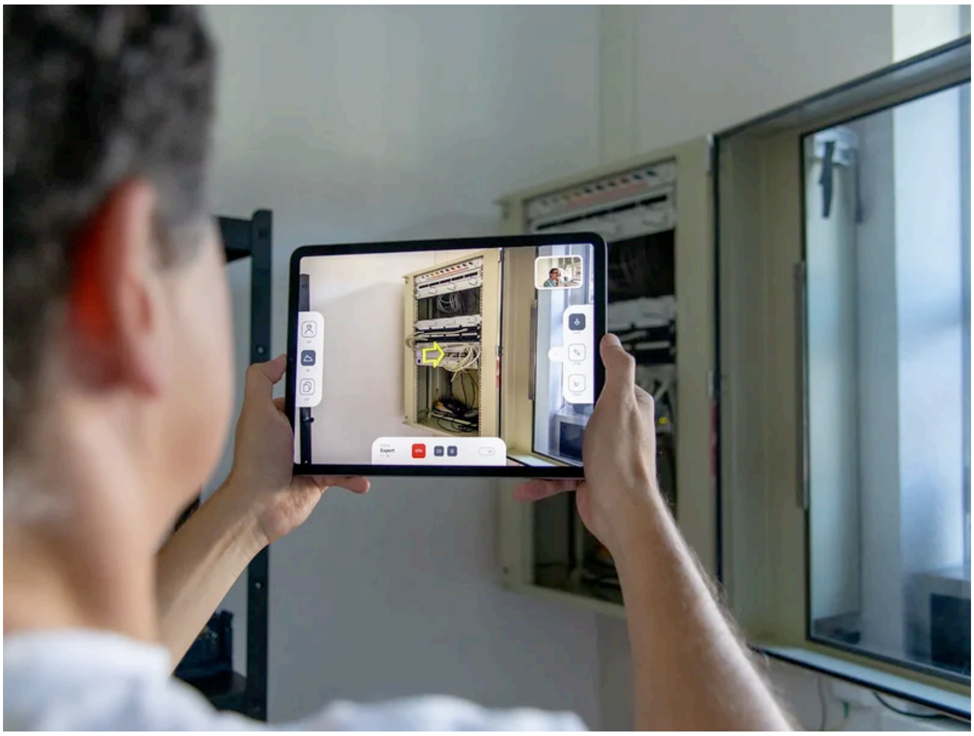


Figure 83: Remote Support Application *Yago*

The usability test was conducted with one on-site user and one remote expert in a laboratory setup. The application is intended to be used on a construction site, but as the test was conducted in a laboratory setup, certain conditions that would be present on a real construction site could not be taken into account. Nevertheless, the framework helped during the usability test by providing a structure for noting relevant points and encouraging reflection on certain elements. The framework was particularly useful in considering the connection between the physical and digital space, where previously this had been done with hand sketches. It helps to structurally note the relationship of the application to the space, interaction and perception, leading to useful insights on how to improve the application and later to analyse how the indented use differs from the real use of the application.

Using the framework provided several insights into how the application could be improved and where the developer's and real user's mental models mismatched.

The framework was explained to the developer before it was used in the usability test with a video call. The explanation of the framework revealed that it is currently too academic, especially the terms of 4E cognition. Therefore, the four dimensions of the framework were renamed to make them easier to understand. Further feedback was that the framework was not self-explanatory and that a short guide on how to use the framework should be included.

12.9.2 Methodological Review

The second review was conducted with Stefan Graser, PhD student in the area of quantitative UX evaluation in AR at the Hochschule RheinMain. The review focused on the methodological basis of the framework and its possible applications. For the review, the entire thesis was sent in advance and the framework was discussed in detail.

As the theory of 4E cognition was new to the reviewer, the four dimensions were discussed especially in combination with AR. Even though the combination is not yet established in the field, the potential of this combination is seen, especially for the holistic analysis of the environment and the situation. However, the addition of a new theory to a new instrument increases the complexity for understanding. Consequently, the burden of use increases and it is seen as important to clearly communicate the benefits of the framework to increase its attractiveness and popularity.

The reviewer is aware of the gap in the state of the art with missing instruments and the knowledge gap in the field. He sees the framework as a useful tool to guide the developer and designer during the usability test on which aspects to focus and analyse. He also mentions that there is no comparable instrument and that it is a novel contribution that could be of great value to the community in terms of knowledge building.

However, the framework is seen as interesting and with a high potential, possible drawbacks and issues could only be assessed if the reviewer would use the framework in a real usability test. He also raises interesting questions as to what applications the framework could not be used for or how certain aspects that are often invisible, such as a user's emotions, could be made visible. He also questions why the 4E cognition theory was used, as it is new to the field and if it seems promising, why no one has made this connection before.

13 Closure

Spatial computing is a new topic in HCI, and there is a shortage of knowledge and research about user experience. This is also reflected in the current state of the art, where there is no guidance on which aspects are relevant for the user experience of spatial computing applications or how to evaluate their UX. Recent publications have highlighted the knowledge gap and the need for new instruments. This thesis aims to be a first step to fill the gap and provide guidance for the UX evaluation of spatial computing applications.

13.1 Conclusion

This thesis investigates how spatial computing applications can be evaluated from an HCI perspective. Most spatial computing applications are evaluated solely from a technical point of view, neglecting the context in which the user is placed. However, it is crucial to consider not only the space, but also the emotions, the connection between the physical and the digital space, the interaction between the two and the surrounding community, as all these elements influence the user experience in spatial computing. Although recent publications highlight this perspective, the evaluation of spatial computing applications is still done with traditional 2D interaction instruments or existing instruments from virtual reality. This not only ignores important aspects of the user experience, but also fails to fully exploit the enormous potential of this technology. Figure 84 illustrates the uncertainty in UX evaluation along the reality-virtuality continuum, where methods and considerations are well established in 2D interaction, where there is a long history, and also at the other end of the continuum in virtual reality, where there is high investment from gaming and the military. However, in the center of the continuum where spatial computing is located uncertainty is high.

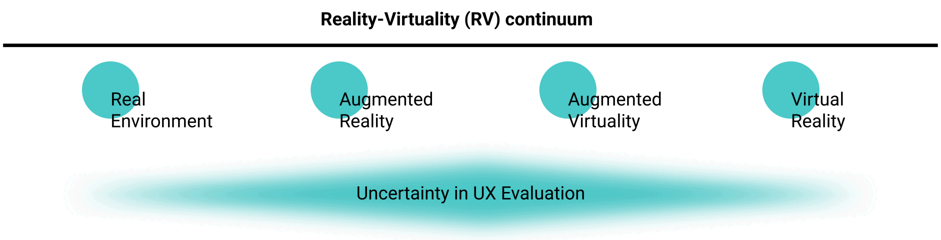


Figure 84: Reality - Virtuality Continuum visualising the amount of uncertainty in UX Evaluation [1]

As the literature review shows, there is a gap in the current state of the art regarding which aspects of the user experience of spatial computing applications should be considered and how they should be evaluated [34]. In order to answer this research question, a literature review was conducted based on three subjects of investigation covering several application domains of spatial computing applications. The three subjects are illustrated in Figure 85, a) Mobile AR: collaborative analysis of a simulation of plants in an online shop, b) Head-Mounted Displays: procedural task for mixing radiopharmaceuticals with a 3D model, and c) Spatial AR: interactive installation combined with physical books in the context of museums.

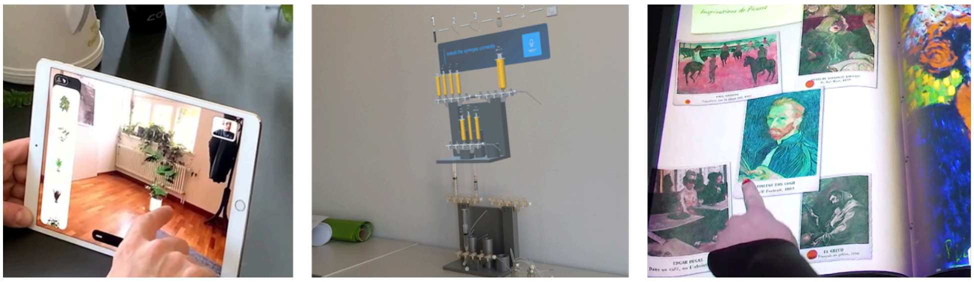


Figure 85: Subjects of Investigation for analysis which aspects may have an influence on the user experience

This review shows that there are many considerations, but they stem from other areas and extensive research is needed to find them. Also, for most of the aspects corresponding instruments exist for evaluation. However, in order to find appropriate evaluation instruments for the relevant aspects in UX, an extensive research is necessary as well. The research shows how time-consuming it is to find relevant aspects and evaluation instruments for spatial computing applications. Most of the people in practice or in research are not able to invest this amount of time and therefore use the established instruments that originate from the 2D interaction or virtual reality.

This conclusion raises the question of how these relevant considerations and evaluation possibilities can be established in practice to support designers and developers in developing spatial computing applications. The creation of a model that shows the influence of the situation poses a promising method. It illustrates in a condensed form the aspects relevant for the user experience in spatial computing and supports designers and developers in deciding which aspects should be considered in the design (Figure 86). The model is further enriched by the results of a user study with the Apple Vision Pro, which analyses the orientation of users in spatial computing.

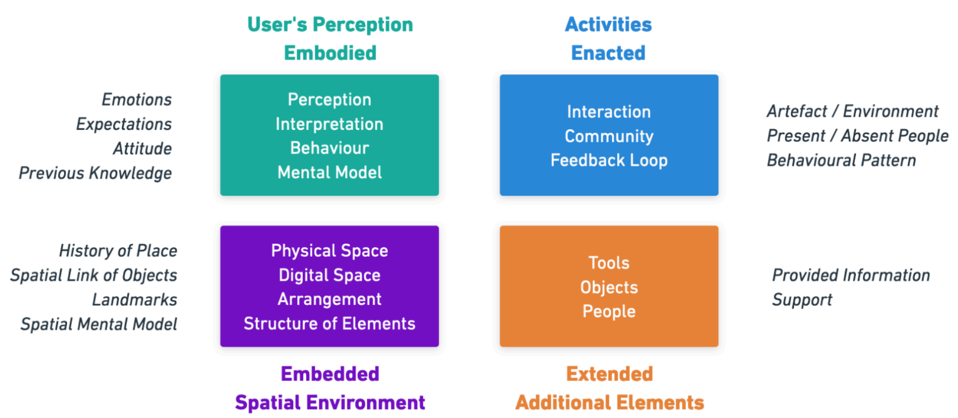
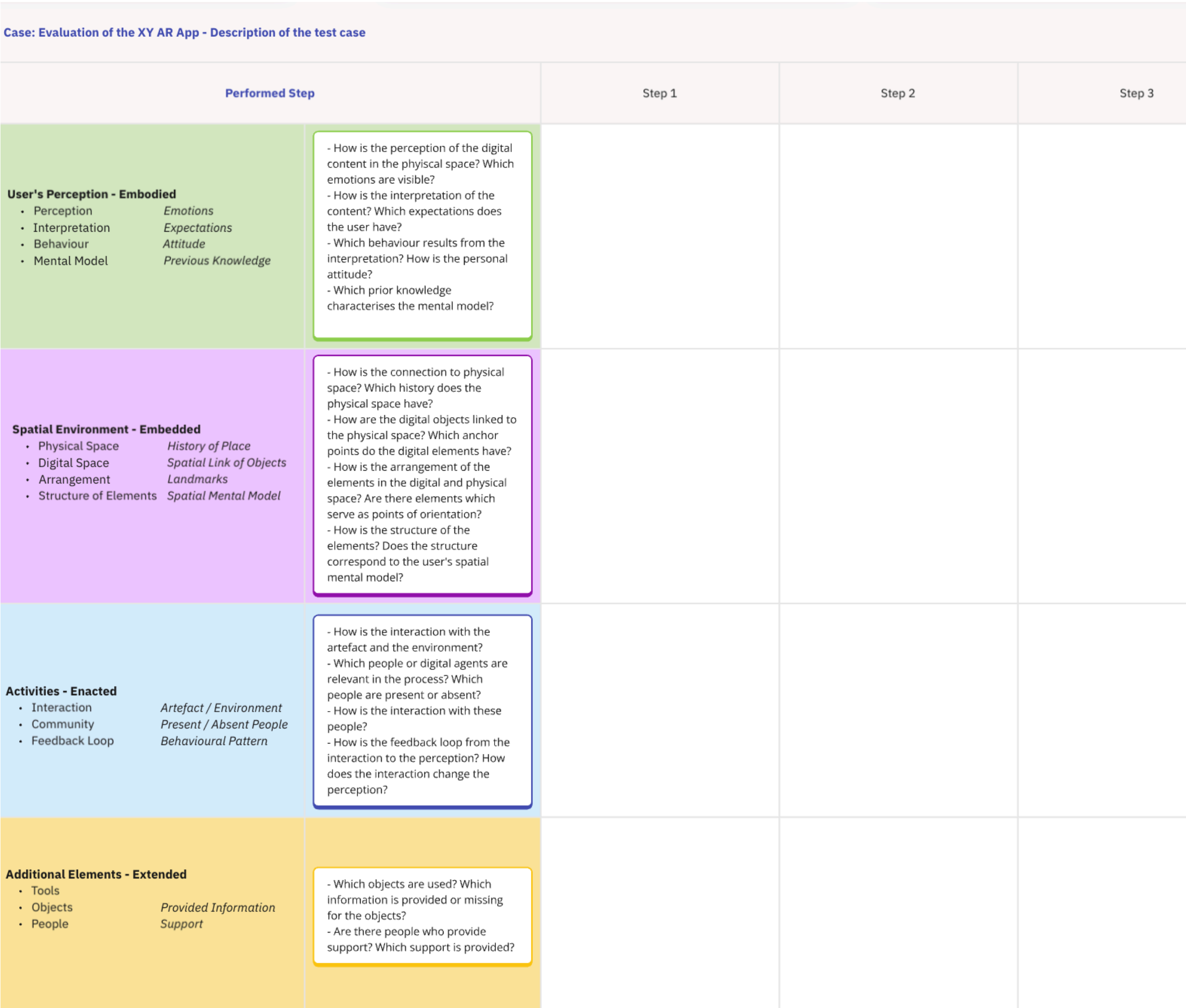


Figure 86: Model of Situational Influence in Spatial Computing based on the Theory of 4E Cognition

The model is based on the theory of 4E cognition, which considers the user's cognition as embodied, strongly connected to the spatial environment, and interacting with the environment and other people. The combination with 4E cognition allows for a holistic analysis of the situation, considering not only the space but also other aspects relevant to the interaction with spatial computing applications.

The literature review also revealed that there is no coherent methodology for the holistic evaluation of spatial computing applications. Therefore, the model showing the situational influence is transformed into an evaluation framework. In this framework, the relevant aspects of the model are formulated as questions for reflection. These questions support the analysis and interpretation of the model. It also incorporates the theory of complexity by breaking down the different steps of the user's interaction into a detailed analysis based on the four dimensions (Figure 87).



Case: Evaluation of the XY AR App - Description of the test case				
Performed Step		Step 1	Step 2	Step 3
User's Perception - Embodied • Perception — *Emotions* • Interpretation — *Expectations* • Behaviour — *Attitude* • Mental Model — *Previous Knowledge*	- How is the perception of the digital content in the phyiscal space? Which emotions are visible? - How is the interpretation of the content? Which expectations does the user have? - Which behaviour results from the interpretation? How is the personal attitude? - Which prior knowledge characterises the mental model?			
Spatial Environment - Embedded • Physical Space — *History of Place* • Digital Space — *Spatial Link of Objects* • Arrangement — *Landmarks* • Structure of Elements — *Spatial Mental Model*	- How is the connection to physical space? Which history does the physical space have? - How are the digital objects linked to the physical space? Which anchor points do the digital elements have? - How is the arrangement of the elements in the digital and physical space? Are there elements which serve as points of orientation? - How is the structure of the elements? Does the structure correspond to the user's spatial mental model?			
Activities - Enacted • Interaction — *Artefact / Environment* • Community — *Present / Absent People* • Feedback Loop — *Behavioural Pattern*	- How is the interaction with the artefact and the environment? - Which people or digital agents are relevant in the process? Which people are present or absent? - How is the interaction with these people? - How is the feedback loop from the interaction to the perception? How does the interaction change the perception?			
Additional Elements - Extended • Tools • Objects — *Provided Information* • People — *Support*	- Which objects are used? Which information is provided or missing for the objects? - Are there people who provide support? Which support is provided?			

Figure 87: Spatial Computing Evaluation Framework

13.2 Discussion

The developed Model showing the situational influence and the according Spatial Computing Evaluation Framework can provide knowledge as to which aspects should be considered in the user experience of spatial computing applications. It provides insight into an area that lacks guidance and research on what elements influence the user experience and increase awareness. Furthermore, the model and framework analyse space not only from a technical point of view, but holistically using the combination of 4E cognition and complexity science. This combination offers a promising way of research as it analyses the situation and interaction with the physical and digital space in which the user is immersed. This perspective is novel and may lead to other research and perspectives that recognise the huge influence of context on the user experience of spatial computing applications. However, it is unclear why the combination of 4E cognition and HCI is novel, as the first mention of it could be found in 2021 by Bennett et al. at the CHI conference [33]. Few other examples could be found in the area of learning with interactive systems [159] or experience of mixed reality in museum [160].

As the model and framework are based on a literature review conducted with three specific examples, the aspects included may differ for other subjects of investigation. Furthermore, not all aspects found are included, as they are often very specific to the use case and not generalisable. The aspects considered in the Model and Framework are held generally to allow for a broad range of application in spatial computing, as these vary heavily with the different technologies. This shows the compromise of a more specific framework that would be more supportive, but only for a specific use case or technology.

Each dimension in the Framework has a limited number of aspects to consider and held rather generally. If more aspects were included, it could be assumed that the framework would become too large and not be used due to its size and increasing complexity. On the other hand, it could also be argued that it is too general to use and therefore not useful. Therefore, the reflection questions are introduced to allow for a wide range of applications to stimulate reflection on the aspects, rather than to provide specific guidance. However, more feedback on whether the questions support reflection and whether the aspects of each dimension are relevant is needed from actual users in practice. Another open question is whether the framework is too academic and whether the terms used in the framework are understandable for people in practice. It is necessary to conduct more tests with actual users from practice to evaluate understandability and practicability.

Furthermore, in a usability test it is difficult to observe or assess some aspects of the framework. For example, a user's emotions or prior knowledge that influences interpretation. The dimension *Embodied* is very subjective and it is not always possible to observe the included aspects in it. Some information could be gathered through an interview, such as the prior knowledge that influences interpretation. However, to find out about the subjective experience, it would be useful to supplement the usability test with a questionnaire. Here, an extension of the framework to a questionnaire would be useful to assess the subjective experience of the same aspects as observed with the framework.

Even if the framework will not be used in actual usability testing, it may encourage reflection on how certain elements are designed in spatial computing applications. The model and framework highlight new considerations and opportunities for UX design. The Model and Framework may also stimulate discussions about UX and UX evaluation in spatial computing, and help to generate new insights in the community. The Model and Framework does not claim to be complete or perfect, rather it is intended to be a living artefact that can be twisted and turned in accordance with needs and raise awareness of the gap in the current state of the art. The Model and Framework are a first step addressing this gap and adding to the body of knowledge in this still small area of research. Hopefully it will spark relevant discussions for further improvement and how to support designers & developers in the creation of outstanding user experiences with spatial computing applications.

13.3 Future Work

The framework is intended to support qualitative usability studies by providing relevant aspects to consider in the user experience of spatial computing. However, usability studies are expensive and time-consuming. Therefore, usability tests are usually conducted at the beginning of a project with a few participants. However, when the product is further developed, usability tests are no longer conducted because they are too expensive. In addition, it is difficult to assess whether the product has improved with further additions. Therefore, it might be useful to transform the qualitative framework into a quantitative questionnaire. This questionnaire could assess the subjective experience of users of spatial computing applications. The questionnaire could also be used in further developments of the product to evaluate whether the quality of the user experience has improved or not. In order to develop the questionnaire, an additional study should be carried out with a specific spatial computing application. A potential application is the implementation of Lusee at the Hypothekarbank Lenzburg. As many branches have a Lusee, a relevant number of users could be found to evaluate the constructed quantitative questionnaire.

Another further development of the framework would be possible by modularisation of it. Currently, the framework is built in four dimensions with general aspects relevant to different applications in spatial computing. To be more specific, a modular structure could be created to build a framework that suits the specific application by enriching the general aspects with relevant terms for specific use cases or applications. However, this approach could raise further issues, as the complexity could increase significantly in the development of the more specific aspects, but also in the use of the framework, where the different modules need to be combined.

The combination with artificial intelligence for the analysis of the results could be a further development as well. Analysing where the UX issues are based on the notes in the framework is time-consuming. This is where the use of AI could be useful in deriving insights as to where the issues are and where the potential root cause is. In addition, AI could be used to derive implications from the issues for UX improvement. However, the feasibility of implementing this cannot be estimated as well as how well the integration with AI would work.

Finally, the framework needs further validation with real users in practice. There are still open questions about the usefulness and comprehensibility of the framework for usability testing of spatial computing applications. Furthermore, the aspects included need to be validated, as well as the questions for reflection.

14 Declaration of independence

I hereby declare that any individual work submitted for assessment is entirely the product of my own effort

- that I have correctly cited all text passages that do not originate from me, in accordance with standard academic citation rules, and that I have clearly mentioned all sources used;
- that I have declared in an “Index of auxiliary tools” all aids used (AI assistance systems such as chatbots [e.g., ChatGPT], translation [e.g., DeepL] paraphrasing [e.g., Quillbot]) or programming applications [e.g., Github Copilot]
- that I have acquired all intangible rights to any materials I may have used, such as images or graphics, or that these materials were created by me;
- that the topic, the thesis or parts of it have not been used in an assessment of another module, unless this has been expressly agreed with the lecturer in advance and is stated as such;
- that I am aware that my work may be checked for plagiarism and for third-party authorship of human or technical origin (artificial intelligence);
- that I am aware that the FHNW School of Engineering will pursue a violation of this declaration of authenticity and that disciplinary consequences (reprimand or expulsion from the study program) may result from this.

Windisch, August 2024

Name: Katja Pott

Unterschrift:

Bibliography

[1] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, “Augmented reality: A class of displays on the reality-virtuality continuum”, *Telemanipulator and telepresence technologies*, vol. 2351, pp. 282–292, 1995.

[2] Greg Madison, “MELODIX • R - Quest3 - Fictional XR App”. Accessed: Feb. 06, 2024. [Online]. Available: https://www.youtube.com/watch?v=pYgHzTdJPHw

[3] W. E. Mackay and A.-L. Fayard, “HCI, natural science and design: a framework for triangulation across disciplines”, *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques*, pp. 223–234, 1997.

[4] Y. Simmen, T. Eggler, A. Legath, D. Agotai, and H. Cords, “Non-overlayed Guidance in Augmented Reality: User Study in Radio-Pharmacy”, *International Conference on Intelligent Human Computer Interaction*, pp. 516–526, 2022.

[5] M. Letter and K. Wolf, “Tangible Version Control: Exploring a Physical Object’s Alternative Versions”, *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pp. 1–7, 2022.

[6] M. Speicher, C. Rosenberg, D. Degraen, F. Daiber, and A. Krúger, “Exploring Visual Guidance in 360-degree Videos”, *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, pp. 1–12, Jun. 2019, doi: 10.1145/3317697.3323350.

[7] A. Flade, *Der rastlose Mensch: Konzepte und Erkenntnisse der Mobilitätspsychologie*. Springer-Verlag, 2013.

[8] “Pokémon GO gets a new and improved augmented reality mode (but only on iOS) | TechCrunch”. Accessed: Feb. 07, 2024. [Online]. Available: https://techcrunch.com/2017/12/20/pokemon-go-gets-a-new-and-improved-augmented-reality-mode-but-only-on-ios/

[9] P. Määttänen, “Action and Experience a Naturalistic Approach to Cognition”, 1993.

[10] P. Määttänen, “Semiotics of space: Peirce and Lefebvre”, 2007.

[11] S. L. Ames, J. S. Wolffsohn, and N. A. Mcbrien, “The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display”, *Optometry and Vision Science*, vol. 82, no. 3, pp. 168–176, 2005.

[12] S. Aromaa, A. Väätänen, E. Kaasinen, M. Uimonen, and S. Siltanen, “Human factors and ergonomics evaluation of a tablet based augmented reality system in maintenance work”, *Proceedings of the 22nd International Academic Mindtrek Conference*, pp. 118–125, 2018.

[13] H. Cho *et al.*, “Time Travellers: An Asynchronous Cross Reality Collaborative System”, *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 848–853, 2023.

[14] R. J. Bailey, C. M. Grimm, C. Davoli, and others, “The real effect of warm-cool colors”, 2006.

[15] L. McAtamney and E. N. Corlett, “RULA: a survey method for the investigation of work-related upper limb disorders”, *Applied ergonomics*, vol. 24, no. 2, pp. 91–99, 1993.

[16] D. Tone, D. Iwai, S. Hiura, and K. Sato, “Fibar: Embedding optical fibers in 3d printed objects for active markers in dynamic projection mapping”, *IEEE transactions on visualization and computer graphics*, vol. 26, no. 5, pp. 2030–2040, 2020.

[17] P. T. Fischer and E. Hornecker, “Urban HCI: spatial aspects in the design of shared encounters for media facades”, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 307–316, 2012.

[18] W. Willett, Y. Jansen, and P. Dragicevic, “Embedded data representations”, *IEEE transactions on visualization and computer graphics*, vol. 23, no. 1, pp. 461–470, 2016.

[19] Z. F. Yong, A. L. Ng, and Y. Nakayama, “The Dyslexperience: Use of Projection Mapping to Simulate Dyslexia”, *2019 International Conference on Multimodal Interaction*, pp. 493–495, 2019.

[20] Y. Zhang, Y. Shen, W. Zhang, Z. Zhu, and P. Ma, “Interactive spatial augmented reality system for Chinese opera”, *ACM SIGGRAPH 2019 Posters*, pp. 1–2, 2019.

[21] S. Hong, Y. Jo, Y. Kang, and H.-K. Lee, “Interactive Experiential Model: Insights from Shadowing Students’ Exhibitory Footprints”, *Journal of Museum Education*, vol. 48, no. 2, pp. 92–108, 2023.

[22] S. Bakker and K. Niemantsverdriet, “The interaction-attention continuum: Considering various levels of human attention in interaction design”, *International Journal of Design*, vol. 10, no. 2, pp. 1–14, 2016.

[23] N. Wouters *et al.*, “Uncovering the honeypot effect: How audiences engage with public interactive systems”, *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, pp. 5–16, 2016.

[24] C. Parker, M. Tomitsch, N. Davies, N. Valkanova, and J. Kay, “Foundations for designing public interactive displays that provide value to users”, *Proceedings of the 2020 chi conference on human factors in computing systems*, pp. 1–12, 2020.

[25] U. Neisser, “Cognition and reality: Principles and implications of cognitive psychology.”, 1976.

[26] N. Bevana, J. Kirakowskib, and J. Maissela, “What is usability”, *Proceedings of the 4th International Conference on HCI*, 1991.

[27] “User Research Landscape by NNG”. Accessed: Jun. 20, 2024. [Online]. Available: https://www.nngroup.com/articles/which-ux-research-methods/

[28] J. Leppink, F. Paas, C. P. Van der Vleuten, T. Van Gog, and J. J. Van Merriënboer, “Development of an instrument for measuring different types of cognitive load”, *Behavior research methods*, vol. 45, pp. 1058–1072, 2013.

[29] C. Harms and F. Biocca, “Internal consistency and reliability of the networked minds measure of social presence”, *Seventh annual international workshop: Presence*, vol. 2004, 2004.

[30] P. Vorderer *et al.*, “Mec spatial presence questionnaire”, *Retrieved Sept*, vol. 18, no. 2015, p. 6, 2004.

[31] M. E. C. Santos, T. Taketomi, C. Sandor, J. Polvi, G. Yamamoto, and H. Kato, “A usability scale for handheld augmented reality”, *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pp. 167–176, 2014.

[32] J. E. Escalas and B. B. Stern, "Sympathy and empathy: Emotional responses to advertising dramas", *Journal of Consumer Research*, vol. 29, no. 4, pp. 566–578, 2003.

[33] D. Bennett *et al.*, "Emergent interaction: Complexity, dynamics, and enaction in HCI", *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–7, 2021.

[34] J. Xu *et al.*, "Spatial Computing: Defining the Vision for the Future", in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–4.

[35] D. Alexandrovsky *et al.*, "Evaluating user experiences in mixed reality", *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–5, 2021.

[36] M. Doulis, D. Agotai, and H. P. Wyss, "Spatial Interface Wahrnehmungsfelder und Gestaltungsansätze im Virtuellen Raum", *Virtuelle Welten als Basistechnologie für Kunst und Kultur?*, p. 55, 2009.

[37] K. Pott and D. Agotai, "Spatial Computing Through an HCI Lens-UX Evaluation Based on Situatedness", in *International Conference on Human-Computer Interaction*, 2024, pp. 102–113.

[38] I. Börsting, M. Heikamp, M. Hesenius, W. Koop, and V. Gruhn, "Software Engineering for Augmented Reality-A Research Agenda", *Proceedings of the ACM on Human-Computer Interaction*, vol. 6, no. EICS, pp. 1–34, 2022.

[39] S. Vi, T. S. da Silva, and F. Maurer, "User experience guidelines for designing hmd extended reality applications", *Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part IV 17*, pp. 319–341, 2019.

[40] H. Stefanidi, A. Leonidis, M. Korozi, and G. Papagiannakis, "The ARgus Designer: Supporting experts while conducting user studies of AR/MR applications", *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 885–890, 2022.

[41] D. Garg, B. Patel, R. Patel, and R. Jani, "Spatial Computing: Next Big Thing of Physical and Digital World", *ICDSMLA 2021: Proceedings of the 3rd International Conference on Data Science, Machine Learning and Applications*, pp. 211–219, 2023.

[42] J. Delmerico *et al.*, "Spatial computing and intuitive interaction: Bringing mixed reality and robotics together", *IEEE Robotics & Automation Magazine*, vol. 29, no. 1, pp. 45–57, 2022.

[43] L. A. Suchman, *Plans and situated actions: The problem of human-machine communication.* Cambridge university press, 1987.

[44] N. Bressa, H. Korsgaard, A. Tabard, S. Houben, and J. Vermeulen, "What's the situation with situated visualization? A survey and perspectives on situatedness", *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 107–117, 2021.

[45] M. Costello, "Situatedness", T. Teo, Ed., New York, NY: Springer New York, 2014, pp. 1757–1762. doi: 10.1007/978-1-4614-5583-7_470.

[46] J. Lowgren, J. M. Carroll, M. Hassenzahl, T. Erickson, and A. Blackwell, "The encyclopedia of human-computer interaction", *Interaction design foundation*, 2019.

[47] S. Bodker, *Through the interface: A human activity approach to user interface design.* CRC Press, 2021.

[48] G. Gurvitch, *The spectrum of social time*, vol. 8. Springer Science & Business Media, 2012.

[49] S. Aromaa, A. Väätänen, I. Aaltonen, V. Goriachev, K. Helin, and J. Karjalainen, “Awareness of the real-world environment when using augmented reality head-mounted display”, *Applied ergonomics*, vol. 88, p. 103145, 2020.

[50] N. Crespi, A. T. Drobot, and R. Minerva, “The Digital Twin: What and Why?”, N. Crespi, A. T. Drobot, and R. Minerva, Eds., Cham: Springer International Publishing, 2023, pp. 3–20. doi: 10.1007/978-3-031-21343-4_1.

[51] I. E. Lokka and A. Cöltekin, “Evaluating route learning performance of older and younger adults in differently-designed virtual environments: a task-differential analysis”, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 383–387, 2018.

[52] F. Meijer, B. L. Geudeke, and E. L. Van den Broek, “Navigating through virtual environments: Visual realism improves spatial cognition”, *CyberPsychology & Behavior*, vol. 12, no. 5, pp. 517–521, 2009.

[53] D. Kahl, M. Ruble, and A. Krüger, “The influence of environmental lighting on size variations in optical see-through tangible augmented reality”, *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 121–129, 2022.

[54] T. Müller, “Challenges in representing information with augmented reality to support manual procedural tasks”, *AIMS Electronics and Electrical Engineering*, vol. 3, no. 1, pp. 71–97, 2019.

[55] J. L. Gabbard, J. E. Swan, and D. Hix, “The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality”, *Presence*, vol. 15, no. 1, pp. 16–32, 2006.

[56] R. Düzgün, P. Mayer, and M. Volkamer, “Shoulder-Surfing Resistant Authentication for Augmented Reality”, *Nordic Human-Computer Interaction Conference*, pp. 1–13, 2022.

[57] M. Funk, K. Marky, I. Mizutani, M. Kritzler, S. Mayer, and F. Michahelles, “Lookunlock: Using spatial-targets for user-authentication on hmds”, *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2019.

[58] R. Duezguen, P. Mayer, S. Das, and M. Volkamer, “Towards secure and usable authentication for augmented and virtual reality head-mounted displays”, *arXiv preprint arXiv: 2007.11663*, 2020.

[59] T. T. M. Tran, S. Brown, O. Weidlich, M. Billinghurst, and C. Parker, “Wearable Augmented Reality: Research Trends and Future Directions from Three Major Venues”, *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–12, 2023, doi: 10.1109/TVCG.2023.3320231.

[60] C. Colby, “Perception of Extrapersonal Space: Psychological and Neural Aspects”. Pergamon, Oxford, pp. 11205–11209, 2001. doi: https://doi.org/10.1016/B0-08-043076-7/03501-4.

[61] J. Zhao, B. E. Riecke, J. W. Kelly, J. Stefanucci, and A. Klippel, "Human spatial perception, cognition, and behaviour in extended reality", *Frontiers in Virtual Reality*, vol. 4, p. 1257230, 2023.

[62] B. Tversky, "Cognitive maps, cognitive collages, and spatial mental models", *European conference on spatial information theory*, pp. 14–24, 1993.

[63] A. Román, A. Flumini, P. Lizano, M. Escobar, and J. Santiago, "Reading direction causes spatial biases in mental model construction in language understanding", *Scientific reports*, vol. 5, no. 1, p. 18248, 2015.

[64] J. Wither, S. DiVerdi, and T. Höllerer, "Annotation in outdoor augmented reality", *Computers & Graphics*, vol. 33, no. 6, pp. 679–689, 2009.

[65] T. Schubert, F. Friedmann, and H. Regenbrecht, "The experience of presence: Factor analytic insights", *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 266–281, 2001.

[66] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire", *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[67] C. D. Wickens and A. L. Alexander, "Attentional tunneling and task management in synthetic vision displays", *The international journal of aviation psychology*, vol. 19, no. 2, pp. 182–199, 2009.

[68] A. Schankin, D. Reichert, M. Berning, and M. Beigl, "[POSTER] The Impact of the Frame of Reference on Attention Shifts Between Augmented Reality and Real-World Environment", *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 25–30, 2017.

[69] Y. Gao, Y. Liu, J.-M. Normand, G. Moreau, X. Gao, and Y. Wang, "A study on differences in human perception between a real and an AR scene viewed in an OST-HMD", *Journal of the Society for Information Display*, vol. 27, no. 3, pp. 155–171, 2019.

[70] G. Ballestin, F. Solari, and M. Chessa, "Perception and action in peripersonal space: A comparison between video and optical see-through augmented reality devices", *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 184–189, 2018.

[71] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness", *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[72] M. Katell, F. Dechesne, B.-J. Koops, and P. Meessen, "Seeing the whole picture: visualising socio-spatial power through augmented reality", *Law, Innovation and Technology*, vol. 11, no. 2, pp. 279–310, 2019.

[73] R. Schmied-Kowarzik, P. Reisewitz, L. Kaschub, R. Rodeck, and G. Wende, "An Approach for Visual Realism Complexity Classification of 3D Models in Virtual and Augmented Reality", *International Conference on Human-Computer Interaction*, pp. 192–202, 2023.

[74] R. L. Schmied-Kowarzik and V. Paelke, "Examining the Importance of Realism in Virtual Reality Therapy Environments for People With Specific Phobias", *GI VR/AR Workshop*, 2021.

[75] A. Skulmowski, “Realistic details impact learners independently of split-attention effects”, *Cognitive Processing*, vol. 24, no. 2, pp. 187–198, 2023.

[76] D. Jonassen and Y. H. Cho, “Externalizing mental models with mindtools”, *Understanding models for learning and instruction*, pp. 145–159, 2008.

[77] S. M. Fiore and J. W. Schooler, “How did you get here from there? Verbal overshadowing of spatial mental models”, *Applied cognitive psychology*, vol. 16, no. 8, pp. 897–910, 2002.

[78] K. A. Hanisch, A. F. Kramer, C. L. Hulin, and R. Schumacher, “Novice-expert differences in the cognitive representation of system features: Mental models and verbalizable knowledge”, *Proceedings of the Human Factors Society Annual Meeting*, vol. 32, no. 4, pp. 219–223, 1988.

[79] F. Ishizawa and T. Nakajima, “An enhanced real space through temporally connecting real and virtual scenes”, *Ambient Intelligence-Software and Applications–7th International Symposium on Ambient Intelligence (ISAmI 2016)*, pp. 11–19, 2016.

[80] F. Ishizawa, M. Takahashi, K. Irie, M. Sakamoto, and T. Nakajima, “Analyzing augmented real spaces gamified through fictionality”, *Proceedings of the 13th international conference on advances in mobile computing and multimedia*, pp. 309–313, 2015.

[81] J. Woodward and J. Ruiz, “Analytic review of using augmented reality for situational awareness”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 4, pp. 2166–2183, 2022.

[82] M. Di Donato, M. Fiorentino, A. E. Uva, M. Gattullo, and G. Monno, “Text legibility for projected Augmented Reality on industrial workbenches”, *Computers in Industry*, vol. 70, pp. 70–78, 2015.

[83] M. Fiorentino, S. Debernardis, A. E. Uva, and G. Monno, “Augmented reality text style readability with see-through head-mounted displays in industrial context”, *Presence: Teleoperators and Virtual Environments*, vol. 22, no. 2, pp. 171–190, 2013.

[84] S. J. Kerr *et al.*, “Wearable mobile augmented reality: evaluating outdoor user experience”, *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*, pp. 209–216, 2011.

[85] B. McGuinness, “Quantitative analysis of situational awareness (QUASA): Applying signal detection theory to true/false probes and self-ratings”, *Command and Control Research and Technology Symposium June*, pp. 15–17, 2004.

[86] S. J. Selcon, R. M. Taylor, and E. Koritsas, “Workload or situational awareness?: TLX vs. SART for aerospace systems design evaluation”, *Proceedings of the Human Factors Society Annual Meeting*, vol. 35, no. 2, pp. 62–66, 1991.

[87] P. M. Salmon *et al.*, “Measuring Situation Awareness in complex systems: Comparison of measures study”, *International Journal of Industrial Ergonomics*, vol. 39, no. 3, pp. 490–500, 2009.

[88] B. V. Syiem, R. M. Kelly, J. Goncalves, E. Velloso, and T. Dingler, “Impact of task on attentional tunneling in handheld augmented reality”, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2021.

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

[90] H. Regenbrecht, C. Botella, R. Baños, and T. Schubert, "Mixed reality experience questionnaire (MREQ)-reference", 2017.

[91] N. C. Martins, B. Marques, J. Alves, T. Araújo, P. Dias, and B. S. Santos, "Augmented reality situated visualization in decision-making", *Multimedia Tools and Applications*, vol. 81, no. 11, pp. 14749–14772, 2022.

[92] A. Mossel, B. Venditti, and H. Kaufmann, "3DTouch and HOMER-S: intuitive manipulation techniques for one-handed handheld augmented reality", *Proceedings of the virtual reality international conference: laval virtual*, pp. 1–10, 2013.

[93] W. Hürst and C. Van Wezel, "Gesture-based interaction via finger tracking for mobile augmented reality", *Multimedia Tools and Applications*, vol. 62, pp. 233–258, 2013.

[94] J. G. Grandi, H. G. Debarba, L. Nedel, and A. Maciel, "Design and evaluation of a handheld-based 3d user interface for collaborative object manipulation", *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 5881–5891, 2017.

[95] A. Cohé, F. Dècle, and M. Hachet, "tbox: a 3d transformation widget designed for touchscreens", *Proceedings of the sigchi conference on human factors in computing systems*, pp. 3005–3008, 2011.

[96] K. Monteiro, R. Vatsal, N. Chulpongsatorn, A. Parnami, and R. Suzuki, "Teachable Reality: Prototyping Tangible Augmented Reality with Everyday Objects by Leveraging Interactive Machine Teaching", *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2023.

[97] T. D. Do, J. J. LaViola, and R. P. McMahan, "The effects of object shape, fidelity, color, and luminance on depth perception in handheld mobile augmented reality", *2020 IEEE International symposium on mixed and augmented reality (ISMAR)*, pp. 64–72, 2020.

[98] J. Heer and M. Agrawala, "Design considerations for collaborative visual analytics", *2007 IEEE Symposium on Visual Analytics Science and Technology*, pp. 171–178, 2007.

[99] Y. Yang *et al.*, "Towards immersive collaborative sensemaking", *Proceedings of the ACM on Human-Computer Interaction*, vol. 6, no. ISS, pp. 722–746, 2022.

[100] N. Reski, A. Alissandrakis, and A. Kerren, "User Preferences of Spatio-Temporal Referencing Approaches For Immersive 3D Radar Charts", *arXiv preprint arXiv:2303.07899*, 2023.

[101] J. G. Shin, G. Ng, and D. Saakes, "Couples designing their living room together: A study with collaborative handheld augmented reality", *Proceedings of the 9th Augmented Human International Conference*, pp. 1–9, 2018.

[102] P. Sasikumar, S. Chittajallu, N. Raj, H. Bai, and M. Billinghurst, "Spatial perception enhancement in assembly training using augmented volumetric playback", *Frontiers in Virtual Reality*, vol. 2, p. 698523, 2021.

[103] Y. Cha, S. Nam, M. Y. Yi, J. Jeong, and W. Woo, "Augmented collaboration in shared space design with shared attention and manipulation", *Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, pp. 13–15, 2018.

[104] Y. S. Chang, B. Nuernberger, B. Luan, and T. Höllerer, "Evaluating gesture-based augmented reality annotation", *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 182–185, 2017.

[105] C. Gutwin and S. Greenberg, "A descriptive framework of workspace awareness for real-time groupware", *Computer Supported Cooperative Work (CSCW)*, vol. 11, pp. 411–446, 2002.

[106] S. Yang, A. A. Poonawala, T.-S. A. Jiang, and B. Schneider, "Can Synchronous Code Editing and Awareness Tools Support Remote Tutoring? Effects on Learning and Teaching", *Proceedings of the ACM on Human-Computer Interaction*, vol. 7, no. CSCW2, pp. 1–30, 2023.

[107] A. R. Smink, E. A. Van Reijmersdal, G. Van Noort, and P. C. Neijens, "Shopping in augmented reality: The effects of spatial presence, personalization and intrusiveness on app and brand responses", *Journal of Business Research*, vol. 118, pp. 474–485, 2020.

[108] T. Hilken, K. de Ruyter, M. Chylinski, D. Mahr, and D. I. Keeling, "Augmenting the eye of the beholder: exploring the strategic potential of augmented reality to enhance online service experiences", *Journal of the Academy of Marketing Science*, vol. 45, pp. 884–905, 2017.

[109] A. R. Smink, S. Frowijn, E. A. van Reijmersdal, G. van Noort, and P. C. Neijens, "Try online before you buy: How does shopping with augmented reality affect brand responses and personal data disclosure", *Electronic Commerce Research and Applications*, vol. 35, p. 100854, 2019.

[110] B. Moncur, M. J. Galvez Trigo, and L. Mortara, "Augmented reality to reduce cognitive load in operational decision-making", *International Conference on Human-Computer Interaction*, pp. 328–346, 2023.

[111] J. Heller, M. Chylinski, K. de Ruyter, D. Mahr, and D. I. Keeling, "Let me imagine that for you: Transforming the retail frontline through augmenting customer mental imagery ability", *Journal of Retailing*, vol. 95, no. 2, pp. 94–114, 2019.

[112] B. Marques, B. S. Santos, T. Araújo, N. C. Martins, J. B. Alves, and P. Dias, "Situated visualization in the decision process through augmented reality", *2019 23rd international conference information visualisation (IV)*, pp. 13–18, 2019.

[113] E. Massa and R. Ladhari, "Augmented reality in marketing: Conceptualization and systematic review", *International Journal of Consumer Studies*, 2023.

[114] S. Barta, R. Gurrea, and C. Flavián, "Using augmented reality to reduce cognitive dissonance and increase purchase intention", *Computers in Human Behavior*, vol. 140, p. 107564, 2023.

[115] S. Petsas, G. E. Raptis, and C. Katsanos, "Turn & Slide: Designing a Puzzle Game to Elicit the Visualizer-Verbalizer Cognitive Style", *IFIP Conference on Human-Computer Interaction*, pp. 46–56, 2023.

[116] M. Kozhevnikov, S. Kosslyn, and J. Shephard, "Spatial versus object visualizers: A new characterization of visual cognitive style", *Memory & cognition*, vol. 33, no. 4, pp. 710–726, 2005.

[117] A. Morrison *et al.*, “Collaborative use of mobile augmented reality with paper maps”, *Computers & Graphics*, vol. 35, no. 4, pp. 789–799, 2011.

[118] S. A. Oloonabadi and P. Baran, “Augmented reality participatory platform: A novel digital participatory planning tool to engage under-resourced communities in improving neighborhood walkability”, *Cities*, vol. 141, p. 104441, 2023.

[119] H. Regenbrecht, S. Zwanenburg, and T. Langlotz, “Pervasive augmented reality—Technology and ethics”, *IEEE Pervasive Computing*, vol. 21, no. 3, pp. 84–91, 2022.

[120] T. Hilken, D. I. Keeling, K. de Ruyter, D. Mahr, and M. Chylinski, “Seeing eye to eye: social augmented reality and shared decision making in the marketplace”, *Journal of the Academy of Marketing Science*, vol. 48, pp. 143–164, 2020.

[121] Z. Wang, C. Nguyen, P. Asente, and J. Dorsey, “Distanciar: Authoring site-specific augmented reality experiences for remote environments”, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2021.

[122] G. Häubl and V. Trifts, “Consumer decision making in online shopping environments: The effects of interactive decision aids”, *Marketing science*, vol. 19, no. 1, pp. 4–21, 2000.

[123] H. Li, S. M. Edwards, and J.-H. Lee, “Measuring the intrusiveness of advertisements: Scale development and validation”, *Journal of advertising*, vol. 31, no. 2, pp. 37–47, 2002.

[124] N. K. Malhotra, S. S. Kim, and J. Agarwal, “Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model”, *Information systems research*, vol. 15, no. 4, pp. 336–355, 2004.

[125] S. Küçük, R. Yilmaz, Ö. Baydas, and Y. Göktas, “Augmented reality applications attitude scale in secondary schools: Validity and reliability study”, *Egitim ve Bilim*, vol. 39, no. 176, 2014.

[126] R. M. Ryan, V. Mims, and R. Koestner, “Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory.”, *Journal of personality and Social Psychology*, vol. 45, no. 4, p. 736, 1983.

[127] C. Boletsis, “The Gaia System: A Tabletop Projection Mapping System for Raising Environmental Awareness in Islands and Coastal Areas”, *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 50–54, 2022.

[128] P. C. Pezzullo, “Between crisis and care: Projection mapping as creative climate advocacy”, *Journal of Environmental Media*, vol. 1, no. 1, pp. 59–77, 2020.

[129] i-art, “The Magic of Interactive Books”. Accessed: Oct. 12, 2023. [Online]. Available: https://iart.ch/en/next/interaktive-buecher

[130] L. J. Perovich, S. A. Wylie, and R. Bongiovanni, “Chemicals in the Creek: designing a situated data physicalization of open government data with the community”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 913–923, 2020.

[131] S. Jung, “The Message Effect of Augmented Health Messages on Body”, *Virtual, Augmented and Mixed Reality: Applications in Health, Cultural Heritage, and Industry: 10th International Conference, VAMR 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part II 10*, pp. 86–93, 2018.

[132] M. Orduna, P. Pérez, J. Gutiérrez, and N. Garc\ia, “Methodology to assess quality, presence, empathy, attitude, and attention in 360-degree videos for immersive communications”, *IEEE Transactions on Affective Computing*, 2022.

[133] L. Besançon, A. Ynnerman, D. F. Keefe, L. Yu, and T. Isenberg, “The state of the art of spatial interfaces for 3D visualization”, *Computer Graphics Forum*, vol. 40, no. 1, pp. 293–326, 2021.

[134] J. A. Walsh, S. Von Itzstein, and B. H. Thomas, “Ephemeral interaction using everyday objects”, *Proceedings of the Fifteenth Australasian User Interface Conference-Volume 150*, pp. 29–37, 2014.

[135] M. Löchtefeld, F. Wiehr, and S. Gehring, “Analysing the effect of tangibile user interfaces on spatial memory”, *Proceedings of the 5th Symposium on Spatial User Interaction*, pp. 78–81, 2017.

[136] M. J. Kim and M. L. Maher, “The impact of tangible user interfaces on designers' spatial cognition”, *Human–Computer Interaction*, vol. 23, no. 2, pp. 101–137, 2008.

[137] D. Duranti, D. Spallazzo, and D. Petrelli, “Smart Objects and Replicas: A Survey of Tangible and Embodied Interactions in Museums and Cultural Heritage Sites”, *ACM Journal on Computing and Cultural Heritage*, 2023.

[138] C. Stephanidis *et al.*, “Seven HCI grand challenges”, *International Journal of Human–Computer Interaction*, vol. 35, no. 14, pp. 1229–1269, 2019.

[139] E. Nofal, R. Stevens, T. Coomans, and A. V. Moere, “Communicating the spatiotemporal transformation of architectural heritage via an in-situ projection mapping installation”, *Digital Applications in Archaeology and Cultural Heritage*, vol. 11, p. e83, 2018.

[140] R. Guarese, E. Pretty, H. Fayek, F. Zambetta, and R. van Schyndel, “Evoking empathy with visually impaired people through an augmented reality embodiment experience”, *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 184–193, 2023.

[141] S. Larsen, “Aspects of a psychology of the tourist experience”, *Scandinavian Journal of Hospitality and Tourism*, vol. 7, no. 1, pp. 7–18, 2007.

[142] L.-W. Mai and G. Schoeller, “Emotions, attitudes and memorability associated with TV commercials”, *Journal of Targeting, Measurement and Analysis for Marketing*, vol. 17, pp. 55–63, 2009.

[143] B. Laugwitz, T. Held, and M. Schrepp, “Construction and evaluation of a user experience questionnaire”, *HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4*, pp. 63–76, 2008.

[144] M. Thüring and S. Mahlke, “Usability, aesthetics and emotions in human–technology interaction”, *International journal of psychology*, vol. 42, no. 4, pp. 253–264, 2007.

[145] L. Merino, M. Schwarzl, M. Kraus, M. Sedlmair, D. Schmalstieg, and D. Weiskopf, “Evaluating mixed and augmented reality: A systematic literature review (2009-2019)”, *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 438–451, 2020.

[146] M. K. Robison, N. T. Diede, J. Nicosia, B. H. Ball, and J. M. Bugg, “A multimodal analysis of sustained attention in younger and older adults.”, *Psychology and aging*, vol. 37, no. 3, p. 307, 2022.

[147] Y.-H. Hung and P. Parsons, “Assessing user engagement in information visualization”, *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*, pp. 1708–1717, 2017.

[148] J. Sterman, “System Dynamics: systems thinking and modeling for a complex world”, 2002.

[149] G. McGowin, S. M. Fiore, and K. Oden, “Me, Myself, and the (Virtual) World: A Review of Learning Research in 4E Cognition and Immersive Virtual Reality”, *International Conference on Human-Computer Interaction*, pp. 59–73, 2022.

[150] F. J. Varela, E. Thompson, and E. Rosch, *The embodied mind, revised edition: Cognitive science and human experience*. MIT press, 2017.

[151] U. Flick, *An introduction to qualitative research*, 7th edition. Los Angeles: SAGE, 2023.

[152] G. Ganis and R. Kievit, “A new set of three-dimensional shapes for investigating mental rotation processes: validation data and stimulus set”, *Journal of Open Psychology Data Files in this Item Files Size Format View 13-116-2-PB. pdf 567.5 Kb PDF View/Open*, 2015.

[153] S. G. Vandenberg and A. R. Kuse, “Mental rotations, a group test of three-dimensional spatial visualization”, *Perceptual and motor skills*, vol. 47, no. 2, pp. 599–604, 1978.

[154] S. Bonfanti, A. Gargantini, G. Esposito, A. Facchin, M. Maffioletti, and S. Maffioletti, “Evaluation of stereoacuity with a digital mobile application”, *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 259, no. 9, pp. 2843–2848, 2021.

[155] J. Blokša, “Design guidelines for user interface for augmented reality”, *Masaryk University*, 2017.

[156] P. E. Kourouthanassis, C. Boletsis, and G. Lekakos, “Demystifying the design of mobile augmented reality applications”, *Multimedia Tools and Applications*, vol. 74, pp. 1045–1066, 2015.

[157] M. Hassenzahl, F. Koller, and M. Burmester, “Der User Experience (UX) auf der Spur: Zum Einsatz von www. attrakdiff. de”, 2008.

[158] R. T. Azuma, “A survey of augmented reality”, *Presence: teleoperators & virtual environments*, vol. 6, no. 4, pp. 355–385, 1997.

[159] O. Christ, M. Sambasivam, A. Roos, and C. Zahn, “Learning in immersive virtual reality: how does the 4e cognition approach fit in virtual didactic settings?”, in *Human Interaction, Emerging Technologies and Future Systems V: Proceedings of the 5th International Virtual Conference on Human Interaction and Emerging Technologies, IHIET 2021, August 27-29, 2021 and the 6th IHIET: Future Systems (IHIET-FS 2021), October 28-30, 2021, France*, 2022, pp. 790–796.

[160] R. Long, W. Han, A. Jiang, and X. Zeng, “Visitor’s Museum Experience Model in Mixed Reality Environment from the Perspective of 4E Cognition”, in *International Conference on Human-Computer Interaction*, 2024, pp. 281–295.

Appendix

A Index of auxiliary tools

- For this thesis the tools *Deepl* for translation and *Deepl Write* to improve grammar and writing style were used through the entire document.

B UI Application

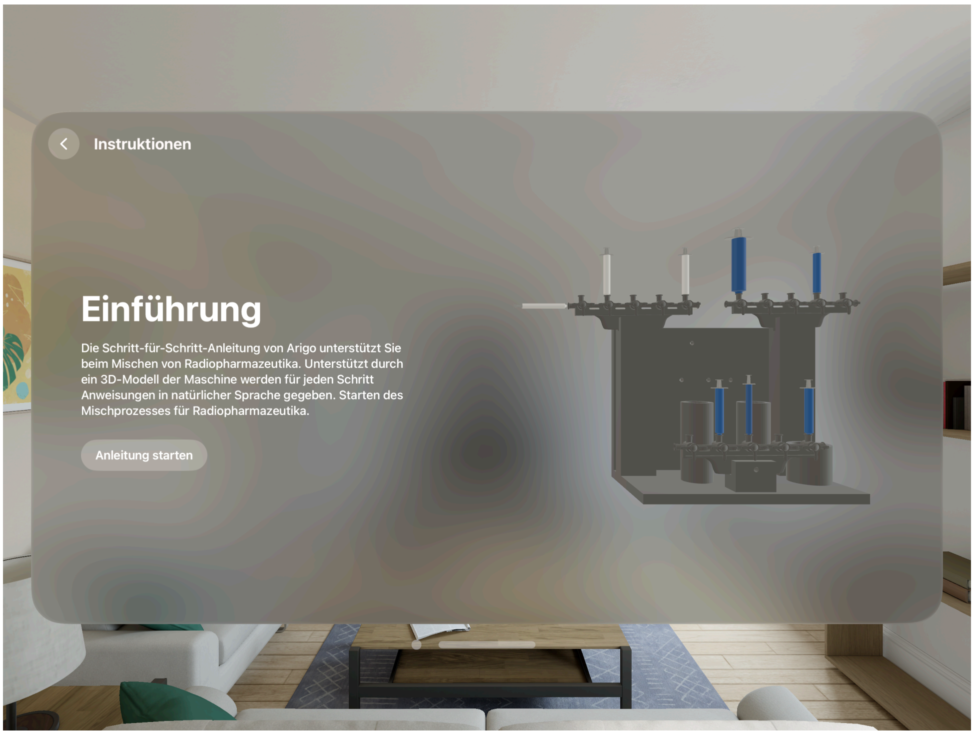


Figure 88: Introduction to the process

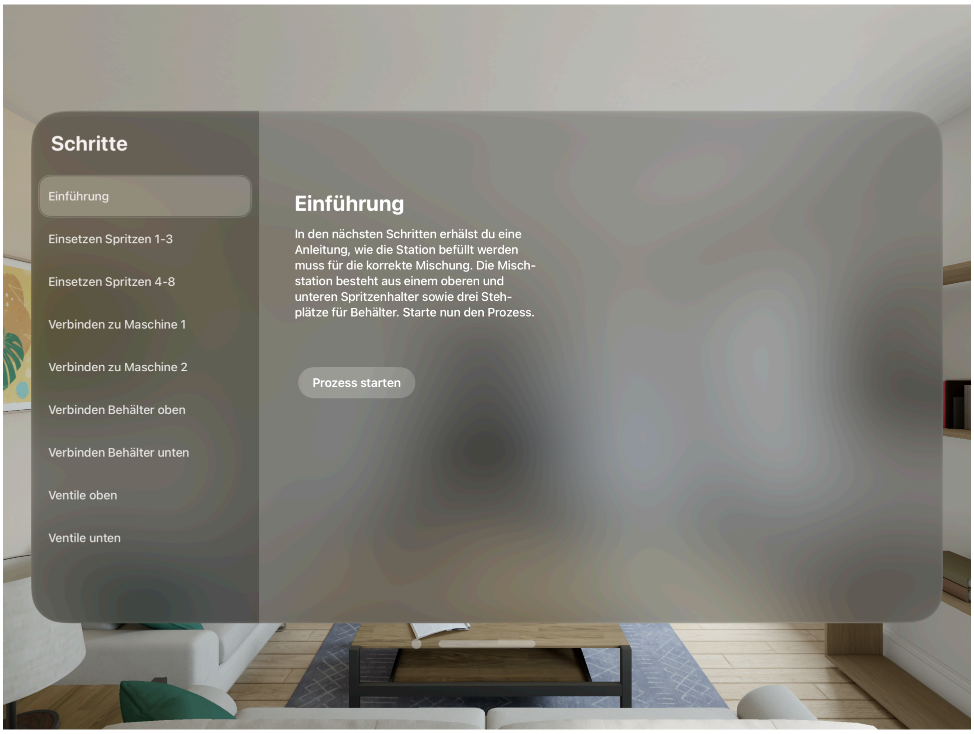


Figure 89: Start information for the process

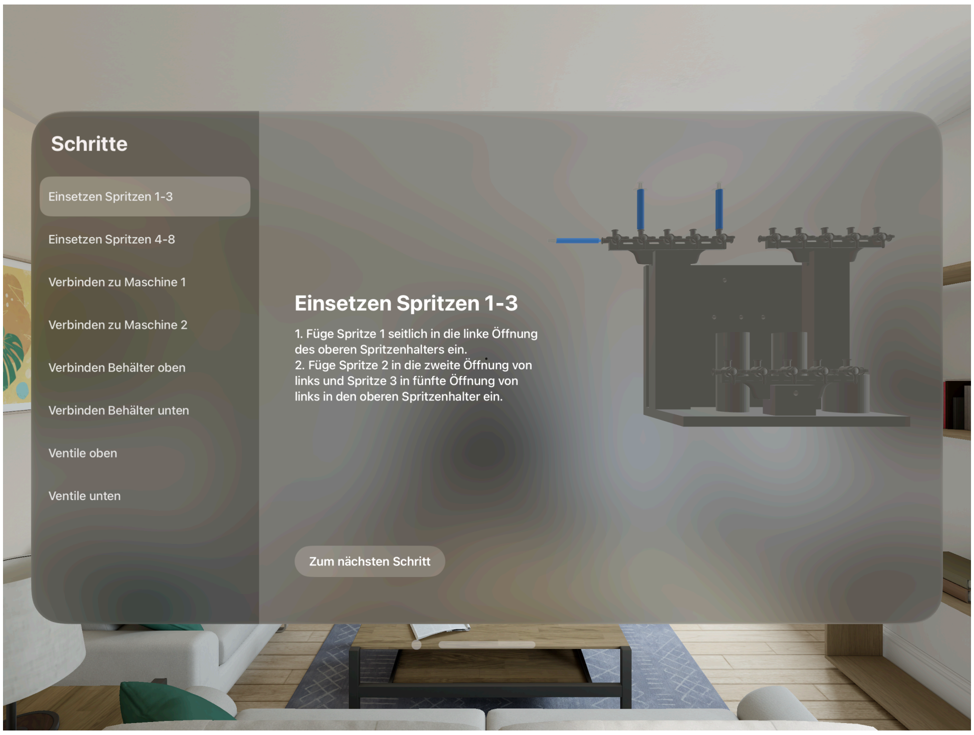


Figure 90: Step 1 in the guidance

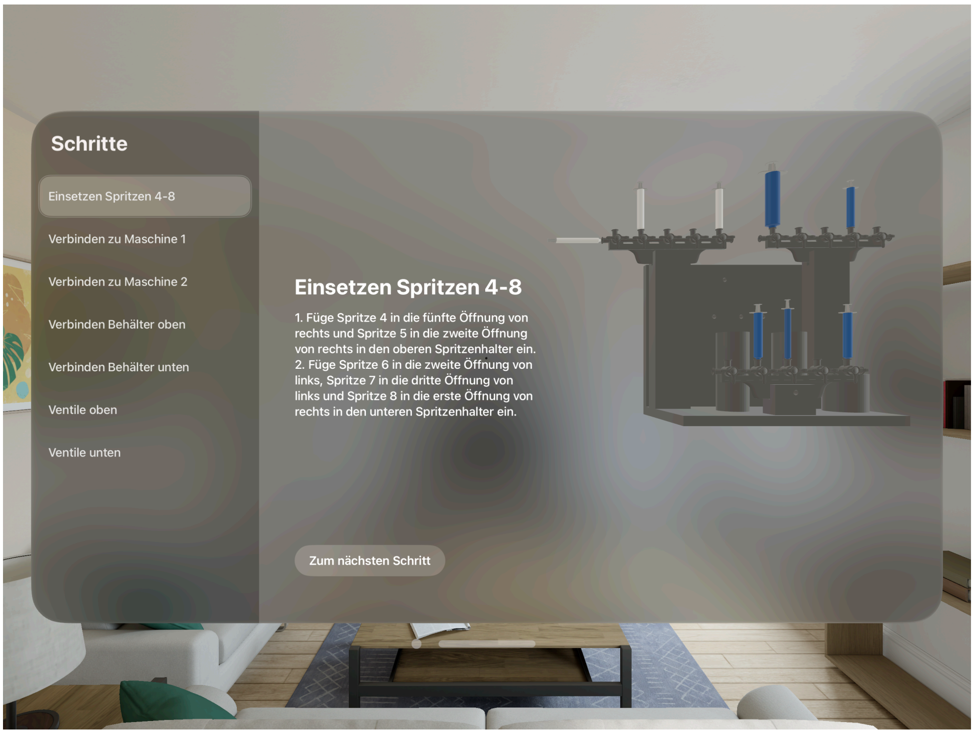


Figure 91: Step 2 in the guidance

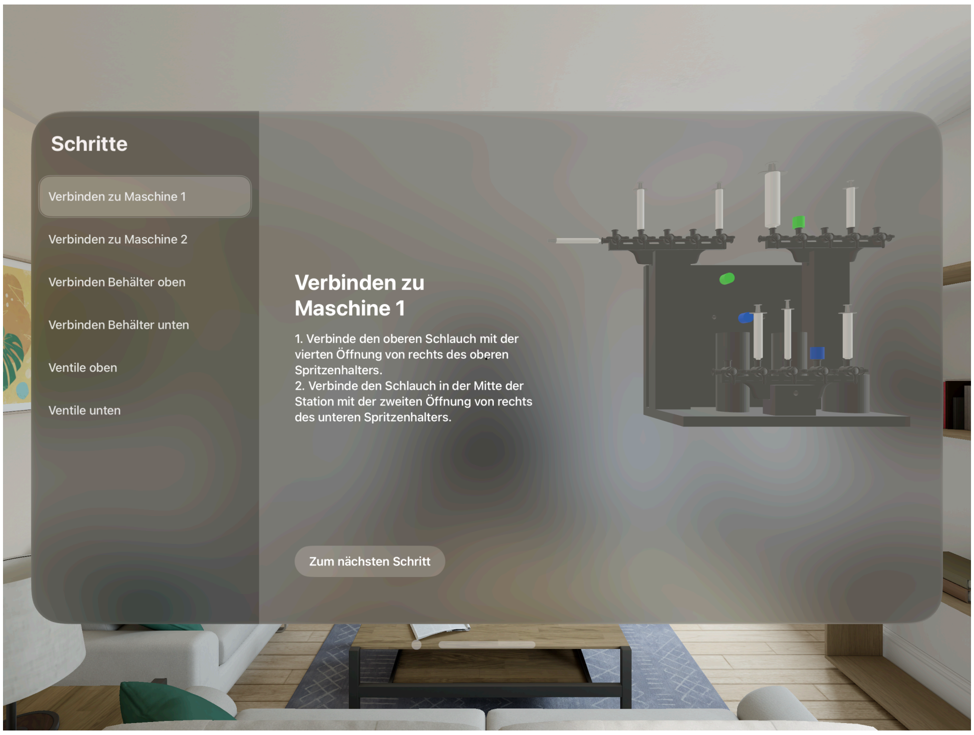


Figure 92: Step 3 in the guidance

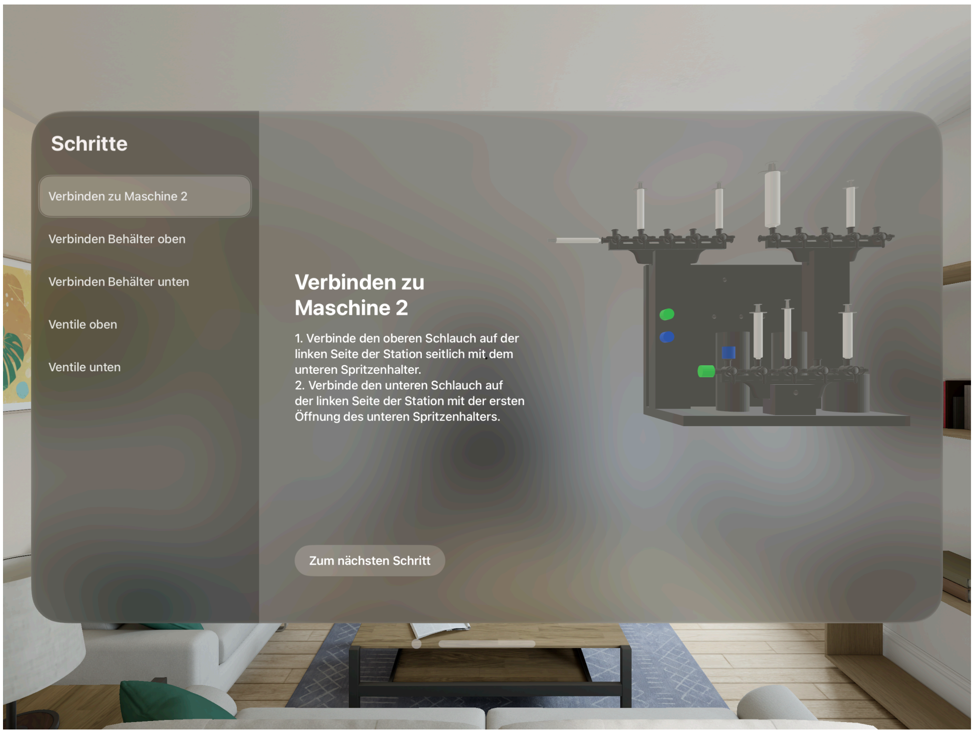


Figure 93: Step 4 in the guidance

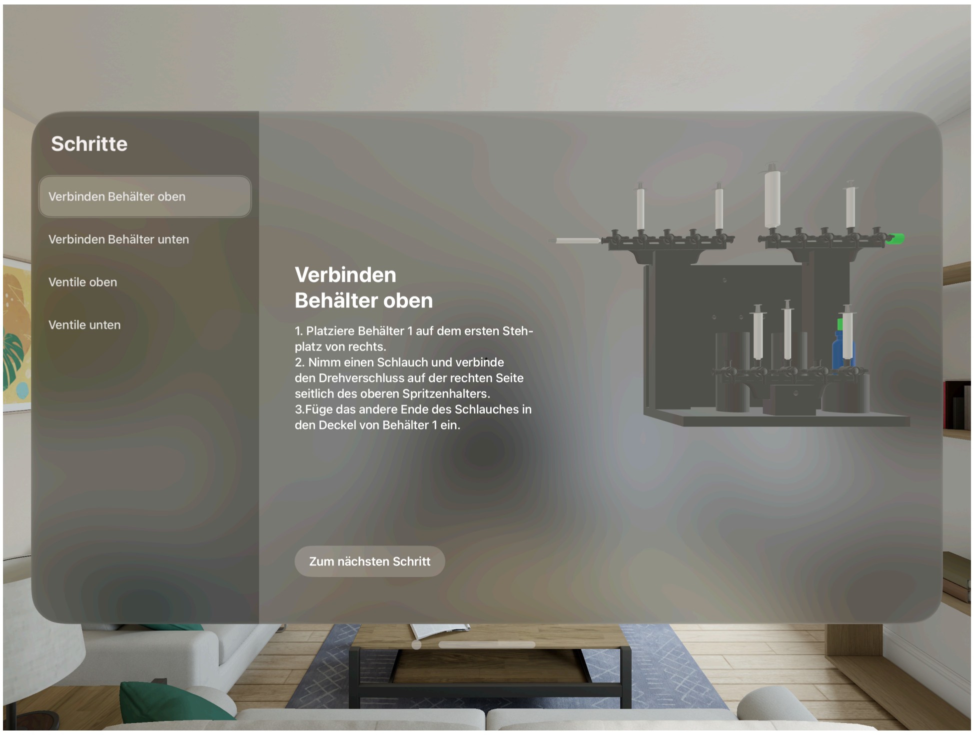


Figure 94: Step 5 in the guidance

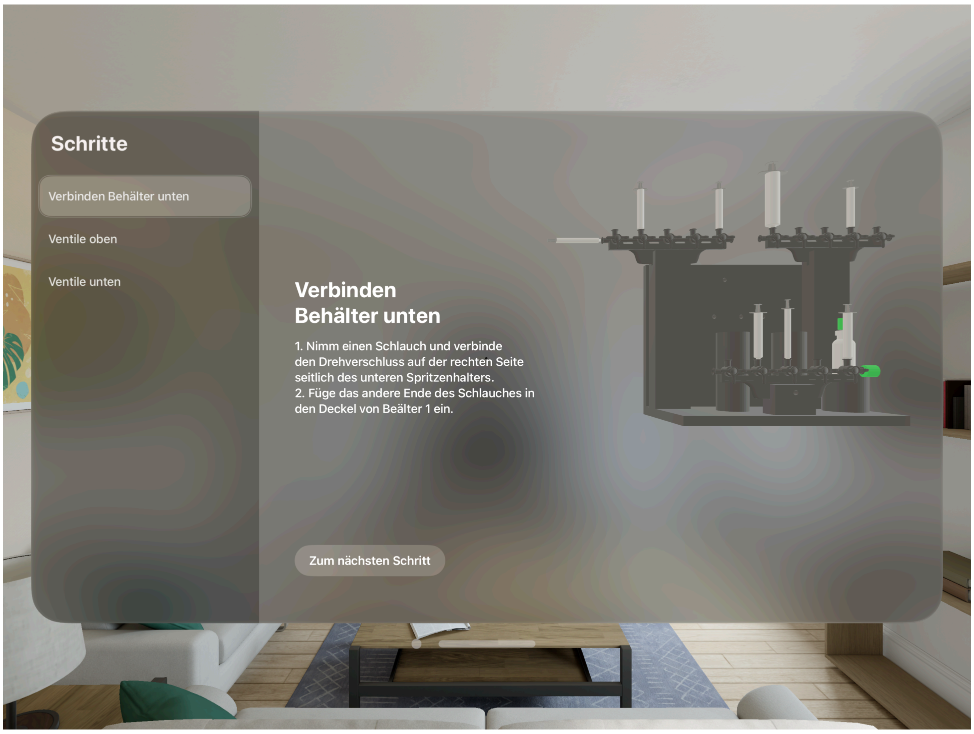


Figure 95: Step 6 in the guidance

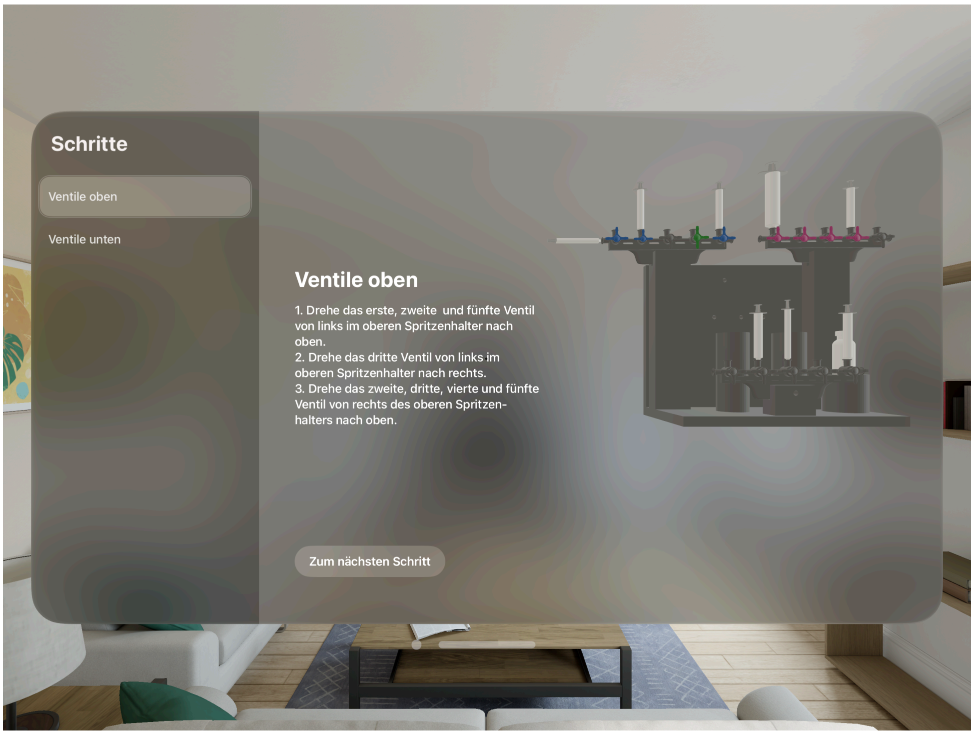


Figure 96: Step 7 in the guidance

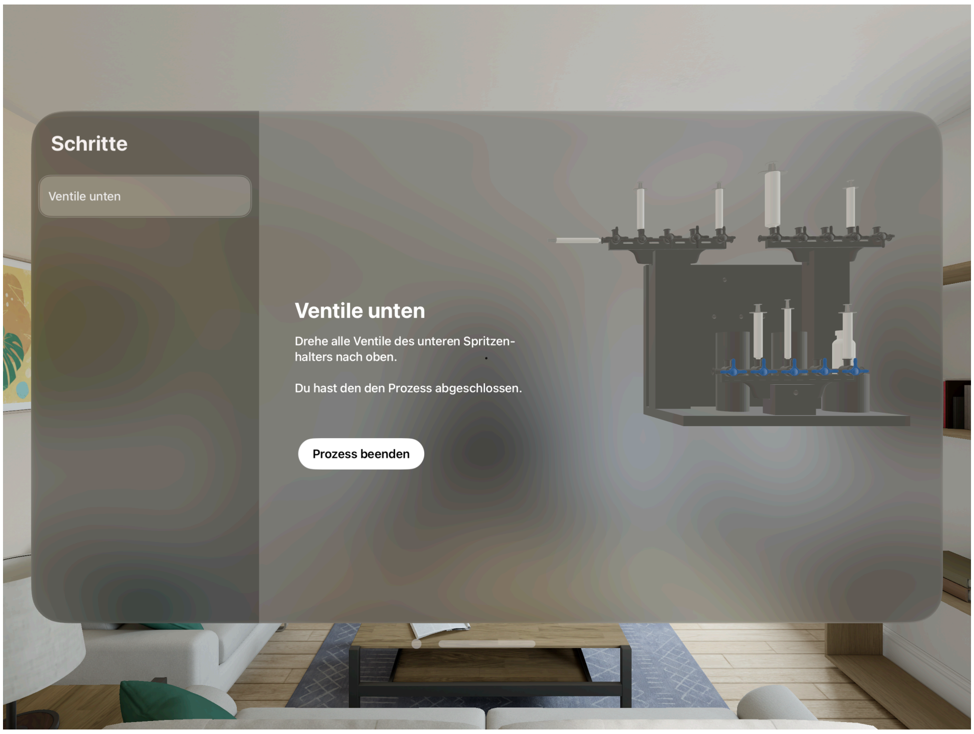


Figure 97: Step 8 in the guidance

C Questions for Interview and Observation

Questions for observation before the study starts

- How is the user feeling?
- Is there any disturbance in the room?

Questions for observation during the performance of the study

- How do emotions change over the process?
- How is the concentration?
- Does the user know what to do?
- Where is the focus during performance of the steps?
- How does the user behave for orientation?
- How is the physical interaction with the elements?
- Is there a pattern in the behaviour?

Interview questions after the study

- How does the user feel when taking off the Apple Vision Pro?
- How were the directional information in the instructions perceived?
- How would they rephrase the instructions for simplification?
- How did the user orient him/herself in the digital model?
- How was the perception of the visual cues in the digital model?
- How would the user simplify the orientation with the digital model?
- How did the user find out where to place the elements?
- Which points were used for orientation in the digital space?
- Which points were used for orientation in the physical space?
- How exhausting was the process?

D Codeline per Participant

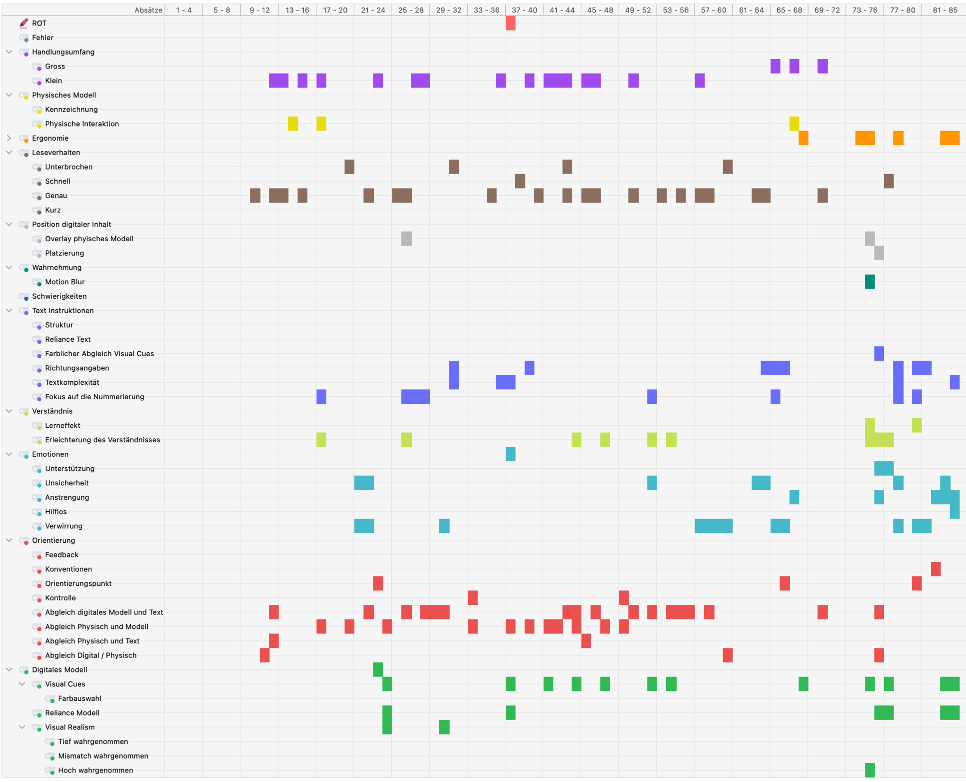


Figure 98: Codeline of User 1

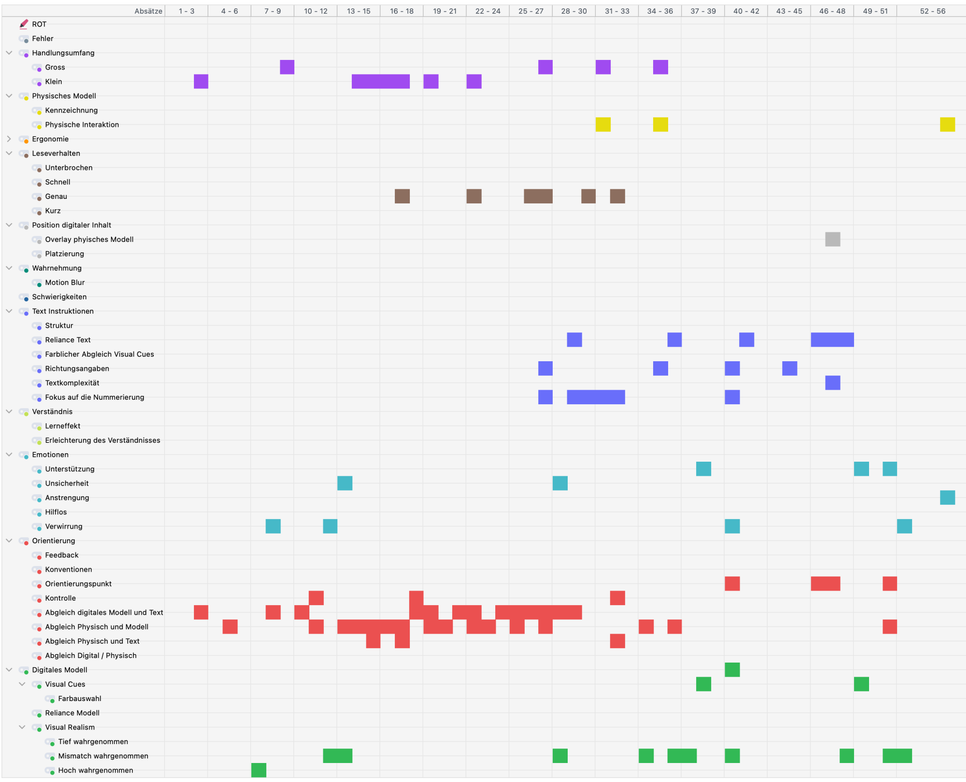


Figure 99: Codeline of User 2

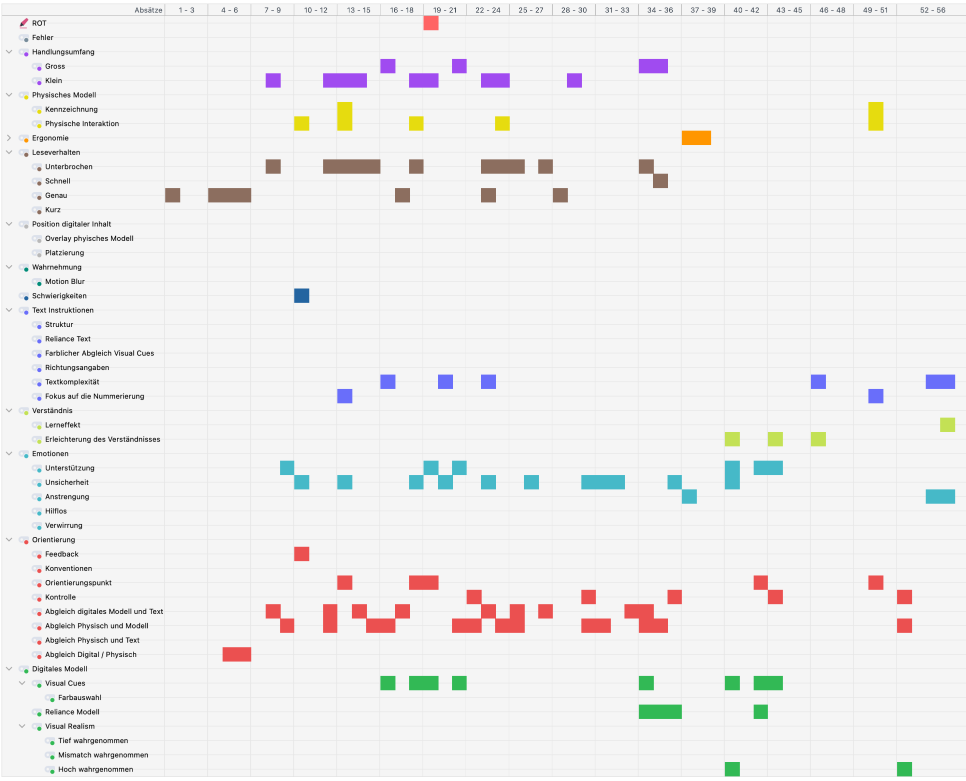


Figure 100: Codeline of User 3

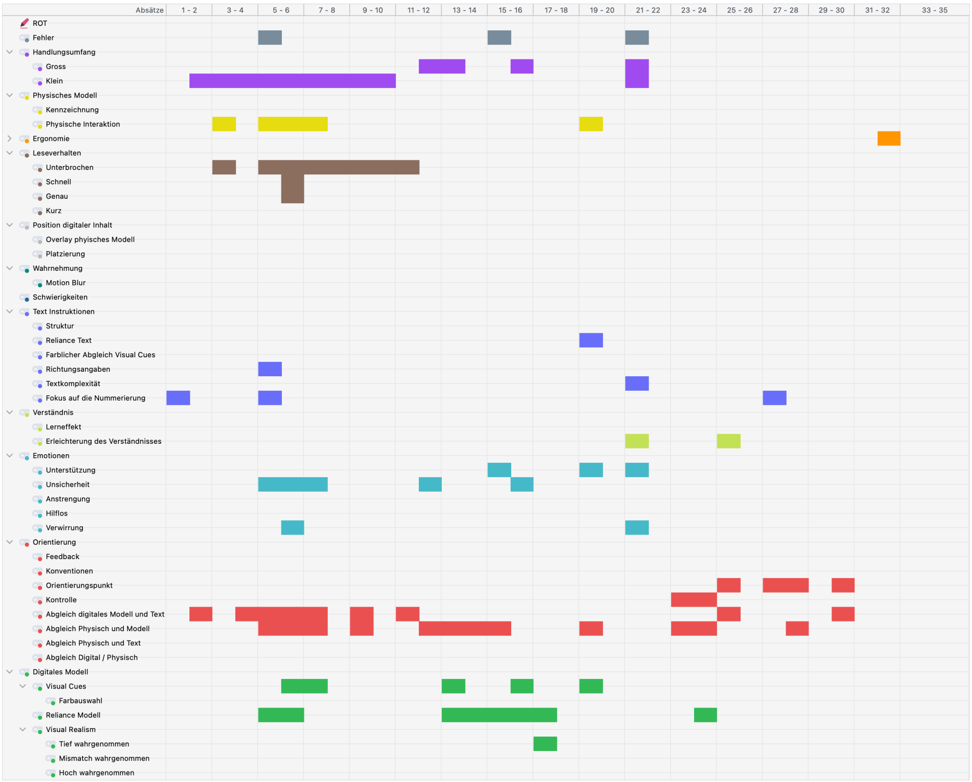


Figure 101: Codeline of User 4

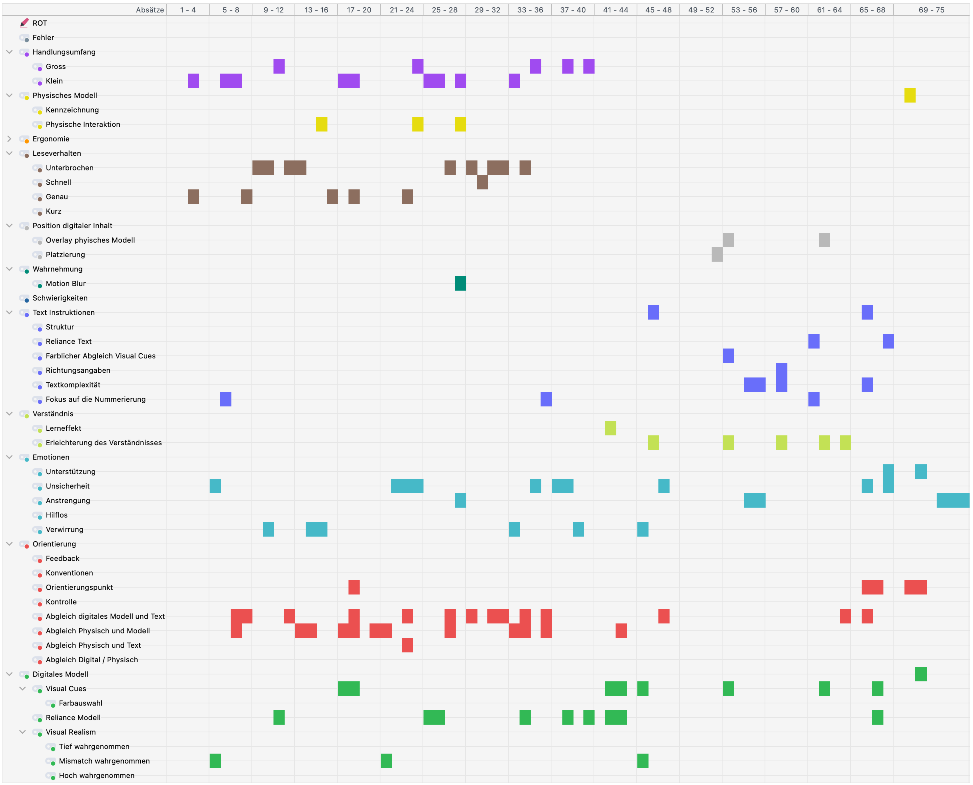


Figure 102: Codeline of User 5

Figure 103: Codeline of User 6

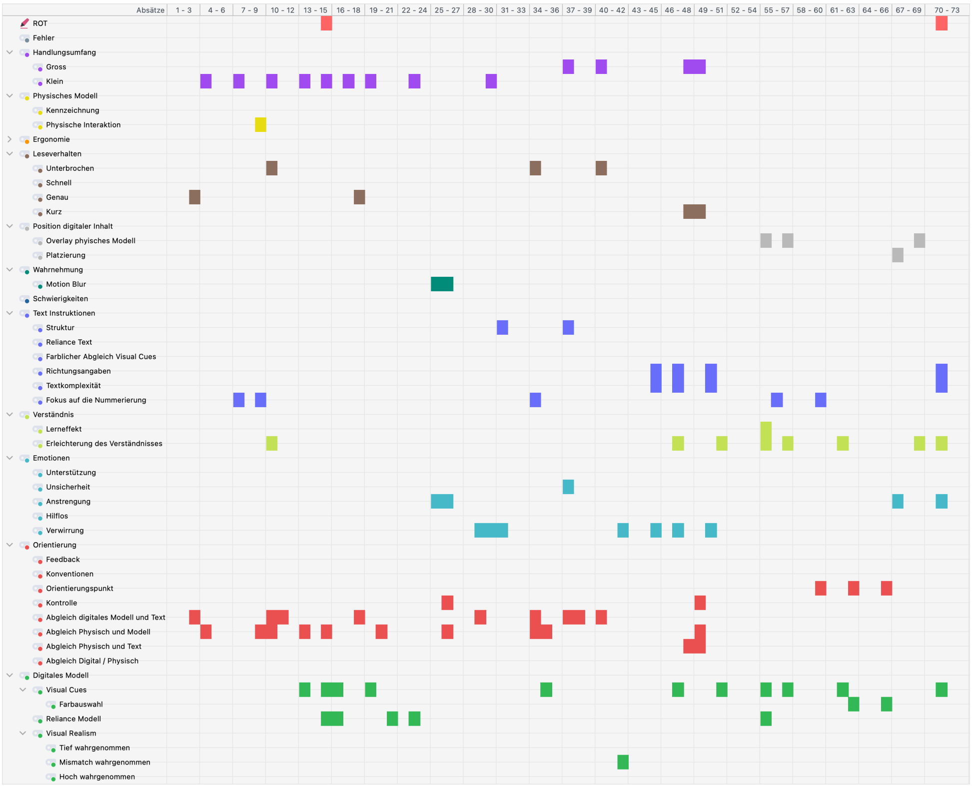


Figure 104: Codeline of User 7

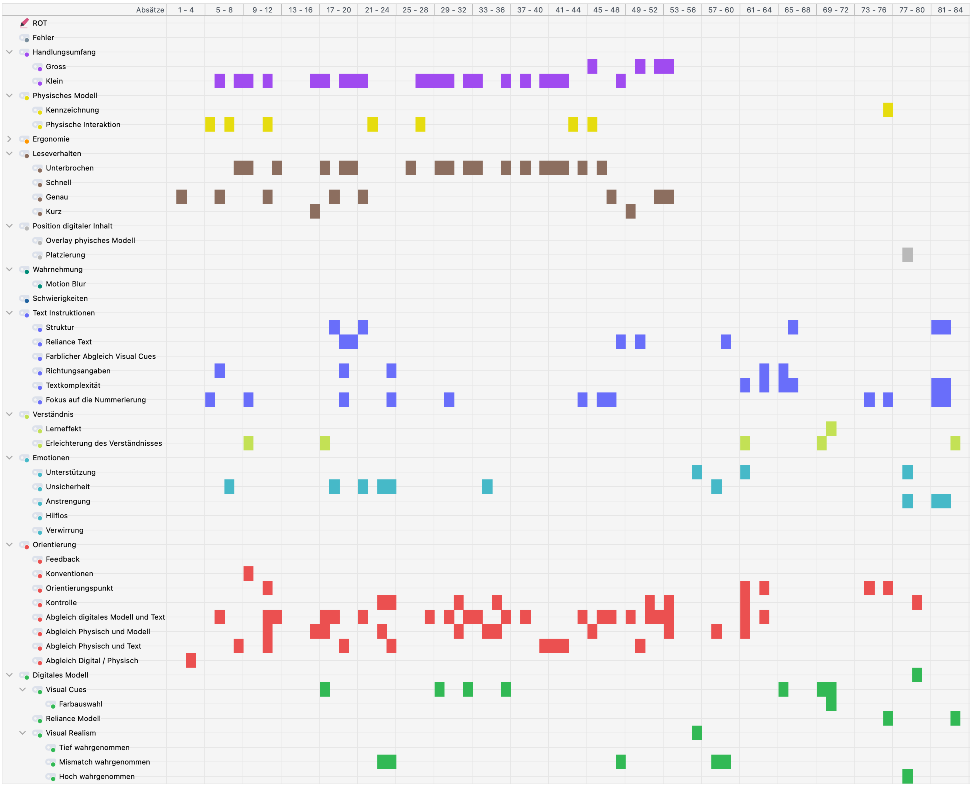


Figure 105: Codeline of User 8

Figure 106: Codeline of User 9