TU WIEN Informatics

Creation, Manipulation, and Interaction with BPMN Models in Augmented Reality

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Wirtschaftsinformatik

eingereicht von

Anna Svoboda, Bsc
Matrikelnummer 11902077

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Associate Prof. Dipl.-Wirtsch.Inf.Univ. Dr.rer.pol. Dominik Bork
Mitwirkung: Univ.Ass. Aleksandar Gavric, MEng MSc BEng

Wien, 1. September 2026

Anna Svoboda	Dominik Bork

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

TU WIEN Informatics

Creation, Manipulation, and Interaction with BPMN Models in Augmented Reality

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Business Informatics

by

Anna Svoboda, Bsc
Registration Number 11902077

to the Faculty of Informatics

at the TU Wien

Advisor: Associate Prof. Dipl.-Wirtsch.Inf.Univ. Dr.rer.pol. Dominik Bork
Assistance: Univ.Ass. Aleksandar Gavric, MEng MSc BEng

Vienna, September 1, 2026

Anna Svoboda | Dominik Bork

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Erklärung zur Verfassung der Arbeit

Anna Svoboda, Bsc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.
Ich erkläre weiters, dass ich mich generativer KI-Tools lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss überwiegt. Im Anhang „Übersicht verwendeter Hilfsmittel“ habe ich alle generativen KI-Tools gelistet, die verwendet wurden, und angegeben, wo und wie sie verwendet wurden. Für Textpassagen, die ohne substantielle Änderungen übernommen wurden, habe ich jeweils die von mir formulierten Eingaben (Prompts) und die verwendete IT-Anwendung mit ihrem Produktnamen und Versionsnummer/Datum angegeben.

Wien, 1. September 2026

Anna Svoboda

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Acknowledgements

First and foremost, I would like to thank my supervisor, Dominik Bork, for his guidance and valuable input throughout the course of this thesis. His feedback during our meetings helped shape the specifications and direction of this work considerably.

I would also like to thank Aleksandar Gavric for his support during the implementation of the prototype and his input on the design decisions. His constructive suggestions were much appreciated.

Finally, I would like to thank my family and friends for their support and encouragement throughout this thesis. A special thank you goes to those who took the time to participate in the user study, as their contribution was a key part of the evaluation of this work.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Kurzfassung

Die Kombination aus Augmented Reality (AR) und konzeptueller Modellierung kann das Verständnis von Diagrammen, die Einbindung der Nutzenden und deren Motivation, sich mit dem Modell auseinanderzusetzen, verbessern. In zukünftigen Anwendungen könnten AR-Modelle Mitarbeitende bei der Prozessausführung unterstüzen, indem Prozessmodelle direkt in die reale Umgebung eingebettet und die als Nächstes auszuführenden Aufgaben angezeigt werden. Des Weiteren könnte AR die Erstellung von Diagrammen unterstützen, indem es Usern ermöglicht, einen Prozess während seiner Ausführung zu modellieren. Dies wäre insbesondere in Umgebungen nützlich, in denen die Hände der Mitarbeitenden verschmutzt sind, oder der Wechsel zu einem herkömmlichen Gerät nicht möglich ist.

Trotz dieser vielversprechenden Möglichkeiten gibt es bislang kaum Forschungsarbeiten zur Erstellung und Bearbeitung von Business Process Model and Notation (BPMN) Modellen mittels Handgesten in immersiven Umgebungen wie Augmented Reality (AR) und Virtual Reality (VR), und es wurden bisher nur wenige geeigneten Interaktionsmodi für diesen Zweck untersucht.

Diese Arbeit stellt den „BPMN Builder“ vor, einen Prototyp, der die Erstellung und Bearbeitung von BPMN Modellen in AR über zwei Interaktionsmodi unterstützt: direkte und indirekte Interaktion. Im direkten Modus wählen User Elemente direkt im AR-Diagramm mittels Handgesten aus, während im indirekten Modus ein Interaktionspanel als Mittel zur Diagrammnavigation und -interaktion dient.

Zur Bewertung des Prototyps wurde eine Benutzerstudie mit einem Within-Subjects-Design durchgeführt, bei der alle Teilnehmenden mittels Aufgaben zur Erstellung und Bearbeitung von Diagrammen beide Interaktionsmodi testeten. Die Interaktionsmodi wurden hinsichtlich Effizienz, Benutzerfreundlichkeit und Nutzerpräferenz verglichen. Insgesamt schnitt der direkte Modus besser ab und erzielte kürzere Bearbeitungszeiten, einen höheren System Usability Scale (SUS) Wert sowie eine höhere Nutzerpräferenz.

Diese Ergebnisse deuten darauf hin, dass die direkte Interaktion der geeignetere Modus für die BPMN-Modellierung in AR ist. Es ist zu erwarten, dass diese Erkenntnisse der gestenbasierten Interaktion auf andere konzeptuelle Modellierungssprachen übertragbar sind.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

The combination of Augmented Reality (AR) and conceptual modeling can improve diagram comprehension, user engagement, and motivation to engage with the model. In future applications, AR models could aid workers during process execution by embedding process models directly in the real environment and displaying the next task to be performed. Furthermore, AR could support diagram creation by allowing users to model a process while performing it. This would be especially useful in environments where workers' hands are dirty or switching to a traditional device is not feasible.

Despite these promising opportunities, little research exists on Business Process Model and Notation (BPMN) model creation and manipulation via hand gestures in immersive environments such as Augmented Reality and Virtual Reality, and few works have investigated suitable interaction modes for this purpose.

This thesis introduces the 'BPMN Builder', a prototype that supports the creation and manipulation of BPMN models in AR via two interaction modes: direct and indirect interaction. In the direct mode, users select elements directly in the AR diagram via hand gestures, whereas in the indirect mode, an interaction panel serves as the means for diagram navigation and interaction.

To evaluate the prototype, a user study with a within-subjects design was conducted, where every participant tested both interaction modes by performing diagram creation and manipulation tasks. The interaction modes were compared in terms of efficiency, usability, and user preference. Overall, the direct mode performed better, achieving shorter task completion times, a higher System Usability Scale (SUS) score, and higher user preference.

These results suggest that direct interaction is the more suitable mode for BPMN modeling in AR. The findings on gesture-based interaction are expected to be generalizable to other conceptual modeling languages.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

Kurzfassung **ix**

Abstract **xi**

Contents **xiii**

1 Introduction **1**
1.1 Motivation and Problem Statement 1
1.2 Research Objective and Focus 3
1.3 Methodology 4
1.4 Structure of the Work 6

2 Research Background **7**
2.1 Extended Reality 7
2.2 Conceptual Modeling 8
2.3 Business Process Management and BPMN 9
2.4 Direct vs. Indirect Interaction 11
2.5 Positioning 12

3 State of the Art **13**
3.1 Methodology of the Literature Review 13
3.2 Results of the Literature Study 15

4 Prototype Design **25**
4.1 Requirements 25
4.2 Architecture 26
4.3 Shared Design 27
4.4 Direct Interaction 37
4.5 Indirect Interaction 40

5 Prototype Evaluation **45**
5.1 Study Design 45
5.2 Results 47

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

6 Conclusion **61**
6.1 Summary 61
6.2 Limitations 62
6.3 Future Work 62

7 Appendix **65**
7.1 Literature Review 65
7.2 User Study 69

Overview of Generative AI Tools Used **75**

List of Figures **77**

List of Tables **79**

Bibliography **81**

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 1

Introduction

1.1 Motivation and Problem Statement

Conceptual modeling is used for abstracting and representing domain knowledge in a communicable form [Muf25, MGG25]. In recent years, a vision has emerged suggesting that conceptual modeling, and enterprise modeling in particular, will become an integral part of everyday work practices [SFH+18]. This has motivated research into new ways of supporting and enhancing conceptual modeling, for example through multimodal data analysis or AI-based assistance, as discussed in more detail in Section 2.2.

Beyond enhancing the modeling process itself, research has also explored how conceptual models can be visualized and interacted with more effectively. Studies have shown that 3D visualizations can improve the layout and information content of process models, as well as help with analyzing and comparing larger and more complex process model structures [DGZ12, RWL16]. Furthermore, 3D model visualizations can provide a better overview, helping to find errors and mistakes [AS19]. Immersive technologies such as Virtual Reality (VR) offer one way to realize such 3D visualizations, and it has been argued that they can improve comprehension and engagement by providing more intuitive 3D representations [WBG+25, MdSAdS26, RHM+19]. The immersive nature of both AR and VR provides additional visualization options such as multiple perspectives within 3D space, which may further increase user motivation [OPM18].

AR specifically offers additional advantages over VR. While VR completely immerses users and disconnects them from the real environment, AR allows users to still perceive the real world, engage in face-to-face collaboration, and experience little to no motion sickness [RGBA19, RA20a, VVW+18]. Through the integration of virtual objects into the real world and the use of hand gestures, AR requires less prior knowledge for interaction, leading to lower individual learning effort [RGBA19, RA20a]. It also allows users to use any space for visualizing AR content while still being able to use physical documents and

objects needed during the task [MHGG17]. Extended Reality (XR) technologies have seen growing adoption across various industries, with AR support for maintenance tasks being one use case identified as offering substantial value [Muf25].

Combining AR and conceptual modeling goes beyond mere visualization. It offers for the first time an explicit and machine-processable connection between physical reality and formal models, going further than prior approaches where models were created with pen and paper and were physically attached to objects. AR allows model information to be dynamically integrated into the real world and enables entirely new ways of interacting with models and their content [FM24]. By embedding model information directly in the real environment, understanding of complex interrelationships in and between different models can be improved. By visualizing BPMN diagrams in AR, users can perform context-aware modeling by spatially connecting process models with their real environments, for example by displaying a factory workflow directly in the production environment, or showing a service process at the location where it is carried out. This linkage of model elements to real objects and locations can improve the process workflow understanding of non-technical stakeholders [MBF26].

One concrete application scenario is the use of AR BPMN models during process execution. If a domain expert working on a task would like to know about the next steps in a process, in a traditional setting, they would have to stop and switch to a classical modeling tool. If instead, they wear a head-mounted display (HMD) that displays the relevant process information embedded directly into their real working environment, the needed information can be accessed without interrupting the process [MF21]. Similarly, workforce learning processes could be minimized by guiding employees step by step through procedures in AR based on conceptual models. Currently, process descriptions for assembly tasks in industry are mostly conveyed through textual descriptions and 2D drawings, making it difficult to imagine the next step in the real-world environment [Muf25].

A further use case is the integration of AR directly into the BPMN modeling process itself through interactive process elicitation. This is particularly useful in environments where the usage of traditional modeling interfaces are not feasible or convenient, such as dirty, hands-free, or remote collaboration scenarios [FM24]. For instance, while modeling a process, a technical expert might wear an AR device during process execution. Camera sensors could automatically detect the activities being executed, which would then be defined as tasks in the BPMN model and labeled via voice input or hand gestures. In the same way, sequences and other BPMN elements could be specified. The created model can then be visualized on the AR device. For checking correctness and completeness, the model can be transferred and displayed in a conventional 2D desktop environment. Later, the model can be visualized directly during process execution to support employees in performing the process correctly. They could be either given an overview of the whole process, or guided through it step by step with instructions on the task that is currently performed [Muf25].

As elaborated above, 3D models offer multiple benefits for model overview and understanding. However, the creation and manipulation of 3D models on traditional display

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

devices remains complex. Performing modeling tasks directly in AR rather than on desktop could eliminate these issues [MBF26].

Despite these promising opportunities, a review of existing literature, elaborated in Section 3, reveals two main gaps in current research on AR and VR for conceptual modeling:

1. Lack of immersive model creation: While existing research highlights promising interaction and visualization capabilities of AR and VR for conceptual modeling, direct model creation and manipulation in immersive environments remain largely unexplored.
2. Missing interaction techniques for BPMN in AR: There is no research on interaction techniques for constructing and manipulating BPMN models in AR, even though these techniques are a prerequisite for future applications such as interactive process elicitation or context-aware process execution support.

This thesis aims to address both of these gaps, as elaborated in the following section.

1.2 Research Objective and Focus

To address the outlined problems, this thesis presents a prototype for creating, manipulating, and interacting with BPMN models in AR via direct and indirect interaction modes. While the prototype targets BPMN specifically, the insights gained regarding direct and indirect interaction are expected to be generalizable to other modeling languages.

The central research question addressed in this thesis is as follows:

> What is an appropriate means to support efficient and usable BPMN model creation and manipulation in AR?

The main objective is to gain insights into AR interaction mechanisms for conceptual modeling, using BPMN as a concrete modeling language to ground the investigation. To this end, an AR-based BPMN modeling tool supporting both direct and indirect interaction is designed, implemented, and evaluated to determine the more suitable interaction mode. The two modes are compared in terms of efficiency, usability, and user preference in a user study.

This overarching research question is addressed through two subquestions:

Subquestion 1 (SQ1): What is an appropriate means to integrate direct and indirect interaction in a BPMN modeling tool to support the creation and manipulation of models in AR?

This question aims to explore and implement feasible AR interaction modes for BPMN model creation and manipulation. The visualization of BPMN elements is kept close to

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

the conventional 2D notation by displaying standard symbols on 3D cubes. In the direct mode, diagram interaction is performed via hand gestures, selecting elements directly in the AR diagram. In the indirect mode, an interaction panel is used for all creation and manipulation functionality. Both modes are operated via hand gestures.

Subquestion 2 (SQ2): How do the implemented direct and indirect interaction modes compare in terms of efficiency, usability, and user preference?

To address this question, a user study was conducted to evaluate and compare the implemented interaction modes. The primary focus was usability, while efficiency served as additional measure of comparison. User preferences for specific interactions as well as overall mode preferences were also assessed.

By answering these subquestions, the thesis aims to contribute a validated AR BPMN modeling prototype and empirical insight into the suitability and user acceptance of direct and indirect AR interaction modes for BPMN modeling tasks. SQ1 is addressed in Chapter 4, which covers the design and implementation of the prototype, while SQ2 is addressed in Chapter 5, which presents the user study and its results.

The three core terms used throughout this thesis, as reflected in its title, are defined as follows:

Creation refers to the construction of new models. It is a formalization task that involves constructing a syntactically correct model that accurately reflects the semantics of the real world and remains understandable to the reader [SFDA+20].

Manipulation refers to the modification of existing models and objects, allowing users to change attributes such as scale and position [MVK13].

Interaction describes the modality through which creation and manipulation are performed. In this thesis, interaction is realized via gesture control, which refers to the recognition and interpretation of hand gestures to interact with and control a computer system without direct physical contact [RA20b]. Specifically, direct and indirect interaction modes are implemented, as elaborated in the following sections.

1.3 Methodology

To address the research question of this thesis and achieve the outlined objectives, this work adopts a structured methodology grounded in Design Science Research [HMPR04]. This approach is well suited for the systematic and rigorous development and evaluation of artifacts designed to address identified problems [HMPR04]. The methodological approach follows the model introduced by Peffers et al., depicted in Figure 1.1.

1. **Literature Review:** The initial phase involved a literature review [GB09], covering both the theoretical foundations and existing research in the field. This step was part of the Design Science Research rigor cycle [Hev07] and was necessary to

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

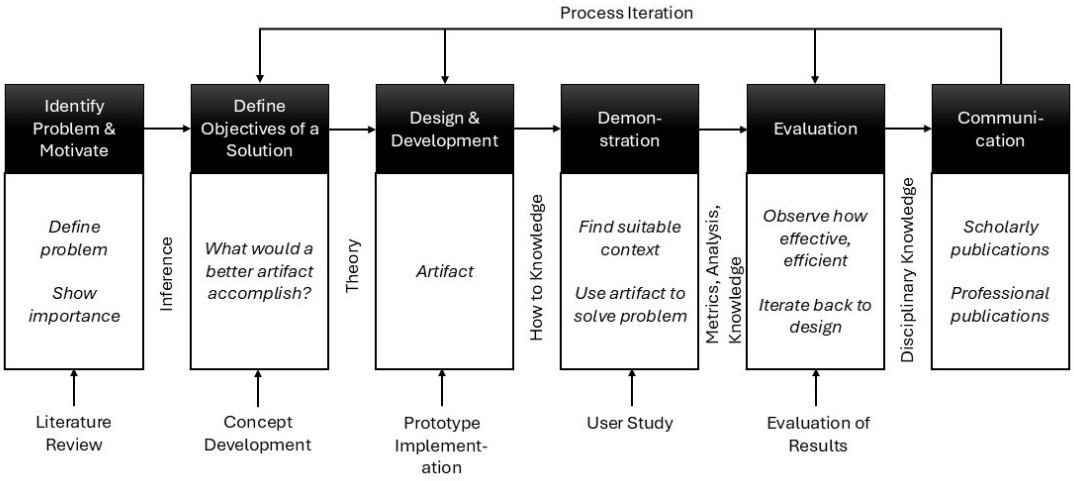


Figure 1.1: Design Science approach (adapted from Peffers et al. [PTRC07])

gain insights into existing approaches concerning the combination of immersive technologies such as AR and VR with conceptual modeling. It helped position the thesis within the current body of knowledge and ensured the novelty of the research. The review followed a reduced version of the guidelines proposed by Kitchenham et al. (2007) [KC07]. The applied search strings, digital libraries, and inclusion and exclusion criteria are documented in Section 3.1.

2. **Concept Development:** Based on the identified research gap, design requirements for the prototype were formulated to ensure goal-oriented development of the software artifact, supporting BPMN model creation and manipulation in AR. Since the goal was to explore and compare the two interaction modes of direct and indirect interaction in AR, the two modes were included in the prototype. This phase laid the groundwork for the subsequent implementation and evaluation, ensuring that the developed artifact addresses the identified requirements, which are outlined in Section 4.1.

3. **Prototype Implementation:** A prototype of the BPMN Builder was implemented as the basis for exploring and evaluating the interaction modes. It enables users to create, manipulate, and interact with BPMN models while being compatible with both desktop and AR environments. This cross-platform compatibility ensures that models can be displayed and edited consistently across both settings. Since two interaction modes were explored, a parallel prototyping approach was applied, meaning that direct and indirect interaction were implemented and evaluated in parallel rather than refining a single solution iteratively [CVL+17].

4. **User Study:** A user study combining both quantitative assessments and qualitative feedback [Bar11] was conducted to gather data concerning the efficiency and

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

usability of the interaction modes, including questions from the System Usability Scale (SUS) [Bro96], as well as user preferences. The study followed a within-subjects design [Bar11] combined with a comparison-of- designs approach. This means that each participant tested and compared both interaction modes, allowing to capture performance differences as well as user preferences. The specifics are elaborated in more detail in Section 5.1.

5. **Evaluation:** Finally, the results of the user study were evaluated. Quantitative data such as task completion time and SUS scores were analyzed using descriptive statistics and non-parametric significance tests, and SUS scores were interpreted according to the classification by Bangor et al. [BKM09]. Qualitative feedback from open-ended survey questions was analyzed to identify recurring themes concerning usability, interaction preferences, and suggestions for improvements. This allowed the two interaction modes to be compared in terms of efficiency, usability, and user preference, and to find the more suitable approach for BPMN modeling in AR, thereby providing an answer to the research question.

1.4 Structure of the Work

The thesis is organized into the following chapters:

Chapter 2: Research Background: This chapter provides the theoretical foundation for this thesis, covering conceptual modeling, BPMN modeling, extended reality, and direct and indirect interaction.

Chapter 3: State of the Art: In this chapter a comprehensive review of existing literature on conceptual modeling in immersive technologies such as AR and VR is provided. The insights from this review establish the foundation for the prototype development.

Chapter 4: Prototype Design: The focus of this chapter is the description of the design requirements and the subsequent design and implementation of the BPMN Builder prototype. It covers the shared design decisions applicable to both interaction modes, as well as the specific design of the direct and indirect interaction modes.

Chapter 5: Evaluation: This chapter outlines the user study design and presents the evaluation results. It presents the efficiency, usability, and user preference findings, as well as qualitative findings and additional observations from the user study.

Chapter 6: Conclusion: The final chapter summarizes the research findings and reflects on how the research objectives were met. Limitations of the work and directions for future research are also discussed.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 2

Research Background

This chapter provides the necessary background knowledge for understanding the content of this thesis. Section 2.1 gives an overview of the Extended Reality (XR) continuum and how AR and VR are distinguished. Section 2.2 explains the meaning of the overarching term conceptual modeling and discusses emerging research directions in this area, and Section 2.3 discusses Business Process Management (BPM) and BPMN modeling in more detail. Section 2.4 explains the difference between direct and indirect interaction, and Section 2.5 positions the research of this thesis relative to the concepts introduced in this chapter.

2.1 Extended Reality

Augmented Reality, Augmented Virtuality, and Virtual Reality, together referred to as Extended Reality, have been increasingly researched and applied in industry over the past years [DBGJ22]. XR is already used in gaming, navigation, maintenance tasks, and training, and its potential for industry applications has been highlighted in multiple studies [CGRR18, Muf25].

According to [LDL21], an AR system is characterized by four key properties: it combines real and virtual objects within a real environment, aligns them spatially with one another, operates in real time, and allows six degrees of freedom for interacting with virtual objects. Six degrees of freedom means that users can freely move and rotate objects along all three spatial axes, controlling both position and orientation.

AR and VR are closely related but differ in one key aspect: VR immerses the user in a fully virtual environment, whereas AR keeps the user grounded in the real environment by displaying virtual content in it [YGWE21]. In VR, both the objects and the surrounding environment are virtualized, creating a simulated experience that can range from closely resembling reality to being entirely different from it [LDL21]. In AR, by contrast, users

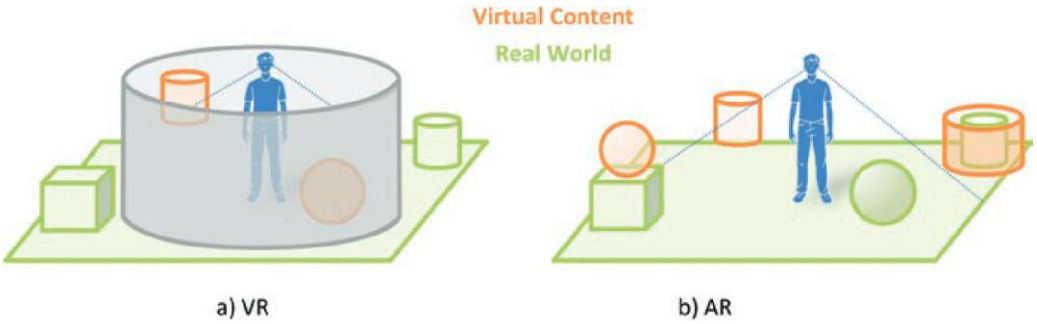


Figure 2.1: AR compared to VR (adopted from [DBGJ22])

interact with both virtual content and the real environment simultaneously, and the two are not strictly separated. Virtual and real elements can overlap, be superimposed onto one another, or interact directly [DBGJ22]. This distinction between AR and VR is illustrated in Figure 2.1.

Mixed Reality (MR) describes a continuum along which real and virtual content are mixed together [LDL21]. Based on Milgram's reality-virtuality continuum [MTUK94], MR extends between reality and virtuality, whereby the share of reality continuously decreases while that of virtuality increases. As far as the share of virtuality is prevailing, without the environment being completely virtual, one speaks of Augmented Virtuality. If on the other hand the share of reality is larger, then it is referred to as AR [LDL21]. The MR continuum is visualized in Figure 2.2, and an Euler diagram showing the relationships between the different XR concepts is shown in Figure 2.3.

2.2 Conceptual Modeling

Conceptual modeling is concerned with representing aspects of the physical or social world through predefined elements and structures to capture the essential structure and behavior of a domain [Myl01, Fil24, PDG21]. Conceptual models do not create new concepts, but are information artifacts that aim to structure existing concepts in a communicable and understandable form [MGG25]. They are used to reduce complexity and abstract the real world. Furthermore, they support communication between different stakeholders, serve as a basis for analysis and validation, and provide the foundation from which systems are implemented [RAB+15].

However, a large share of human knowledge is tacit, reflecting what experts actually do rather than what they explicitly articulate, and often resisting formal codification [CBB17]. Traditional modeling practices are based on what experts communicate in interviews, documentation, and explanations, leading to a gap between reported and actual behavior [LNR+25, GBP24]. To close this gap, research has explored enhancing conceptual models with the tacit knowledge that experts rely on but rarely verbalize

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

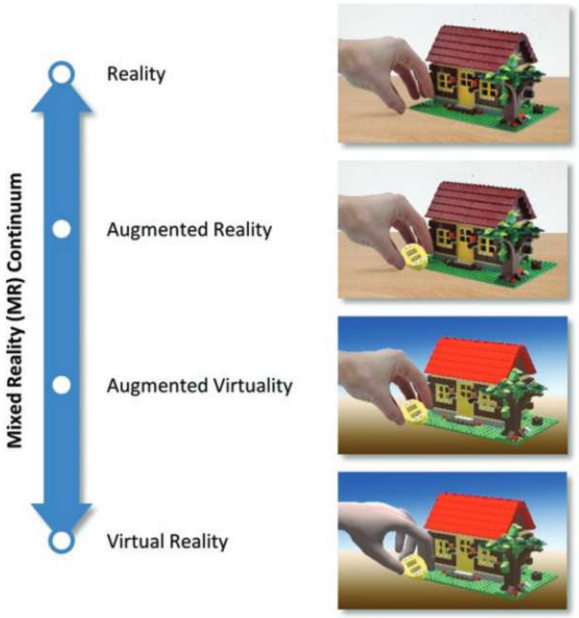


Figure 2.2: Reality-virtuality continuum (adopted from [DBGJ22])

directly [Gav26]. Capturing this relies on multimodal data, such as video, audio, spatial movement, and event logs, often combined with Artificial Intelligence (AI) based analysis [GBP25, GBP24, CPR+25, CPR+26, BSV+25]. AI has also been researched in combination with conceptual modeling by assisting modelers or automating modeling activities [BAR23]. This assistance can take the form of structuring conceptual schemas, as well as describing, extending, or creating conceptual models [FFK23, FFK25].

In the context of enterprise modeling, conceptual modeling is used, for example, for the representation of business processes, organizational structures, and enterprise architectures [Fil24, SFH+18]. A range of standardized conceptual modeling languages have been developed for this purpose, including BPMN [Gro11] for business process modeling, ArchiMate [Gro26b] for enterprise architecture modeling, and the Unified Modeling Language (UML) [Gro26a] for software and systems modeling [Fil24, Gav26]. This thesis focuses on BPMN as the target modeling language, while related work also covers other conceptual modeling languages such as UML and ArchiMate, as elaborated in Section 3.

2.3 Business Process Management and BPMN

Business Process Management addresses how work is performed within an organization. It aims to ensure consistent outcomes and identify opportunities for improvement by focusing on chains of events, activities, and decisions, referred to as processes, that create

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

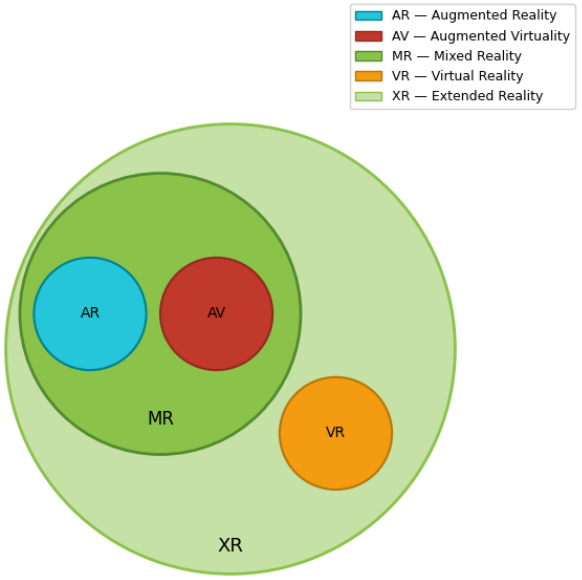


Figure 2.3: Extended Reality Euler diagram (adapted from [DBGJ22])

value for both the organization and its customers. BPM spans the full lifecycle of a business process. This lifecycle typically consists of six phases [DLRMR18]:

1. *Process identification*, where a business problem is posed and relevant processes are identified,
2. *Process discovery*, where the current state of the identified as-is processes is documented,
3. *Process analysis*, where the issues with the current as-is processes are analyzed and documented,
4. *Process redesign*, where to-be process models are created reflecting the process changes needed to overcome the previously identified issues,
5. *Process implementation*, where the necessary changes to transition from the as-is process to the to-be process are made, and
6. *Process monitoring*, where the running redesigned process is monitored and analyzed regarding its performance.

BPMN [Gro11] is a widely used and standardized notation for BPM. It is primarily situated within the process discovery and process redesign phases of the BPM lifecycle,

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

where it is used to document as-is processes and specify to-be process models. BPMN is designed to be understandable for all business stakeholders, while capturing process complexity and execution details [OPM18]. Over the years, the models have become more complex and integrated with both business and IT, leading to reflections on novel ways of modeling, including the integration of BPMN into immersive environments [OPM18].

In particular, this thesis focuses on BPMN 2.0. The BPMN elements supported in the prototype and their visual representation are shown in Table 4.1 in Section 4.

BPMN includes various element types. This thesis does not use them all, but mainly focuses on nodes, connections, and data objects. Nodes are split into *Activities* representing tasks that are performed, *Events* conveying that something happens during the course of a process, and *Gateways* splitting the control flow. Parallel tasks can be modeled through *Parallel Gateways*, and alternative paths through *Exclusive Gateways*. Events are split into start, intermediate, and end events: the *Start Event* marks the beginning of a process, the *Intermediate Event* occurs between the start and end and affects the process flow, and the *End Event* indicates where a process ends. Regarding connections, *Sequence Flows* show the order in which tasks are performed, while *Associations* link *Data Objects* with tasks. Data Objects represent required input or output elements [SFDA+20, Gro11].

2.4 Direct vs. Indirect Interaction

Direct and indirect interaction are two fundamental paradigms in XR interface design. Direct interaction allows users to manipulate virtual objects through natural, real-world-like hand movements, which creates a strong sense of control and presence due to the immediate and intuitive feedback it provides. Indirect interaction, on the other hand, relies on interface elements such as buttons and menus as intermediaries. While this introduces an additional layer of cognitive effort, it can offer greater precision and deliberate control over virtual objects. Both approaches contribute to a sense of agency, which refers to the user's perception of being the cause of actions and changes in the virtual environment. When users observe a clear and consistent correlation between their inputs and the resulting system response, this strengthens their feeling of intentionality and control, regardless of whether the interaction is direct or indirect [DG24].

Dastmalchi et al. [DG24] compared direct and indirect interaction techniques in VR, where participants were asked to design a new office space. Their use of both methods for copying, rotating, moving, and scaling objects was analyzed. The study suggests that indirect interaction may offer greater accuracy when executing commands to manipulate objects in the virtual environment, while direct interaction involves less time for users to select and manipulate objects. These findings motivated the implementation of both interaction modes in the BPMN Builder prototype presented in this thesis, with the goal of comparing them in a user study to determine which mode is more suitable for model creation and manipulation in AR.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

2.5 Positioning

This thesis is positioned at the intersection of conceptual modeling, BPM, and immersive technologies. It supports the process discovery and process redesign phases of the BPM lifecycle, as introduced in Section 2.3, by enabling BPMN model creation and manipulation directly in AR. This connects to the broader body of research combining AR and VR with conceptual modeling, as introduced in Section 1.1 and systematically reviewed in Chapter 3. Within this body of research, this thesis specifically investigates and compares direct and indirect interaction modes, a dimension not yet thoroughly addressed by existing work.

While research on enriching conceptual models with multimodal data and AI-based analysis, as discussed in Section 2.2, has primarily focused on improving model accuracy and automating parts of the modeling process, this thesis instead addresses a complementary challenge: how models can be created and manipulated within immersive environments. Even though AI-assisted modeling is a relevant and active research direction, the prototype developed in this thesis does not incorporate AI-based interaction. Instead, it investigates the foundations of gesture-based interaction for conceptual modeling tools, providing empirical insights into direct and indirect interaction that could complement AI-assisted approaches in future work.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 3

State of the Art

This chapter elaborates on the current state of the art regarding the integration of conceptual modeling into immersive environments such as AR and VR. After an explanation in Section 3.1 of the methodology applied for the literature review, the results of the literature study are presented in Section 3.2, and selected approaches are discussed in more detail.

3.1 Methodology of the Literature Review

Relevant literature was identified using a search string derived from keywords and search criteria. To ensure only relevant studies were included, exclusion criteria were defined and applied to filter the results.

Although this thesis focuses on AR, approaches involving VR and MR were also included to obtain a comprehensive overview of the current research status on modeling in immersive environments. Similarly, keywords were not limited to a single conceptual modeling language, since insights from one conceptual modeling language are expected to be generalizable to others. Keywords for AR, VR, MR, and XR were therefore combined with conceptual modeling languages such as BPMN, UML, and ArchiMate, as well as terms related to visualization, creation, manipulation, and interaction.

The search was executed on the scientific databases Scopus[1], IEEE Xplore[2], and ACM Digital Library[3], adapted to the syntax of each database. It was applied to the *title*, *abstract*, and *keywords* of each publication.

Three groups of keywords were combined using Boolean operators to form the search string. The XR group (XR) covers terms related to immersive technologies, the modeling

[1] https://www.scopus.com/
[2] https://ieeexplore.ieee.org/Xplore/home.jsp
[3] https://dl.acm.org/

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

group (MK) covers conceptual and enterprise modeling languages and notations, and the interaction group (IK) covers terms related to visualization and interaction. The final search string takes the form:

$$Q = XR \ \wedge \ MK \ \wedge IK$$

where terms within each group are combined with OR, and the three groups are combined with AND. The complete query is defined as follows:

```
TITLE-ABS-KEY((AR OR VR OR MR OR XR OR 'augmented reality' OR
'virtual reality' OR 'mixed reality' OR 'extended reality') AND
(BPMN OR UML OR ArchiMate OR EA OR 'business process model' OR
'business process' OR 'conceptual modeling' OR 'conceptual
modelling' OR 'enterprise architecture') AND (visualization OR
visualisation OR interaction OR manipulation OR creation OR
editing OR modeling OR modelling))
```

The following exclusion criteria were applied to identify and filter out irrelevant publications:

- **EC-1**: Publications not written in English
- **EC-2**: Publications not related to the subject areas Computer Science, Engineering, Business Management and Accounting, or Decision Sciences
- **EC-3**: Publications published in 2016 or later
- **EC-4**: Publications not available as full text, as well as non-scientific publications such as posters and extended abstracts

This led to the following number of identified publications:

- Scopus: 320 results
- IEEE Xplore: 111 results
- ACM Digital Library: 6 results

The results from the three scientific databases were merged and duplicates were eliminated. As a next step, the abstracts of the retrieved publications were screened to identify those most relevant to the research scope. Publications that were irrelevant to the research question or did not fit the purpose of this review were excluded. Where no clear decision could be made based on the abstract alone, the full text was skimmed to determine inclusion or exclusion. This led to a total of 31 relevant papers, which are discussed in the following section.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

3.2 Results of the Literature Study

This section presents the results of the performed literature review. After an overview of the identified approaches, selected approaches are elaborated in more detail, categorized into two groups: those that cover model visualization and simple interactions, and those that support model creation or manipulation in immersive environments. Finally, the results of the literature review are discussed.

3.2.1 Overview

Research on conceptual modeling in immersive environments like AR and VR has been conducted for diverse modeling languages. In the following, an overview of the current research status is presented and selected approaches are elaborated in more detail. These approaches were selected based on their closeness to the specific gap addressed by this thesis, namely the creation and manipulation of process models in immersive environments. Table 3.1 summarizes key aspects of selected approaches, while the full list of the 31 approaches identified through the literature search is provided in Tables 7.1, 7.2, and 7.3. Of the 31 approaches, the majority focused on VR, while 10 incorporated AR or a combination of AR and VR, and one used a 3D desktop environment. Regarding modeling languages, 10 implemented UML and 8 focused on BPMN. As can be seen from the tables, most works focued on the visualization of diagrams and only allowed for simple interactions. Only a few supported diagram creation and manipulation in AR or VR. Furthermore, most approaches relied on VR controllers, whereas only three integrated hand gestures for diagram interaction.

3.2.2 Visualization and Simple Interactions

The majority of the identified research approaches focus on the visualization of conceptual models in an immersive environment [MF21, RWL16, Muf25]. Some allow for limited diagram manipulations or interactions without supporting full model creation, such as changing the visualization or color of model elements [ACR+19, MHGG17], manipulating properties in diagrams, like the visibility of attributes [RHM+19], zooming into models [BKvdW24], or navigating through the diagram [MdSAdS26]. Grambow et al. [GHOP22] used BPMN to support IIoT-aware processes, augmenting them with additional information in AR, though the BPMN models themselves are not displayed in AR.

Oberhauser et al. investigated the feasibility of visualizing and interacting with diagrams in VR space. In their research, they investigated diverse models and modeling languages including BPMN, ArchiMate, and UML. Their research comprises VR-BPMN [OPM18], VR-EA [OP19], VR-EAT [OSM20], VR-UML [Obe21], VR-EA+TCK [OBS22], VR-EvoEA+BP [OBS23], and VR-SDLC [Obe24]. Across all their implementations, they used fly-through and teleportation to navigate through the models in a VR space and used VR controllers as interaction modality. The following subsections

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Table 3.1: Overview of selected works on conceptual modeling in immersive environments

Reference	Tech.	Notation	Creat.	Manip.	Interaction
[OPM18]	VR	BPMN	N	N	Controllers
[OP19]	VR	ArchiMate, BPMN	N	N	Controllers
[Obe21]	VR	UML	Y	Y	Controllers
[GRM20]	VR	BPMN	N	Y	Controllers
[SKKA21]	MR	IoT processes	Y	Y	Gestures
[PSLT20]	VR	Process models	Y	Y	Controllers
[PSMT21]	VR	Process models	Y	Y	Controllers
[YGWE21]	VR	UML	Y	Y	Controllers
[YGWE23]	VR	UML	Y	Y	Controllers
[WBG+25]	VR	UML	Y	Y	Controllers
[GSD24]	AR	Causal maps	N	Y	Gestures
[RGBA19]	AR	EA	N	N	Controller, voice
[MBF26]	AR	Conceptual models	N	Y	Gestures

Creat. = creation, Manip. = manipulation, Y = supported, N = not supported

elaborate on three of their solution concepts in more detail: VR-BPMN, VR-EA, and VR-UML.

VR-BPMN [OPM18] addresses the connection between BPMN and VR and investigates visualization options to enhance model comprehension. To represent BPMN elements, the authors used different 3D shapes with BPMN symbols projected onto every side of the geometry. Furthermore, they explored novel ways of visualizing traditionally 2D BPMN models in a 3D space, leveraging the enhanced visualization capabilities of VR. For annotating BPMN elements, they used colors and popups to reduce the amount of information visible in the scene, making it less cluttered and easier to grasp. Subprocesses were visualized as pyramids displayed beneath elements, introducing a new concept to the conventional 2D BPMN visualization, as depicted in Figure 3.1. Model interaction was performed via VR controllers, supplemented by a mixed reality keyboard for element tagging. To evaluate their prototype, the researchers conducted a user study with eight computer science students with no or limited prior experience with BPMN. They compared their VR solution with paper-based and PC-based BPMN models to analyze how VR affects process analysis, comprehension, and interaction. Students were shown BPMN models across the three environments and asked to complete tasks such as identifying modeling errors. The results showed that process comprehension was similar across all modes. Compared to the PC-based version, VR-BPMN was more efficient, but slower than the paper-based mode, and users found the VR controls intuitive. However, users preferred the PC-based tool over the VR solution. While this paper investigates

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

novel ways of BPMN model visualization, it does not offer model creation or manipulation functionality.

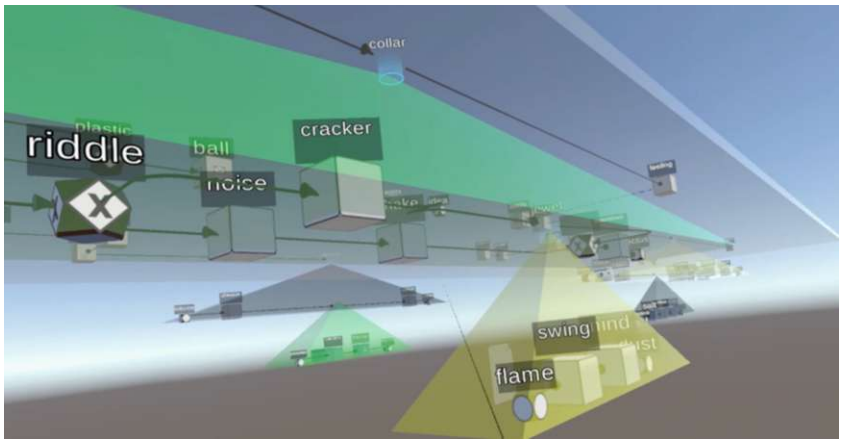


Figure 3.1: VR-BPMN prototype [OPM18]

In **VR-EA** [OP19], a VR hypermodel solution was presented for visualizing, navigating, and interacting with ArchiMate and BPMN models in VR. The solution allows ArchiMate and BPMN models to be visualized in the same field of view, providing a deeper understanding of business process context and potential dependencies. The researchers aimed to improve model comprehension in VR space by keeping the design close to the characteristics of the original 2D ArchiMate and BPMN elements. Oberhauser et al. projected images of the 2D elements onto generic 3D cubes and added textual labels, making their approach compatible with different modeling notations. They decided to make their elements opaque since partially transparent elements could become visually confusing, even though that could lead to the issue of not knowing whether an element hides behind another one. Labels are semi-transparent and rotate automatically toward the camera to remain readable from all angles. Users can interact with the diagram by making hyperplanes visible or invisible by selecting the plane, enabling or disabling inter-view connections, and selecting connections to emphasize them. The solution is shown in Figure 3.2. It supports visualization of ArchiMate and BPMN diagrams and some simple diagram interactions, but does not support diagram creation or manipulation. For evaluation, a case study for the ArchiMate modeling language was conducted.

Gall and Rinderle-Ma [GRM20] explored how to visualize BPMN in 3D and developed their 3DViz language. In a controlled experiment, they compared 2D, 3D, and 3D with VR BPMN focusing on efficiency. Their 3D visualization of BPMN elements resembles the standard notation. However, start and end events are visualized as black balls, distinguished by a grid of three lines on the end event. Exclusive, and parallel gateways are visualized as cubes with a plus and an X on them. To distinguish the two, the exclusive gateway additionally has a line on the bottom of the X. Activities are visualized as cuboids, where each long side can display its on label, for example the name of the activity on the front, the ID on top, and the IT service on the bottom.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

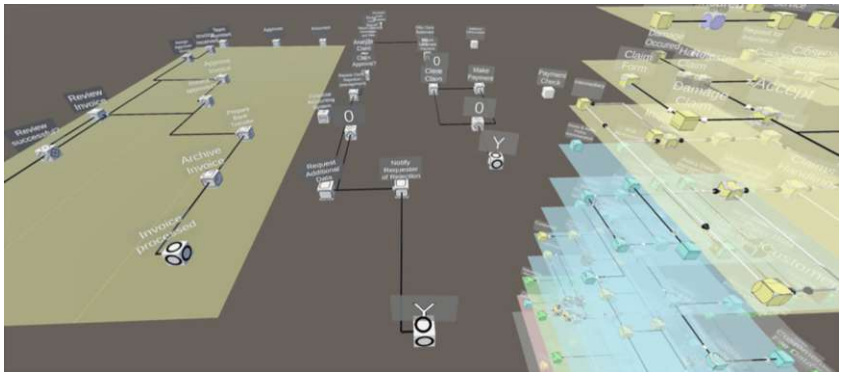


Figure 3.2: VR-EA prototype [OP19]

3DViz allows three visualization options for activities: positioning on the axes, scaling, and labeling. Depending on the placement and scaling on the axes, different attributes can be communicated (e.g. activities with higher costs are positioned further back on y or z-axis, or scaled larger). Both numerical attributes (e.g. costs) and non-numerical attributes (e.g. roles) can be visualized this way. For non-numerical attributes, swimlanes and different positioning on the z-axis is used. Overall, up to nine attributes can be encoded in a single 3D BPMN diagram. The solution is shown in Figure 3.3. Diagrams are generated from XML files and can be viewed on both desktop and VR devices, with zooming, panning, rotating, and clicking available in both environments. On desktop, these are performed via mouse interactions, while in VR, head movements are used for navigation and VR controllers for clicking. In VR, diagrams are displayed on a conference table in a virtual conference room. During the experiment, participants selected elements based on given questions (e.g. 'select the activity with the longest duration'). VR achieved the lowest number of wrong answers and the best task completion time, followed by the 2D tool and then 3D without VR. When asked about preferences, the majority favored the 2D tool for simple tasks but preferred the 3D VR version for more complex tasks. This indicates that VR can offer benefits for conveying richer attribute information compared to traditional 2D modeling. However, the solution does not support diagram creation, manipulation, or direct interaction with diagrams in VR.

Rehring et al. [RGBA19] explored the visualization of and interaction with Enterprise Architecture (EA) models in AR, displaying models with three layers. Users can select different functions to change the visualization of the model. For example, selecting an object in analysis mode causes lines to appear that visualize connections to other EA objects across layers, helping to identify relationships. Users can also change size or color of objects to communicate for instance the importance, or positive or negative assessments. A filter option additionally allows users to show or hide selected layers or object types. Function selection is performed using the cursor of the AR glasses or via voice control. While the implementation supports various interactions with EA models, it does not offer diagram creation functionality. The subsequent evaluation of the solution [RA20a] indicated that AR is suitable for EA analysis.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

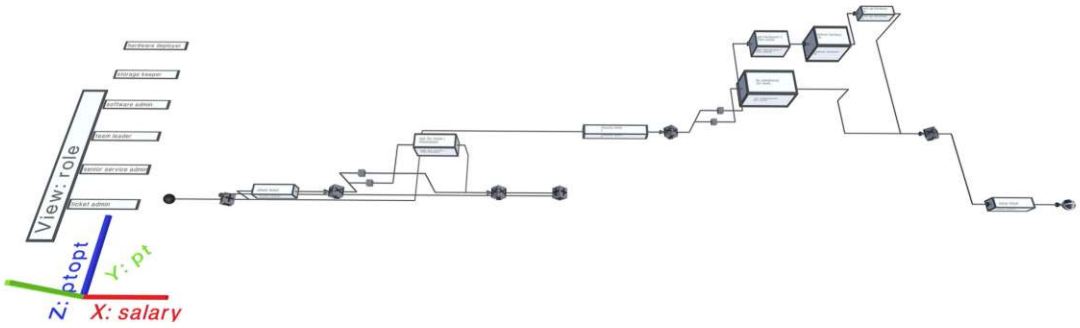


Figure 3.3: 3DViz prototype [GRM20]

Research has also been conducted in the direction of gamification in combination with VR modeling [YSB+22, YSB+24a, YSB+24b]. **Yigitbas et al.** investigated the gamification potential of UML diagrams in VR, implementing different game modes like hangman, where users could perform a small set of diagram manipulations such as completing and correcting diagrams. While this work supports some model creation and manipulation, the interactions are constrained to a limited set of actions rather than offering general-purpose modeling functionality.

3.2.3 Model Creation and Manipulation

Some of the identified approaches support model creation and manipulation for different modeling notations, as elaborated in the following. Oberhauser et al. introduced in their solution concepts VR-UML [Obe21], VR-EA+TCK [OBS22], VR-EvoEA+BP [OBS23], and VR-SDLC [Obe24], a VR-Tablet as an indirect interaction technique for diagram interaction. The VR-Tablet resembles a conventional tablet with different options and tabs. In [OBS22, OBS23, Obe24], the VR-Tablet is used for browsing, filtering, and searching for nodes.

In **VR-UML** [Obe21], UML diagrams are visualized as 3D hyperplanes in VR space. The solution supports hypermodeling, meaning that UML, ArchiMate, BPMN, and EA tool models can be visualized simultaneously in the same VR space, enabling cross-model analysis. Diagrams can be moved to view them side by side, or hidden to reduce visual clutter. UML elements are visualized as shallow boxes with generic UML icons placed on the upper right and lower left on top of the element. Interaction is performed via the VR controllers and a virtual tablet, which provides CRUD (create, retrieve, update, delete) capabilities for each element and a virtual keyboard for text entry. The VR-Tablet was chosen as the primary interaction mechanism since tablet usage was considered familiar and intuitive. Diagram planes are made slightly transparent to avoid visually occluding other diagrams. Figure 3.4 depicts the virtual tablet. The study indicated that VR can support modeling and that the solution is particularly advantageous for complex model structures with inter-diagram relations. While VR-UML supports model creation and

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

manipulation, interaction relies on VR controllers rather than hand gestures.

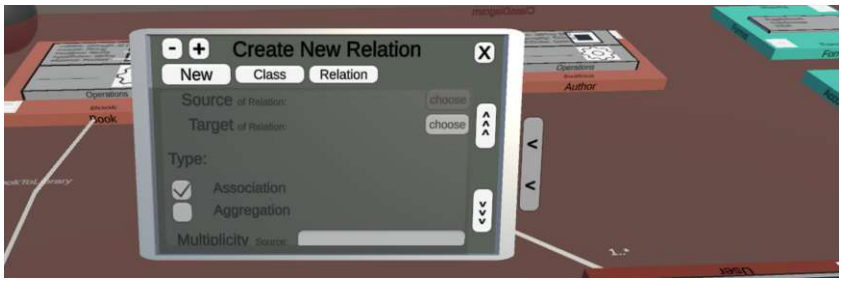


Figure 3.4: VR-UML virtual tablet [Obe21]

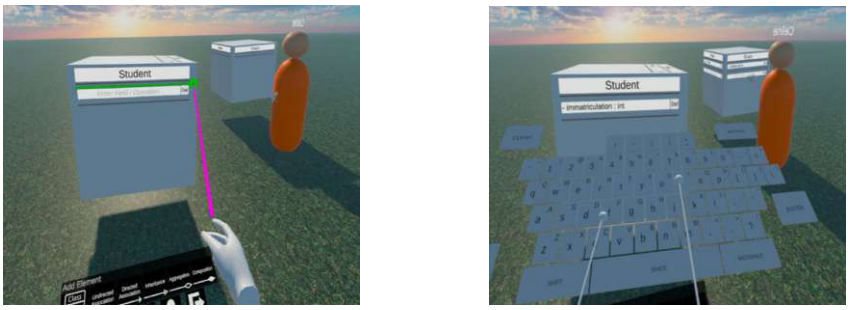
In **MM-AR** [MBF26], Muff et al. presented a web-based modeling client that allows users to create conceptual 2D or 3D models from predefined metamodels. The modeling client has a three-dimensional drawing area and runs on desktop, with the resulting models subsequently visualized in AR space. Since the software supports diverse model types, BPMN diagrams can also be modeled and later visualized. Direct hand interaction is supported for actions such as moving elements. However, model creation is performed on desktop rather than directly in AR.

Seiger et al. [SKKA21] presented HoloFlows, a Mixed Reality system for modeling Internet of Things (IoT) processes. Their process models are designed similarly to BPMN, with gateways and connections. A virtual mouse pointer is controlled via head and eye movements, and pinching gestures (pressing index finger and thumb) are used to perform a click. Users can create connections between IoT devices by pinching source and target elements, and can select whether an OR or AND gateway should be added when creating a connection. While this solution implements elements of model creation similar to BPMN, it does not support full BPMN model creation since it focuses specifically on IoT processes.

Yigitbas et al. [YGWE21, YGWE23] focused on collaborative UML modeling in VR. To visualize 2D UML elements in a 3D environment, rectangles were transformed into 3D cubes and cuboids, and lines into thin tubes. Class names are displayed on each side of class elements to ensure visibility from all perspectives. Elements can be moved using a grabbing interaction with VR controllers, and connections can be adjusted by grabbing and moving their start and end points. For text input, a virtual keyboard is used. The authors decided to avoid speech recognition, since UML class diagrams often include special characters and capitalization, making this approach impractical. For model editing, each collaborating user has a Creation Menu following their virtual avatar. To keep the menu compact, only a few icons are displayed, while additional panels can be opened or closed as needed. Elements can be added by using the creation menu and deleted by switching to delete mode, pointing at an element with a laser pointer, selecting it, and pressing a button to delete it. In VR, users are represented as 3D avatars and can navigate either through natural physical movement (the user moves in the real room and the avatar moves in VR space) or teleportation. Figure 3.5 depicts

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

their solution. To evaluate the prototype, the researchers compared it to PC-based collaborative modeling. They assessed both modes with respect to efficiency, effectiveness, and user satisfaction by conducting a user study with 24 participants, where participants had to perform two different small UML class diagram modeling tasks. The results showed that while efficiency and effectiveness were lower in the VR setting, users had more fun and were more motivated. Even though the solution offers model creation and manipulation functionality, it relies on VR controllers rather than hand gestures, and focuses on VR rather than AR, which are the key aspects investigated in this thesis.



(a) (b)

Figure 3.5: Collaborative UML modeling in VR [YGWE23]

White et al. [WBG+25] adapted the application from [YGWE23] for a study comparing UML learning in VR with a standard 2D desktop application. Their results showed that long-term retention was comparable for both modes, but VR participants performed better for application-based diagram questions and reported higher confidence in their ability to model in UML. The authors suggested that immersive VR environments may support long-term retention and increase learner confidence when working with complex technical subjects.

Giabbanelli et al. [GSD24] developed a system to support multiple users in collaboratively creating a joint graph map in AR. Each user starts with their own predefined graph, and the goal is to reach a consensus. The graph maps can be visualized in either 2D or 3D depending on user preference. In the graphs, nodes are visualized as spheres and edges as straight lines. Users can move, delete, and merge nodes, as well as create and delete edges. Creation of nodes is not supported, since the objective is for users to reach a consensus based on their existing individual graphs rather than building a new graph from scratch. When users click on a node, the options to delete it, create a new edge, or merge nodes appear. Nodes can be moved by pointing at them, grabbing them by performing a pinch gesture, and then relocating them. The solution supports some model manipulation but does not offer full model creation. The subsequent evaluation [SG25] revealed that all participants were able to complete the given modeling tasks in AR and that usability ratings were generally positive. The authors argued that this suggests that AR is a feasible approach for supporting essential collaborative mapping tasks such as finding, deleting, and merging nodes, creating edges, and correcting errors. While this work

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

demonstrates the feasibility of hand gesture-based interaction for model manipulation in AR, it targets causal maps rather than BPMN and does not support model creation from scratch, which is the focus of this thesis.

Pöhler et al. [PSLT20] developed a system for participatory process modeling in VR. They used an immersive VR environment, where processes can be modeled in different virtual office rooms. Two roles are supported: a VR user who interacts with virtual objects and moves within the VR space, and a desktop user who controls an avatar via mouse and keyboard. Communication between users is done via a headset. A simplified, semi-formal modeling notation is used, where events and functions are displayed as rectangles that can be linked to represent sequence flows. Process field colors can be modified and departments can be added to define involved stakeholders. Supported functions include adding process descriptions and organizational units, selecting colors, deleting process functions, and creating and accepting edges. Since the user study indicated that the environment should also support AND and OR connectors, the solution was further developed in [PSMT21].

The follow-up solution [PSMT21] similarly allows users to create VR working environments, where process steps are placed at locations corresponding to where they are executed in the real world, as shown in Figure 3.6. The system supports multi-user access with both a VR and a desktop user collaborating on process modeling. Users can add activities, gateways, connectors, start events, and end events, all of which can be created, deleted, and repositioned. Activities are visualized as blue balls and positioned by pointing a controller at the desired location and pressing the trigger. A trash can icon beneath each activity allows deletion, and elements can be repositioned by holding the trigger while pointing at them. Activities consist of descriptions and the execution roles. Connections are created by targeting a ring at the bottom of an element with the cursor, which causes an arrow to appear that can be dragged and connected to other elements. The system also allows graphical notes such as pictures of 2D BPMN diagrams to be opened in VR for reference. The evaluation showed that the system motivated employees to participate in process management activities, enhanced the process experience compared to 2D representations, and that participants enjoyed connecting activities to their real-world locations. However, the required effort was considered too high. While the diagram elements are inspired by BPMN, they do not closely resemble standard BPMN notation, and the implementation focuses on spatializing process steps rather than providing a general model overview.

3.2.4 Discussion of Findings

Currently, there is very little research on AR/VR-assisted modeling where model creation and manipulation are performed directly in immersive environments [Muf25]. While some approaches support model creation and manipulation, these are implemented in VR rather than AR, and rely on controllers rather than hand gestures. For BPMN specifically, no identified approach supports full model creation and manipulation directly in AR. Furthermore, none of the identified approaches compared different interaction modes.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

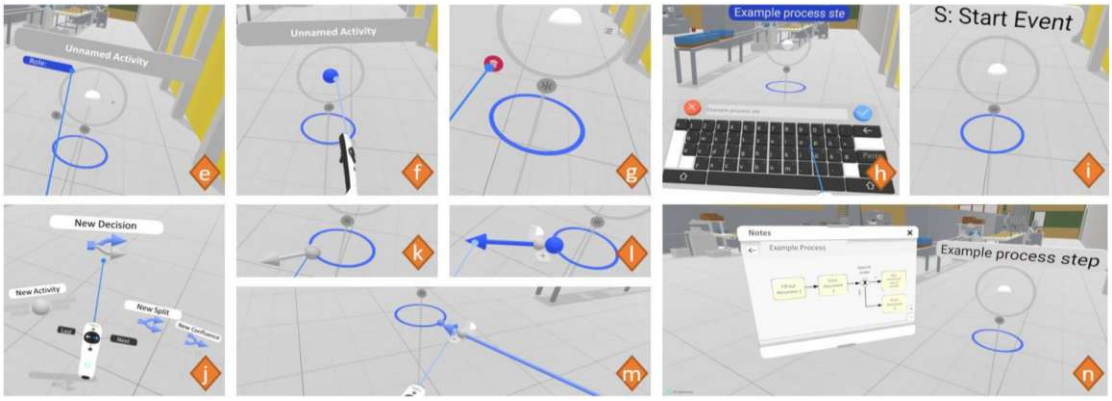


Figure 3.6: Participatory process modeling in VR [PSMT21]

Existing work focuses on presenting single proof-of-concept solutions demonstrating that model visualization or interaction is possible in immersive environments, without systematically evaluating and comparing alternative interaction techniques.

This thesis directly addresses this gap by designing, implementing, and evaluating direct and indirect hand gesture-based interaction modes for BPMN modeling in AR.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 4

Prototype Design

The prototype consists of three parts: a desktop version, the AR direct interaction mode, and the AR indirect interaction mode. In AR, users can interact with BPMN models in two ways: directly, by manipulating elements through hand tracking and pinch gestures (pressing the index finger and thumb of either hand together), or indirectly, through a floating AR panel operated via buttons. The following sections describe the architecture and design of the prototype in more detail.

4.1 Requirements

Before the implementation of the prototype, functional requirements were defined to guide the development. These cover model creation and manipulation functionality, cross-platform compatibility, and interaction modality.

- **R1 – Model Creation:** The prototype must support the creation of new BPMN models, including adding elements (tasks, events, gateways, and data objects), labeling elements, and connecting elements via sequence flows and associations.
- **R2 – Model Manipulation:** The prototype must support the manipulation of existing models, including deleting elements, relabeling elements, and transforming elements via translation, scaling, and rotation.
- **R3 – Cross-Platform Compatibility:** The prototype must be compatible with both desktop and AR environments, allowing models to be created and manipulated in either setting and transferred between them.
- **R4 – Persistence:** Models must be synchronized with a database to enable seamless switching between environments. Additionally, diagrams should be saved automatically at regular intervals to prevent data loss.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

- **R5 – Mode Switching:** Seamless switching between direct and indirect interaction modes must be possible without loss of model state, allowing users to continue working on the same diagram after switching modes.

- **R6 – Diagram Management:** Users must be able to save diagrams to the database, load existing diagrams from the database, and export and import diagrams as local files.

- **R7 – Gesture-Based Interaction:** Both interaction modes must be operable via hand gestures, enabling model creation and manipulation without requiring additional devices.

- **R8 – Undo/Redo:** The prototype must support undoing and redoing actions, allowing users to correct mistakes during model creation and manipulation.

- **R9 – Selection and Hover Feedback:** The prototype must provide visual feedback when elements are hovered over or selected, to signal the current interaction state to the user.

- **R10 – Context-Sensitive Controls:** Only functions applicable to the currently selected element type should be displayed, reducing cognitive load and preventing invalid interactions.

4.2 Architecture

The prototype is a web-based application built in JavaScript, using a Node.js[1]/Express[2] backend and a Vite-bundled frontend. The JavaScript 3D library Three.js[3] handles 3D rendering. AR functionality is enabled through the WebXR Device API, a web standard supported by all major browsers that allows access to AR and VR devices without additional software installation [MBF26]. This allows the prototype to run directly in the browser on both a standard desktop computer and the Meta Quest 3 used in the user study, without requiring separate builds, thereby fulfilling the cross-platform compatibility requirement of this thesis. The backend exposes a RESTful API for managing BPMN diagrams. Diagram data, including all elements and connections, is stored as JSON documents in a PostgreSQL[4] database. On desktop, diagrams can also be exported and imported as local JSON files, providing an alternative to the database for saving and loading work. Figure 4.1 presents an overview of the architecture of the prototype. Both backend and frontend are hosted on Render[5]. A demonstration of the frontend is

[1] https://github.com/nodejs/node
[2] https://github.com/expressjs/express
[3] https://github.com/mrdoob/three.js
[4] https://www.postgresql.org/docs/
[5] https://render.com/

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

available at the following URL[6], and the full code of the prototype can be accessed on GitHub[7].

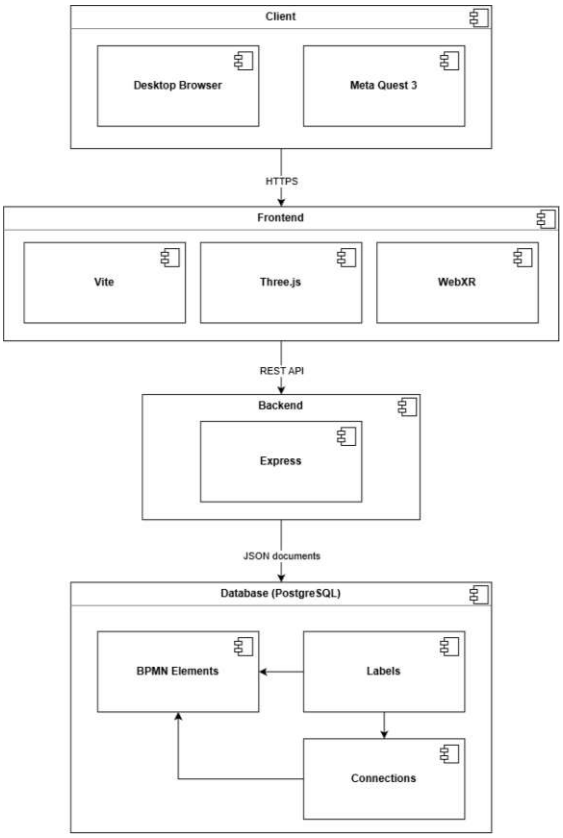


Figure 4.1: Overview of the architecture of the BPMN Builder prototype

The JSON documents store different information depending on the element type. For BPMN elements such as tasks, gateways, and events, a unique UUID, the element type, position, rotation, scale, and the associated label are stored, as shown in Figure 4.2 for a task element. For labels, a unique UUID, the label text, position, rotation, and scale are stored. For connections, the connection type (sequence flow or association), source ID, target ID, and the associated label are stored.

4.3 Shared Design

This section presents the shared design of the direct and indirect modes. It elaborates on the interaction design, the included design principles, the implemented desktop mode, and how pointing and element selection are integrated in both modes. Furthermore, the shared element visualizations, the logic for adding and connecting elements, the

[6] https://bpmn-builder.onrender.com/
[7] https://github.com/anna596/BPMN_Builder

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

```
{
  "type": "task",
  "uuid": "968093dc-bf51-485e-a3b4-9f2f6cfe311b",
  "position": [-382.61, 135.58, 120],
  "rotation": [0, 0, 0, "XYZ"],
  "scale": [1, 1, 1],
  "label": {
    "text": "Validate Request"
  }
}
```

Figure 4.2: JSON representation of a task element

implemented selection feedback, the integration of element manipulation, and the controls and functions available in both modes are discussed.

4.3.1 Interaction Design

The interaction design of the prototype is based on several XR interaction patterns described in [Jer16]. For selecting elements and buttons in both interaction modes, the *Pointing Pattern* is used. In this pattern, a ray is extended from the hand into the scene, and the first intersected object can be selected via a trigger. In the prototype, this is done via a pinch gesture. According to the author, pointing is preferable when interactions do not need to be of a realistic nature, as it is faster than directly grabbing objects and can be performed from a distance [Jer16]. This is relevant for the implementation of the BPMN builder prototype, since users should also be able to select elements or buttons when standing further away from the diagram.

For manipulating elements in the direct interaction mode, the *Direct Hand Manipulation Pattern* is applied. It is similar to the way objects are manipulated in the real world. An object is attached to the hand upon selection and moves together with it until the pinch is released [Jer16]. Rather than grabbing with a full hand, selection and manipulation are initiated via a pinch gesture, as described in Section 4.3.4.

For the indirect interaction mode, the *Indirect Control Pattern* is used, in which control is provided through an intermediary to manipulate an object [Jer16]. Specifically, the *Widgets and Panels* sub-pattern is applied. Widgets are geometric user interface elements that can be directly interacted with. Panels are container structures on which multiple widgets can be placed. In the BPMN Builder prototype, this is realized through a floating AR panel containing interactive buttons, through which all diagram interactions are performed. According to the author, the placement of panels is important for ensuring easy access [Jer16].

The design of the indirect interaction panel was further motivated by Brasier et al. [BCF+20], who explore ARPads as an indirect interaction method in AR as an alternative

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

to gaze or direct hand input. Their study found that indirect input can achieve the same performance as direct input while causing less fatigue, and that positioning a pad close to the wrist or waist is more efficient than thigh positioning. They also found that a vertical orientation, parallel to the user, is preferred over a horizontal one, and that users perform better when the pad's orientation matches that of the headset screen. Furthermore, their study showed that direct hand input suffers from precision issues and can be physically tiring, suggesting that indirect input may be better suited for interactions with small objects.

These findings directly inspired the design of the indirect interaction panel in the prototype. Similarly to an ARPad, the panel is placed at approximately waist level with a vertical orientation, slightly tilted to allow easier reading of the buttons. For the direct interaction mode, some buttons are also placed close to the body, specifically at the wrist, which was the second preferred positioning in Brasier et al.'s study [BCF+20]. These wrist-based buttons primarily handle general, diagram-independent interactions, while the primary interactions in the direct mode are performed through pointing and selecting elements directly in the diagram. These design decisions laid the foundation for the two distinct interaction approaches described in Sections 4.4 and 4.5.

4.3.2 Design Principles

The prototype design is guided by selected interaction design principles proposed by Jerald [Jer16].

Intuitiveness: Jerald recommends designing interfaces that can be quickly understood and accurately predicted, using interaction metaphors that exploit knowledge users already have from other domains [Jer16]. To promote intuitiveness, familiar interaction metaphors were used where possible. For example, the indirect mode uses arrow buttons and single pinches for navigation and selection, drawing on conventions users already know from desktop applications.

The following Norman principles of interaction design [Nor13] were additionally applied, based on Jerald's recommendations for XR environments:

Affordances and Signifiers: Affordances define possible actions and how an object can be interacted with by a user [Jer16]. The 3D cube representation of BPMN elements affords physical manipulation, suggesting to users that elements can be grabbed and moved. According to Jerald, the current interaction state should further be made obvious to the user through signifiers [Jer16]. In the prototype, the available actions are communicated through buttons displayed in the scene. In the direct mode, these appear as wrist menus and context-sensitive button groups attached to the selected elements, while in the indirect mode, the interaction panel displays the currently available functions. Since the button sets change depending on the selected element type and current mode, they always reflect the possible interactions.

Constraints: Jerald suggests adding constraints where appropriate to limit possible actions and improve accuracy, and making constraints consistent so that learning can

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

be transferred across tasks [Jer16]. In the prototype, only functions applicable to the currently selected element type are displayed. For example, rotation is not available for labels, and connection options are not shown when a connection is selected. This prevents invalid interactions and reduces cognitive load.

Feedback: Jerald recommends using highlighting to signify the touching of objects [Jer16]. Following this, visual feedback is provided immediately upon interaction in the prototype. For example, hover effects and selection boxes on elements and buttons confirm that an element has been pointed at or selected.

Mappings: Maintaining directional compliance is recommended so that users can anticipate motion in response to their physical input [Jer16]. In the direct interaction mode, hand movement direction corresponds directly to the resulting transformation direction for translation and rotation, allowing users to intuitively anticipate the outcome of their gestures.

4.3.3 Desktop Mode and Mode Switching

The desktop mode allows users to create or manipulate models, as well as download and upload models in the described JSON format. The database can also be accessed from the desktop mode, allowing models to be saved, manipulated, or deleted.

The desktop mode can be used to create models that are later visualized and, if needed, manipulated in AR, or conversely, models created in AR can later be visualized and manipulated on desktop. The design of the desktop mode can be seen in Figure 4.3. However, the desktop mode was implemented only as an additional tool to facilitate model management. It is not part of the research question of this thesis and was therefore not evaluated in the user study.

The AR modes can be accessed by opening the prototype in the browser and pinching the 'Start AR' button at the bottom of the screen. The default mode in AR is the indirect interaction mode, and users can switch between direct and indirect mode by pinching the corresponding buttons. This feature is further elaborated later.

4.3.4 Pointing and Selection

For hand-tracking interaction, the prototype uses Three.js's OculusHandPointerModel[9], a reference implementation for WebXR hand-tracking interaction provided by Three.js. It provides two core functions. First, raycasting casts a ray from the hand's pointer position and orientation to determine what the user is aiming at in the scene. Second, pinch detection determines whether the hand is currently performing a pinch, which is executed by pressing the index finger and thumb of either hand together. The ray itself is not

[8]The text labels in this and all subsequent figures showing model screenshots may not be legible when printed. This is intentional, as legibility of the labels is not necessary for understanding the structure of the models.

[9]`https://threejs.org/docs/pages/OculusHandPointerModel.html`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

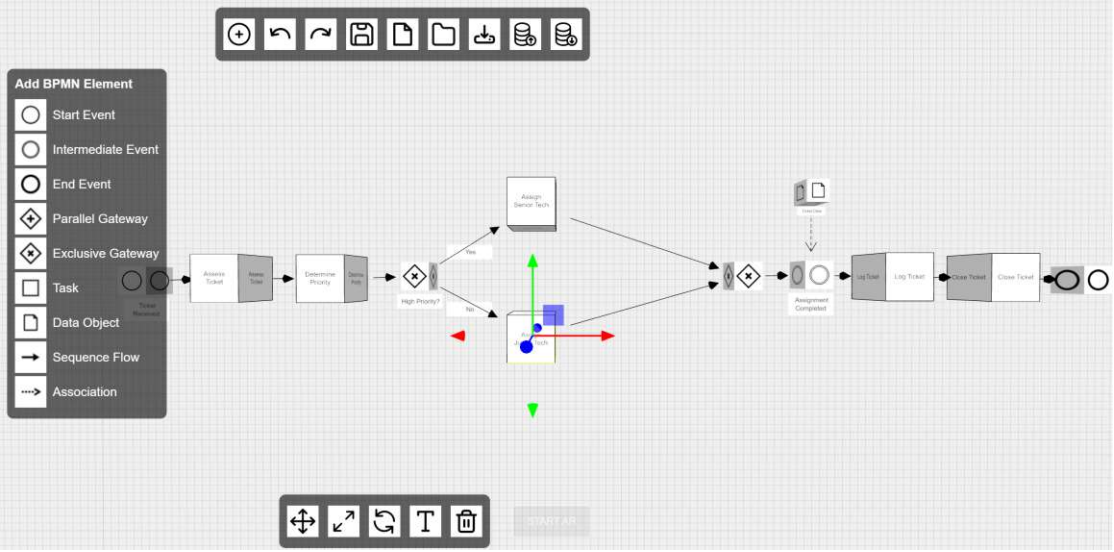


Figure 4.3: Desktop mode[8]

visually displayed, instead, a small triangle indicates the current pointing direction and a hit circle marks where the pointer intersects with an element or button in the scene. As the OculusHandPointerModel is compatible with any WebXR-conformant hand-tracking device, it could be directly used with the Meta Quest 3 in this implementation.

For all AR interactions, the combination of raycasting and pinching was used as the basis for selection. This combination can be classified as a *Virtual Pointer Metaphor*, where ray direction defines the virtual pointer [PIWB98]. The pinch gesture was chosen for all AR interactions, following Jerald's recommendation to use a small number of well-defined, natural gestures that are easy for users to remember and easy for the system to recognize [Jer16]. Furthermore, users have a preference for traditional interactions, and pinch is the standard built-in gesture for Meta Quest devices[10], as described in Meta's official hand tracking documentation [GNM+23].

4.3.5 Visualization

Similarly to [OP19], BPMN elements are visualized as cubes with the typical BPMN element icon displayed on each side of the cube as its texture. This allows the user to recognize the elements from every angle. The authors argue that diverging too far from the visual conventions of the specification would reduce the recognition and standardization benefits that the specification provides [OP19]. The authors also discuss the issue that elements in 3D space can be hidden behind other elements. One option to address this would be to make elements partially transparent, but in their implementation they decided to keep elements opaque to avoid visual confusion, allowing users to check for

[10] https://www.meta.com/de-de/help/quest/290147772643252

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

hidden elements by briefly adjusting their perspective [OP19]. For the BPMN Builder prototype, a similar visualization approach was used, with elements kept opaque. Tasks are also displayed as cubes, with their labels written on every side as texture of the cube. Labels of other BPMN elements are rendered onto a 2D canvas and are displayed underneath the element. Similarly to [OP19], they rotate automatically toward the camera to remain readable from every angle. This visualization allows users familiar with BPMN to recognize elements easily and read labels from any angle. When a new task is created, it displays 'Task' as its default label. Other BPMN elements have no default label.

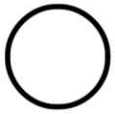
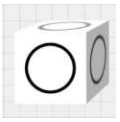
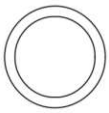
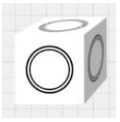
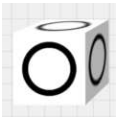
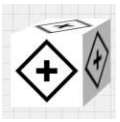
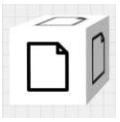
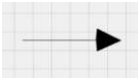
It was decided to write the labels of tasks directly on all sides of the element rather than displaying them as a separate object in order to keep the number of elements in the scene low while staying close to the BPMN standard. This approach is supported by [YGWE21], who note that a user can see at most three sides of a cuboid at any given time, meaning a label displayed on only one side could become invisible from certain angles. Displaying the label on all sides therefore ensures readability from every perspective. Following the same reasoning, task cubes do not rotate toward the camera. The authors argue that automatically rotating elements would change the model's appearance without user input, reducing the sense of stability and naturalness. For other BPMN elements, the label is a separate flat plane geometry displayed underneath the element that rotates toward the camera to ensure readability. Representing labels as cuboids was deliberately avoided, as this would make them unreadable when a user looks down on the diagram and would clutter the scene with unnecessary three-dimensional objects. Finally, rotating all element cubes toward the camera was also ruled out, since it would always present the same face to the user regardless of viewing angle, reducing the benefit of the three-dimensional representation, while rotating only task cubes could be visually irritating given that other elements remain static.

When elements are added to the diagram, they receive a default size, which users can adjust afterwards. By default, tasks are larger than events, gateways, and data objects. Labels also receive a fixed width and font size. Text that exceeds this width wraps onto additional lines, growing the canvas height dynamically. However, the maximum number of characters and the maximum number of lines are capped in order to avoid labels of infinite size. When the text of a label is changed, both the texture and the plane geometry of the label are updated to match the new dimensions, preventing skewing of the text. Furthermore, the label shifts downward to maintain a constant gap to its parent element. For connection visualization, the standard BPMN notation was followed by using straight rather than curved connections. This is additionally supported by the empirical findings of [BVD19], which showed that straight edges achieved higher user acceptance than curved edges in AR.

Table 4.1 compares the conventional visualization of BPMN elements from to their representation in the BPMN Builder.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Table 4.1: Comparison of standard BPMN element visualization to the visualization of BPMN elements in the BPMN Builder prototype

Element	**Standard BPMN**	**BPMN Builder**
Task		
Start Event		
Intermediate Event		
End Event		
Exclusive Gateway		
Parallel Gateway		
Data Object		
Sequence Flow		
Association		

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

4.3.6 Adding and Connecting Elements

Connection creation (sequence flow or association) is performed differently in desktop mode and in AR. In desktop mode, connection creation is implemented as follows: the user pinches the add button, then pinches the connection symbol, and then pinches on the two elements to be connected. In the two AR modes, the user first selects a source element, then pinches the add button, selects the connection type, and finally selects the target element. The implementation differs in AR in order to simplify the interaction, since selecting multiple elements in sequence in AR is more challenging than on a desktop interface. The desktop BPMN modeling tool bpmn.io[11] implements both of these connection creation methods. For the prototype, it was decided to incorporate only one method per mode to avoid confusing users with multiple ways of performing the same action.

Adding elements can follow two different workflows in AR, depending on whether a BPMN element is previously selected or not.

1. No BPMN element is previously selected: the user pinches the add button (depending on whether the user is currently in direct or indirect mode, the add button is accessed differently), and then selects the element to be added. In direct mode, the user then pinches the spot in the scene where the new BPMN element should be added. In indirect mode, after selecting the element type to be added, the element is automatically placed in the middle of the diagram. The specifics of both modes are elaborated in Sections 4.4 and 4.5.

2. A BPMN element is previously selected: if a user has already selected an element and wants to add a new element connected to it, the user can pinch the add button, then select the symbol of the element to be added. In direct mode, the user then pinches the spot in the scene, where the new BPMN element should be added. In indirect mode, after selecting the element to be added, the element is automatically placed to the right of the previously selected element. In both modes, the new element is automatically connected to the previously selected element.

This differentiation of workflows was implemented since it cannot always be assumed that users start with a pre-existing diagram in which elements can be selected, as the prototype should also support the creation of diagrams from scratch. Furthermore, BPMN allows for different roles and parallel flows where not all elements are connected. Therefore, users must be able to add unconnected elements to the scene.

4.3.7 Selection Feedback and Element Manipulation

The prototype includes hover and selection feedback for all interactable objects in the scene, including BPMN elements and UI buttons. When the user hovers over an

[11] https://bpmn.io/

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

interactable object, it becomes enlarged. When the user selects an object, a yellow selection box is displayed around the element and its label, if a label exists. Tasks, gateways, events, data objects, and connections can be selected together with their label, if one exists. The selection box then surrounds both the element and the label. Additionally, labels can also be selected separately from their parent BPMN element and manipulated independently, for example by moving the label without moving the element it belongs to.

Tasks, gateways, events, and data objects can be transformed via translation, scaling, and rotation. These transform controls cannot be applied to connections, and labels can only be translated and scaled. Translation and rotation are possible in all six directions in space for both interaction modes. Elements can only be scaled uniformly, so that labels and icons do not become skewed.

For both AR interaction types, it was ensured that when a user selects an element, only the functions available for that element are displayed. When a label is selected, neither the rotation option nor the option to add elements is displayed, since rotation is not implemented for labels, and new BPMN elements cannot be connected directly to labels but only to their parent elements. When a connection is selected, only the options to delete or label the connection are visible, since no BPMN elements can be added to connections and connections cannot be transformed.

4.3.8 Shared Controls

Both interaction modes share the same functionality for adding, manipulating, and deleting BPMN elements, as well as opening new or existing diagrams, exiting the scene, and switching between modes. An overview of all shared controls in the BPMN Builder is provided in Table 4.2. How these controls are accessed differs between the two modes and is described in the following sections.

The undo option undoes the last step, while the redo option reapplies the last undone step. Multiple rotations of the same element, for example, count as a single step, so the entire rotation would be reverted if the undo option is selected.

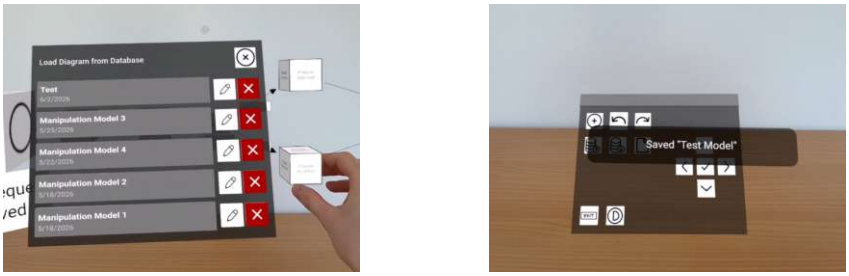
The option to load from the database opens the database menu showing all saved diagrams, as can be seen in Figure 4.4a. Selecting a diagram loads it in the currently active interaction mode. The database menu also allows users to rename or delete diagrams from the database. The option to save to the database allows users to save the current diagram to the database, with a success message confirming the action, as depicted in Figure 4.4b. To avoid data loss, the diagram is autosaved to the database every 30 seconds.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Table 4.2: Overview of shared UI controls available in both interaction modes

Symbol	Meaning
	Add: Opens the view for adding BPMN elements or connections to the diagram.
	Delete: Deletes the currently selected element from the diagram.
	Label: Opens a text input for adding or editing the label of the currently selected element.
	Translate: Activates translation mode, allowing the selected element to be moved in any of the six spatial directions.
	Scale: Activates scaling mode, allowing the selected element to be uniformly scaled up or down.
	Rotate: Activates rotation mode, allowing the selected element to be rotated in any of the six spatial directions.
	Undo: Reverts the last action in the diagram.
	Redo: Reapplies the last undone action to the diagram.
	New Diagram: Creates a new, empty diagram, clearing the current scene.
	Save to Database: Saves the current diagram to the database.
	Load from Database: Opens the database to load a previously saved diagram.
	Switch to Direct: Switches to direct interaction mode, hiding the AR interaction panel.
	Switch to Indirect: Switches to indirect interaction mode, displaying the AR interaction panel.
	Exit: Exits the AR scene and returns to the desktop view.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) Open saved diagrams (b) Save to database

Figure 4.4: Database interaction in the prototype

4.4 Direct Interaction

This section presents the specifics of the direct interaction mode. It elaborates on the implemented pinch gestures, the integrated element and hand menus, as well as the integration of transform controls in this mode.

4.4.1 Pinch Gestures

In the direct interaction mode, three different types of pinches are implemented: a *short single pinch*, a *double pinch*, and a *long pinch*. The double pinch is recognized when two pinches on the same element occur within a 500ms window. The first pinch starts a timer, and if a second pinch is registered on the same target before the timer expires, it is treated as a double pinch.

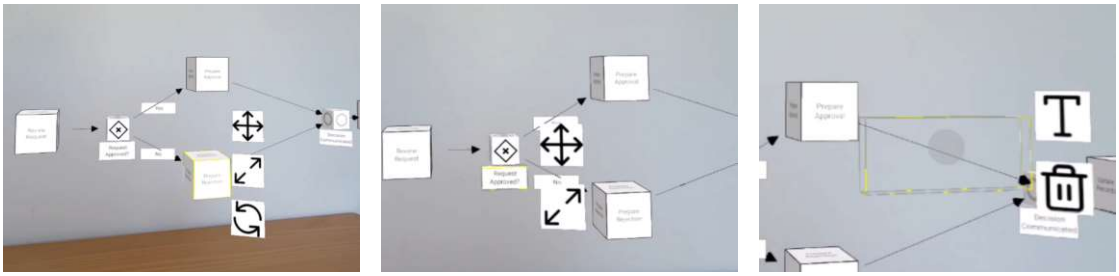
A long pinch is recognized when a pinch on empty space is held continuously for at least 500ms. If the pinch is released before this delay elapses, the timer is canceled and no long-pinch action is triggered. How these pinches integrate with the functionality of the BPMN Builder is elaborated in the following.

4.4.2 Element Menus

A single pinch on a BPMN element or label opens the transform control options for that element. As elaborated in Section 4.3.7, for BPMN elements the transform controls translation, scaling, and rotation are available, while for labels, only translation and scaling are supported. Figures 4.5a and 4.5b show these transform options.

A double pinch on an element opens the 'edit options' for that element, which allow users to add elements that are automatically connected to the selected element, label the selected element, or delete the selected element. Since transform controls do not apply to connections, a single pinch already opens the 'edit options' for connections, as can be seen in Figure 4.5c. Figure 4.6a shows the 'edit options'. When the user pinches the 'add' button, new buttons appear, replacing the 'edit options' and showing the BPMN elements and the two connection types (sequence flow or association) that can be added. These options are visualized in Figure 4.6b. If a user selects the option to label an element, the

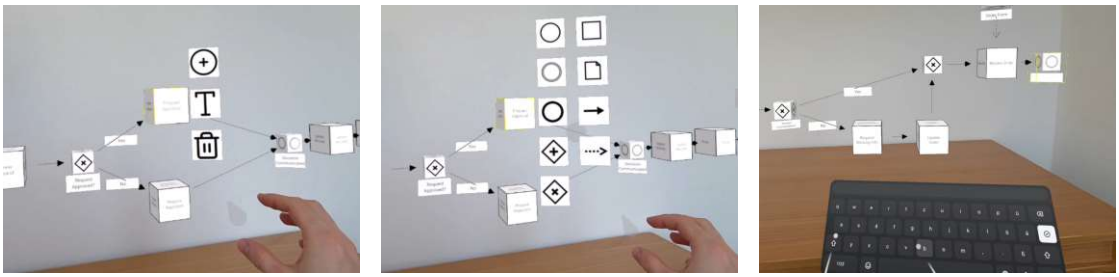
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Selected BPMN element (b) Selected label (c) Selected connection

Figure 4.5: Element selection

system keyboard opens and allows them to type the desired label text. If a label already exists and the user pinches the label button, the existing label is selected and its text can be edited. For all three button types (transform, edit, and add buttons), the buttons disappear when the user pinches anywhere else in the scene, rather than on a button.



(a) Double pinch (b) Selected add (c) Adding label

Figure 4.6: Edit options

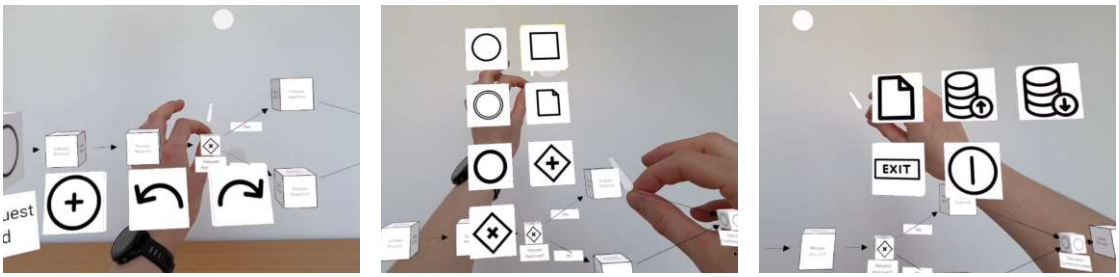
4.4.3 Hand Menus

The long pinch is utilized for the left and right hand menus. These are buttons attached to the left or right wrist of the user, following the recommendation by Jerald [Jer16] to place commonly used tools relative to the body. When the user moves their hand, the buttons follow the movement. As soon as the pinch is released, the buttons disappear. Performing a long pinch with the left hand opens the 'left hand menu' with the options to add, undo, or redo, as displayed in Figure 4.7a. The add button allows adding elements when no BPMN element is previously selected, as elaborated in Section 4.3.6. Pinching the add button replaces the 'left hand menu' with a selection of BPMN elements that can be added to the diagram (Figure 4.7b). Connections are not displayed as options, since they cannot be added via the 'left hand menu' but only after a BPMN element has been previously selected. The undo button undoes the last step and the redo button reapplies it, as described in Section 4.3.8. Performing a long pinch with the right hand opens options for creating a new diagram, opening the database to load an existing diagram,

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

saving the current diagram to the database, exiting the AR scene, and switching to indirect interaction (Figure 4.7c), as described in Table 4.2.

The interaction with the left and right hand menus is as follows: the user performs a long pinch with the left or right hand, making the left or right hand menus appear. The user then pinches the desired button with the other hand. For example, when adding a new, unconnected element, the user performs a long pinch with the left hand and then pinches the add button with the right hand, making the 'add buttons' appear that display the BPMN elements that can be added. Selecting the desired BPMN element type highlights the button with a yellow selection box. The user then pinches the spot in the scene where the new element should be added. The new element is added to the scene, and the 'left hand buttons' disappear automatically.



(a) Left hand menu (b) Left hand menu add (c) Right hand menu

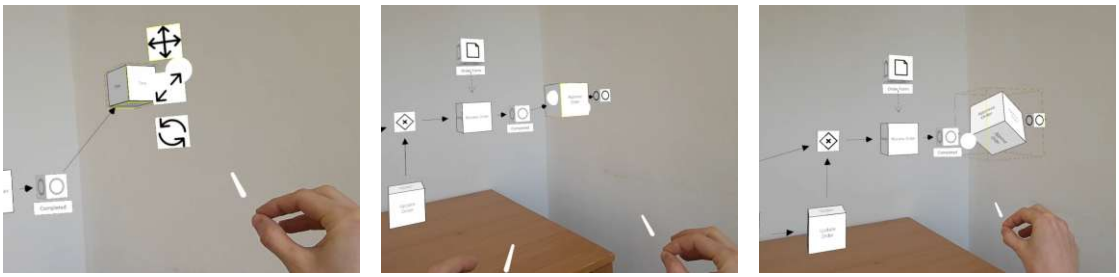
Figure 4.7: Hand menus

4.4.4 Transform Controls

When the transform control buttons are displayed, the user can select the transform mode to be applied. As a result, the selected transform mode is highlighted (Figure 4.8a). For translating an element, the user pinches the corresponding button, then pinches the element again and performs a sustained pinch while moving the hand to where the element should be placed. While the pinch is held, the element moves together with the hand. When the pinch is released, the element is placed at the new position. For scaling an element, after selecting the corresponding button, the user moves both pointers to the element to be scaled and performs a sustained pinch with both hands, as depicted in Figure 4.8b. While holding the pinches, the user moves the hands closer together to make the element smaller, or further apart to make it bigger. When the pinches are released, scaling stops and the element retains its current size. For rotating an element, after selecting the corresponding button, the user moves one pointer to the element to be rotated and performs a sustained pinch (Figure 4.8c). Then, the user moves their hand up-down, left-right, or in-out to rotate the element in the corresponding direction. The rotation direction with the largest hand movement is taken as the rotation axis. To rotate an element on several axes, the user has to release the sustained pinch and pinch the element again, then perform the next rotation.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

During all transform actions, the transform buttons disappear. Once the action is finished, the buttons reappear with the element still selected and the previous transform mode still active. The user can transform the element again in the selected mode by performing a sustained pinch on the element once more.



(a) Transform control options (b) Scaling an element (c) Rotating an element

Figure 4.8: Direct interaction transform controls

4.5 Indirect Interaction

This section presents the specifics of the indirect interaction mode. The panel layout is elaborated, as well as how diagram navigation is implemented. Furthermore, the element views, the integration of transform controls, and the addition and connection of elements in this mode are explained.

In indirect interaction, all actions are performed via an AR panel that is displayed in the scene. The panel contains buttons for all diagram interactions. The idea of using a panel similar to a tablet, also appeared in the works [Obe21, Obe24, OBS22, OBS23], as elaborated in Section 3.

4.5.1 Panel Layout and Initialization

When the scene is initialized, the indirect interaction panel is oriented toward the user's camera and placed in the lower third of the scene, roughly at hip height for a standing user. It is positioned an arm's length away to make button selection easier, and slightly tilted so that all buttons remain clearly visible despite the panel sitting below eye level.

The panel consists of two sides. On the right-hand side, there is a star of arrows with a checkmark in the middle, used for navigating the diagram and confirming element selection. On the left-hand side, a set of buttons allows for various actions. These buttons differ depending on the current view (no element selected, element selected, transform mode, or add mode) and on the type of element currently selected. All buttons on the panel are operated using a single pinch.

The panel can be translated and scaled in a similar way as elements are transformed in the direct interaction mode. At the top of the panel there is a light gray bar that

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

becomes slightly enlarged when hovered over. Performing a sustained pinch on this bar and moving the hand drags the panel along with it. Releasing the pinch leaves the panel at its new position, reoriented to face the user. The panel was made stationary rather than continuously following the camera, since this could become distracting if the user frequently looks around the diagram. To scale the panel, the user places both pointers on it, performs a sustained pinch with both hands, and moves the hands closer together to shrink the panel or further apart to enlarge it.

4.5.2 Navigation

On the right side of the panel, four directional navigation buttons are displayed, visualized as arrows pointing up, down, left, and right. These are used to navigate between elements in the diagram. The current position in the diagram is indicated by a blue navigation box. When an element is selected via the checkmark button, which is located in the middle of the navigation buttons, the box turns yellow.

When no element is highlighted and the user pinches one of the navigation buttons, the first press selects the topmost element in the diagram, regardless of direction. If multiple elements share the same vertical position, the leftmost among them is selected and highlighted with the blue navigation box.

From a highlighted element, pressing a directional button moves the navigation box to another element. The pressed direction defines a 90° sector centered on that axis. Only elements whose displacement from the current element is greater along the intended axis than along the perpendicular axis qualify as candidates, which is equivalent to falling within ± 45° of the intended direction. This divides the space into four non-overlapping sectors covering all directions, ensuring each element belongs to exactly one directional neighbourhood, as visualized in Figure 4.9. Among all qualifying elements, the nearest one by Euclidean distance is selected, ensuring that the spatially closest element in the intended direction is always chosen, rather than one that happens to be further away within the same sector.

If no element falls within the sector, which can occur at the edges of the diagram, the navigation wraps to the element at the opposite extreme along that axis. For example, pinching right when at the rightmost element jumps to the leftmost.

It should be noted that the algorithm selects the nearest element by spatial distance rather than following connection paths.

4.5.3 Element Views

When users select an element, they enter the 'element selected view', which allows them to add elements or connections to the selected element, add or edit its label, delete the selected element, or transform it, as can be seen in Figure 4.10a. When a label is selected, the options to add elements or rotate the label are not displayed, and when a connection is selected, the options to add elements or transform the connection are not displayed,

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

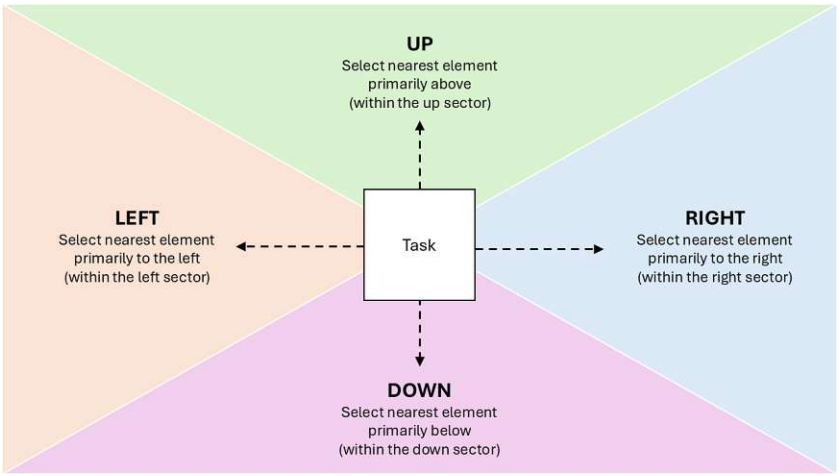


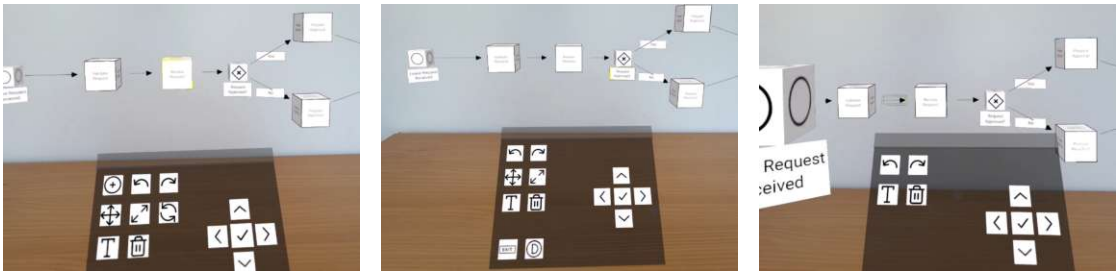
Figure 4.9: Navigation directions

as can be seen in Figures 4.10b and 4.10c. This is similar to the label and connection handling explained in Section 4.3.7.

The 'no element selected view' offers different functionality. Users can create new diagrams, load diagrams from and save diagrams to the database, exit the scene, and switch to the direct interaction mode. This is depicted in Figure 4.13a.

In both the 'no element selected' and the 'element selected' views, the undo, redo, mode switch, and exit buttons are visible. In the 'transform views', elaborated in Section 4.5.4, only the mode switch and exit buttons are displayed. Pinching the mode switch button switches the user to direct mode, upon which the panel disappears and the user can interact directly with the diagram elements through pointing and pinching, rather than through the panel.

The database buttons are only visible in the 'no element selected' view.



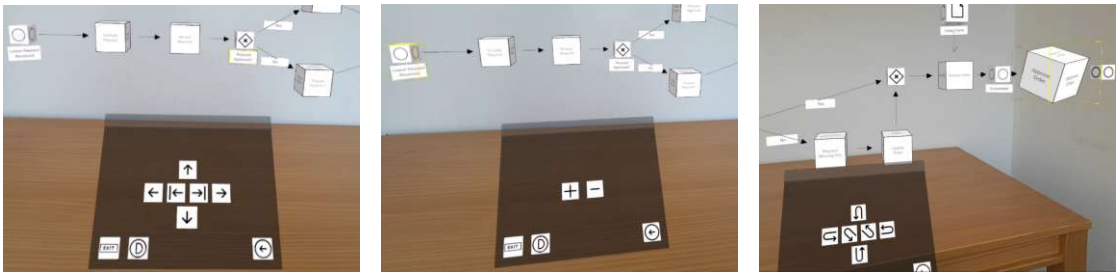
(a) Element selected view (b) Label selected view (c) Connection selected view

Figure 4.10: Element selected views

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

4.5.4 Transform Controls

Similarly to the direct mode, there are three different transform buttons for translating, scaling, and rotating. These can be accessed in the 'element selected view'. When the user pinches one of the transform buttons, they enter a new view of the panel with the transform buttons corresponding to the selected mode. The different views are depicted in Figure 4.11. For translating and rotating, these are arrows representing all six directions in which an element can be translated or rotated. For scaling, a plus button and a minus button are displayed, used to increase or decrease the element's size. In each of these views, a small circled arrow appears, allowing the user to return to the 'element selected view'. Each transformation is performed in fixed steps, determining how much an element moves, scales, or rotates per pinch. To perform larger transformations, such as moving an element a greater distance, the user has to pinch the corresponding button multiple times. Continuous transformation by pinching and holding a button was not implemented, in order to allow for more precise transformations.



(a) Translation view (b) Scaling view (c) Rotation view

Figure 4.11: Transform controls in indirect mode

4.5.5 Adding and Connecting Elements

If the user is currently in the 'element selected view' and pinches the add button, they can add elements to the diagram or connect the selected element to another element. When the user selects a new element to add, the element is automatically connected to the previously selected element, added next to it, and automatically selected for further manipulation, indicated by a yellow selection box. This is depicted in Figure 4.12.

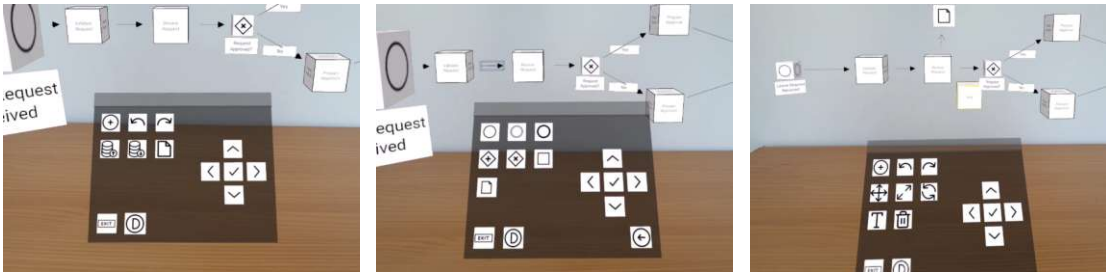
Users can also add elements in the 'no element selected view' via the plus icon. After selecting the add option, they can select the element type they want to add. This view does not offer the option to add a connection (sequence flow or association), since connections are only available when an element is selected. Once the user selects the element type they want to add, the element appears in the middle of the diagram and is automatically preselected to allow direct manipulation, such as moving or connecting it, without requiring the user to navigate to it and select it as a separate action after element creation. This is shown in Figure 4.13.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

(a) Add menu with element selected

(b) Element added and connected

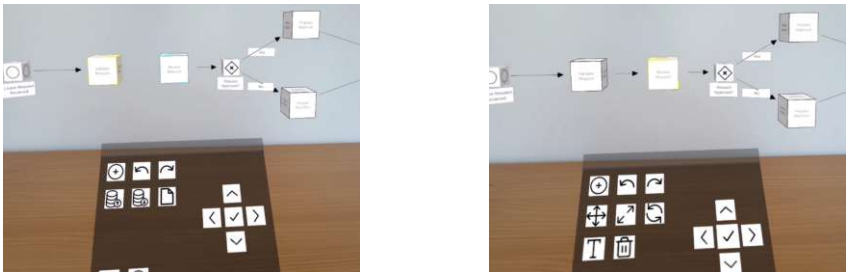
Figure 4.12: Adding an element with prior selection



(a) No element selected view (b) Add menu, no selection (c) Element added to center

Figure 4.13: Adding an element without prior selection

When the user chooses to add a sequence flow or association, the previously selected element remains visually selected via the yellow selection box, and the user can navigate via a blue navigation box to the element to which the previously selected element should be connected. When the user pinches the checkmark button, the connection appears and the newly connected element is selected, as visualized in Figure 4.14.



(a) Selecting connection target

(b) Connection created

Figure 4.14: Connection creation in indirect mode

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 5

Prototype Evaluation

To evaluate the two interaction modes of the prototype presented in Chapter 4, a user study was conducted assessing efficiency, usability, and user preferences. This marks the fourth step of the Design Science Research methodology outlined in Section 1.3.

5.1 Study Design

Similarly to [GRM20, SKKA21, WBG+25, YSB+24a, YGWE23, YSB+24a], a within-subjects design [Bar11] was used for the user study, combined with a comparison-of-designs approach. This means that each participant tested and compared both interaction modes of the BPMN Builder prototype, allowing to capture efficiency and usability differences as well as user preferences. The study followed a 2×2 structure, where participants used both direct and indirect interaction modes for two task types, namely model creation and model manipulation.

For the user study, a convenience sample of 20 participants was selected. The majority had prior modeling experience and worked or studied in the computer science or business informatics field. The sessions were conducted with one to three participants at a time. Before performing the tasks, participants were asked to familiarize themselves with the BPMN modeling language to ensure a common minimum level of knowledge for the subsequent tasks. At the beginning of the session, participants were also asked to try out the web-based BPMN modeling tool bpmn.io[1]for about two minutes, to give them a feel for how 2D BPMN modeling works and how common functions are implemented. Subsequently, the prototype was introduced via a short PowerPoint presentation, including videos demonstrating the different actions and functions available in both direct and indirect interaction modes. Each participant could try out each mode for about five minutes before starting with the tasks.

[1] https://bpmn.io/

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Overall, the participants were split into four groups: A, B, C, and D, each performing two sessions covering model creation and manipulation tasks. The exact split is depicted in Table 5.1. This allowed varying the order in which tasks were performed in each mode, eliminating potential bias introduced by always starting with the same mode or task type. Every group created the same model in both direct and indirect mode. This model was kept simple and did not carry semantic meaning, in order to minimize the number of elements to be added and labeled, while still covering all core interactions of the prototype: adding, labeling, connecting, and translating elements. The model that had to be recreated in direct and indirect mode can be found in Appendix 7.2.1.

For the manipulation tasks, participants received different models and task sets for each mode. In total, four different manipulation models with four corresponding sets of tasks were used. All task sets were designed to include the same number of actions of each type: deleting, adding, renaming, connecting, and transforming elements. The exact number of individual actions performed varied depending on how participants chose to complete the tasks. For instance, adding an element while another element is already selected automatically connects the two, whereas adding an element without a prior selection requires the connection to be added separately. All manipulation tasks can be found in Appendix 7.2.2. During task execution, assistance was provided whenever needed, as both the AR environment and the prototype were new to most participants. Hand positioning and gesture recognition posed challenges for most participants at the beginning of the session, but improved with guidance over time.

After performing the tasks, the participants were asked to scan one of two QR codes with their phones to open the questionnaire, hosted via Google Forms [2]. Two different QR codes were used so that the questionnaire presented the evaluation of the mode the participant started with first, followed by the second mode.

The questionnaire included three parts: 1) background questions on prior AR/VR and modeling experience, 2) questions specifically concerning direct and indirect interaction modes, and 3) questions regarding preferences and the comparison of both modes. The full questionnaire can be found in the Appendix 7.2.3.

The first part of the survey asked about prior BPMN experience, prior AR/VR experience, and experience with gesture-based interactions.

In part two of the questionnaire, the usability of both modes was evaluated. Similarly to related work [WBG+25, YGWE21, YGWE23], the System Usability Scale (SUS) [Bro96] and custom 5-point Likert items were used. The SUS is a widely used tool for measuring the usability of products and services, consisting of ten statements rated on a five-point scale from 'Strongly Disagree' to 'Strongly Agree', alternating between positive and negative statements [Bro96]. A mean SUS score of around 70 is considered the threshold for acceptable usability [BKM09]. The custom Likert items also used a 5-point scale to stay consistent with the SUS.

[2] https://docs.google.com/forms

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

The last part of the survey included questions concerning mode preferences for specific operations: connecting, transforming, labeling, adding elements, and navigating to and selecting elements, as well as open-ended questions concerning overall mode preferences, positive and negative aspects of each mode, and suggestions for improvement.

Table 5.1: Group split for user study

Group	Session 1	Session 2	Session 3	Session 4
A	Indirect Creation Creation Model	Indirect Manipulation Model 2	Direct Creation Creation Model	Direct Manipulation Model 4
B	Indirect Manipulation Model 3	Indirect Creation Creation Model	Direct Manipulation Model 1	Direct Creation Creation Model
C	Direct Creation Creation Model	Direct Manipulation Model 1	Indirect Creation Creation Model	Indirect Manipulation Model 3
D	Direct Manipulation Model 4	Direct Creation Creation Model	Indirect Manipulation Model 2	Indirect Creation Creation Model

5.2 Results

This section presents the results of the conducted user study. It covers the previous experience and background knowledge of the participants, efficiency measures of the direct and indirect modes, usability of both modes, qualitative findings, and additional observations made during the user study.

5.2.1 Background

The participants reported rather low prior experience in BPMN modeling ($median$: 1.5, $IQR : 1.0 - 3.25$), experience with AR/VR devices ($median : 1.0$, $IQR : 1.0 - 2.0$), and gesture-based interaction ($median$: 1.0, IQR : $1.0 - 2.0$). This can be seen in Figure 5.1.

5.2.2 Efficiency

During all individual user study tasks, the time participants needed to complete each task was measured. The results can be seen in Figure 5.2 and are summarized in Table 5.2. The mean task completion time for model creation tasks was 290.0 seconds ($SD = 81.6$) for direct and 397.6 seconds ($SD = 123.4$) for indirect. For manipulation tasks, the mean time was 313.4 seconds ($SD = 96.0$) for direct and 388.5 seconds ($SD = 94.8$) for indirect,

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

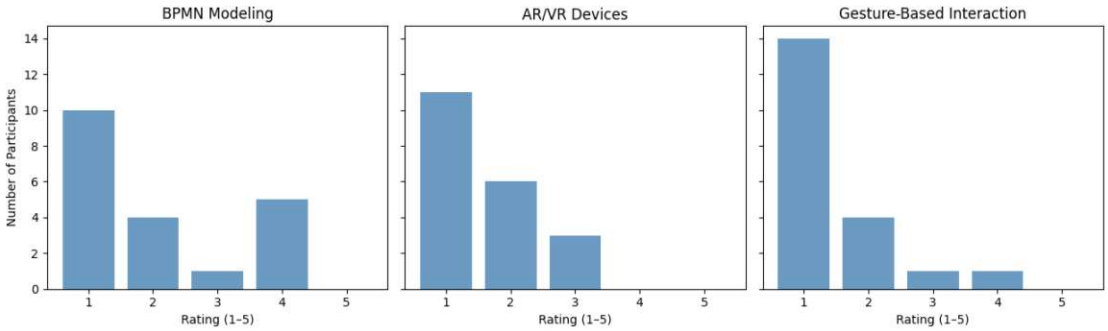


Figure 5.1: Prior experience

and overall, 301.7 seconds ($SD = 88.7$) for direct and 393.0 seconds ($SD = 108.7$) for indirect.

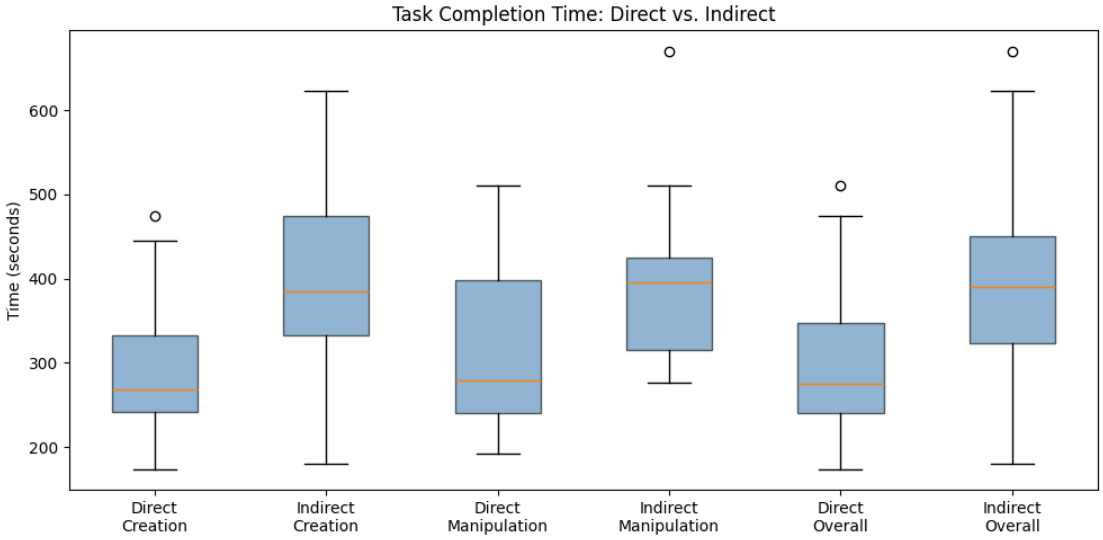


Figure 5.2: Efficiency

Table 5.2: Task completion time (in seconds) by mode and task type

	Direct		Indirect	
Task	**Mean (SD)**	**Median**	**Mean (SD)**	**Median**
Creation	290.0 (81.6)	268.0	397.6 (123.4)	385.0
Manipulation	313.4 (96.0)	279.0	388.5 (94.8)	396.0
Overall	301.7 (88.7)	274.5	393.0 (108.7)	390.0

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

For significance testing, the Wilcoxon signed-rank test [Ari17] was used, since a within-subjects design was applied, where each participant tested both modes, leading to possibly correlated observations. Furthermore, due to the small sample size, normality could not be assumed. A significance level of $\alpha = 0.05$ was used for all statistical tests. It should be noted that the sign of the Z-value depends on the order in which groups or conditions are entered into the statistical software and does not by itself indicate which group or condition scored higher. The test reported $Z = -2.277,\ p = 0.021$ for creation, $Z = -2.133,\ p = 0.033$ for manipulation, and $Z = -2.576,\ p = 0.008$ overall. This means that all three results are statistically significant (all $p < 0.05$) and that indirect interaction was consistently slower than direct interaction across all task types. On average, direct interaction was 27.06% faster for creation, 19.33% faster for manipulation, and 23.41% faster overall than indirect interaction.

To investigate potential order effects, it was tested whether mode order, starting mode, or task order had an influence on task completion time. For significance testing, the Mann-Whitney U test [Ari17] was used, since comparing task completion times between groups that performed the tasks or used the modes in different orders constitutes a between-subjects comparison. Furthermore, due to the small sample size, normality could not be assumed.

The mode order effect tested whether participants were faster in a specific mode when they used it as their second mode. Participants who used the direct mode as their second mode (255.3 seconds) were significantly faster than those who used it as their first mode (348.1 seconds), a difference of 92.8 seconds ($U = 82.0,\ Z = 2.419,\ p = 0.017$). Similarly, participants who used the indirect mode as their second mode (334.7 seconds) were significantly faster than those who used it as their first mode (451.4 seconds), a difference of 116.7 seconds ($U = 89.0,\ Z = 2.948,\ p = 0.004$). These results suggest that a mode order effect exists: participants performed significantly faster in a mode when it was used second, likely due to increased familiarity with the AR environment and prototype gained during the first mode.

The starting mode effect tested whether beginning with the direct or indirect interaction mode gave participants an overall speed advantage across all tasks and both modes. In other words, it was analyzed whether participants who used the direct mode first ended up with a different mean per-task pace across their entire session (both modes combined) than participants who used the indirect mode first, and vice versa. Participants who started with the indirect mode had an overall mean completion time of 353.3 seconds, compared to 341.4 seconds for those who started with the direct mode. This difference was not statistically significant ($U = 44.0,\ Z = -0.454,\ p = 0.678$), suggesting that the choice of starting mode did not affect overall task performance.

The task order effect looked at whether the order in which participants performed creation and manipulation tasks affected their speed. For creation tasks, participants who completed them first were slower (376.6 seconds) than those who completed them second (310.9 seconds), but this difference was not statistically significant ($U = 71.5,\ Z = 1.625,\ p = 0.112$). Similarly, manipulation tasks showed a similar trend (first: 375.3

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

seconds vs. second: 326.6 seconds), but also did not reach significance ($U = 64.0,\ Z = 1.058,\ p = 0.307$). These results suggest that task order did not significantly influence completion time, though the consistent trend toward faster performance when a task type was completed second may reflect general familiarization with the AR environment, as also observed for the mode order effect.

It was also tested whether prior AR/VR experience affected task completion time. Spearman's rank correlation [Sed14] was used, since AR/VR experience is an ordinal variable and the sample size is small. The relationship between AR/VR experience and task completion time was not statistically significant in either the direct mode ($p = 0.902$) or the indirect mode ($p = 0.270$), suggesting no meaningful association between participants' prior AR/VR experience and task completion time.

5.2.3 Usability

The following presents the results of the user study concerning the usability of both modes implemented in the prototype, covering the System Usability Scale (SUS), the custom Likert items, and user preferences.

System Usability Scale

The first usability measure applied was the SUS. The results can be seen in Figure 5.3 and Table 5.3. The direct mode scored considerably better, with a mean SUS score of 75.00 (*median* : 73.75), compared to the indirect mode, which reached a mean of 59.12 (*median* : 62.50). According to the classification by Bangor et al. [BKM08, BKM09], this corresponds to an adjective rating of 'Good' for the direct mode and 'OK' for the indirect mode. With a mean SUS score of 70 as the threshold for acceptable usability [BKM09], the direct mode is classified as 'Acceptable', while the indirect mode falls into the 'Marginal Low' category. In terms of letter grades, the direct mode receives a 'C', while the indirect mode receives an 'F', since its mean SUS score of 59.12 falls slightly below the threshold of 60.00 required for a 'D'. It should be noted that the letter grade and the adjective rating are based on different cutoff values. A SUS score of around 51 corresponds to the adjective rating 'OK' and a value of around 71 corresponds to 'Good'. The letter grade, in contrast, follows a stricter cutoff based on the traditional school grading scale, in which scores below 60 are assigned an 'F' [BKM09]. A Wilcoxon signed-rank test confirmed a statistically significant difference in usability scores between the two modes ($Z = -2.699,\ p = 0.007$).

Table 5.3: SUS results by interaction mode

Mode	Mean	SD	Median	Grade	Adjective	Acceptability
Direct	75.00	13.79	73.75	C	Good	Acceptable
Indirect	59.12	20.51	62.50	F	OK	Marginal Low

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

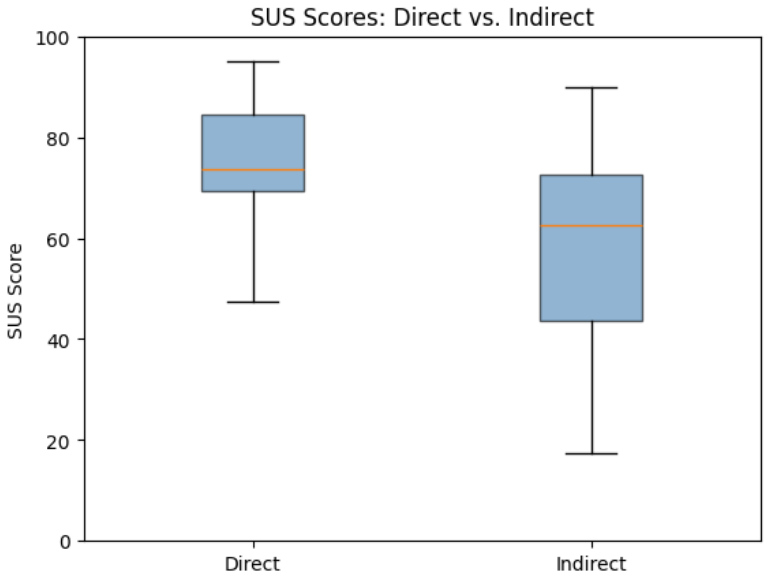


Figure 5.3: SUS score distribution by interaction mode

The Mann-Whitney U test was also used to test the relationship between SUS and mode order. The mean SUS score is slightly higher when the direct mode was used second ($mean = 77.5$) than when it was used first ($mean = 72.5$). The same trend was observable for the indirect mode, with a higher score when used second ($mean = 64.5$) compared to first ($mean = 53.8$). However, neither the result for the direct ($U = 43.0,\ Z = -0.529,\ p = 0.621$) nor for the indirect mode ($U = 35.5,\ Z = -1.096,\ p = 0.287$) reached significance. These results suggest that mode order did not significantly affect perceived usability.

It was also investigated via Spearman's rank correlation [Sed14] whether AR/VR experience or task completion time affected the SUS score. No significant relationship was found between AR/VR experience and the SUS score for either the direct mode ($p = 0.290$) or the indirect mode ($p = 0.497$), suggesting that prior AR/VR experience did not influence how usable participants found either mode. Regarding the relationship between task completion time and SUS score, a significant negative correlation was found for the direct mode ($r_s = -0.508,\ p = 0.022$), meaning that participants who completed tasks faster also tended to rate the direct mode as more usable. For the indirect mode, a similar negative trend was observed ($r_s = -0.403,\ p = 0.078$), but this did not reach significance. These results suggest that efficiency and perceived usability were more closely linked in the direct mode than in the indirect mode. It was also analyzed via the Mann-Whitney U test whether a relationship exists between the SUS score and the preferred mode. Participants who preferred the direct mode rated the direct mode higher on the SUS ($mean = 76.2$) than those who preferred the indirect mode ($mean = 65.8$). The same pattern was observable for the indirect SUS

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

scores, where participants who preferred the indirect mode rated it higher ($mean = 80.8$) than those who preferred the direct mode ($mean = 54.3$). These results are neither significant for the direct mode ($U = 29.0,\ Z = 0.770,\ p = 0.474$), nor for the indirect mode ($U = 6.0,\ z = -1.955,\ p = 0.057$), though the result for the indirect mode nearly reaches significance. This suggests that preference and perceived usability were aligned, particularly for the indirect mode. However, with only three participants preferring the indirect mode, the test was strongly underpowered and no strong conclusions can be drawn from these results.

Custom Likert Items

Usability was also assessed via four custom 5-point Likert items. Each item was included twice in the questionnaire, once per interaction mode, with *[interaction mode]* replaced by 'direct interaction mode' or 'indirect interaction mode' respectively:

- I always knew how to perform the action I wanted in the [interaction mode]. (1 = strongly disagree, 5 = strongly agree)
- I felt in control while interacting with the diagram in the [interaction mode]. (1 = strongly disagree, 5 = strongly agree)
- The visual presentation in the [interaction mode] felt cluttered or hard to read. (1 = strongly disagree, 5 = strongly agree)
- Using the [interaction mode] was physically comfortable (gestures, hand position, fatigue). (1 = strongly disagree, 5 = strongly agree)

The results can be seen in Figure 5.4 and Table 5.4. The median was used as measure for comparing the two modes. For all items, the direct mode showed a higher or better score than the indirect mode. For learnability (i.e., whether participants always knew how to perform the action they wanted), both modes scored well, with the direct mode reaching the maximum score of 5.0 and the indirect mode a score of 4.0. The scores for perceived control were lower, with a median of 4.0 for the direct mode and 3.0 for the indirect mode. For physical comfort, both modes scored similarly, with a median of 3.5 for the direct mode and 3.0 for the indirect mode. The clutter item is reverse coded, meaning that a lower score indicates less perceived clutter. Here, the direct mode achieved the best possible median score of 1.0, while the indirect mode also scored relatively low at 2.0. The Wilcoxon Signed-Rank Test was applied to assess the significance of these differences. Only the clutter item reached significance ($p < 0.05$), with participants rating the direct mode as significantly less cluttered than the indirect mode.

User Preferences

When asked about their preferred mode, the majority of the user study participants (85%) replied that they preferred the direct mode. Only 3 out of 20 participants preferred the indirect mode over the direct mode. This result is depicted in Figure 5.5.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

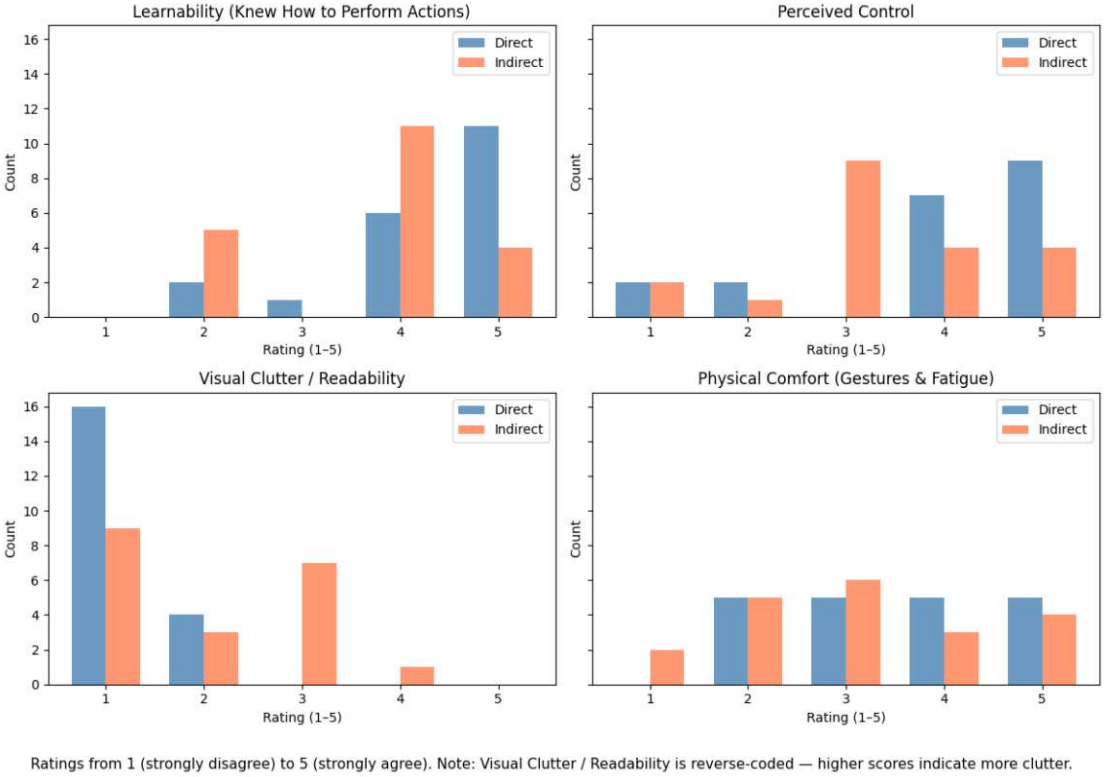


Figure 5.4: Usability: custom Likert items

Table 5.4: Wilcoxon signed-rank test results for custom Likert items

Item	**Direct Median**	**Indirect Median**	**Z**	**p-value**
Learnability	5.0	4.0	-1.572	0.110
Control	4.0	3.0	-1.633	0.097
Clutter	1.0	2.0	-2.621	0.007
Physical Comfort	3.5	3.0	-1.098	0.266

Furthermore, participants were asked to rate their preferred interaction mode per operation. They could give answers from 1, meaning they strongly preferred the direct mode, to 5, meaning they strongly preferred the indirect mode. The results are shown in Figure 5.6 and Table 5.5. For all operations (connecting, transforming, labeling, adding, and navigating to and selecting elements), the direct mode was preferred over the indirect mode. Furthermore, the IQR span from a strong preference for the direct mode to at most no preference across all operations, indicating that the majority of participants did not lean toward the indirect mode for any operation. This was also in line with the majority

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

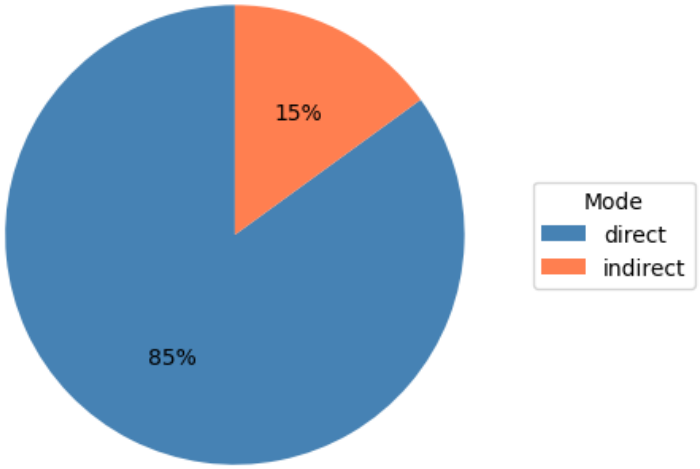


Figure 5.5: Overall user preference

of participants favoring the direct mode overall. The operations navigating/selecting, transforming, and connecting elements showed the clearest signal with a median of 1.0, indicating a strong preference for the direct mode. For adding and labeling, the median was 2.0, suggesting a moderate preference for the direct mode. To assess the significance of these results, Wilcoxon Signed-Rank Tests were applied, testing against a neutral preference value of 3.0. As can be seen in Table 5.5, the results were significant ($p < 0.05$) for all operations, except labeling, where the p-value ($p = 0.054$) narrowly exceeded the significance threshold.

Table 5.5: Operation-specific interaction mode preferences and Wilcoxon signed-rank test results

Operation	**Median**	**IQR**	**Z**	**p-value**
Navigating/Selecting	1.0	1.0–1.25	-3.201	< 0.001
Adding	2.0	1.0–3.0	-3.051	0.002
Labeling	2.0	1.0–3.0	-1.874	0.054
Transforming	1.0	1.0–3.25	-2.172	0.020
Connecting	1.0	1.0–2.0	-3.114	0.001

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

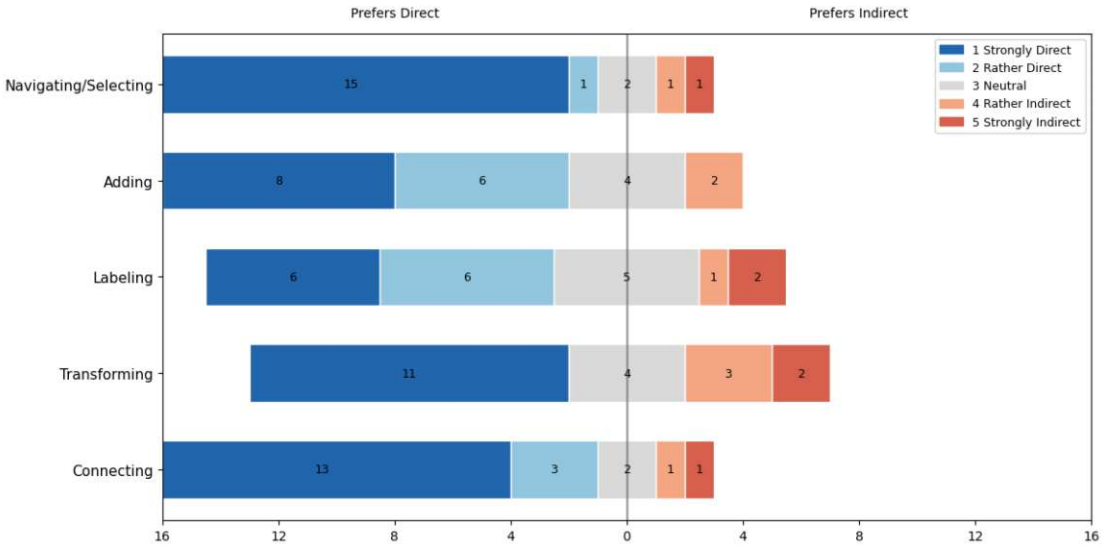


Figure 5.6: Operation-specific mode preference

5.2.4 Qualitative Findings

The user study questionnaire also included four open-ended questions:

- Overall, which interaction mode did you prefer (direct, indirect, or no preference)? If you have a preference, please briefly explain why.
- What did you like about the interaction modes?
- What did you dislike about the interaction modes?
- Do you have any suggestions for improving either interaction mode?

Seven recurring themes were identified in the participants' responses, of which themes 1–3 were the most prominent.

Theme 1: Navigation and Element Selection in Indirect Mode Was Cumbersome

The majority of participants (14 out of 20) found navigating to elements and subsequently selecting them frustrating and cumbersome. Using the arrows for navigation *'felt unnecessarily complicated'* and harder than directly selecting elements in the diagram. Furthermore, participants mentioned that it was hard to see where the navigation box currently was in the diagram while simultaneously operating the panel: *'sometimes it was hard to see the diagram while navigating it like you had to glance at the diagram and see where you are while still keeping the hand on the buttons'*. The navigation also felt unintuitive, since the navigation box sometimes moved in unexpected ways, particularly

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

when elements where at different heights. This is also reflected in the quantitative results for operation-specific mode preferences. For navigation and selection, participants strongly preferred the direct mode, as discussed in Section 5.2.3.

Theme 2: Gesture Recognition Was Unreliable

Almost half of the participants (9 out of 20) experienced issues with unreliable gesture recognition. The pinch gesture was registered inconsistently, and the double pinch in particular posed problems: *'Very hard to do the double pinch'*, *'Pinching sometimes didn't register'*, *'Pinch gesture was registered really inconsistently, which was annoying'*.

Theme 3: Direct Interaction Was Perceived as More Intuitive

Seven of the 20 participants mentioned that direct interaction felt more intuitive than indirect interaction: *'Direct interaction was well thought out and intuitive to use'*, *'Direct felt very intuitive'*, *'direct: intuitive, quick, ease of use'*. Participants also mentioned that they liked resizing elements in the direct mode and noted that navigating to elements, selecting them, and translating them felt quick. However, some found that *'rotating and resizing felt too insensitive (you had to perform a lot of physical motion for very little change in the model)'* and it was sometimes difficult to select tools and options since one had to use both hands and remember which tools were located at which hand menu.

Theme 4: Indirect Mode Offers Stability and Advantages for Certain Operations

Six participants pointed out that the indirect mode offered advantages for certain operations. Some found *'It was easier to select items and use interactions with the indirect mode'* and that *'selection of objects was more stable'*. Furthermore, participants mentioned that the indirect mode made it easier to create diagrams from scratch and add new elements: *'indirect felt faster for creating new elements especially'*, *'better for creating from scratch'*. Some also preferred the indirect mode for transforming objects: *'transforming objects is better'*. It was also noted that the indirect mode was less fatiguing and straightforward to understand. However, moving objects over longer distances in the indirect mode was perceived as cumbersome. Interestingly, this preference for certain operations is not fully reflected in the operation-specific preference results discussed in Section 5.2.3, where participants overall still preferred the direct mode across all operations.

Theme 5: Visibility and Feedback Issues

Five of the 20 participants mentioned difficulty seeing which element was currently selected, particularly in indirect mode where the blue selection box was hard to spot: *'it was sometimes hard to see which item was selected'*, *'it was hard to see which object in the diagram was selected with the blue frame while controlling it with the 'menu keyboard' to jump to a different object in the diagram'*. The visibility of the pointers also posed

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

problems and participants wished for *'better visibility of the pointers for indirect mode'* and *'a visible pointer to see where you are moving exactly'*.

Theme 6: Desire for a Combined Mode

Three participants suggested combining both modes: *'mix of both modes would be best'*. For example, keeping the panel for managing functions and transformations while selecting elements directly in the diagram. It was also noted that the direct interaction mode was better suited for manipulating existing diagrams, while the indirect mode was better for creating new diagrams from scratch.

Theme 7: Learning Curve and Improvement with Practice

Lastly, three participants reported a steep learning curve and noted that the performance improved during the session: *'strong learning curve. It took little time to know everything and to perform well on all tasks'*, *'the double pinch was hard for me to perform, although that got better while performing the tasks'*. This aligns with the quantitative mode order effect described in Section 5.2.2.

Suggestions for Improvement

In addition to the desire for a combined mode discussed above, participants suggested several further improvements. Feedback and visibility of the selection box and pointers were mentioned as areas for improvement. Hand gestures should be fine-tuned for more reliable recognition, and text size on elements should be increased. It was also suggested that the visualization of BPMN elements could be altered, for example by using *'a ball for start and end'* events. Furthermore, participants suggested allowing BPMN elements to be replaced directly with other element types, rather than having to delete an element and add a new one in its place. Other suggestions included the ability to align elements in direct mode and to add tooltips for buttons. One participant also suggested changing the tool selection mechanism in direct mode, proposing a radial menu *'where you motion to the tool you would like to select and have to keep your hand there for a second to select the tool'* instead of relying on pinching.

5.2.5 Additional Observations

During the user study, several issues experienced by participants were directly observed, some of which are also reflected in the qualitative findings discussed above. The observations can be grouped into the following areas.

Gesture Recognition

As mentioned in Section 5.2.4, participants experienced issues with pinch recognition and double pinching. Sometimes this was due to improper execution of the pinch gesture, such as holding the hand in a way that the gesture could not be identified by the headset

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

(e.g. turned away or held too low), or not moving the pointer close enough to the target element in the scene. Some participants also had difficulty understanding how fast or slow a double pinch should be executed, or accidentally pinched unintentionally, leading to unintended interactions. For example, accidentally pinching an element or somewhere in empty space interrupted their double pinch gesture or caused buttons to disappear. These issues could partly be mitigated by spending more time with the prototype and becoming familiar with the gestures. During the user study, it was observable that the longer participants interacted with the modes, the easier the handling became. They grew more aware of the sensitivity of the pinch detection and adjusted their gestures accordingly, making them more deliberate and precise. However, the gesture mechanics could also be improved on the implementation side, for example by reducing sensitivity and fine-tuning the double pinch timer. Issues with pinch gesture recognition have also been identified in prior AR research [RGBA19], suggesting that this remains an open challenge in hand gesture-based AR interaction.

Spatial and Visibility Issues

Some participants moved outside the AR session's tracking area, causing the diagram to disappear until they returned to their original position, or they did not move close enough to the diagram, which made element selection harder. Particularly in the direct mode, selection was easier when participants moved closer to the elements they wanted to select rather than remaining stationary. Some participants also placed the indirect interaction panel in an unfavorable position or orientation, such as too far to the right or too low, or they reduced its size too much, making it harder to hit the buttons with the pointers.

Furthermore, as discussed in Section 5.2.4, participants had difficulty seeing where they were currently pointing. This could be improved by displaying a colored ray for each pointer and changing the hit indicator to a more visible color. In the version used for the user study, the hit circle is gray and pointers have no ray, only a small triangle indicating the pointing direction.

Indirect Mode Navigation

As elaborated in Section 5.2.4, participants struggled with navigation in the indirect mode. This was also directly observable during the user study, and frustration was noticeable. The difficulties were partly due to the blue navigation box not being visible enough, and partly due to the challenge of simultaneously tracking the navigation box in the diagram and hitting the navigation buttons on the panel. Based on observable improvements during the session, particularly when participants repositioned the interaction panel, it can be assumed that these difficulties would improve with more practice and better panel placement. Furthermore, the algorithm for navigating through the diagram was not always predictable to participants, since it favoured elements positioned closer in space rather than connected elements, which was what participants expected. Having to separately confirm the selection of an element via the checkmark button to see the

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

'element selected buttons', even though the navigation box was already placed on it, caused confusion. Participants were sometimes unsure whether they had already selected an element or connection. Some suggested text prompts like *'you selected the source of the connection, now select the target'*, which could be helpful when getting familiar with the prototype, but could feel noisy for more experienced users. A mode with prompts that can be toggled on or off would be a possible improvement.

Furthermore, it was observable that participants who had started with the direct mode also attempted to select elements directly in the diagram while in the indirect mode. This suggests that, as discussed in Section 5.2.4, direct selection in the diagram feels more intuitive than using an intermediary such as the interaction panel.

Missing Functionality

Visible frustration was observed when some interactions were not supported in either mode. This included moving the entire diagram as a whole rather than individual elements separately. When a diagram is loaded from the database, it is placed stationary in the center of the scene. Being able to move the whole diagram would allow users to place it in a more convenient position without having to move each element individually or reposition themselves. Further missing functionality that caused frustration included changing element types directly (e.g. converting a parallel gateway to an exclusive gateway without having to delete and re-add it) and redirecting connections (i.e. changing the source or target of an existing connection without having to delete it and create a new one separately).

Element Overlap and Scaling

It led to confusion when newly added elements appeared inside or overlapping with existing elements in the diagram, making it unclear to participants whether their action had worked. This was particularly problematic in the indirect mode, where new elements are automatically placed in the center of the diagram when nothing is selected, or directly to the right of the selected element.

Finally, in the direct mode, some participants had difficulty with scaling, since it was hard for them to move both pointers onto the element to be scaled. This was partly due to not fully understanding the required gesture, as participants performed the pinching motion in empty space rather than placing both pointers on the element first, and partly due to the pointers not being visible enough.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 6

Conclusion

6.1 Summary

This thesis addressed BPMN model creation and manipulation in AR environments, supported through direct and indirect interaction modes. The primary aim was to conceptualize, implement, and evaluate efficient and usable means to create and manipulate conceptual models in AR, using BPMN as a concrete example of a modeling language. This was addressed through two subquestions: SQ1 investigated how direct and indirect interaction modes can be designed and implemented in a prototypical AR BPMN modeling tool, while SQ2 examined how the two modes compare in terms of efficiency, usability, and user preference. The insights gained are expected to be generalizable to other modeling notations.

The research followed the Design Science Research Methodology, beginning with a literature review to ground the work in the existing body of knowledge. After an introduction to BPMN, Extended Reality, and direct and indirect interaction, the literature review in Chapter 3 identified relevant approaches focusing on conceptual model visualization, creation, and interaction in immersive environments such as AR and VR. While many works addressed different aspects of this domain across diverse modeling notations, the support of immersive model creation and manipulation via hand gestures was identified as a research gap. Furthermore, no research on interaction modes for diagram creation and manipulation in AR could be identified.

Based on these insights, Chapter 4 presented the design of the BPMN Builder prototype, which supports BPMN model creation and manipulation via two interaction modes: direct and indirect interaction. In the direct mode, elements are selected directly in the diagram via pinch gestures, while in the indirect mode, an interaction panel serves as the means for navigation and interaction.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

The evaluation of the prototype was presented in Chapter 5. A user study with 20 participants was conducted to compare the interaction modes in terms of efficiency, usability, and user preference. The user study followed a within-subjects design, where participants tested both modes by performing creation and manipulation tasks. The results showed higher efficiency of the direct mode, compared to the indirect mode, as well as a higher SUS score. Furthermore, 85% of participants preferred the direct mode over the indirect mode and it was perceived as less cluttered. Concerning operation-specific preferences, the direct mode was preferred for navigation and selection, adding elements, transforming elements, and connecting elements. A mode order effect was also identified, meaning that participants were faster in whichever mode they used second, likely due to increased familiarity with the AR environment. Qualitative findings further revealed that participants found direct interaction more intuitive, while indirect mode navigation was perceived as cumbersome and gesture recognition unreliability was a recurring challenge.

Overall, this thesis contributes to the research field by providing a validated prototype for immersive BPMN model creation and manipulation, and by offering empirical insights into the applicability of direct and indirect interaction modes in AR-based conceptual modeling.

6.2 Limitations

Limitations of the user study include the small number of participants as well as the limited prior AR or VR experience among them. The identified difficulties with hand tracking and pinch recognition could be minimized by having more participants with prior experience in AR or VR and with gesture-based controls.

The within-subjects design allowed participants to directly compare both modes, but introduced a learning effect. Participants who used a mode second performed faster, likely due to increased familiarity with the AR environment and prototype gained during interaction with the first mode.

Concerning the prototype, both hands need to be free to execute certain functions in both interaction modes (indirect: scaling the panel; direct: scaling elements and using the hand menus), which could be problematic in scenarios where one hand is occupied, for example during process execution in a real-world environment. Furthermore, the restriction to a small subset of BPMN elements is a limitation, as the prototype does not support the full BPMN 2.0 notation.

6.3 Future Work

Future work could further develop the prototype by incorporating all existing BPMN elements and by combining the direct and indirect interaction modes, as suggested by some of the user study participants. Different selection mechanisms that not solely rely on pinch gestures, such as additional hand gestures, eye-based selection, or radial selection menus,

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

could also be investigated. Another valuable addition could be to allow users to modify the visualization of BPMN elements, for example by coloring elements, adjusting font size, or applying customized element visualizations, as investigated in [ACR+19]. Supporting collaboration between multiple users for shared diagram creation and manipulation, as researched in [YGWE23, YGWE21, GSD24, PSLT20, PSMT21], would be another interesting direction.

Furthermore, AI could be incorporated into the BPMN Builder to assist users during modeling, for example by suggesting diagram additions, checking label grammar, and validating diagram correctness. AI could even generate initial model drafts from natural language descriptions or multimodal data, building on the broader trend toward AI-assisted and multimodal conceptual modeling discussed in Section 2.2. This could be combined with speech recognition to allow fully hands-free diagram creation and manipulation.

It would also be valuable to support the direct import of existing 2D BPMN diagrams into the BPMN Builder, without requiring users to recreate them entirely in AR. Finally, it would be interesting to analyze the applicability of the BPMN Builder in process execution contexts, for example to support process modeling or process execution guidance.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 7

Appendix

This appendix presents an overview table of the literature identified during the literature review, the tasks given to user study participants across the different groups, and the full survey participants completed after performing the tasks.

7.1 Literature Review

This section gives an overview of the literature found during the literature review, detailed in Section 3. It classifies the identified sources concerning title, authors, year, immersive technology, model notation, whether model creation is supported, whether model manipulation is supported, and integrated interaction modalities.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Table 7.1: Literature overview part 1 of 3 (Abdul – Muff)

Title	Authors	Year	Immersive Tech.	Notation	Creat.	Manip.	Interaction
UBBA: Unity Based BPMN Animator [ACR+19]	Abdul et al.	2019	3D	BPMN	N	N	Desktop
Proteus: A Framework for Enhancing Model-Based Systems Engineering [BKvdW24]	Barosan et al.	2024	VR	System models	N	N	Desktop
Assessing Process Attribute Visualization and Interaction Approaches [GRM20]	Gall & Rinderle-Ma	2020	VR	BPMN	N	Y	Controllers
Design and Development of a Collaborative AR Environment for Systems Science [GSD24]	Giabbanelli et al.	2024	AR	Causal maps	N	Y	Gestures
A Context and AR BPMN and BPMS Extension for IIoT Processes [GHOP22]	Grambow et al.	2022	AR	BPMN	N	N	Desktop
The Effectiveness of 3D Virtual Environments for BPMN Teaching [MdSAdS26]	Matias et al.	2026	VR	BPMN	N	N	Controllers
Exploring Microsoft HoloLens for Interactive Visualization of UML Diagrams [MHGG17]	Mikkelsen et al.	2017	VR	UML	Y	Y	Gestures
Metamodeling for Extended Reality [Muf25]	Muff	2025	AR/VR	Conceptual models	Y	Y	Desktop
MM-AR: A Web-Based Open-Source Metamodeling Platform [MBF26]	Muff et al.	2026	AR	Conceptual models	N	Y	Gestures
Initial Concepts for AR and VR-based Enterprise Modeling [MF21]	Muff & Fill	2021	AR/VR	BPMN, EA	N	N	Desktop

Creat. = creation, Manip. = manipulation, Y = supported, N = not supported

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Table 7.2: Literature overview part 2 of 3 (Oberhauser – Rehring)

Title	Authors	Year	Immersive Tech.	Notation	Creat.	Manip.	Interaction
VR-BPMN: Visualizing BPMN Models in Virtual Reality [OPM18]	Oberhauser et al.	2018	VR	BPMN	N	N	Controllers
VR-EA: Virtual Reality Visualization of Enterprise Architecture Models [OP19]	Oberhauser & Pogolski	2019	VR	ArchiMate, BPMN	N	N	Controllers
VR-EAT: Visualization of Enterprise Architecture Tool Diagrams in VR [OSM20]	Oberhauser et al.	2020	VR	ArchiMate, BPMN	N	N	Controllers
VR-UML: The Unified Modeling Language in Virtual Reality [Obe21]	Oberhauser	2021	VR	UML	Y	Y	Controllers
VR-EA+TCK: Visualizing Enterprise Architecture, Content, and Knowledge in VR [OBS22]	Oberhauser et al.	2022	VR	ArchiMate, BPMN, KMS, ECMS	N	N	Controllers
VR-EvoEA+BP: Using Virtual Reality to Visualize Enterprise Context Dynamics [OBS23]	Oberhauser et al.	2023	VR	ArchiMate, BPMN, KMS, ECMS	N	N	Controllers
VR-SDLC: A Context-Enhanced Life Cycle Visualization in VR [Obe24]	Oberhauser	2024	VR	LML	N	N	Controllers
Enabling Collaborative Business Process Elicitation in Virtual Environments [PSLT20]	Pöhler et al.	2020	VR	Process models	Y	Y	Controllers
Let's Get Immersive: How VR Can Encourage User Engagement in Process Modeling [PSMT21]	Pöhler et al.	2021	VR	Process models	Y	Y	Controllers
Let's Get in Touch: Decision Making about EA Using 3D Visualization in AR [RGBA19]	Rehring et al.	2019	AR	EA	N	N	Controllers, voice

Table 7.3: Literature overview part 3 of 3 (Rehring – Yigitbas)

Title	Authors	Year	Immersive Tech.	Notation	Creat.	Manip.	Interaction
Evaluating the User Experience of an AR Prototype for Enterprise Architecture [RA20a]	Rehring & Ahlemann	2020	AR	EA	N	Y	Gestures
Using Augmented Reality in Software Engineering Education? [RHM+19]	Reuter et al.	2019	AR	UML	N	Y	Gestures
VisAr3D: An Innovative 3D Visualization of UML Models [RWL16]	Rodrigues et al.	2016	VR	UML	N	N	Desktop
HoloFlows: Modelling of Processes for the IoT in Mixed Reality [SKKA21]	Seiger et al.	2021	MR	IoT processes	Y	Y	Gestures
Are We Ready for Synchronous Conceptual Modeling in AR? [SG25]	Shrestha & Giabbanelli	2025	AR	Causal maps	N	Y	Gestures
Learning Software Modeling with Virtual Reality [WBG+25]	White et al.	2025	VR	UML	Y	Y	Controllers
Collaborative Software Modeling in Virtual Reality [YGWE21]	Yigitbas et al.	2021	VR	UML	Y	Y	Controllers
Gamification-based UML Learning Environment in VR [YSB+22]	Yigitbas et al.	2022	VR	UML	N	Y	Controllers
Design and Evaluation of a Collaborative UML Modeling Environment in VR [YGWE23]	Yigitbas et al.	2023	VR	UML	Y	Y	Controllers
GaMoVR: Gamification-based UML learning environment in virtual reality [YSB+24b]	Yigitbas et al.	2024	VR	UML	N	Y	Controllers
Gamification- and Virtual Reality-Based Learning Environment for UML Class Diagram Modeling [YSB+24a]	Yigitbas et al.	2024	VR	UML	N	Y	Controllers

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

7.2 User Study

The following presents the tasks that participants were asked to perform during the user study. For the creation task, they had to recreate the Creation Model shown in Figure 7.1 in both modes. For the manipulation tasks, they were given Model 1 (Figure 7.2), Model 2 (Figure 7.3), Model 3 (Figure 7.4), or Model 4 (Figure 7.5) and were asked to manipulate it according to the tasks listed in Table 7.4, Table 7.5, Table 7.6, and Table 7.7, respectively.

7.2.1 Creation Task

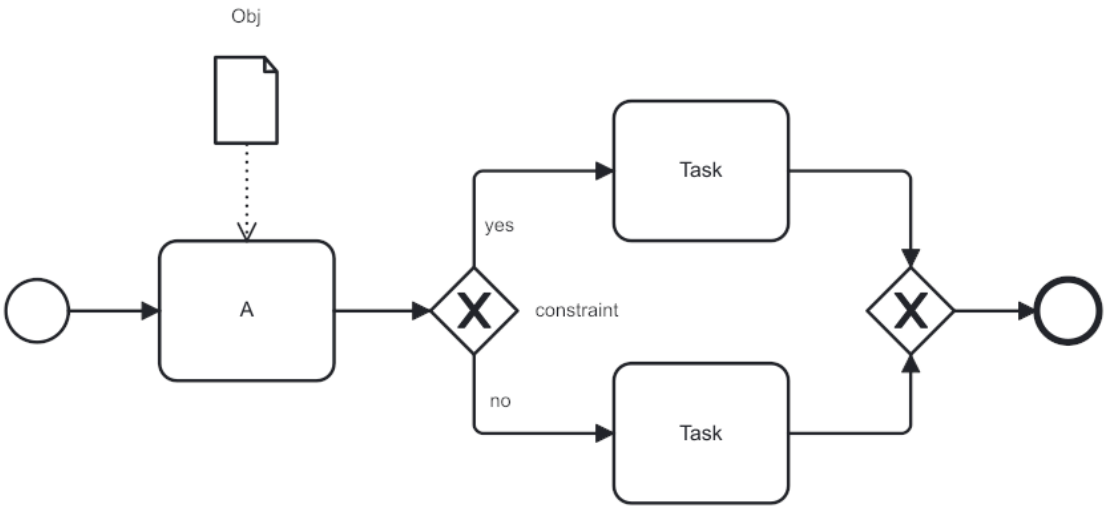


Figure 7.1: Model for creation task

7.2.2 Manipulation Tasks

Model 1

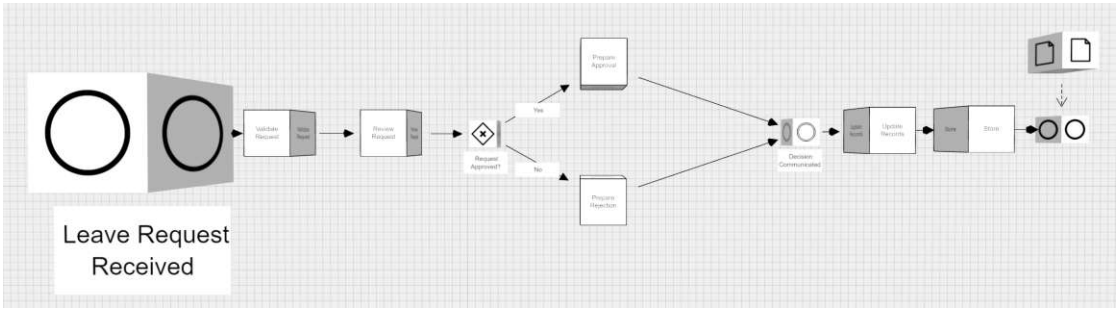


Figure 7.2: Model 1 for manipulation tasks

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Table 7.4: Manipulation tasks for manipulation model 1

No.	Task
1	Add the missing XOR merge gateway and reconnect the two dead-ending branches to it
2	Rename 'Store' to 'Archive Request'
3	Move the document artifact currently linked to the end event to 'Validate Request' and label it 'Leave Request Form'
4	Scale the start event to make it slightly smaller

Model 2

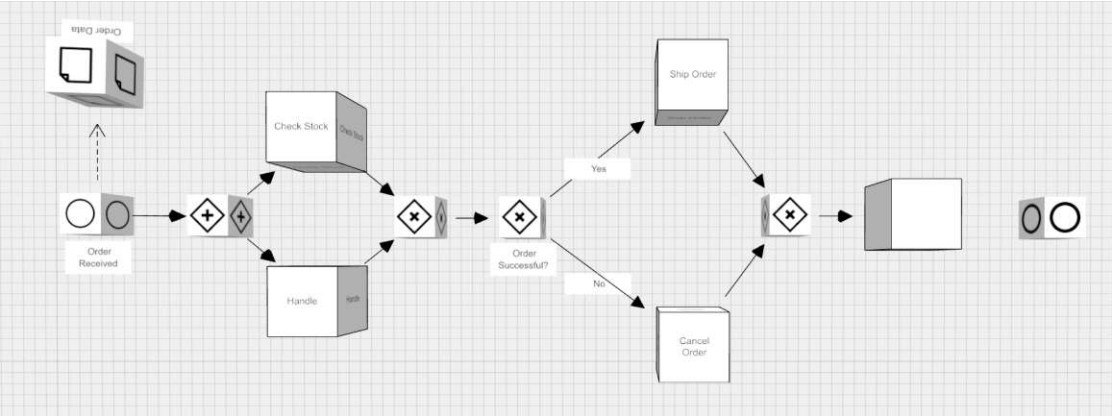


Figure 7.3: Model 2 for manipulation tasks

Table 7.5: Manipulation tasks for manipulation model 2

No.	Task
1	Replace the XOR gateway after the parallel split with a parallel join
2	Connect the unlabeled task to the end event and rename it 'Close Order'
3	Rename 'Handle' to 'Process Payment'
4	Rotate the document artifact 180 degrees so the folded corner moves to the top right

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Model 3

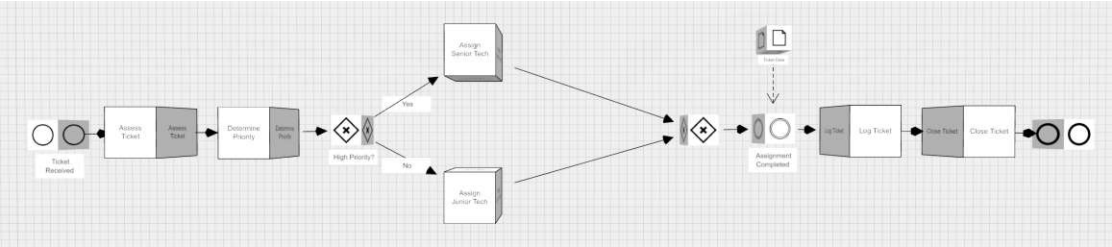


Figure 7.4: Model 3 for manipulation tasks

Table 7.6: Manipulation tasks for manipulation model 3

No.	Task
1	Add a new task 'Escalate to Manager' between 'Assign Senior Tech' and the XOR merge gateway
2	Move the document artifact currently associated with 'Assignment Completed' and associate it with 'Assess Ticket' instead
3	Rename 'Determine Priority' to 'Classify Ticket'
4	Scale the document artifact to make it slightly bigger

Model 4

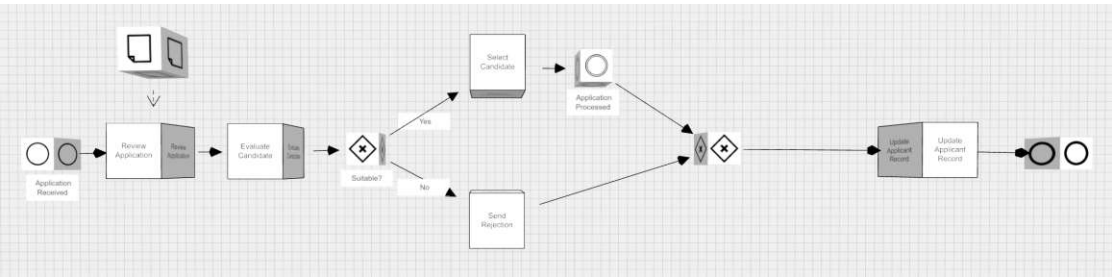


Figure 7.5: Model 4 for manipulation tasks

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Table 7.7: Manipulation tasks for manipulation model 4

No.	Task
1	Move the intermediate event 'Application Processed' to after the XOR merge gateway
2	Add a new task 'Schedule Interview' between 'Select Candidate' and the XOR merge gateway
3	Rotate the document artifact 180 degrees so the folded corner moves to the top right
4	Label the document artifact 'Application Data'

7.2.3 User Study Questionnaire

This section presents the questionnaire that participants filled out after completing all tasks. The questionnaire was structured into three parts: background, the usability of each mode, and user preferences. Usability was assessed based on the SUS and custom Likert items.

Part 1: Background

B1: How would you rate your experience with BPMN modeling? (1 = no experience, 5 = very experienced)

B2: How would you rate your experience with AR or VR devices? (1 = no experience, 5 = very experienced)

B3: How would you rate your experience with gesture-based interaction (e.g., hand tracking, motion controls)? (1 = no experience, 5 = very experienced)

Part 2: Usability of Direct and Indirect Mode

SUS1: I think that I would like to use this system frequently (1 = strongly disagree, 5 = strongly agree)

SUS2: I found the system unnecessarily complex (1 = strongly disagree, 5 = strongly agree)

SUS3: I thought the system was easy to use (1 = strongly disagree, 5 = strongly agree)

SUS4: I think that I would need the support of a technical person to be able to use this system (1 = strongly disagree, 5 = strongly agree)

SUS5: I found the various functions in this system were well integrated (1 = strongly disagree, 5 = strongly agree)

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

SUS6: I thought there was too much inconsistency in this system (1 = strongly disagree, 5 = strongly agree)

SUS7: I would imagine that most people would learn to use this system very quickly (1 = strongly disagree, 5 = strongly agree)

SUS8: I found the system very cumbersome to use (1 = strongly disagree, 5 = strongly agree)

SUS9: I felt very confident using the system (1 = strongly disagree, 5 = strongly agree)

SUS10: I needed to learn a lot of things before I could get going with this system (1 = strongly disagree, 5 = strongly agree)

U11: I always knew how to perform the action I wanted in the *[interaction mode]*. (1 = strongly disagree, 5 = strongly agree)

U12: I felt in control while interacting with the diagram in the *[interaction mode]*. (1 = strongly disagree, 5 = strongly agree)

U13: The visual presentation in the *[interaction mode]* felt cluttered or hard to read. (1 = strongly disagree, 5 = strongly agree)

U14: Using the *[interaction mode]* was physically comfortable (gestures, hand position, fatigue). (1 = strongly disagree, 5 = strongly agree)

Part 3: User Preferences

C1: What is your preferred mode for navigating to elements and selecting them? (direct = strongly prefer, indirect = strongly prefer)

C2: What is your preferred mode for adding elements? (direct = strongly prefer, indirect = strongly prefer)

C3: What is your preferred mode for labeling elements? (direct = strongly prefer, indirect = strongly prefer)

C4: What is your preferred mode for transforming (moving, scaling, rotating) elements? (direct = strongly prefer, indirect = strongly prefer)

C5: What is your preferred mode for connecting elements? (direct = strongly prefer, indirect = strongly prefer)

C6: Overall, which interaction mode did you prefer? If you have a preference, briefly explain why. (Direct / Indirect / No preference)

O1: What did you like in the interaction modes?

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

O2: What did you dislike in the interaction modes?

O3: Do you have ideas on how to improve either mode?

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Overview of Generative AI Tools Used

Claude[1] and GitHub Copilot[2] were used for coding support during the implementation of the prototype. Claude was additionally used to support language clarity, spell checking, grammatical corrections, and the structuring of chapters and sections of this thesis. DeepL Write[3] was used to aid the German translation of the abstract.

The originality and academic integrity of this work were not influenced by AI usage. All conceptual development, analysis, and final formulations were carried out independently by the author.

[1] https://claude.ai
[2] https://github.com/copilot
[3] https://www.deepl.com/write

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

List of Figures

1.1 Design Science approach (adapted from Peffers et al. [PTRC07]) 5

2.1 AR compared to VR (adopted from [DBGJ22]) 8
2.2 Reality-virtuality continuum (adopted from [DBGJ22]) 9
2.3 Extended Reality Euler diagram (adapted from [DBGJ22]) 10

3.1 VR-BPMN prototype [OPM18] 17
3.2 VR-EA prototype [OP19] 18
3.3 3DViz prototype [GRM20] 19
3.4 VR-UML virtual tablet [Obe21] 20
3.5 Collaborative UML modeling in VR [YGWE23] 21
3.6 Participatory process modeling in VR [PSMT21] 23

4.1 Overview of the architecture of the BPMN Builder prototype 27
4.2 JSON representation of a task element 28
4.3 Desktop mode 31
4.4 Database interaction in the prototype 37
4.5 Element selection 38
4.6 Edit options 38
4.7 Hand menus 39
4.8 Direct interaction transform controls 40
4.9 Navigation directions 42
4.10 Element selected views 42
4.11 Transform controls in indirect mode 43
4.12 Adding an element with prior selection 44
4.13 Adding an element without prior selection 44
4.14 Connection creation in indirect mode 44

5.1 Prior experience 48
5.2 Efficiency 48
5.3 SUS score distribution by interaction mode 51
5.4 Usability: custom Likert items 53
5.5 Overall user preference 54
5.6 Operation-specific mode preference 55

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

7.1 Model for creation task 69
7.2 Model 1 for manipulation tasks 69
7.3 Model 2 for manipulation tasks 70
7.4 Model 3 for manipulation tasks 71
7.5 Model 4 for manipulation tasks 71

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

List of Tables

3.1 Overview of selected works on conceptual modeling in immersive environments 16

4.1 Comparison of standard BPMN element visualization to the visualization of BPMN elements in the BPMN Builder prototype 33
4.2 Overview of shared UI controls available in both interaction modes 36

5.1 Group split for user study 47
5.2 Task completion time (in seconds) by mode and task type 48
5.3 SUS results by interaction mode 50
5.4 Wilcoxon signed-rank test results for custom Likert items 53
5.5 Operation-specific interaction mode preferences and Wilcoxon signed-rank test results 54

7.1 Literature overview part 1 of 3 (Abdul – Muff) 66
7.2 Literature overview part 2 of 3 (Oberhauser – Rehring) 67
7.3 Literature overview part 3 of 3 (Rehring – Yigitbas) 68
7.4 Manipulation tasks for manipulation model 1 70
7.5 Manipulation tasks for manipulation model 2 70
7.6 Manipulation tasks for manipulation model 3 71
7.7 Manipulation tasks for manipulation model 4 72

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Bibliography

[ACR+19] Basit Mubeen Abdul, Flavio Corradini, Barbara Re, Lorenzo Rossi, and Francesco Tiezzi. UBBA: Unity Based BPMN Animator. In Cinzia Cappiello and Marcela Ruiz, editors, *Information Systems Engineering in Responsible Information Systems*, pages 1–9, Cham, 2019. Springer International Publishing.

[Ari17] Ahmed Sabbir Arif. A Brief Note on Selecting and Reporting the Right Statistical Test, 2017.

[AS19] Ikpe Justice Akpan and Murali Shanker. A comparative evaluation of the effectiveness of virtual reality, 3D visualization and 2D visual interactive simulation: an exploratory meta-analysis. *SIMULATION*, 95(2):145–170, February 2019.

[Bar11] Carol M. Barnum. Planning for usability testing. In *Usability Testing Essentials*, pages 105–155. Elsevier, 2011.

[BAR23] Dominik Bork, Syed Juned Ali, and Ben Roelens. Conceptual Modeling and Artificial Intelligence: A Systematic Mapping Study, March 2023.

[BCF+20] Eugenie Brasier, Olivier Chapuis, Nicolas Ferey, Jeanne Vezien, and Caroline Appert. ARPads: Mid-air Indirect Input for Augmented Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 332–343, November 2020. ISSN: 1554-7868.

[BKM08] Aaron Bangor, Philip T. Kortum, and James T. Miller. An Empirical Evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction*, 24(6):574–594, July 2008.

[BKM09] Aaron Bangor, Philip Kortum, and James Miller. Determining what individual SUS scores mean: adding an adjective rating scale. *J. Usability Studies*, 4(3):114–123, 2009.

[BKvdW24] Ion Barosan, Thijs Koenraadt, and Huub van de Wetering. Proteus: A Framework for Enhancing Model-Based Systems Engineering Through 3D Visualization of System Models. In *2024 International Conference on*

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Electrical, Computer and Energy Technologies (ICECET, pages 1–9, July 2024.

[Bro96] John Brooke. SUS: A 'Quick and Dirty' Usability Scale. In *Usability Evaluation in Industry*, pages 189–194. Taylor and Francis, London, 1996.

[BSV+25] Vladimir Berezovsky, Natalia Sluzova, Artem Vahrushev, Daniil Gubkin, and Vladislav Potyomin. Intelligent Support and Modeling of Organizational Structure Redesign for Micro and Small Businesses based on Large Language Models and Fuzzy Rules. In Cengiz Kahraman, Selcuk Cebi, Basar Oztaysi, Sezi Cevik Onar, Cagrı Tolga, Irem Ucal Sari, and Irem Otay, editors, *Intelligent and Fuzzy Systems*, pages 535–543, Cham, 2025. Springer Nature Switzerland.

[BVD19] Wolfgang Büschel, Stefan Vogt, and Raimund Dachselt. Augmented Reality Graph Visualizations. *IEEE Computer Graphics and Applications*, 39(3):29–40, May 2019.

[CBB17] Osvaldo Cairó Battistutti and Dominik Bork. Tacit to explicit knowledge conversion. *Cognitive Processing*, 18(4):461–477, November 2017.

[CGRR18] Pietro Cipresso, Irene Alice Chicchi Giglioli, Mariano Alcañiz Raya, and Giuseppe Riva. The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature. *Frontiers in Psychology*, 9, November 2018.

[CPR+25] Flavio Corradini, Sara Pettinari, Barbara Re, Lorenzo Rossi, and Massimiliano Sampaolo. Enabling Process Mining on Multimodal Robotic Data. In Jānis Grabis and Yves Wautelet, editors, *Advanced Information Systems Engineering Workshops*, pages 257–269, Cham, 2025. Springer Nature Switzerland.

[CPR+26] Flavio Corradini, Sara Pettinari, Barbara Re, Lorenzo Rossi, and Massimiliano Sampaolo. Multimodal Zero-Shot Activity Recognition for Process Mining of Robotic Systems. In Mahendrawathi ER, Avigdor Gal, Thomas Grisold, Flavia Santoro, Mathias Weske, Remco M. Dijkman, Dimka Karastoyanova, Banu Aysolmaz, Wasana Bandara, and Kate Revoredo, editors, *Business Process Management: Responsible BPM Forum, Process Technology Forum, Educators Forum*, pages 263–275, Cham, 2026. Springer Nature Switzerland.

[CVL+17] Bradley Camburn, Vimal Viswanathan, Julie Linsey, David Anderson, Daniel Jensen, Richard Crawford, Kevin Otto, and Kristin Wood. Design prototyping methods: state of the art in strategies, techniques, and guidelines. *Design Science*, 3:e13, January 2017.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[DBGJ22] Ralf Doerner, Wolfgang Broll, Paul Grimm, and Bernhard Jung, editors. *Virtual and Augmented Reality (VR/AR): Foundations and Methods of Extended Realities (XR)*. Springer International Publishing, Cham, 2022.

[DG24] Mohammad Reza Dastmalchi and Amir Goli. Embodied Learning in Virtual Reality: Comparing Direct and Indirect Interaction Effects on Educational Outcomes. In *2024 IEEE Frontiers in Education Conference (FIE)*, pages 1–7, October 2024. ISSN: 2377-634X.

[DGZ12] Xiaoming Du, Chuang Gu, and Ning Zhu. A survey of business process simulation visualization. In *2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, pages 43–48, June 2012.

[DLRMR18] Marlon Dumas, Marcello La Rosa, Jan Mendling, and Hajo A. Reijers. *Fundamentals of Business Process Management*. Springer, Berlin, Heidelberg, 2018.

[FFK23] Hans-Georg Fill, Peter Fettke, and Julius Köpke. Conceptual Modeling and Large Language Models: Impressions From First Experiments With ChatGPT. *Enterprise Modelling and Information Systems Architectures (EMISAJ)*, pages 3:1–15 Pages, April 2023.

[FFK25] Peter Fettke, Hans-Georg Fill, and Julius Köpke. LLM, LAM, LxM Agent: From Talking to Acting Machines: Insights from the Perspective of Conceptual Modeling. *Enterprise Modelling and Information Systems Architectures (EMISAJ) - International Journal of Conceptual Modeling*, 20, May 2025.

[Fil24] Hans-Georg Fill. Spatial Conceptual Modeling: Anchoring Knowledge in the Real World. In Hans-Georg Fill and Harald Kühn, editors, *Metamodeling: Applications and Trajectories to the Future: Essays in Honor of Dimitris Karagiannis*, pages 35–50. Springer Nature Switzerland, Cham, 2024.

[FM24] Hans Georg Fill and Fabian Muff. Bridging the mental and the physical world: Conceptual modeling and augmented reality. In *Informing Possible Future Worlds: Essays in Honour of Ulrich Frank*, pages 197–212. 2024.

[Gav26] Aleksandar Gavric. *Enhancing conceptual modeling through multimodal data analysis and mixed reality*. PhD thesis, TU Wien, Vienna, March 2026. Unpublished.

[GB09] Maria J. Grant and Andrew Booth. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2):91–108, 2009.

[GBP24] Aleksandar Gavric, Dominik Bork, and Henderik Proper. Multimodal Process Mining. In *2024 26th International Conference on Business Informatics (CBI)*, pages 99–108, September 2024. ISSN: 2378-1971.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[GBP25] Aleksandar Gavric, Dominik Bork, and Henderik A. Proper. Enriching Business Process Event Logs with Multimodal Evidence. In Elda Paja, Jelena Zdravkovic, Evangelia Kavakli, and Janis Stirna, editors, *The Practice of Enterprise Modeling*, pages 175–191, Cham, 2025. Springer Nature Switzerland.

[GHOP22] Gregor Grambow, Daniel Hieber, Roy Oberhauser, and Camil Pogolski. A Context and Augmented Reality BPMN and BPMS Extension for Industrial Internet of Things Processes. In Andrea Marrella and Barbara Weber, editors, *Business Process Management Workshops*, pages 379–390, Cham, 2022. Springer International Publishing.

[GNM+23] Despoina Gavgiotaki, Stavroula Ntoa, George Margetis, Konstantinos C. Apostolakis, and Constantine Stephanidis. Gesture-based Interaction for AR Systems: A Short Review. In *Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments*, PETRA '23, pages 284–292, New York, NY, USA, August 2023. Association for Computing Machinery.

[GRM20] Manuel Gall and Stefanie Rinderle-Ma. Assessing Process Attribute Visualization and Interaction Approaches Based on a Controlled Experiment. *International Journal of Cooperative Information Systems*, 29(04):2050007, December 2020.

[Gro11] Object Management Group. Business Process Model and Notation (BPMN) Version 2.0, 2011.

[Gro26a] Object Management Group. About the Unified Modeling Language Specification Version 2.5.1, 2026.

[Gro26b] The Open Group. Archimate® 3.2 specification, July 2026.

[GSD24] Philippe Giabbanelli, Anish Shrestha, and Loic Demay. Design and Development of a Collaborative Augmented Reality Environment for Systems Science. 2024.

[Hev07] Alan Hevner. A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, 19, January 2007.

[HMPR04] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. Design Science in Information Systems Research. *MIS Quarterly*, 28(1):75–105, 2004.

[Jer16] Jason Jerald. *The VR book: human-centered design for virtual reality.* ACM books. Morgan & Claypool, [San Rafael], first edition edition, 2016.

[KC07] Barbara Kitchenham and Stuart Charters. Guidelines for performing Systematic Literature Reviews in Software Engineering, January 2007.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[LDL21] Ze-Nian Li, Mark S. Drew, and Jiangchuan Liu. Augmented Reality and Virtual Reality. In Ze-Nian Li, Mark S. Drew, and Jiangchuan Liu, editors, *Fundamentals of Multimedia*, pages 737–761. Springer International Publishing, Cham, 2021.

[LNR+25] Lien P. Le, Thu Nguyen, Michael A. Riegler, Pål Halvorsen, and Binh T. Nguyen. Multimodal missing data in healthcare: A comprehensive review and future directions. *Computer Science Review*, 56:100720, May 2025.

[MBF26] Fabian Muff, Daniel Borcard, and Hans Georg Fill. MM-AR: a web-based open-source metamodeling platform for spatial conceptual modeling, 2026.

[MdSAdS26] Diógenes Carvalho Matias, Lenylda Maria de Souza Albuquerque, and Denis Silva da Silveira. The Effectiveness of 3D Virtual Environments for BPMN Teaching: A Quasi-Experimental Study. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 71–87, Cham, 2026. Springer Nature Switzerland.

[MF21] Fabian Muff and Hans-Georg Fill. Initial Concepts for Augmented and Virtual Reality-based Enterprise Modeling. In Roman Lukyanenko, Binny M. Samuel, and Arnon Sturm, editors, *Proceedings of the ER Demos and Posters 2021 co-located with 40th International Conference on Conceptual Modeling (ER 2021), St. John's, NL, Canada, October 18-21, 2021*, volume 2958 of *CEUR Workshop Proceedings*, pages 49–54. CEUR-WS.org, 2021.

[MGG25] John Mylopoulos, Giancarlo Guizzardi, and Nicola Guarino. Conceptual modeling: Foundations, a historical perspective, and a vision for the future. *Data & Knowledge Engineering*, 160:102483, November 2025.

[MHGG17] Anders Mikkelsen, Sondre Honningsøy, Tor-Morten Grønli, and George Ghinea. Exploring Microsoft Hololens for Interactive Visualization of UML Diagrams. In *Proceedings of the 9th International Conference on Management of Digital EcoSystems*, MEDES '17, pages 121–127, New York, NY, USA, November 2017. Association for Computing Machinery.

[MTUK94] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies*, (2351), 1994.

[Muf25] Fabian Muff. *Metamodeling for Extended Reality*. Springer Nature Switzerland, Cham, 2025.

[MVK13] Annette Mossel, Benjamin Venditti, and Hannes Kaufmann. 3DTouch and HOMER-S: intuitive manipulation techniques for one-handed handheld augmented reality. In *Proceedings of the Virtual Reality International Conference: Laval Virtual*, VRIC '13, pages 1–10, New York, NY, USA, 2013. Association for Computing Machinery.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[Myl01] John Mylopoulos. Conceptual Modelling and Telos. 2001.

[Nor13] Donald A. Norman. *The Design of Everyday Things*. Basic Books, New York, NY, USA, 2013.

[Obe21] Roy Oberhauser. VR-UML: The Unified Modeling Language in Virtual Reality – An Immersive Modeling Experience. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 40–58, Cham, 2021. Springer International Publishing.

[Obe24] Roy Oberhauser. VR-SDLC: A Context-Enhanced Life Cycle Visualization of Software-or-Systems Development in Virtual Reality. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 112–129, Cham, 2024. Springer Nature Switzerland.

[OBS22] Roy Oberhauser, Marie Baehre, and Pedro Sousa. VR-EA+TCK: Visualizing Enterprise Architecture, Content, and Knowledge in Virtual Reality. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 122–140, Cham, 2022. Springer International Publishing.

[OBS23] Roy Oberhauser, Marie Baehre, and Pedro Sousa. VR-EvoEA+BP: Using Virtual Reality to Visualize Enterprise Context Dynamics Related to Enterprise Evolution and Business Processes. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 110–128, Cham, 2023. Springer Nature Switzerland.

[OP19] Roy Oberhauser and Camil Pogolski. VR-EA: Virtual Reality Visualization of Enterprise Architecture Models with ArchiMate and BPMN. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 170–187, Cham, 2019. Springer International Publishing.

[OPM18] Roy Oberhauser, Camil Pogolski, and Alexandre Matic. VR-BPMN: Visualizing BPMN Models in Virtual Reality. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 83–97, Cham, 2018. Springer International Publishing.

[OSM20] Roy Oberhauser, Pedro Sousa, and Florian Michel. VR-EAT: Visualization of Enterprise Architecture Tool Diagrams in Virtual Reality. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 221–239, Cham, 2020. Springer International Publishing.

[PDG21] S. P. F. Peters, R. M. Dijkman, and P. W. P. J. Grefen. Resource Optimization in Business Processes. In *2021 IEEE 25th International Enterprise Distributed Object Computing Conference (EDOC)*, pages 104–113, October 2021. ISSN: 2325-6362.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[PIWB98] I. Poupyrev, T. Ichikawa, S. Weghorst, and M. Billinghurst. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. *Computer Graphics Forum*, 17(3):41–52, 1998.

[PSLT20] Ludger Pöhler, Julian Schuir, Simon Lübbers, and Frank Teuteberg. Enabling Collaborative Business Process Elicitation in Virtual Environments. In Boris Shishkov, editor, *Business Modeling and Software Design*, pages 375–385, Cham, 2020. Springer International Publishing.

[PSMT21] Ludger Pöhler, Julian Schuir, Pascal Meier, and Frank Teuteberg. Let's Get Immersive: How Virtual Reality Can Encourage User Engagement in Process Modeling. *ICIS 2021 Proceedings*, December 2021.

[PTRC07] Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, and Samir Chatterjee. A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3):45–77, 2007.

[RA20a] Kevin Rehring and Frederik Ahlemann. Evaluating the User Experience of an Augmented Reality Prototype for Enterprise Architecture. In *ResearchGate*, Potsdam, Germany, 2020.

[RA20b] Kasper Rise and Ole Andreas Alsos. The Potential of Gesture-Based Interaction. In Masaaki Kurosu, editor, *Human-Computer Interaction. Multimodal and Natural Interaction*, pages 125–136, Cham, 2020. Springer International Publishing.

[RAB+15] Stewart Robinson, Gilbert Arbez, Louis G. Birta, Andreas Tolk, and Gerd Wagner. Conceptual modeling: Definition, purpose and benefits. In *2015 Winter Simulation Conference (WSC)*, pages 2812–2826, December 2015. ISSN: 1558-4305.

[RGBA19] Kevin Rehring, Malte Greulich, Laurenz Bredenfeld, and Frederik Ahlemann. Let's Get in Touch - Decision Making about Enterprise Architecture Using 3D Visualization in Augmented Reality. January 2019.

[RHM+19] Rebecca Reuter, Florian Hauser, Daniel Muckelbauer, Theresa Stark, Erika Antoni, Jürgen Mottok, and Christian Wolff. Using Augmented Reality in Software Engineering Education? First insights to a comparative study of 2D and AR UML modeling. 2019.

[RWL16] Claudia Susie C. Rodrigues, Cláudia M. L. Werner, and Luiz Landau. VisAr3D: an innovative 3D visualization of UML models. In *Proceedings of the 38th International Conference on Software Engineering Companion*, ICSE '16, pages 451–460, New York, NY, USA, 2016. Association for Computing Machinery.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[Sed14] Philip Sedgwick. Spearman's rank correlation coefficient. *BMJ: British Medical Journal*, 349, 2014.

[SFDA+20] Josep Sànchez-Ferreres, Luis Delicado, Amine Abbab Andaloussi, Andrea Burattin, Guillermo Calderón-Ruiz, Barbara Weber, Josep Carmona, and Lluís Padró. Supporting the Process of Learning and Teaching Process Models. *IEEE Transactions on Learning Technologies*, 13(3):552–566, July 2020.

[SFH+18] Kurt Sandkuhl, Hans-Georg Fill, Stijn Hoppenbrouwers, John Krogstie, Florian Matthes, Andreas Opdahl, Gerhard Schwabe, Ömer Uludag, and Robert Winter. From Expert Discipline to Common Practice: A Vision and Research Agenda for Extending the Reach of Enterprise Modeling. *Business & Information Systems Engineering*, 60(1):69–80, February 2018.

[SG25] Anish Shrestha and Philippe J. Giabbanelli. Are We Ready for Synchronous Conceptual Modeling in Augmented Reality? A Usability Study on Causal Maps with HoloLens 2. *Information*, 16(11):952, November 2025.

[SKKA21] Ronny Seiger, Romina Kühn, Mandy Korzetz, and Uwe Aßmann. HoloFlows: modelling of processes for the Internet of Things in mixed reality. *Software and Systems Modeling*, 20(5):1465–1489, October 2021.

[VVW+18] Alla Vovk, View Profile, Fridolin Wild, View Profile, Will Guest, View Profile, Timo Kuula, and View Profile. Simulator Sickness in Augmented Reality Training Using the Microsoft HoloLens. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM Conferences, pages 1–9. April 2018.

[WBG+25] Maryangela White, Jaya Surya Bontha, Sudipto Ghosh, Marcia Moraes, Anne M. Cleary, and Mohammed Safayet Arefin. Learning Software Modeling with Virtual Reality. In *2025 ACM/IEEE 28th International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, pages 11–18, October 2025.

[YGWE21] Enes Yigitbas, Simon Gorissen, Nils Weidmann, and Gregor Engels. Collaborative Software Modeling in Virtual Reality. In *2021 ACM/IEEE 24th International Conference on Model Driven Engineering Languages and Systems (MODELS)*, pages 261–272, October 2021.

[YGWE23] Enes Yigitbas, Simon Gorissen, Nils Weidmann, and Gregor Engels. Design and evaluation of a collaborative UML modeling environment in virtual reality. *Software and Systems Modeling*, 22(5):1397–1425, October 2023.

[YSB+22] Enes Yigitbas, Maximilian Schmidt, Antonio Bucchiarone, Sebastian Gottschalk, and Gregor Engels. Gamification-based UML learning environment in virtual reality. In *Proceedings of the 25th International Conference*

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

on Model Driven Engineering Languages and Systems: Companion Proceedings, MODELS '22, pages 27–31, New York, NY, USA, November 2022. Association for Computing Machinery.

[YSB+24a] Enes Yigitbas, Maximilian Schmidt, Antonio Bucchiarone, Simone Bassanelli, and Gregor Engels. Gamification- and Virtual Reality-Based Learning Environment for UML Class Diagram Modeling. In *2024 IEEE Global Engineering Education Conference (EDUCON)*, pages 1–10, May 2024. ISSN: 2165-9567.

[YSB+24b] Enes Yigitbas, Maximilian Schmidt, Antonio Bucchiarone, Sebastian Gottschalk, and Gregor Engels. GaMoVR: Gamification-based UML learning environment in virtual reality. *Science of Computer Programming*, 231:103029, January 2024.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.