XR-based technical instructions in organized testing: User behaviour and design performance

Satu Rantakokko (satu.rantakokko@uwasa.fi)
University of Vaasa
Rebekah Rousi
University of Vaasa

Research Article

Keywords: Technical instructions, Extended reality, Virtual reality, User-centered design, TIER model

Posted Date: June 9th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3037074/v1

Abstract

Extended reality (XR) is changing the way humans interact with information. Developers are currently experiencing a need to understand more about the use and design of XR-delivered information. Humans are active agents within human-technology interaction. There is a need to systematically study how human agency affects the nature of XR content and how this knowledge can be applied within the content design process. The present article focuses on the context of an iterative co-design process for XR-based technical instructions. The research question is: In what ways can technology-based user behavior inform the design of XR-based technical instructions? Studies exist concerning XR-based technical instructions, yet the comparability of these studies is low. User characteristics among other variables influence the results of using XR-based technical instructions. The method comprises two experimental studies that aimed to identify how users notice and act upon or ignore both designed and unintended features of XR-based technical instructions. A framework for systematically testing XR-based technical instructions was developed and applied. The results indicate that verbal evaluations of the XR-based technical instructions may not be reliable. A systematic mixed method test process is crucial transitioning the design of technical instructions to a new medium. Conclusion: There is a need for a systematic method to test how users notice and act upon the intended features of XR-based instructions. The specific target users need to be involved in this iterative process.

1. INTRODUCTION

Radical changes are occurring in contemporary society, many of which are influenced by exponential technological developments. The past three years alone have seen work environments change from central offices to remote and hybrid work settings, and Generative Artificial Intelligence (AI) sending knowledge workers into existential dilemma (see e.g., Shukla 2023). The rhetoric of the Metaverse (virtual reality (VR) and extended reality (XR) has taken a stronghold on revolutionizing the ways in which we experience space, time, each other, and reality itself (see e.g., Brambilla Hall & Baier-Lentz 2022). With this in mind, it should not be surprising that these technological game-changers may provide numerous pragmatic advantages when re-thinking traditional forms of information presentation. One area of information presentation and design with a reputation for being awkward to use is that of technical instructions. Arguably, instructions (manuals), and particularly technical instructions, while intended to assist users and consumers, have been the bane of object engagement, assembly, and understanding particularly in the age of personal computing and self-assemblage of products (see e.g., Novick & Ward 2006; Woodson 1995).

Professionals such as Marrazzo (2018) explain the Active Users Paradox, in which users optimize the time it takes to achieve use goals through short-cutting instruction engagement. Marrazzo highlights the need to incorporate alternate learning styles in the instruction design as well as minimize reading time by emphasizing affordances, or pictorial information showing what can be done with what component to achieve specific results. For this reason, the area of XR shows great promise via its immediacy, proximity, and fidelity in the domain of technical instructions (Fiorentino et al. 2014).

1.1 XR and technical instructions

XR has increasingly been used to deliver technical instructions in industrial contexts (Gattullo et al., 2019; Vanneste et al. 2020). The ability to offer 3D, situation-related, and interactive instructions can significantly improve the efficiency of learning about and operating devices (Doolani et al. 2020). XR is an umbrella term that refers to 3D, interactive and real-time environments that include artificial, virtual elements (see, e.g., Fast-Berglund et al. 2018, p. 32). XR consists of virtual reality (VR) and mixed reality (MR). MR can be further divided into augmented reality (AR) and augmented virtuality (AV) (Milgram and Kishino, 1994). However, technological and social barriers have slowed XR from becoming widespread. These barriers are now starting to be resolved in various applications (Fast-Berglund et al. 2018, pp. 31–32). Yet, technological maturity and mass industrial uptake has taken some time. With this said, there have been several more advanced companies that have been forerunners in developing and utilizing XR for technical instructions, such as Mobidev, XRMeet, TechSee, WhaTech, and EDIIIE to name some. EDIIIE's[1] applications for instance, extend to the fields of aviation, automotive, education, military, consumer goods, aerospace, and entertainment. In particular, the application of XR in instructional design has been studied more in the fields of education (see, e.g., Ding et al. 2020; Etambakonga 2021) and medical services (see, e.g., Bao & Hurriyet 2021; Ara et al. 2021).

1.2 XR-enabled technical instructions – the missing links

The public knowledge of business customers in this area is not extensive. Thus, insight into the role of users within the development process, how users experience XR-enabled technical instructions (in industrial settings and otherwise), and how the systems have been iteratively developed based on user feedback is not widely available. This supports motivation for engaging in systematic user-centered inquiry of XR-based technical instruction design processes. For, in order to design for optimal fluency of information presentation, comprehension and application, there is the need to understand more about how the behavior and specifications of users impact the effectiveness of the instructions themselves. In moments of human-technology interaction (HTI), this behavior is often in relation to the technology. Meaning that there is a true interactive process between what is presented in the XR technology, how people receive and respond to the information, and how in turn human responses impact the instruction design.

While there are numerous usability testing and design methods in circulation, no extensive model has been designed to date, with the specific purpose of analyzing this interactional reactive-proactive process of humans in XR-based technical instruction use. The affordance model of technical instructions in extended reality (TIER) was developed to assist in the dissection of human-technical components involved in XR-based instruction design. The research question is: in what ways can technology-based user behavior inform the design of XR-based technical instructions? By technology-based the authors refer to behavior that is specifically influenced by the technology, its affordances and limitations. For instance, while wearing VR glasses users are not encouraged to physically move around physical space. Affordances of the XR displays such as 3D details and interactivity with the virtual environment might encourage users to explore in ways they would not have necessarily imagined in relation to previous paper or video versions of instructions. This renders the testing phase within the design process particularly important, especially from the perspective of ascertaining how humans behave when receiving instructions via XR.

This article focuses on the testing process of XR-based technical instructions via the TIER model to explore how both designed and unintended affordances of the instructions are either noticed or not noticed, actualized (or ignored) by the users. The study represents a meta-experimental approach in which the testing, its method and data collection are also under the microscope. The study was undertaken in the tradition of user-centered design (UCD). According to Abras, Maloney-Krichmar and Preece (2014), in UCD “the end-users influence how a design takes shape.” Users are involved in UCD in some form or another, “typically during requirements gathering and usability testing” (ibid. p. 1). In the context of this research, this means focusing on the users’ actions and experiences when they start using XR-based technical instructions.

1.3 Technical Communication Professionals and User-Centered Design

Rising concern across fields has been in relation to human users and their roles in relation to the design process. The goal of UCD is to hold users – people who engage with and utilize digital designs – at the center of design processes in order to ensure relevance, usability, desire to use, and alliance with values to name some (see e.g., Dopp et al. 2019; Karat 1997).

In the tradition of UCD, a central facet of the work of technical communication professionals Over the past few decades, increasing emphasis has been placed on human users, and UCD has been one of the methodological approaches that technical communication has engaged in accordingly (Abras et al. 2014; LaRoche & Traynor 2010; Salvo 2001). Designing technical instructions for a new medium, such as XR, requires both theory and methods of alternate ways to present instructions within the environment[2]. Thus, as in any other domain of multimedia design, XR offers possibilities for text, images (3D models), videos, sound and increasingly more multisensory information (i.e., haptic sensations and perhaps even smell and taste in the future[3]).

In order to heighten fidelity, accuracy and even personalization of the product and experience within the individual interacting with the XR-based instructions, increasing amounts of data need to be collected from multiple sources (sensory technology as well as user-fed information and user accounts). For this reason, technical instruction design is becoming ever more complex, as no longer is the human-instruction interaction a case of the individual engaging and utilizing the instructions, but it is necessary for the instructions and supporting technology to *use the user*. Thus, issues such as privacy and ethics are now also entering the discussion. As, although artificial intelligence (AI) is not the focus of this article, it must be acknowledged that the XR systems are increasing powered by AI, which in turn, relies on large data for machine learning (ML)

The current research reports the results of two experiments. In the experiments, participants were presented with a simple set of instructions for building an abstract LEGO structure delivered via a VR head-mounted display. The tests sought to: 1) discern whether participants noticed the designed affordances; and 2) how they acted (real use) upon or reacted to those noticed affordances in the experiment. Thus, the main research question of the study is: in what ways can technology-based user behavior inform the design of XR-based technical instructions?

In order to understand the level of reliance on the details of the information delivered within the XR-based instructions, we pose a sub-question: To what extent do users rely on the details of the instructions – how much do individuals compensate for inadequacies by use of mental information contents (already learned and stored mentally bound information – i.e., memory, see e.g., Rousi, Saariluoma & Leikas 2010; Silvennoinen, Rousi, Jokinen & Perälä 2015?

Thus, the XR-based instructions included a designed-in glitch: one LEGO block was replaced with a similar block of a different color. The glitch was included to investigate how the users acted when confronted by flawed technical instructions to see how the affordances of these instructions support cognitive fluency (Unkelbach 2006), supported by high level cognitive processing (top-down – from thought to perception) when perceiving information in the environment (Belke, Leder, Strobach & Carbon 2010).

This article is structured as follows: A literature review introduces research concerning XR-based technical instructions and their benefits. The results indicated that even very simple XR-based instructions could lead to various unexpected outcomes, highlighting the importance of an organized test process with the intended users of the instructions when starting to deliver them with a new medium.
[1] See more at: https://www.ediiie.com/cases/

[2] To exemplify the novelty of this area of scholarship, studies discussing the relationship between mode of information delivery or representation in XR and the user are concentrated in the area of learning and educational science (see e.g., Islam 2019).

[3] Significant advances have occurred in the field of multisensory human-computer interaction (see e.g., Spence et al. 2016).

2. PREVIOUS RESEARCH

XR as an environment for technical instructions is still an emerging domain that is severely lacking in standardized empirical methods and theoretical models of the core factors concerning the design, experience and implications of this area of design. It is imperative to formulate a robust experiment design when studying testing XR-based solutions. This is due to the nature of XR itself. In addition to the AI and data-driven issues that emerge, communication modes and standards also pose challenges. For example, Verhulsdonck and Morie (2009, p. 8) raised awareness of a dilemma caused by the psycho-physiology of human beings when engaged in communication. In face-to-face contexts, many non-verbal elements and movements, such as eye gaze, are unintentional. In XR environments these movements may impact interaction, i.e., accidentally giving an eye-based command or selection, which thus affects the display. It may be difficult to design the instructions to account for unintentional user movements.

In the context of work, this requires special consideration. For example, from the point of view of how the hands of the user may or may not be available to make selections in the XR environment while performing different tasks in the real environment. Another example is the ability to warn users of critical work-related situations. As Burova et al. (2020) noted, previous professional experience influences the user's performance in specialized contexts. In their research, novice technicians required more interaction with warnings and needed situation-related warnings in order to catch their attention. Experienced technicians focused more on task performance and productivity, and considered extensive notifications irrelevant.

It is important to identify the characteristics that impact the user's safety, performance, and efficiency etc. In particular, it is crucial to pinpoint exactly what factors XR introduces to the context, in order to fully utilize the benefits of XR, and avoid the risks. However, target user groups, company culture, and the combination of equipment, technical devices, environments, tasks etc., form a unique composition of elements, or assemblage (see e.g., Delanda 2016). By this we acknowledge that use contexts are varied and dynamic, where every aspect within the context has some form of impact. There appears to be a lack of research concerning the actual test process and tools or models utilized. The need for a systematic experimental design has been recognized in previous research. Funk et al. (2015) studied AR-based assembly instructions and noted that variations in tasks and instruction modes rendered comparison difficult if not impossible between different use cases. Due to this dilemma, Funk et al. see the need for theoretical models and standardized means for designing XR-based instructions for assembly tasks. Their answer to this need is the General Assembly Task Model.

That model proposed “a benchmark experiment design consisting of two cheap and easily reproducible assembly tasks” (Funk et al. 2015, p. 253). The tasks were an assembly task with Duplo bricks and a drilling task on a wooden board. The aim was to compare different approaches to instructions with a uniform benchmark. Funk et al.’s research indicates a problem in separate, task-oriented XR studies, and that is the difficulty in comparing results. For this reason, studies that examine the effectiveness and technological performance for instance, possess limitations, particularly if there are any inconsistencies between tasks, contexts, and purpose.

There are many usability studies of different types of technical instructions (see e.g., Alexander 2013; Sapienza 2004). A plethora of studies also focus on different types of XR employed in delivering various modes of instruction (see e.g., Tzimas et al. 2019; Doshi et al. 2017; Zauner et al. 2003; Hoedt et al. 2017). However, when it comes to usability tests for XR-based technical instructions, the amount of research decreases. The usability tests concerning XR-based solutions tend to focus on the hardware and software enabling the virtual or augmented experience (see e.g., An et al. 2020), rather than on the technical instructions themselves. There are some exceptions, however, like Fussell et al. (2019). Fussell and colleagues adopted a mixed methods approach to study the learnability, satisfaction and effectiveness of a VR-enabled tutorial for aviation training. Their results revealed that while engaging and usable, the design was cognitively loading particularly for novices. There was promise however, in that those who made errors and found the tutorials mentally taxing, still managed to correct their mistakes without external guidance.

As these studies indicate, organized and comparable experimental design with suitable models and tools for XR-based technical instructions are needed, to support in finding ideal solutions in regards to specific assemblages (cases). The following article reports an experiment in which TIER has been systematically applied to study usability and affordances within XR-based technical instructions. In particular, it focuses on how the affordance, or qualities of the VR, encouraged participants to progress with their tasks despite a designed-in glitch within the instructions.

3. METHOD

The research goal of this paper was to identify how the design of XR-based technical instructions could be informed by user behavior, especially from the perspective of technological novelty. As previously noted, there are multiple variables to consider when both designing the technical instructions themselves, as well as when planning a UCD experiment. While it is difficult to account for all variables across diverse use contexts, it is important to utilize comprehensive models for even basic testing and experimentation of XR-based technical instructions. Users and use situations should always be at the center of iterative design and testing cycles in order to isolate the consistent factors that emerge across contexts. The aim of the study as to explore the effects when technical instruction users encounter and interact with novel XR-based instructions, while being expected to complete a task using those instructions.

3.1 Procedure

The first part of the study involved the formulation of XR-based technical instructions that would assist users (experiment participants) in assembling a LEGO construction. The rationale behind the choice of task hinges upon its simplicity – both ease in design of XR-based instruction, and simplicity that enables users to easily interpret and if necessary, ‘work around’ (modify instructions ad hoc if needed). This approach is in line with Funk et al. (2015) as well as Korn et al. (2013). Korn and colleagues’ work adopted a simplistic approach to ensure that the complexity of instructions would not be the confounding factor that hindered the effectiveness of the XR-enabled technical instructions. Their study was designed specifically with impaired production workers in mind, and utilized the Assistive Systems Experiment Designer (ASED) software to aid in the assembly process. The experimental set-up used in their examination comprised automatic logging of successfully completed assembly tasks.

Thus, the experiments in this current study examined the effectiveness of XR (or VR) as a medium for mediating technical instructions. The actual instructions played less of a role, despite the element of the built-in glitch. Focus was on to identify features when designing XR-enabled instructions, based on user responses to designed-in or accidental affordances. The researchers observed how participants acted upon or ignored functionalities offered. The instructions were developed especially for this experiment in order to avoid bias through familiarity (Casaló, Flavián & Guinalíu 2008). The chosen form of XR was VR due to its simplicity and relatively cost and time efficient set-up, while still offering a fully immersive experience. The instructions were recorded with Insta360 One X camera. The end result was a one-minute long 360-video that showed the participants how to construct the

LEGO shape, block-by-block. No speech nor virtual elements were added. This is similar to the approach adopted by Funk and colleagues (2015) and Korn and colleagues (2013).

In this study the affordance model of technical instructions in extended reality (TIER) model was applied as an analytical framework and technique for conducting the experiments (see Fig. 1). The TIER model has been piloted in (Rantakokko, submitted) in a similar test. Piloting results have shown that TIER is a suitable framework for this type of study, since it directs the focus toward the properties that should be delivered to users, and accounts for considerations of device and media types that are optimal for supporting effective delivery. Additionally, this is the first article in which the TIER model is systematically utilized in the testing phase of a design process for XR-based technical instructions. The focus is on the users, and in a normal design process of instructions, the feedback received from the testing phase would be instrumental for an iterative design process, indicating the types of changes the instructions may need before the next testing phase. This agile approach enables the detection of unintended affordances that could result in undesired outcomes.

The methodological approach is based on experimental research (see, i.e., Ross & Morrison 2004) aimed to identify whether the participants noticed the designed affordances, and how they acted upon or reacted to those affordances. The research consisted of two tests that were based on specific tasks that were mediated via the VR design. The data was collected via questionnaires, interviews, and observations of the experiment situations. The TIER model was used to both structure the experiment, as well as analyze the results. Application of TIER took place through dividing the properties of the XR-enabled instructions into two main domains: 1) affordances – what the technical instructions offer the user (how they support the user's actions) and what the technology itself offers the user, the central possible functionalities that must be fulfilled for technical instructions (Rantakokko & Nuopponen 2019 - accessing, finding, understanding, and relying on); and 2) data handling – considering the types of data and methods implicated in operating this type of design combined with the phases of the data handling process in XR (collection, processing, storage, transfer, presentation, and combining data). The affordances are further evident via four levels in the TIER model: rules, design, possibilities, and actualizing. (Rantakokko 2022, p. 28–29)

The five experimental steps correspond with the affordance-based analytical framework that involves: accessing – gaining access to the instructions (equipment available and functions, WIFI available etc.); finding – orienting oneself between the instructions and VR environment, as well as the real world with physical components; understanding and in turn, actualizing the assembly process; and relying on – coming to depend on the instructions that are mediated through the technology rather than utilizing intuition and memorized actions from previous experience (Rantakokko 2022). The understanding and actualizing components were operationalized in Step 2. Steps 3 and 4 sought subjective reporting from participants, and Step 5 was determinant on research observations.

3.2 Research design

The experiment involved the construction of a LEGO structure, as shown in Fig. 2. The task instructions were delivered via 360-degree VR-based video in which the assembly procedure was shown step-by-step. No linguistically-based instructions (i.e., written or verbal) were used to ensure fidelity between image and real world. The researcher introduced the task to the participants. They were then requested to independently complete the task without interacting with the researcher, and then to complete a questionnaire. The participants were interviewed after completing the questionnaire.

The instruction video showed 20 LEGO blocks. Five of these blocks were correct, and were used in the demonstration of how to assemble the construction in the right order. Before commencement, the blocks were arranged on a black surface. Again, black was chosen as the surface color due to its simple shade and high contrast. The initial arrangement of pieces can be seen in Fig. 3.

The participants were not told that the experiment would include a designed-in glitch. The piece circled in Fig. 4 was replaced with another piece among the LEGO pieces to be utilized for the actual experiment. This was to investigate how much the participants relied on the instructions themselves (as seen relation to the TIER affordances), and how much the possibilities of the technology allowed participants to improvise (combining constructional memory with initiative in light of available building blocks). Otherwise, the instructions were consistent with the real-world components. The purpose of this glitch was to observe how the users would act in a situation where the task was simple but flawed. The situation was intended to reflect work situations in which real world conditions do not match the details of the technical instructions. The idea was to observe the level of fixation with detail between the technical instructions and the participant's real-world resources and actions. On this note, Table 1 presents the participants, materials, and instruments involved in the experiments.

Table 1
Experimental variables in Test 1 and Test 2 – participants, materials, instruments

	Test 1	**Test 2**
Participants: **Little or no experience with XR**		
Number	4	4
Gender	3 male, 1 female	4 female
Age	27–72	23–39
Recruitment	Convenience sampling	Convenience sampling
Material:		
Guidance for test procedure	Verbal instructions, printed procedural instructions	Verbal instructions, printed procedural instructions
Equipment	VR Shinecon + Fairphone 3	Oculus Quest VR + Fairphone 3
Task instructions	VR-based video	VR-based video
Environment:	Home environment	Laboratory
Instruments for data gathering:		
	Questionnaire, interview, observation	Questionnaire, interview, observation

The main differences between these two experiments was the environment, and the equipment used. In test 1, the enviroment was cozy, home-like room, and the equipment a low-quality budget VR glasses. This was to build an amateurish and less formal impression, which may influence the participants' reliance on the instructions, but also make them more relaxed. The second experiment was conducted in a lab environment with standard-quality VR glasses. This could give more professional impression of the experiment.

3.2.1 Participants

The participants were recruited via convenience sampling. The research design received approval from the university's ethics committee. Overall, there were eight participants in the tests, ages 23 – 72, of which three were male and five were female. All of the participants possessed either no or little experience with XR technology. Those with a little experience had used the technology once or twice previously. The number of participants was limited because the research did not aim to evaluate designed instructions themselves. Rather, the goal here was to produce information on how the user's actions and reactions could be used to inform the design of XR-based technical instructions. That is, through focusing on aspects related to the TIER model in terms of accessing, finding, understanding and relying on, it was hoped that fruitful information would be derived on what properties enhanced which factor and why.

Convenience sampling was chosen due to the challenges that the COVID-19 pandemic conditions placed on recruitment and participation in empirical research. The study participants represented diverse backgrounds and ages. While small, the sample was intended to be pre-emptive of the variation among users in real work environments. A generic approach was taken to the type of instructions and use context applied in the experiment. Yet, in future research when specific technical instruction design is developed and tested, it will be important to recruit professionals and users from the specific areas in question.

3.2.2 Material and equipment

The material consisted of the experiment guidance, equipment, and instructions to complete the task. The guidance for the test procedure was both verbal and supplied in printed form. The instructions included practical guidance for the test, its purpose, and how the collected data would be used. Furthermore, the participants were informed that they are allowed to withdraw from the test at any time if they chose to do so. They were also informed that the test was not focused on them, but the instructions and how they supported the construction process. Therefore, if they would not succeed to complete the task according to the instructions, this would indicate usability issues regarding the instructions.

The equipment mediating the technical instructions in Test 1 comprised the VR Shinecon virtual glasses (see Fig. 4). These are low-quality budget glasses designed to be used with a smartphone placed inside. The smartphone used was a Fairphone 3. In Test 2, the instructions were mediated via Oculus Quest VR glasses. These are standard-quality VR glasses providing a wireless VR experience that include two controls. The quality of the Oculus Quest is significantly better than the VR Shinecon. The choice of equipment was two-fold. Firstly, these devices were available and used within the context of the research facilities, as well as during specific points of the pandemic. Secondly, the devices offered insight into the cross-platform effectiveness of the technical instructions in VR. In particular, it is noted that internationally the demographics of potential users will vary greatly, meaning that while one user group will have access to the likes of Oculus Quest, another group will be limited to devices similar to VR Shinecon. Therefore, the researchers were interested in understanding how the instructions operate in general across devices.

3.2.3 Data collection

A mixed method approach was adopted in this study. The data collection was conducted with questionnaires, observation, and interviews. Thus, data includes quantitative and qualitative data. While the participants were engaging with the task, the situation was video recorded and subsequently underwent observational analysis. After the participants had finished the task, they were asked to complete the questionnaire. The questions and the answers of the questionnaire are introduced and analyzed in Results section, and listed in Tables 3–6. Mainly, the questionnaire included questions with answers on a Likert scale. This was supplemented with open questions.

Upon completing the questionnaire each participant was interviewed. The interviews were exploratory and tailored based on what happened during the experiment as well as the answers they had given in the questionnaire. The participants were also encouraged to discuss anything they felt important.

3.3 Procedure

The participants were instructed both verbally and via a printed information sheet. The technical instructions in Test 1 played in a continuous loop. The participants were allowed to watch them as many times as they felt necessary before and during the building task. In Test 2, it was possible to pause and re-start the instructions at any time. In Test 1, the instructions were watched from a seated position next to the table with the LEGO blocks. In Test 2, the instructions were viewed in a designated observation area while standing. Outside the observation area, the instructions became translucent. In Test 1, the environment was an informal, simple household room. Test 2 was conducted in a lab environment. The central elements can be seen in Table 2.

Table 2
Central elements of the test procedure

Test 1	**Test 2**
VR-based instructions for the building task:	
Shown via continuous loop	Ability to pause and resume viewing
Test environment:	
Informal environment	Laboratory environment
Variables (Devices):	
VR Shinecon + Fairphone	Oculus Quest
Manipulation: Designed-in glitch	
Observing: Recording of the test situation + live observations	
Questionnaires: Pre-interview questionnaire (Questions represented in Tables 3–6 in Results section)	
Interviews: Based on observations of the test situations and answers to questionnaires	

The researcher undertook live observations in addition to recording the sessions via video (on a smartphone device). These videos were later transcribed. Participants were not to interact with the researcher while performing tasks. Despite this instruction, all participants did speak to the researcher. It was decided during the experiments that the tasks would still be completed despite this issue, as the information offered freely by participants during the tasks served as thinking aloud data. The post-questionnaire and post-experiment interview focused on what was observed during the experiment, as well as the answers to the questionnaire. One standard question that was asked to all participants regarded whether or not they had noticed the designed-in glitch.

3.4 Data analysis

The interviews and experiments were video recorded and later transcribed, with every action timestamped. The transcriptions were read through several times, and only relevant observations were retained in the data. Exclusion criteria included unrelated actions such as scratching one's nose. The chosen data was then combined in matrix tables that separately represented the data from each experiment and each test group. The data of both experiments was combined to create overall findings tables.

The tables were then analyzed in light of the TIER model (Rantakokko 2022). The TIER model was utilized according to its levels of affordances that represent different phases of the design process. Therefore, as this research focuses on the testing phase, relevant parts of the TIER model were chosen (see Fig. 5).

In this research, all the affordances of technical instructions were relevant. Due to the fact that the research focused on the users' actions and how the designed or unintended affordances actualized while users were engaging in the tasks, the focus was on the levels of possibilities and actualizing. The users evaluated the XR-based instructions and used them to complete the task, therefore, focus was on the presentation phase of XR data handling.

In order to operationalize the TIER model for the analysis, all observational notes were inserted into the matrix tables. The tables are presented in the results section. These notes were then categorized according to affordances and the levels that the notes were related to. For example, some notes focused on the affordance of accessing, particularly since the XR technology used in this study was VR. This meant that, participants could only see the instructions while wearing the head-mounted display (HMD). Yet, similarly, participants needed to remove the HMD to undertake the tasks in the real-world. If the user could not use the HMD or the technical instructions due to how they were mediated via the devices, the affordance of accessing would not be actualized.

The analysis considered the affordances noticed by the participants. The approach permitted an organized comparison of the results and what they indicated regarding the affordances featured in the tests.

In a design process of actual XR-based instructions, the TIER model could be used as follows. First, relevant laws, regulations and guidance should be taken into consideration, as well as other desired features of the XR-based technical instructions to be designed. These features would be placed into the matrix tables of the TIER model. When the first version of the instructions is ready to be tested, the results of each feature would be easy to evaluate, as it is readily seen from the table. If the designed affordances are noticed by users (level of possibilities) and if users act upon them (level of actualizing), these are entered into the TIER table. In the experiment design of the current research, the designed-in affordances pertained to the technical instructions and how they could be accessed, found and understood, in order to construct a LEGO form. The instructions included a designed-in glitch that rendered it impossible to build an identical construction. This in turned tested the error handling and relying on factors of interaction between users and XR-enabled technical instructions.

4. RESULTS

The research question of this paper was: In what ways can technology-based user behavior inform the design of XR-based technical instructions? To answer this question, data was gathered from two experiments focused on testing XR-based technical instructions. In these experiments users were studied in terms of how they acted upon the affordances they noticed during test situations. Attention was also placed on how this information could be applied to the design of XR-based technical instructions.

The experiments mainly focused on the level of actualizing. Participants tried to complete the task with the guidance of XR-based technical instructions and evaluated the instructions by following the guidance. Although only a few variables were designed into the tests, and the instructions were made as simple as possible, anomalies were still observed during the test processes. Anomalies are characterized based on if they were noticed (level of possibilities) by the participants, and what, if any, actions followed (level of actualizing). Combined notes from the tests and the observed anomalies are introduced in Tables 3–6. First, Table 3, presents the combined key notes concerning the affordance of *accessing*.

4.1 The affordance of accessing

Table 3
Key notes from the experiments related to the affordance of accessing

Accessing	
Questionnaire responses	
Test 1	**Test 2**
Q: *How did you experience using the VR glasses?*	
Extremely easy (1), relatively easy (2), neither easy nor difficult (1)	Relatively easy (4)
Q: *Did you experience problems using the VR glasses during the test?*	
Sharpness (2), loose glasses (1), no problems (1)	Wi-Fi problem (1), problems adapting to VR glasses (1), no problems (2)
Q: *Did problems occur during the test that you would have needed assistance with?*	
No problems (2), sharpness (1), uncertainty around block replacement (1)	No problems (3), Wi-Fi (1)
Anomalies observed	
Level of possiblities	**Level of actualizing**
Difficulties in fitting the VR glasses	
Noticed	Succeeding to fit, or holding them with one hand
Difficulties to see	
Noticed	Finishes with insecurities
Glasses opened at one point while put on the table	
Unclear if noticed	Was not possible to continue before fixing
Smartphone heated up badly during the tests	
Noticed by the observant	Did not complicate the tests, possible problem in long-time work situations
Online connection failed	
Noticed	Was not possible to start before connection restored

Equipment-related notes are marked under the affordance of accessing. This is because XR-based technical instructions cannot be accessed without the equipment. Using XR (in this case, VR) glasses was evaluated positively in both tests despite the differences in devices. In Test 1, the answers varied from *neither easy nor difficult* to *extremely easy*, while in Test 2, everyone answered *relatively easy*. However, when participants were asked specifically about possible problems they may have experienced during the test with XR equipment, they listed several issues.

The questionnaire included a question about the problems faced during the test concerning XR equipment in general, and another on problems that participants would have needed help with during the test. Despite being instructed not to interact with the researcher, all participants did either speak or ask for assistance during the procedure. Thus, the second question was automatically answered. Yes, there were aspects that participants needed help with. In order to have discouraged interaction, the researcher should have been located further away from the test area. This was however, not possible. The decision to allow communication during the experiments enabled continuity of the tests that otherwise may have been disrupted. More tests with users working on their own would be required if developing actual XR-based technical instructions in order to see how the users would operate in situations of independent work.

Problems reported by participants in Test 1 included poor-fitting glasses (HMD) and poor sharpness. Assistance would have been needed with sharpness issues. Uncertainty of where in the construction certain pieces should be placed was mentioned as a problem that would have benefited with assistance. Problems reported in Test 2 included Wi-Fi problems, taking time to realize how the system works (i.e., warm-up tests), and adjusting to the conditions of the HMD in order to identify objects clearly. One participant required assistance with the Wi-Fi problem. Half of the participants in Test 1 and three out of four in Test 2 mentioned that there was no need for assistance. There were other issues that would have needed assistance. Therefore, it would be important to first assist the users and later on instruct with the issues that may occur more often. More tests would show which issues keep repeating and which ones seldom occur.

The Wi-Fi failure is worth highlighting. Concerning this experiment, it would have been possible to have only watched the instructions via regular video. However, leaving the experiment at the mere watching of a video would not have probed the reality of interacting with technical instructions in VR. In the work environment, failure of the Wi-Fi connection could mean a disruption to the work process if there is no back-up (i.e., offline mode and functionalities). If the work task is a response to a critical situation at a distant location, Wi-Fi failure can be catastrophic. Therefore, offline-mode and hard copy backups should be required to guarantee access to crucial instructions at all times.

Interestingly, the participants reported that they found using the XR equipment relatively easy, even though most of them faced at least some problems during the experiment, mostly with the equipment. In Test 1, trying to adjust sharpness was problematic and not usually resolved fully. In addition, fitting the headset was challenging. The fitting problems were usually resolved by adjusting the glasses, but sometimes the participants had to hold the glasses with one hand. In Test 2, the participants had more control and could start and pause the instructions via the equipment's user interface. However, the basic usage of the equipment was relatively convenient and effective, and most people got the headwear to fit correctly. Nobody had issues with understanding the basic functions. As mentioned before, observing the testing of different equipment was useful as participants did not actively mention the problems themselves. With a new technology, it may not be easy to notice problems until having more experience. For this reason, giving participants the opportunity to test different devices may aid in deepening understanding of what device is more effective for what purpose. This adds an extra yet necessary layer to testing XR-based technical instructions, as the experience does not purely hinge upon one technical factor, rather several (see studies such as, Bowman et al. 2002; Kaewkannate & Kim 2016).

While there may be situations in which professionals cannot influence the device used by their employer, there may be opportunities in some cases to participate in the organization's purchasing decisions. This in itself is an important part of user experience, employee engagement, organizational citizenship, and sustainability adoption of information technology (see, Nikas & Poulymenakou 2008). Thus, via experience professionals strengthen knowledge of optimal devices for particular technical instruction modes and purposes.

Anomalies not noticed or mentioned by the participants, were the glasses opening at one point during Test 1 with one participant, and the smart phone inside the headset heating substantially. While the glasses opened, they did not break, so the test could continue after the observer closed them. However, if this happened in a critical work situation the results could be detrimental. It remains uncertain as to whether the participant noticed the situation, because it seemed that the participant's focus was on the task at hand, rather than trying to fix the problem. The heating of the smart phone seemed to not be noticed by participants. This was due to the fact that they did not have to touch the smart phone inside the glasses, and it did not affect the tests. However, in general, the overheating issue would probably prohibit use of that technological solution daily for long periods. In the design process of real XR-based technical instructions, these would be some of the issues that would need to be considered. This has often been noted in research focusing on the design of healthcare systems (Bitkina, Kim & Park 2020).

4.2 The affordance of finding

Notes on the affordance of *finding* are introduced in Table 4. Completion times mostly relate to how well the participants could find the right information. The instructions were in 360-degree video form that played on continuous loop in Test 1. Here, participants could not pause it at any time. In Test 2, when the participants needed to re-visit information, they either located the spot in the video where it was shown by pausing and scrolling back, or let the video roll until the right spot was reached again. These results were not compared with other media. It would be surprising however, if this solution would make completion times shorter than other types of instructions. If the participants were fast, the instruction video was still one minute long in both experiments, and therefore all the participants had to use at least this amount of time with the instructions before or during the task. If they were insecure, they checked the instructions multiple times. Had the participants used instructions delivered via any other medium (i.e., AR, paper instructions or even television – no requiring full eye coverage), they would have had the option to start the task while watching the instructions.

Three participants developed a workaround (a way of continuing work flow despite usability and technical flaws, see Ejnefjäll & Ågerfalk 2019) for this medium as well.

Table 4
Observation findings from Tests 1 and 2 – affordance of finding

Finding							
Questionnaire responses							
Test 1				**Test 2**			
Task completion time (from start to completion)							
0:24	2:38	7:22	4:27	2:28	2:51	1:58	5:49
First checks							
1	8–9	1	1	1	1	1	1
Re-checks							
1	1	Ongoing	Ongoing	Ongoing	1	2	3
Q: *How many times did you need to watch the instructions?*							
2, 3, 4, 3				5, 2, 3 times + an extra time to make sure everything is correct, 1 time completely plus two separate sections			
Anomalies observed							
Level of possiblities				**Level of actualizing**			
VR does not allow users to see actual reality							
Noticed				Having to put glasses on and off during the task			
Three-page form misunderstood as a one-page form							
Not noticed				Participants were informed about the page account, so they filled in all the pages			
Forgetting (or fear of forgetting) instructions read or seen							
Noticed / not noticed				Re-checks for confirmation			
Slowness of watching the VR video (no rewind)							
Noticed				Completed at a slower pace			
Requires getting used to initially (Participant 1.2)							
Noticed				Insecurities in the beginning			
Direction of the video had a glitch and kept reversing (Test 2)							
Noticed				Necessity to continuously turn			
Having to move while wearing the glasses can cause dangerous situations							
Not noticed				Kept moving, did not react to warning			
Focusing only on one direction							
Not noticed				If the relevant information was situated in more than one spot in the video, it might not be noticed			

Anomalies concerning the affordance of finding are detailed here. VR is completely immersive and removes users from ocular (sight) contact with the real world. This is already a known property regarding the system. However, in the context of technical instructions this poses an extra challenge to successfully using the system to carry out actions related to the instructions. In a task that requires assembling a physical structure, complete immersion of VR is an issue unless the use purpose is for remote assembling, whereby a robot or other augmented technology undertakes the physical actions. Ordinarily, removal from the immediate physical environment and negotiation between the two spaces (virtual and real) increases cognitive workload (Andersen et al. 2016; Kosch et al. 2023). This was exacerbated since the participants were either forced to remember the instructions, or to re-check them.

Fear of forgetting created further delays. Since the instructions were based on the video structure, watching the instructions was a slow process, especially in Test 1, where participants could not control the video. One participant in Test 1 wished for a rewind function. In Test 2, participants were able to control the video with pause and replay etc. Even though this issue was expected, this result highlights the importance of testing.

For one participant in Test 2, the VR system had a glitch and the direction of the video kept reversing 180 degrees. Therefore, the participant needed to continuously move around in order to be able to watch the instructions. Sudden movements posed a risk because the participant nearly bumped into objects in the real-world whilst not being able to see the objects. Furthermore, the participant did not hear when warned not to move so much. Physical safety during the use of equipment is crucial at all times, as witnessed in similar XR cases such as Pokémon GO (Serino et al. 2016). The environment needs to be designed so that there are no immediate physical dangers. In work environments, VR-based solutions that completely block the view of physical objects should be positioned in controlled spaces.

Only a few of the problems were mentioned by the participants. Of these problems, visual difficulties were the most mentioned in Test 1. The issue was caused by the poor quality of the VR HMD used. This HMD did not accommodate for participant eye glasses. In this category, there was the factor of it sometimes being difficult to see from the instruction video where exactly the blocks were situated in the structure. These visibility issues did not cause any of the participants to exit the task. Instead, they created workarounds by guessing, reasoning, and trying to mathematically estimate the locations and appropriate pieces to be used. Difficulties to see also resulted in longer task times and some degree of frustration.

It was noted that none of the participants using VR glasses took time to look around the virtual environment. They all focused directly on the target area. Had there been important elements in more than one location, participants may not have noticed them. If total environments are used in the instructions, it could be relevant to inform the users that they should look around. Furthermore, other multisensory elements such as sound and vibration etc., could be used to enhance the visual elements and indicate locations of information.

4.3 The affordance of understanding

The affordance of *understanding* was reported to be *very easy* by six of the participants. Two reported understanding to be *relatively easy*. Regarding possible flaws, seven of eight participants found the instructions to be functional. However, the instructions were indeed not functional because the designed-in glitch made it impossible to build the exact same structure. Despite this, each participant answered that they successfully completed the task with the help of the instructions. Most of the participants built almost the same form, just replacing the irregular piece with a similar shaped piece of a different color. The key notes regarding affordance of understanding are presented in Table 5.

Table 5
Observations from Tests 1 and 2 related to the affordance of understanding

Understanding	
Questionnaire responses	
Test 1	**Test 2**
Q: *Did you find the instructions understandable?*	
Very easy to understand (2), relatively easy to understand (2)	Very easy to understand (4)
Q: *In general, did you find the instructions functional, or were there flaws?*	
Functional (4) with one comment about rewinding functionality	Functional (3), Flawed with a mention of the designed-in glitch (1)
Q: *Did you succeed in building the LEGO construction by following the instructions?*	
Yes (4)	Yes (4)
Anomalies observed	
Level of possiblities	**Level of actualizing**
The designed-in glitch	
Noticed by some	Finished with the wrong colored block
Voice narration would have been beneficial	
Noticed	Finished without voice narration
Characters in the video too big (Test 2)	
Noticed	Having to look up and down

Anomalies concerning the affordance of understanding consisted of the designed-in glitch and unintended scaling issues of the instructions and real-world conditions. One participant mentioned that audible instructions would have been beneficial.

Participants tended not to mention the glitch. When asked later, most of them had noticed the glitch but did not find it relevant or worth mentioning. This finding might signal an issue in real work situations. It does however, at least indicate that sometimes people reason that minor flaws are unimportant and can be worked around (see e.g., Fuller & Arnold 2019). Thus, when flaws could be ignored by slight modification, participants did not explicitly verbally articulate the flaws. Only one participant mentioned the glitch and was disturbed by it, but even that participant finished the task. Following the work flow during the test periods would be helpful. If it cannot be trusted that the test participants either notice or voice flaws, then observations are instrumental in detecting subconscious behavior (Awan, Esteve & van Witteloostujn 2020). Following eye gaze paths could indicate whether or not they have noticed flaws they assume are not important.

All the participants justified how the task was possible to finish regardless of the glitch. Participant 1 from Test 1 ignored the glitch because, *"it was LEGO, and the color didn't matter."* Participant 4 of Test 1 explained:

"The shape was correct. In such a situation, you look at the blocks on the table and when the video ends here, then you know that the wrongly colored block can be used. But if the video continues, there is no way to know if that block will be needed later. So, then you can say that now you cannot continue since there is not a block of that color and shape."

In the opinion of this participant, *"the general idea is that if you do something, you use it if it fits."* None of the participants left the task unfinished on account of the glitch.

In Test 2, the instruction video featured an unintended surprise that was not present in Test 1. Using Oculus Quest headsets made the images and contents of the virtual environment extremely large, which forced participants to continuous look up and down. This can be likened to watching a giant building with huge LEGO pieces. The size difference between the pieces in the VR instructions and the real-world pieces on the table was disturbing for some of the participants. Elements such as voice instructions would have been beneficial, especially for the participants who struggled to see well. Choosing the right equipment and adjusting the design of the instructions for that equipment is pertinent. Changing the equipment at a later stage may result in distortion of visual outputs.

4.4 The affordance of relying on

With the affordance of *relying on*, the differences in Tests 1 and 2 were remarkable. In Test 1, all four participants would have chosen VR-based instructions over paper instructions. However, in Test 2, only one participant preferred VR-based instructions over a paper version (see Table 5). In a more complex task, two of the participants in Test 1 would have preferred VR-based instructions over other forms. In Test 2, all participants would have preferred other forms. The reliance experienced was high in general, especially in Test 1, where three out of four participants found the XR-based instructions to be *very reliable* and one described them as *relatively reliable*. In Test 2, one informant found the instructions to be *very reliable*, and three reported them to be *relatively reliable*. Observations are reported in Table 6.

Table 6
Observations of Tests 1 and 2 related to the affordance of relying on

Relying on	
Questionnaire responses	
Test 1	**Test 2**
Q: *Would you rather use paper instructions (possibly with pictures) or VR instructions like these?*	
Rather VR instructions (4)	Rather paper (3), rather VR (1)
Q: *Would you rather use Youtube instruction videos, or VR instructions like these?*	
Rather VR (3), Blank (1)	Rather video (3), rather VR (1)
Q: *In the case of a more complex task, e.g., LEGO construction of 50 pieces, what type of instructions would you rather use?*	
VR (2) (1 stated that it would be better to see through the HMD, thus preferring AR even though not naming it), Video (1), Video/VR (1)	Paper (2), video (1), paper / video (1)
Q: *Did you find the instructions reliable? (referring to instructional content and VR HMD)*	
Very (3), Relatively (1)	Very (1), Relatively (3)
Anomalies observed	
Level of possiblities	**Level of actualizing**
Insecurity about instructions and how to act	
Noticed	Sometimes hesitating to ask when needing help, making decisions to act or attempting to respond in a way they feel the researcher is wanting.
Not relying on the technology (Participant 4)	
Noticed	Based on previous experiences, reluctance to engage in regular use.

In the experiments, the answers concerning the reliance on XR technology are not completely related to participants' preconceptions. Instead, the experiences during the experiments may have contributed either positively or negatively to how reliable participants believed XR equipment in general was as a medium for technical instructions.

Anomalies that arose regarding the affordance of relying on were insecurity about the instructions and how to act, and not relying on the technology in general. The insecurities seemed to slow down the participants' performance, sometimes remarkably, while they gathered the courage to ask for further information from the researcher. One participant had negative experiences in the past that related to the causing of headaches and eye fatigue. This influenced the participant's opinion that the technology was not reliable.

The participant in question was willing to undertake the test, but expressed reluctance to use XR-enabled instructions in regular bases, such as in work or studies. This was despite the fact that the symptoms did not reoccur during this study. Allowing the participants to rehearse usage of the system would have reduced insecurities[4]. When new technological solutions become familiar, it is easier to give feedback and see which features would be important to improve, and which features are well-functioning as they are[5]. Therefore, the testing phase should include longer use periods of the equipment, and several occasions to use. Feedback should be iteratively collected.

If this would be a design process of actual XR-based technical instructions, there would be the need to engage in another iteration sprint. After possible improvements, the next round of tests would be conducted, reported, and compared to see if there is improvement. This article reports one way to complete the process via the assistance of the TIER model. It shows how the actions of the users in response to new technology informs the design of XR-based technical instructions via feedback.

[4] This is commonly used in driving simulation studies (see e.g., Godley, Triggs & Fildes 2002).

[5] This has been noted in relation to concept familiarity for instance (see e.g., Grooms, Sampson & Enderle 2018).

5. CONCLUSIONS

This research aimed to identify how the design of XR-based technical instructions is informed when the actions of users in relation to new technology are considered. The answer was sought via two empirical tests, and the analysis of the results was structured according to the TIER model. The tests were conducted using simple VR-based instructions. The participants' task was to complete the construction of a LEGO structure by following the instructions. The data was gathered via questionnaires, interviews, and observations. The TIER model has been designed to support and analyze the organized, step-by-step design process of XR-based technical instructions. It begins with laws and regulations, continued by designing the desired affordances, and moves towards testing and re-designing or re-iterating the technical instruction design.

Since XR changes the design process and the use situations of technical instructions that are delivered via the medium, it is important to conduct a thorough and iterative test process of the instructions, while involving users in the process. One way to do this is by testing users' interaction with the system (Abras et al. 2014, p. 9), as is done in this research. The focus of the current paper is on an organized and comparable test process of XR-based technical instructions.

The test results indicated that even when using simple XR-based instructions, several unexpected events can occur. Most of the problems or inconveniences were found via observations and discussions based on the observations. Only a few problems were mentioned in the questionnaire responses. Furthermore, the participants often did not mention the flaws they noticed, if they were not directly asked about the specific flaw in question. In these tests, a glitch was deliberately designed into the instructions. The glitch entailed that one LEGO piece present in the instructions was replaced with a similar piece of a different color. Most of the participants noticed the glitch, but did not mention it. They simply replaced the wrongly colored piece with another of another color. All the participants reported they had completed the task.

These findings highlight the importance of observation rather than basing the test process merely on the users' evaluations (see, e.g., Gram 2010). This is because the evaluations can be flawed in several ways, such as not noticing the problems or deciding that the problems noticed are not relevant. There may be several reasons for participants not reporting the problems that occurred during the tests, or the designed-in glitch even if they noticed it. One possible reason seemed to be a willingness to disregard issues to be polite to the researcher. That could be an issue in work environments. Participants also seemed to be willing to do as they perceived that they were supposed to. Even though their capabilities were not the subject of the research, people often tend to try to succeed in a test situation (see, e.g., Nederhof 1985; Ming et al. 2021). The inclination could be socially accepted by the researcher, is classified as "the tendency of some respondents to report an answer in a way they deem to be more socially acceptable than would be their "true" answer. They do this to project a favorable image of themselves and to avoid receiving negative evaluations." (Callegaro 2008, p. 826)

Reluctance to mention or focus on problems may originate also from the new experience in general. Norman (2002, p. 41) noted:

Any pleasure derivable from the appearance or functioning of the tool increases positive affect, broadening the creativity and increasing the tolerance for minor difficulties and blockages. Minor problems in the design are overlooked. The changes in processing style released by positive affect aid in creative problem solving that is apt to overcome both difficulties encountered in the activity and those created by the interface design. In other words, when we feel good, we overlook design faults.

Because all the participants were rather unexperienced with XR, there may have been a tendency to overlook minor problems while enjoying the new experience. This needs to be taken into consideration in real tests in the work environment as well. Further testing is needed for longer periods to ensure that this bias does not affect evaluations. This phenomenon is also related to the halo effect, in which emotions and attitudes towards phenomena may bias judgements (Minge & Thüring 2018, p. 23).

There are some limitations to this research. First, the instructions were perhaps too simple, and did not include interactive elements. Therefore, it is debatable whether they can be counted as XR-based instructions. This is because the aim was to keep them as simple as possible and focus on the affordances that the users noticed in order to illustrate how the TIER model functions within user testing of XR-based technical instructions. Having more variables in the instructions might have compromised the understanding of the results of this particular research design.

Second, pondering over the reliability of the experiment design in light of the research question there may be debate. While only using LEGO pieces in this research, many people have prior experience in LEGO construction. This may have influenced their attitudes towards color accuracy. Therefore, it could be understood that using LEGO is possibly not so effective in the study of technical instructions. However, this solution has also been used before (see e.g., Funk et al. 2015; Korn et al. 2013).

Third, there was no clear inclusion or exclusion criteria of participants. Thus, the characteristics of participants were not decided or classified prior to the experiments. This led to factor-related uncertainty regarding the participants within the groups of Test 1 and 2. For instance, numerous variables may have influenced the participants' experience of the designs. These include, educational background, experience with technology, age profiles, test environments, and even the types of equipment utilized.

Fourth, control methods of usability testing were not used in this research in an organized matter. This is because in the current research, the purpose was not to develop functional technical instructions, the design process did not continue with an iterative process as it normally would. It was already anticipated however, that feedback would be given by participants regarding its unsuitability for the use context in question. This pertains to the fact that users of assembly instructions often interact with the real physical environment at the same time as following instructions. Thus, in an actual use environment, AR would be the ideal XR method. It was assumed in this study that if participants would not verbally mention this aspect, then potentially other crucial information would be missing from their interviews.

Fifth, in this research it was not possible to conduct a comparison of different types of XR-based solutions. As Fast-Berglund et al. (2018, p. 37) state, different types of XR are effective in different manufacturing phases and various solutions even with the same type of XR may lead to very different outcomes. Therefore, a preliminary comparison before choosing a certain type of XR would be crucial to conduct. Conducting preliminary comparisons could place strain on resources (time and financial).

Therefore, future research should concentrate on investigating alternative means and methods for testing and comparing XR-based technical solutions systematically. Moreover, it would be important to test the TIER model in an actual design process of XR-based instructions. Other research directions may include the comparison of XR-based technical instructions and their usability via methods that enable the analysis and comparison, such as seen in the TIER model. This would offer more information in terms of what solutions may be most relevant and successful for intended purposes. This research is an attempt to organize the test process in a way that would be useful in actual design processes of real XR-based technical instructions.

Declarations

Competing interests: Authors disclose having no financial or non-financial interests directly or indirectly related to the work submitted for publication.

Research involving human participants: The research design received approval from the university of Vaasa's ethics committee

Informed consent: All the participants of these experiments were informed that participating is voluntary, and they have a right to withdraw from participation at any time they would choose to.

Data availability: The datasets generated during and/or analyzed during the current study are not publicly available due the protection of the participants' privacy and security. All the data available is published in this article.

Funding: Academy of Finland; Finnish Cultural Foundation

Author contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Satu Rantakokko. The first draft of the manuscript was written by Satu Rantakokko and both the authors participated in background information gathering, and commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

1. Abras, C., Maloney-Krichmar, D. & Preece, J. (2014). User-Centered Design. In Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications. (in press) https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.381&rep=rep1&type=pdf
2. Alexander, K. P. (2013). The usability of print and online video instructions. Technical Communication Quarterly. Vol. 22 (3), pp. 237–259. https://doi.org/10.1080/10572252.2013.775628
3. An, J., Poly, L-P. & Holme, T. A. (2020). Usability testing and the development of an augmented reality application for laboratory learning. Journal of Chemical Education. Vol 97 (1), pp. 97–105. https://doi.org/10.1021/acs.jchemed.9b00453
4. Andersen, S. A. W., Mikkelsen, P. T., Konge, L., Cayé-Thomasen, P., & Sørensen, M. S. (2016). The effect of implementing cognitive load theory-based design principles in virtual reality simulation training of surgical skills: a randomized controlled trial. Advances in Simulation, 1(1), 1–8. https://advancesinsimulation.biomedcentral.com/articles/10.1186/s41077-016-0022-1
5. Ara, J., Karim, F.B., Alsubaie, M.S.A., Bhuiyan, Y.A., Bhuiyan, M.I., Bhyan, S.B. & Bhuiyan, H. (2021). Comprehensive Analysis of Augmented Reality Technology in Modern Healthcare System. International Journal of Advanced Computer Science and Applications(IJACSA), 12(6), 2021. http://dx.doi.org/10.14569/IJACSA.2021.0120698
6. Awan, S., Esteve, M., & van Witteloostuijn, A. (2020). Talking the talk, but not walking the walk: A comparison of self-reported and observed prosocial behaviour. Public Administration, 98(4), 995–1010. https://doi.org/10.1111/padm.12664
7. Bao, T. & Hurriyet, O. (2021). Secure Augmented Reality (AR) for Telehealth and Emergency Medical Services (EMS): A Survey. https://cspri.seas.gwu.edu/sites/g/files/zaxdzs1446/f/downloads/Secure%20AR%20Literature%20Survey_20210722.pdf
8. Belke, B., Leder, H., Strobach, T., & Carbon, C. C. (2010). Cognitive fluency: high-level processing dynamics in art appreciation. Psychology of Aesthetics, Creativity, and the Arts, 4(4), 214. https://psycnet.apa.org/doi/10.1037/a0019648
9. Bitkina, O. V., Kim, H. K., & Park, J. (2020). Usability and user experience of medical devices: An overview of the current state, analysis methodologies, and future challenges. International Journal of Industrial Ergonomics, 76, 102932. https://doi.org/10.1016/j.ergon.2020.102932
10. Bowman, D. A., Datey, A., Ryu, Y. S., Farooq, U., & Vasnaik, O. (2002). Empirical comparison of human behavior and performance with different display devices for virtual environments. Proceedings of the human factors and ergonomics society annual meeting (Vol. 46, No. 26, pp. 2134–2138). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/154193120204602607
11. Brambilla Hall, S. & Baier-Lentz, M. (2022). 3 technologies that will shape the future of the metaverse – and the human experience https://www.weforum.org/agenda/2022/02/future-of-the-metaverse-vr-ar-and-brain-computer/
12. Burova, A., Mäkelä, J., Hakulinen, J. Keskinen, T. Heinonen, H. Siltanen, S. & Turunen, M. (2020). Utilizing VR and gaze tracking to develop AR solutions for industrial maintenance. CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing System, Honolulu HI, April 2020, pp. 1–13, https://doi.org/10.1145/3313831.3376405
13. Callegaro, M. (2008). Social Desirability. Encyclopedia of Survey Research Methods. https://methods.sagepub.com/reference/encyclopedia-of-survey-research-methods/n537.xml
14. Casaló, L., Flavián, C., & Guinalíu, M. (2008). The role of perceived usability, reputation, satisfaction and consumer familiarity on the website loyalty formation process. Computers in Human behavior, 24(2), 325–345. https://doi.org/10.1016/j.chb.2007.01.017
15. DeLanda, M. (2016). Assemblage theory. Edinburgh University Press.

16. Ding, Y. Li, Y. & Cheng, L. (2020). Application of Internet of Things and Virtual Reality Technology in College Physical Education. IEEE Access, vol. 8, pp. 96065–96074, 2020, https://doi.org/10.1109/ACCESS.2020.2992283.
17. Doolani, S., Wessels, C., Kanal, V., Sevastopoulos, C., Jaiswal, A., Nambiappan, H., & Makedon, F. (2020). A review of extended reality (xr) technologies for manufacturing training. Technologies, 8(4), 77. https://doi.org/10.3390/technologies8040077
18. Dopp, A. R., Parisi, K. E., Munson, S. A., & Lyon, A. R. (2019). A glossary of user-centered design strategies for implementation experts. Translational behavioral medicine, 9(6), 1057–1064. https://doi.org/10.1093/tbm/iby119
19. Doshi, A., Smith, R. T., Thomas, B. H. & Bouras, C. (2017). Use of projector based augmented reality to improve manual spot-welding precision and accuracy for automotive manufacturing. The International Journal of Advanced Manufacturing Technology, vol. 89, no. 5–8, pp. 1279–1293. https://doi.org/10.1007/s00170-016-9164-5.
20. Ejnefjäll, T., & Ågerfalk, P. J. (2019). Conceptualizing workarounds: Meanings and manifestations in information systems research. Communications of the Association for Information Systems, 45(1), 20. https://aisel.aisnet.org/cais/vol45/iss1/20/
21. Etambakonga, C. L. (2021). The Rise of Virtual Reality in Online Courses: Ethical Issues and Policy Recommendations. IntechOpen. https://doi.org/10.5772/intechopen.97516
22. Fast-Berglund, Åsa, Gong, Liang and Li, Dan (2018) Testing and validating Extended Reality (xR) technologies in manufacturing. Procedia Manufacturing, Vol 25, 2018, pp. 31–38. https://doi.org/10.1016/j.promfg.2018.06.054
23. Fiorentino, M., Uva, A. E., Gattullo, M., Debernardis, S., & Monno, G. (2014). Augmented reality on large screen for interactive maintenance instructions. Computers in Industry, Vol 65, No 2, pp. 270–278. https://doi.org/10.1016/j.compind.2013.11.004
24. Fuller, H. J., & Arnold, T. (2019, September). Identifying Medical Equipment Usability Issues from Social Media Reports. Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care (Vol. 8, No. 1, pp. 217–221). Sage CA: Los Angeles, CA: SAGE Publications. https://journals.sagepub.com/doi/pdf/10.1177/2327857919081055
25. Funk, M., Kosch, T., Greenwald, S.W. & Schmidt, A. (2015). A benchmark for interactive augmented reality instructions for assembly tasks. Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia (MUM '15). Association for Computing Machinery, New York, NY, USA, 253–257. https://doi.org/10.1145/2836041.2836067
26. Fussell, S. G., Derby, J. L., Smith, J. K., Shelstad, W. J., Benedict, J. D., Chaparro, B. S., Thomas, R. & Dattell, A. R. (2019). Usability testing of a virtual reality tutorial. Proceedings of the Human Factors and Ergonomics Society 2019 Annual Meeting. Vol. 63, No. 1, pp. 2303–2307. https://journals.sagepub.com/doi/pdf/10.1177/1071181319631494
27. Gattullo, M., Scurati, G. W., Fiorentino, M., Uva, A. E., Ferrise, F., & Bordegoni, M. (2019). Towards augmented reality manuals for industry 4.0: A methodology. Robotics and Computer-Integrated Manufacturing, 56, 276–286. https://doi.org/10.1016/j.rcim.2018.10.001
28. Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. Accident analysis & prevention, 34(5), 589–600. https://doi.org/10.1016/S0001-4575(01)00056-2
29. Gram, M. (2010). Self-reporting vs. observation: some cautionary examples from parent/child food shopping behaviour. International Journal of Consumer Studies, 34(4), 394–399. https://doi.org/10.1111/j.1470-6431.2010.00879.x
30. Grooms, J., Sampson, V., & Enderle, P. (2018). How concept familiarity and experience with scientific argumentation are related to the way groups participate in an episode of argumentation. Journal of Research in Science Teaching, 55(9), 1264–1286. https://doi.org/10.1002/tea.21451
31. Hoedt, S., Clayes, A., Van Landeghem, H. & Cottyn, J. (2017). The evaluation of an elementary virtual training system for manual assembly. International Journal of Production Research, vol. 55, no. 24, pp. 7496–7508. https://doi.org/10.1080/00207543.2017.1374572
32. Islam, Z. (2019). Constructivist Digital Design Studio with Extended Reality for Effective Design Pedagogy. Design and Technology Education: an International Journal, 24(3), 52–76. https://ariadneproduction.lboro.ac.uk/DATE/article/view/2651
33. Kaewkannate, K., & Kim, S. (2016). A comparison of wearable fitness devices. BMC public health, 16, 1–16. https://doi.org/10.1186/s12889-016-3059-0
34. Karat, J. (1997). Evolving the scope of user-centered design. Communications of the ACM, 40(7), 33–38. https://dl.acm.org/doi/pdf/10.1145/256175.256181
35. Korn, O., Schmidt, A., & Hörz. T. (2013). Augmented manufacturing: a study with impaired persons on assistive systems using in-situ projection. Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments

(PETRA '13). Association for Computing Machinery, New York, NY, USA, Article 21, 1–8. https://doi.org/10.1145/2504335.2504356

36. Kosch, T., Karolus, J., Zagermann, J., Reiterer, H., Schmidt, A., & Woźniak, P. W. (2023). A Survey on Measuring Cognitive Workload in Human-Computer Interaction. ACM Comput. Surv. https://doi.org/10.1145/3582272
37. LaRoche, C. S., & Traynor, B. (2010, July). User-centered design (UCD) and technical communication: The inevitable marriage. 2010 IEEE International Professional Comunication Conference (pp. 113–116). IEEE. https://doi.org/10.1109/IPCC.2010.5529821.
38. Marrazzo, M. (2018). People don’t read manuals. https://www.researchgate.net/publication/324106690_People_don't_read_manuals/link/5abdf19145851584fa6fee94/download
39. Milgram, P. & Kishino, F (1994). A Taxonomy of Mixed Reality Visual Displays. IEICE Transactions on Information and Systems, vol. E77-D, no. 12, pp. 1321–1329, December 1994. https://search.ieice.org/bin/summary.php?id=e77-d_12_1321
40. Ming, J., Heung, S., Azenkot, S., & Vashistha, A. (2021). Accept or address? Researchers’ perspectives on response bias in accessibility research. Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21). Association for Computing Machinery, New York, NY, USA, Article 20, 1–13. https://doi.org/10.1145/3441852.3471216
41. Minge, M., & Thüring, M. (2018). Hedonic and pragmatic halo effects at early stages of user experience. International Journal of Human-Computer Studies, 109, 13–25. https://doi.org/10.1016/j.ijhcs.2017.07.007
42. Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. European journal of social psychology, 15(3), 263–280. https://doi.org/10.1002/ejsp.2420150303
43. Nikas, A. & Poulymenakou, A. (2008). Technology Adaptation: Capturing the Appropriation Dynamics of Web-Based Collaboration Support in a Project Team. International Journal of e-Collaboration (IJeC), 4(2), 1–28. http://doi.org/10.4018/jec.2008040101
44. Norman, D. (2002). Emotion & design: attractive things work better. Interactions 9, 4 (July 2002), 36–42. https://doi.org/10.1145/543434.543435
45. Rantakokko, S. (2022). Creating a Model for Developing and Evaluating Technical Instructions that use Extended Reality. Technical Communication, 69(3), 24–39. https://doi.org/10.55177/tc734125
46. Rantakokko, S. (Submitted).
47. Rantakokko, S. & Nuopponen, A. (2019). Laajennetun todellisuuden tarjoumat tekniselle viestinnälle – kohti teoreettista mallia. VAKKI Publications, 10, 53–66. http://www.vakki.net/publications/2019/VAKKI2019_Rantakokko&Nuopponen.pdf
48. Ross, S. M. & Morrison, G. R. (2004). Esperimental Research Methods. Handbook of Research on Educational Communications and Technology. https://www.taylorfrancis.com/chapters/edit/10.4324/9781410609519-51/experimental-research-methods-steven-ross-gary-morrison?context=ubx
49. Rousi, R., Saariluoma, P., & Leikas, J. (2010). Mental contents in user experience. Proceedings of MSE2010, 2, 204–206. https://www.researchgate.net/profile/Rebekah-Rousi/publication/48330680_Mental_contents_in_user_experience/links/544e60b70cf2bca5ce90b0ba/Mental-contents-in-user-experience.pdf
50. Salvo, M. J. (2001). Ethics of engagement: User-centered design and rhetorical methodology. Technical communication quarterly, 10(3), 273–290. https://doi.org/10.1207/s15427625tcq1003_3
51. Sapienza, F. (2004). Usability, Structured Content, and Single Sourcing with XML. Technical Communication. Vol. 51 (3), pp. 399–408. https://www.ingentaconnect.com/content/stc/tc/2004/00000051/00000003/art00006#Supp
52. Serino, M., Cordrey, K., McLaughlin, L., & Milanaik, R. L. (2016). Pokémon Go and augmented virtual reality games: a cautionary commentary for parents and pediatricians. Current opinion in pediatrics, 28(5), 673–677. https://doi.org/10.1097/MOP.0000000000000409
53. Shukla, J. (2023). How Generative AI and ChatGPT might impact knowledge work. https://www.gravityunion.com/blog/2023/3/generative-ai
54. Silvennoinen, J. M., Rousi, R., Jokinen, J. P., & Perälä, P. M. (2015). Apperception as a multisensory process in material experience. Proceedings of the 19th International Academic Mindtrek Conference (pp. 144–151).

https://doi.org/10.1145/2818187.2818285

55. Spence, C., Okajima, K., Cheok, A. D., Petit, O., & Michel, C. (2016). Eating with our eyes: From visual hunger to digital satiation. Brain and cognition, 110, 53–63. https://doi.org/10.1016/j.bandc.2015.08.006
56. Tzimas, E., Vosniakos, G-C. & Matsas, E. (2019). Machine tool setup instructions in the smart factory using augmented reality: a system construction perspective. International Journal on Interactive Design and Manufacturing (IJIDeM), vol. 13, no. 1, pp. 121–136. https://doi.org/10.1007/s12008-018-0470-z.
57. Unkelbach, C. (2006). The learned interpretation of cognitive fluency. Psychological Science, 17(4), 339–345. https://doi.org/10.1111/j.1467-9280.2006.0170
58. Vanneste, P., Huang, Y., Park, J. Y., Cornillie, F., Decloedt, B., & Van den Noortgate, W. (2020). Cognitive support for assembly operations by means of augmented reality: an exploratory study. International Journal of Human-Computer Studies, 143, 102480. https://doi.org/10.1016/j.ijhcs.2020.102480
59. Verhulsdonck, G., & Morie, J. F. (2009). Virtual chironomia: Developing standards for non-verbal communication in virtual worlds. Journal of Virtual Worlds Research, vol 2, no 3. https://doi.org/10.4101/jvwr.v2i3.657
60. Zauner, J., Haller, M., Brandl, A. & Hartman, W. (2003). Authoring of a mixed reality assembly instructor for hierarchical structures. The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings., Tokyo, Japan, 2003, pp. 237–246, https://doi.org/10.1109/ISMAR.2003.1240707.

Figures

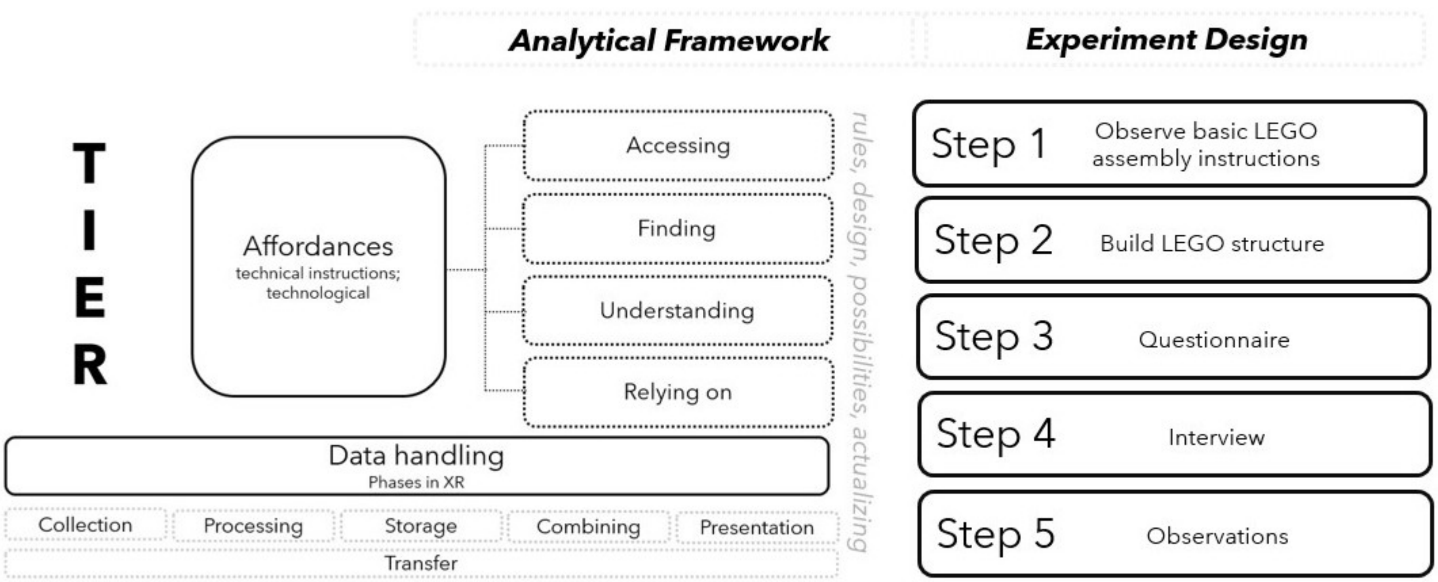

Figure 1

TIER and its application as an analytical framework & experimental technique

Figure 2

LEGO construction from the instruction video

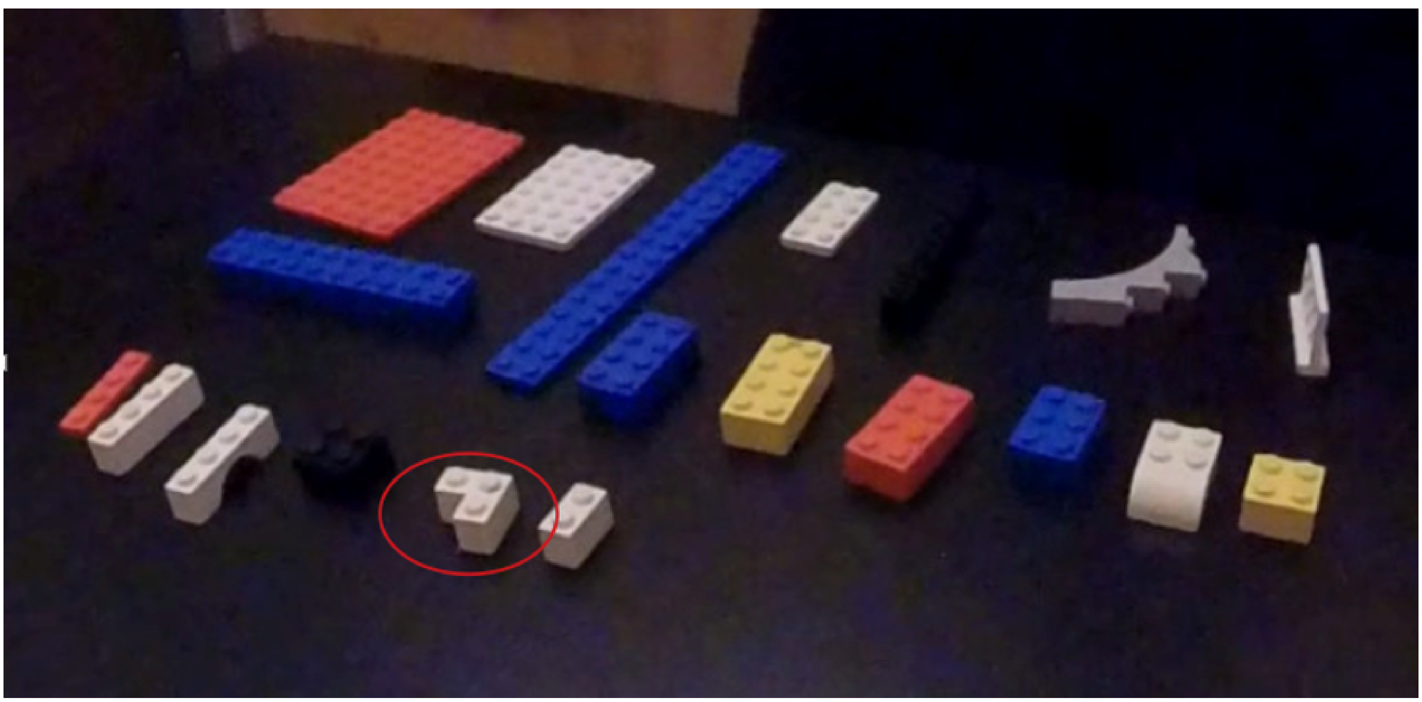

Figure 3

Arrangement of LEGO pieces in the 360-degree video instructions

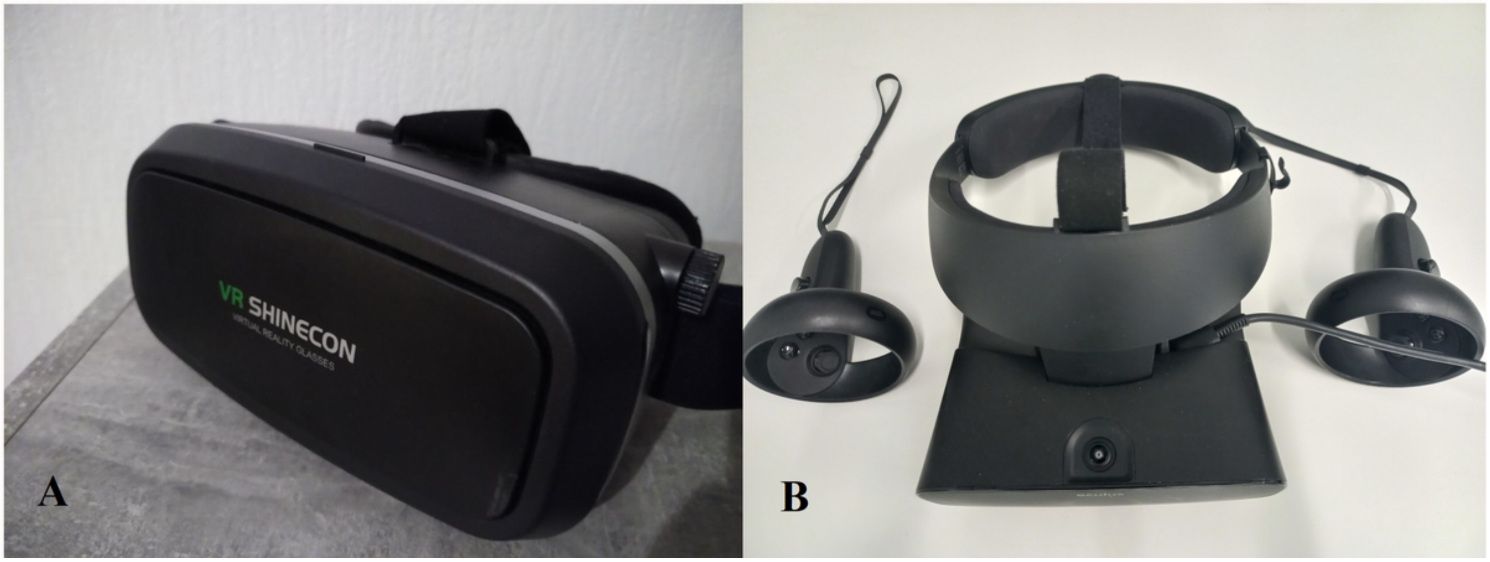

Figure 4

VR Shinecon glasses (4a); Oculus Quest (4b)

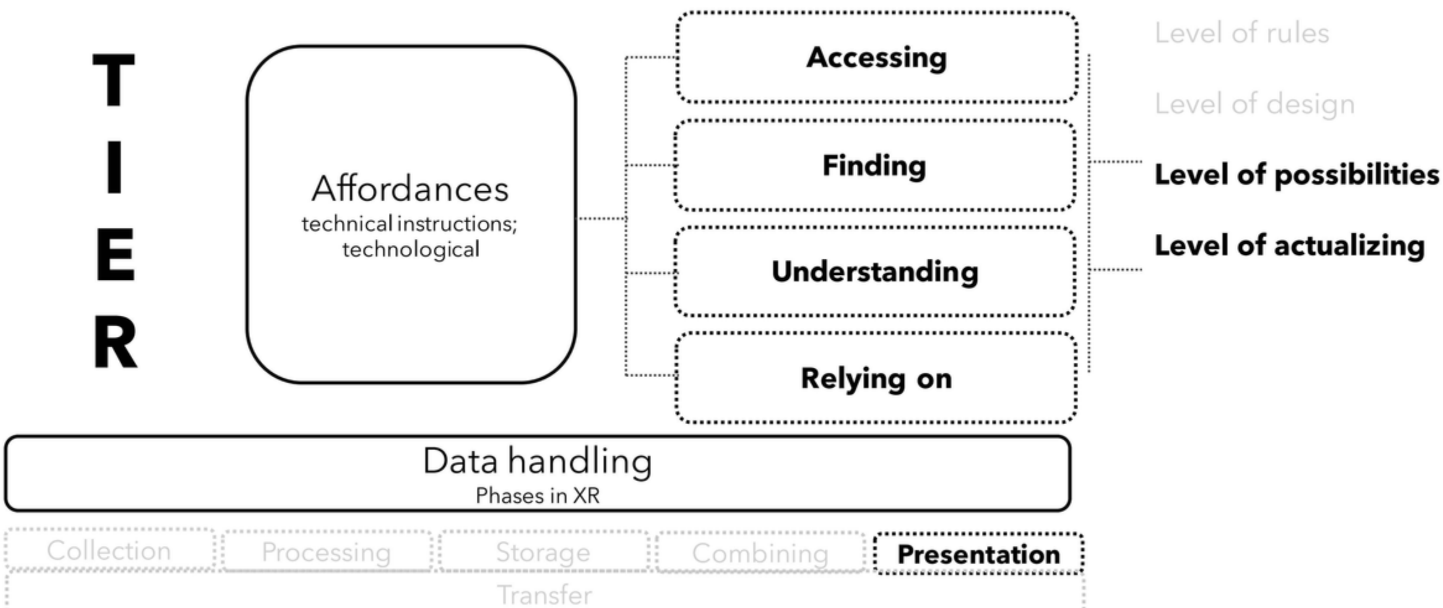

Figure 5

Selected facets of TIER model utilized in the study