EXAMENSARBETE INOM INFORMATIONS- OCH KOMMUNIKATIONSTEKNIK,
AVANCERAD NIVÅ, 30 HP
STOCKHOLM, SVERIGE 2018

Evaluation of a Multi-User Virtual Reality System for Collaborative Layout Planning Processes

Utvärdering av ett virtual reality-system med flera användare för samarbete i layoutplanering

JIM TOLMAN

Sammanfattning

Denna studie beskriver tillämpningen och utvärderingen av ett system för användning av Virtual Reality (VR) i samband med layoutplanering av Scania-fabriker. Målet är att utvärdera samarbetet inom systemet samt att bedöma användarvänligheten. Studien använder befintliga metoder på nya sätt. 16 deltagare filmas när de utför en gemensam uppgift och kodas sedan för Collaborative Joint Attention (CJA). Utvärderingen använder sig även av System Usability Scale (SUS) och Nielsens Heuristics. SUS-poängen var över genomsnittet, men deltagare med tidigare erfarenhet av layoutplanering gav systemet ett högre betyg. Det fanns många problem relaterade till att det fysiska rummet var mindre än det virtuella rummet, begränsad användarkontroll och att gestaltningen av brukarens avatar visade sig vara distraherande. Resultaten har konsekvenser för byggare och utvärderare av VR-fleranvändarsystem för samarbete. En rekommendation till utvärderare är att överväga användning av CJA som en beroende variabel.

KTH
SKOLAN FÖR ELEKTROTEKNIK OCH DATAVETENSKAP

Evaluation of a Multi-User Virtual Reality System for Collaborative Layout Planning Processes

Jim Tolman
KTH Royal Institute of Technology
Stockholm, Sweden
tolman@kth.se

ABSTRACT

This paper discusses the application of a tool for experiencing the usage of Virtual Reality (VR) in the factory layout planning process of Scania. The goal is to evaluate the system's collaborative capabilities and to assess the usability. The study combines existing methodologies in a novel way. The method consists of recording 16 participants in performing a collaborative task, and then coding for Collaborative Joint Attention (CJA). Furthermore the evaluation makes use of the System Usability Scale (SUS) and Nielsen's Heuristics. The system's score on the SUS appeared to be above average, but participants with higher experience in factory planning gave higher scores. There were numerous problems related to the physical room being smaller than the virtual room, user control was limited and the embodiment of the users (avatars) proved to be distracting. The findings have implications for builders and evaluators of multiparty VR systems, that allow for collaboration. The evaluators need to consider including CJA as one of their dependent variables.

CCS CONCEPTS

•**Human-centered computing → Virtual reality;** Visualization theory, concepts and paradigms; *Computer supported cooperative work;* **Heuristic evaluations;**

KEYWORDS

Virtual Reality, Smart Factory, Human-Computer Interaction Design, Joint Attention, Mediated Collaboration

1 INTRODUCTION

Classic industries are increasingly investing in smart, networked systems in order to increase quality and production. This phenomenon is described as Industry 4.0 [18]. Factories are becoming more complex and data-driven. A growing number of industries find utility in visualising and manipulating their data using Virtual Reality (VR) systems [2].

The industry's rising interest in VR systems is accelerated by their increased capability and affordability. The incremental problem with this development is that there is no operating system for VR, so a common graphical user interface is nonexisting. There is a substantial amount of studies being conducted on defining design principles of a user interface for VR. Even though VR was invented before the graphical user interface, most VR projects have an experimental Human-Computer Interaction (HCI) design. Summarizing; unlike the interface designs in the domain of the screen-keyboard-and-mouse paradigm, user interfaces for VR can still provide many interesting HCI questions.

Within this playing field, Scania provides a use case for creating and evaluating a multi-user VR system for "factory layout planning". Constructing a new factory is a significant investment. Hence, careful planning is of high importance before the construction or reconstruction of a factory. In addition to that, the planning process involves many stakeholders and disciplines that use different tools. The latter can impose communication and collaboration problems.

This Master thesis describes a study that measures and reflects on the usability of a proposed VR system for collaborative factory layout planning, developed with Scania AB Södertälje. The vision is of a system that can bring together the parties involved in the planning processes.

The requirements from Scania were to create and evaluate a multi-user VR system, that will be used by layout-, human factors- and maintenance experts, as well as process engineers, for planning and evaluating future factory layouts. At the start of the project, a single-user VR experience that contained a detailed model of a factory was already available. The goal of this project was to adapt this VR system in order to make it a planning and discussion tool that can be used by two or more people at the same time. It was required that both people have the same visual feedback and the same controls in terms of navigating and altering the VR environment.

1.1 Research Question

Following Scania's use case, a research question with scientific relevance was formulated. The research question that guided this thesis project is:

"*What are the affordances and limitations of a multi-user virtual reality (VR) system for supporting layout experts, project managers, and project members with expertise in logistics, human factors and maintenance in their collaborative task of planning and evaluating factory layouts of vehicle manufacturers as measured through usability inspection and collaborative joint attention (CJA)?*"

The work is scientifically relevant as there is a lot of recent innovation and development in the area of VR experiences [6, 7, 23, 37, 38], and it is important to explore and define what interactions are best suited when creating a multi-user variant of such an experience [8, 16]. Furthermore, it is important to do an extensive user

DOI: 10.1145/nnnnnnn.nnnnnnn

evaluation, in order to make sure the innovation and development contribute to more effective and pleasant collaboration.

1.2 Report Organisation

The Master thesis follows the structure of a scientific paper. The introduction (Section 1) contains the motivation for the research and ends with a narrow research question, that states what the research is about. Theoretical background (Section 2) shows that the research question is both relevant and novel, by pointing towards the work that is built upon. The method section (Section 3) is the core of the document, as it describes how the research question will be answered. The results section (Section 4) contains both the plain results in the form of untouched data. Analysis (Section 5) tries to answer the research question. The discussion (Section 6) and conclusion (Section 7) contains a reflection on the relation between the research, the related work and real-world practices, as well as a summary and future work (Subsection 7.1).

2 THEORETICAL BACKGROUND

This section contains related work and foundational work that informed this research. Relevant Digital Factory- (Subsection 2.2), Mediated Collaboration- (Subsection 2.3), VR- (Subsection 2.4) and User Evaluation (Subsection 2.5) research is discussed.

2.1 Definition of Terms

The work is looking through the lens of Human-Computer Interaction in the field of Production and Manufacturing Engineering. It is highly likely that this thesis will be read by HCI experts that need to be introduced to concepts from production engineering and vice versa, thus it proves to be useful to define terms.

Coding or Annotating borrowed from the behavioural sciences, is an analytical process in which data, in both quantitative or qualitative form, is categorized to facilitate further analysis. Oftentimes the process helps to subtract quantitative data from qualitative data.

Collaboration the action of working with someone to achieve a goal. Exchanging information is essential for collaboration. Explaining an idea is a building block of collaboration.

Experience mostly referring to user experience, meaning the overall experience of a person using a product such as a website or a computer application, especially in terms of how easy or pleasing it is to use.

Gaze the act of seeing and being seen; a steady intent look or stare. In the context of this research, the user's gaze concerns both the position of the head and the field of view (perspective) of the user, i.e. the position of the eyes is not considered.

Joint attention or jointed engagement is the shared focus of two individuals on an object. It is achieved when one individual alerts another to an object by means of eye-gazing, pointing or other verbal or non-verbal indications.

Learner the complementary expert, the listener. The critical opponent to the teacher. (Participant B)

Teacher the explainer, the layout expert. The person that both has to prove a point as well as ask for feedback. (Participant A)

Tool primarily referring to a piece of software, as opposed to a physical tool, like a hammer.

Virtual Reality in 1994, defined by Milgram and Kishino [24] as a digital environment in which the participant-observer is totally immersed in, and able to interact with, a completely synthetic world. VR is a subset of *Mixed Reality* of which *Augmented Reality* is also part.

2.2 Digital Factory

The digital factory is a model of a planned or existing factory used for design, planning and operations. Research [18] points towards related concepts of Industry 4.0, Industrial Internet, Cyber-physical Systems and Smart Factory. The increasing integration of the Internet of Everything into the industrial value chain has built the foundation for the next industrial revolution called Industry 4.0 [18]. Although Industry 4.0 is currently a top priority for many companies, research centres, and universities, there is no generally accepted definition of the term. The constructs of a Smart Factory and a Digital Factory arise from the Industry 4.0 literature. The fourth industrial revolution has been called an unhelpful construct because it would be more of an evolution than a revolution. It is said that heads of state are using the term to focus too much on technological advancement, and too little on social and political enlightenment[1]. Many governments and institutions embrace the term.

Hermann et al. [18] describe guiding principles for Industry 4.0 that prove to be helpful in understanding the role of VR in Smart Factory. VR appears to be useful for virtual assistance [13], collaboration [14, 37] and transparent information [30, 32]. Related work is, for example, done in finding VR simulation methods of human-robot collaboration (HRC) in production engineering [6]. In this project, a practical implementation in Unity is provided as proof of concept.

New implementations of VR systems can be used to annotate spacial data [33]. This demonstrates that it is possible to label and inspect large sets of room-scale point clouds in a fast and intuitive way.

2.3 Mediated Collaboration

Large multinational organisations, like Scania AB, generally have factories and/or offices that are geographically distributed. The organisation's teams are therefore often culturally diverse and their communication is digitally mediated. The field that researches digital mediated communication and collaboration is long-standing, seen in the light of technological advancement since the invention of the internet.

Related work in the field can consist of assessing remote collaboration (e.i. through Google Docs), computer-mediated collaboration in the same room (i.e. through a smartboard), collaboration through email, instant messaging, voice call, video conference and mixed reality (i.e. VR and AR) [21]. In 1999, gaze direction in multiparty video conferencing using eye tracking was studied [36]. The GAZE Groupware System is developed and used to simulate a four-way

[1] "This Is Not the Fourth Industrial Revolution". 29 January 2016 - via Slate - Accessed on April 21, 2018 http://www.slate.com/articles/technology/future_tense/2016/01/the_world_economic_forum_is_wrong_this_isn_t_the_fourth_industrial_revolution.html

round-table meeting by placing a 2D image (or persona) of each participant around a table in a virtual room, at a position that would otherwise be held by that remote participant. The authors show that gaze direction can help in establishing who is talking or listening to whom. Through empirical methods, they identified this as one of the more important problems in multiparty mediated communication.

Researchers developed a truly mixed reality system for remote collaboration [29]. In the novel implementation, an AR user can capture and share their local environment with a remote user in VR to collaborate on spatial tasks in shared space. The system enriches collaboration through combinations of gestures, head gaze, and eye gaze input, and provides visual cues to improve awareness of a remote collaborator. They shared a video at SIGGRAPH Asia 2017[2]. The system shows the boundary of what the users can see through their display to inform a user of what their partner can see. To share the exact gaze location a perpendicular ray can be projected from a user's pupil gaze direction. Publishing an extensive user evaluation remains future work. A study with 21 primary school pupils writing a school play through an online word processor with a chat function, saw how this allowed them to give rise to hybrid spaces where the discourses of *schooling* and *everyday life* intersected (Boundary Crossing) [19]. The study enriches the present-day understanding of the social, emotional, and cultural dimensions of chat interaction in computer-mediated collaboration. It shows that the field of Mediated Collaboration is broad and diverse.

2.4 Virtual Reality and Presence

The conventionally held view of a VR environment is one in which the participant-observer is totally immersed in, and able to interact with, a completely synthetic world [24]. Such a world may mimic the properties of some real-world environments, either existing or fictional; however, it can also exceed the bounds of physical reality by creating a world in which the physical laws ordinarily governing space, time, mechanics, material properties, etc. no longer hold.

One of the challenges of VR is to create a virtual world that gives the user the impression that it is a real-life environment. The human factors of this challenge are described on a conceptual level[5]. The foundational work explains that users understand the layout of a cluttered natural environment through the use of nine or more sources of information: occlusion, height in the visual field, relative size, relative density, aerial perspective, binocular disparities, accommodation, convergence, and motion perspective. On a more practical note, it was found that users can experience symptoms that are similar to motion sickness symptoms when exposed to a virtual environment [20]. A theory on how to deal with cybersickness is for the user's experience to approach "presence". Presence, a term derived from the shortening of the original "telepresence," is a phenomenon enabling people to interact with and feel connected to the world outside their physical bodies via technology [39]. To achieve presence all the components of the system have to work together exceptionally well. The following aspects have to be taken into account: the field of view, the screen resolution, the pixel persistence, the height of the refresh rate, the (simultaneous) illumination of the pixels of the display, the optics, the calibration, the tracking and the latency.

Using VR in a manufacturing environment has been done before. A solution was implemented that makes use of both assembly simulation and virtual reality, to create interactive and immersive visualisations of factories [32]. They mention the potential for factory layout planning, but a user evaluation is lacking. VR evaluation systems are focusing on the design of products, but not manufacturing systems [38]. Noticing this, a collaborative design platform and create three applications were developed, which addressed different levels of a manufacturing system design. The framework and application can improve the planning quality, accelerate the planning velocity and avoid production shutdowns.

Collaboration in VR, outside the realm of the production and manufacturing, can be found in teacher-learner scenarios like EVA; a concept for mediated teaching and learning that sits at the intersection of exploratory learning, telepresence, and attention awareness [15]. In order to evaluate the system they statistically compared the NASA-TLX score and the results of a five-question questionnaire between four apps for EVA. Then they gathered qualitative results from interviews and compared outcomes from people that used the VR variant and people that used the AR variant. There are numerous ways to evaluate (mixed reality) computer systems; the next section lists a number of them.

2.5 User Evaluation

Several qualitative and quantitative ways of evaluation will be discussed in order to give insight into what can be measured through user studies and how to do it. This will inform the method of the study.

The aforementioned NASA Task Load Index is a subjective, multidimensional scale designed to obtain workload estimates from one or more operators while they are performing a task or immediately afterwards [17]. After more than 30 years of use, NASA-TLX has achieved the status of a benchmark against which the efficacy of other factors is compared. The questionnaire consists of five questions that have to be assessed on a 21-point scale from *very low* to *very high*, with questions like "*How mentally demanding was the task?*".

Another validated questionnaire is the System Usability Scale (SUS) [4, 35] that can be used to take a quick measurement of how people perceived the usability of computer systems on which they were working. It has reached the status of a de facto standard. It assesses usability on ten 5-point scales from strongly disagree to strongly agree, with questions like "*I felt very confident using the system*".

In addition to quantifying a subjective construct using a questionnaire, experienced evaluators like to use methods that involve expertise besides solely relying on the law of large numbers. A set of methods with which an evaluator inspects a the usability of a user interface is called a usability inspection. This is in contrast to usability testing where the usability of the interface is evaluated by testing it on real users. An example of this is the heuristic evaluation. Nielsen (1994) [27, 28] developed this method on the basis of several years of experience in teaching and consulting about usability engineering. It is a rather informal method and involves

[2] The video at SIGGRAPH Asia 2017 was downloadable and the authors of this paper repost it here: https://goo.gl/HAoske

having usability specialists judge whether each dialogue element follows established usability principles. The heuristics provide ten 'themes' to discuss, during two phases that take several hours each. There are several sets of heuristics, the evaluator can use them as inspiration and stepping stone, thus selecting heuristics that are found to be fitting. 'Themes' that would inform design heuristics for future the evaluation of VR systems were brought forth and selected for the research [25].

It is suggested that using questionnaires, interviews or expert evaluations, can generate noise and is scientifically speaking less pure than observations. In addition to questioning the user and thinking like the user, watching an interaction between a user and a computer system unfold can bring valuable insight. To operationalise and analyse observations into a research method, an evaluator can use coding. After videoing sessions where users interact with a system, the evaluator can use grounded or a priori coding. Grounded coding refers to allowing notable themes and patterns emerge from the recordings themselves, whereas a priori coding requires the researcher to apply pre-existing theoretical frameworks to analyze the recordings. An example of an a priori coding scheme (annotation scheme) is Facial Action Coding System [11], which exists of 98 *action units* that describe something that could be observed about a person's face. For example, action unit 11 nasolabial deepener, is the code for *smile lines* and action unit 55 simply describes that the head of the subject is tilted left.

Oftentimes coding schemes emerge from the type of data that is gathered during an HCI study. For example, the observational studies annotated transcripts of sessions in which groups subjects were asked to collaborate[34]. A line from the transcript would be "*Ok, ok let's do slide and tap*", and the accompanying annotation would be 'Writes "*slide and tap*"'. If this type of annotation comes up more often, it might be incorporated in a coding scheme that later is used to compare collaborative tasks with certain independent variables.

Joint attention is a measure that is common in research that concerns autistic children since the timing of the development of joint attention skills can be an indicator for diagnosis of autism. It is also proposed as a measure in autism therapies. Studies try to examine the effects of social stimulation on the joint attention behavior of different groups of children and thus develop a coding scheme with categories like: Supported Joint Engagement, *Child is actively involved with toy that adult manipulates in such a way as to alter child's experience with that object (e.g., child laughs at adult toy demonstration and reaches for toy)* [22]. The construct of joint attention is later used in HCI research to evaluate User-Designer Collaboration [26].

Explaining an idea is a core component of the type of collaborative decision-making process discussed in this paper. Similar to a doctor explaining relevant information to a patient in a shared decision-making process [9] and children taking part in collaborative learning activities [10], this study observes a layout expert (teacher) and a complementary (learner) in their collaborative process.

3 METHOD

Part of Scania's Digital Factory efforts is to work with simulation-driven production development. When constructing or reconstructing a facility, there is an opportunity to employ VR technology, in order for better communication between all the stakeholders, for example, human factors experts. A VR system is developed following the given requirements.

The mixed method used in this research is informed by methods of time-motion study (observational studies) and heuristic evaluation (usability inspection). Although the research is mostly qualitative, standard metrics of user experience and satisfaction are used to gather some quantitative results (benchmarking). The goal of the VR system is to support the project members in their collaborative task of planning and evaluating factory layouts. The comparative part was conducted to contrast a multi-user VR system with a desktop monitor showing the same 3D representation of the design of a future factory. The desktop condition mimics current working practices.

3.1 Study Design

The within-subject study invited 8 pairs of participants to perform the collaborative task of increasing the capacity of the pedal car production line by adding one workstation and evaluating the factory layout. Each pair consisted of one layout planning expert and one other expert (Manager/Logistics/Safety/Human Factors/etc). All eight pairs were asked to use the VR system for their collaboration. A between-subject study would allow for more comparative methods, but given that only domain experts (with very limited availability) are recruited, a within-subject design was chosen in order to render deep and broad understanding in the collaborative capabilities of the system.

A substantial part of collaboration is explaining an idea to peers. Exchanging information is an essential part of any collaborative process. The main task that was chosen was for one subject to explain the design decisions of the factory layout that is present in the VR environment to a second subject.

During the session, the participants were given time to get used to the VR system, the point of view, the camera & location control, and the mechanics of the handheld devices. Then they had to perform the task, during which the body positions and the field of view of the participant were recorded by means of a video camera and screen capturing software. As a post-experience questionnaire, the System Usability Scale (SUS) was used (Appendix D). In the video analysis, the time-motion study is conducted and the interaction will be reflected upon using heuristic brought forth by Nielsen, complemented by a set of VR specific heuristics.

To compare the VR system with current practices is to compare with LayCAD; a software tool for factory layout design. LayCAD has a tremendous amount of functionalities and has a steep learning curve. The time it takes to fulfil a task is measured in VR and in LayCAD.

3.2 System Implementation and Setup

The VR implementation uses Unity and VRTK, which enables for combining different 3D models into one and creating a VR experience. The developed tool uses two HTC VIVE headsets and two

Figure 1: Screenshot of the VR environment: The pedal car line consisting of interactable objects. The layout is based on the Scania pedal car factory line, as it is set up inside the building of KTH Södertälje.

sets of controllers. The virtual environment (see fig. 1) is a representation of an existing location, with interactable objects and the possibility to measure distances, make annotations and move objects. To orientate, the users can see and hear each other. For physical safety, the users did not share the same physical environment while immersed in VR and blind to each other (see fig. 2).

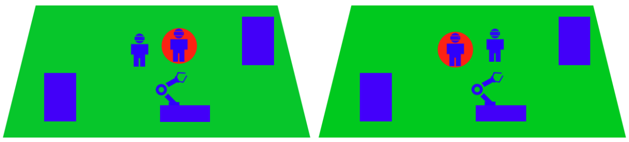


Figure 2: The VR environment consists of players and 3D models. The users are in the same virtual space, but do not share the same physical environment.

3.3 Participants

The 16 participants were all Scania employees, with an average age of 42. A quarter of the participants identified themselves as females and the rest identified themselves as males. Most participants were born in Sweden; a quarter of the participants were born elsewhere, naming India, Brazil and Iraq as their countries of birth. Half of the participants indicated their preferred language in computer systems is English, the other half filled in Swedish. 62,5% of the participants indicated they had used a VR headset before, 3 participants indicated they work with VR occasionally and 1 selected to be working with VR weekly. From the layout experts, 7 followed a LayCAD course and they amount to an average 7.64 years of experience in the software tool.

3.4 Independent Variables

The quantitative part of the study compares the action of moving an object within the VR system to the action of doing the same within LayCAD, the layout planning desktop application. The changing condition can thus be described as VR system / Desktop system. The qualitative part of this study comprises a tremendously high amount of degrees of freedom (controllers, avatars, pop-up menus, etc). The age, gender identity and previous VR experience of the participants was recorded.

3.5 Manipulation and Legibility Check

To check if this setup tests the collaborative capabilities of the VR system, the participants were asked to fill in a pre-session questionnaire (See Appendix C) with questions concerning the previous experience in VR and in LayCAD in order to find any biases of the participants towards VR or LayCAD. In addition, demographic information was requested in order to be transparent about the cultural-, gender- and age-differences that are found in the pool of participants.

3.6 Dependent Variables

The dependent variables for the quantitative part were the time it takes to move the object within either the VR or the desktop environment. For the qualitative part, the SUS served as a dependent variable. Then, from the further analysis, the successfulness of the collaborative interaction appeared as a qualitative, dependent variable.

3.7 Procedure

A controlled lab environment in the offices of Scania served as the location for the experiment. The participants were welcomed and introduced to the research area. They were asked to fill in basic demographic information, as well as information regarding their previous experience in VR and LayCAD using a digital form. During the following training session, the participants were invited to try all the buttons and learn how to grab, teleport and measure. As these basic actions were mastered, the participant became increasingly proficient in the use of the factory planning tool as a whole. Also during this time, participant A was shown around the logistics room of the factory and was taught 3 design decisions of the layout.

The goal of task 1 was to learn about the collaborative capabilities of the system. Inspired by the Telephone Game (Chinese Whispers)[3], participant A has to explain the design decisions of the logistics room to participant B. As part of this explanation, participant A has to use a measuring tool (See fig. 3 and 4), to substantiate the decisions. Participant A measured the distance from the logistics rack to the wall (Distance 1). During and after this exercise the collaborative behaviour of the participants was analyzed using a coding scheme for collaborative joint attention (CJA).

As task 2, the participants were asked to re-arrange the boxes at the Red Rack, change the position of the Yellow Rack next to the Blue rack, and change the height of the level of the boxes in the Blue Rack. The user experience for this type of action is an indicator of the usability of the VR tool.

In order to compare the user experience in VR with the LayCAD experience, participant A was asked to measure the path in the logistics room inside LayCAD as well.

[3]Definition of Chinese whispers in English - via Oxford dictionaries - Accessed on June 29, 2018 https://en.oxforddictionaries.com/definition/chinese_whispers

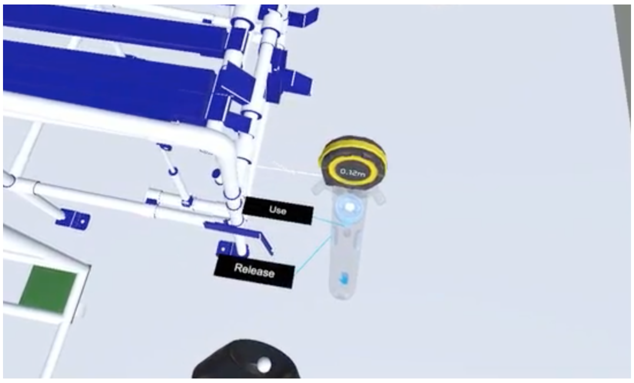


Figure 3: Screenshot of the measuring tape in use, from the teacher's perspective. The hints appear on the representation of the controller, every time the user can perform special actions, like *Grab*, *Use* and *Release*.

Figure 4: Screenshot of the measuring tape in use, from the learner's perspective within the VR experience. The avatar is animated on top of the tracked locations of the VR headset and the controllers.

3.8 Measurements

The behaviour of the participants was videoed and analysed using an annotation scheme for CJA. The coding scheme contained the following leading definition of CJA: "*Participant B is actively involved with the measuring-tape (and process) that participant A manipulates in such a way as to explain something to participant B (e.g., Participant B asks an in-depth question that shows understanding).*" Coding was based on the approximate direction the participants were looking and the verbal cues they gave. CJA was considered to be established when participants are either involved in mutual gaze (looking at each other) or object engagement (looking at and talking about the measurement (tape) and the task at hand). From this coding, the total time that a pair of participants was in engaged in CJA was found. Also, the participants filled in a usability questionnaire at the end of the session (See the SUS questions in Appendix D).

3.9 Data Overload

The study gathered a lot of data in the form of video and audio material from 4 different angles on 8 sessions. In addition, participants filled in 2 questionnaires (SUS in Appendix D and previous experience in Appendix C) and one open-ended question asking for their general opinion of the system. In order to perform a proper analysis, it is decided to focus on **one** task very closely and qualitatively evaluate the VR based on that. In addition, the quantitative measures are reported and discussed.

4 RESULTS

With an overall average of 69.8 and a standard deviation of 13.3, the evaluation would prove the VR system to have an above average SUS-score. Participants in the category of layout expert turned out to give higher scores (layout expert avg = 72.8; complementary expert avg = 66.9). The calculation of the SUS-scores can be found in Appendix B.

Other questions included in the post-experience questionnaire (See Appendix D) were concerning the perceived joint-attention and the mediation by the VR system. The majority of the participants report that they could see and understand the other participant, as well as, they felt seen and understood by the other participant. Figure 5 shows the outcomes of the Likert scales as they were filled in.

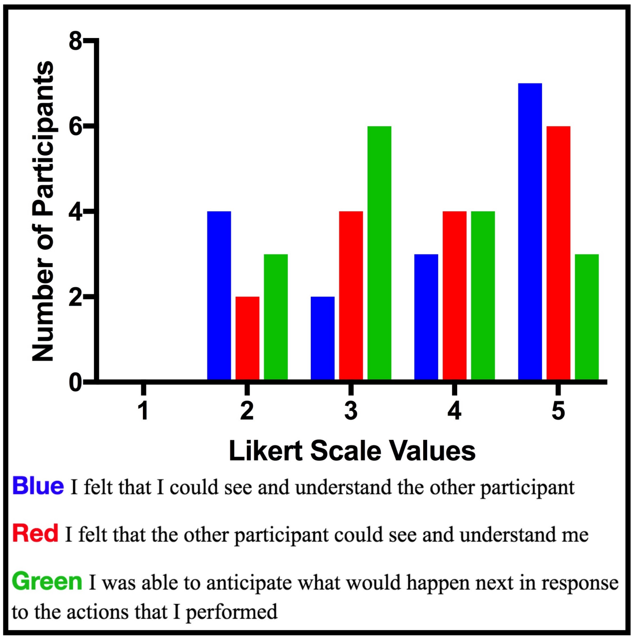


Figure 5: Results of three questions from the post-session questionnaire; these questions were asked in addition to the questions related to SUS. On the x-axis, the points on the Likert scale (1-5) are shown, on the y-axis the number of people that gave that score is shown.

5 ANALYSIS

This section consists of three subsections. Firstly, in Subsection 5.1 the coding efforts (qualitative analysis) of the recorded sessions can be found. Secondly, the analysis of the SUS scores can be found in Subsection 5.2. Then, the comparative experiment is reflected upon in Subsection 5.3. Lastly, in the instances the collaborative interaction broke, Heuristics categorize UI problems in 5.4.

5.1 Coding for CJA

The moment that the participants measure the distance between the kitting rack and the wall (review fig. 3 and 4) is the most interesting concerning collaboration and joint-attention. The video, audio and screen recordings were compiled into one synchronised video per session. These eight videos were annotated using a coding scheme for CJA.

Measuring Distance 1 required fulfilling seven collaborative actions. (1) establish communication / ask for attention, (2) grab + hold measuring tape, (3) pull the trigger to start, (4) move with arm, (5) release trigger to end, (7) communicate the outcome. During these steps, the coders looked for mutual gaze (looking at each other) or object engagement (both looking and talking about the measuring (tape) in relation to the task at hand). Mutual gaze happens when participant B watches participant A or the activity of participant A and vice versa. Object engagement occurs as participant B is actively engaged solely with the measuring-tape (e.g., following the hands of participant A very closely).

Some pairs of participants did not go through all seven steps. When action 1 and action 7 did not occur, the collaboration was labelled unsuccessful. Successful collaboration consisted of interactions where participant B verbally *("Aha yes")* or physically (nodding) confirms participant A and where participant B shows interest in the outcome of the measurement.

The total duration of CJA throughout the interaction for the successful collaborations was compared with the not successful ones. Table 1 shows the outcomes per pair. Figure 6 shows the average of all pairs. The total average CJA for the successful condition had a duration of 62.75 seconds (sd = 18.95), where unsuccessful took 32 seconds (sd = 27.75). These results showed no significance between the two conditions (p = .112).

In the pre-session questionnaire (see Appendix C), users were asked to fill in information about their previous experience with VR. From the answers, a VR score is given. The question "*Did you ever use a Virtual Reality headset before?*" gave 2 (yes) or 0 (no) points. The question "*How often do you work with Virtual Reality?*" gave 0 (never), 1 (occasionally), 2 (monthly), 3 (weekly), or 4 (daily) points. "*Are you likely is it that you get motion sickness?*" subtracted 0 to 4 points, mapped on the 5 points Likert scale. The sum of these scores is normalised by adding 2 points across the board (so there are no negative values). This gives the individual VR scores (IVR). The sum of the IVR of a pair of participants gives the combined VR score (CVR).

The total CJA is compared with the CVR in a scatter plot (see fig. 7) and a positive correlation is found between CJA and CVR. It should be noted that the R-squared value is low (.439). Figure 8 shows a scatterplot comparing the age of participant A and B with the total duration of CJA. The age of participant A shows a negative correlation with the CJA duration (R^2 = .473). The R-squared value for the age of participant B is so low (.075) that it is safe to say that the model used is not helpful.

Table 1: Total CJA and successfulness of the collaboration per pair of participants.

	Total CJA in Seconds	Successful Collaboration
Pair 1	73s	No
Pair 2	16s	No
Pair 3	14s	No
Pair 4	77s	Yes
Pair 5	25s	No
Pair 6	72s	Yes
Pair 7	67s	Yes
Pair 8	35s	Yes

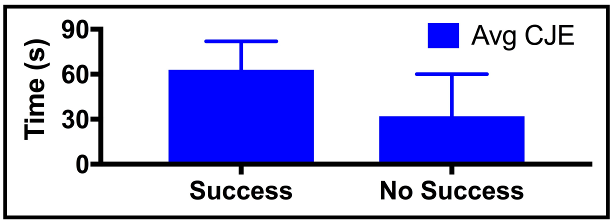


Figure 6: The average total CJA in seconds. The total duration of CJA was on average lower when the collaboration was not successful.

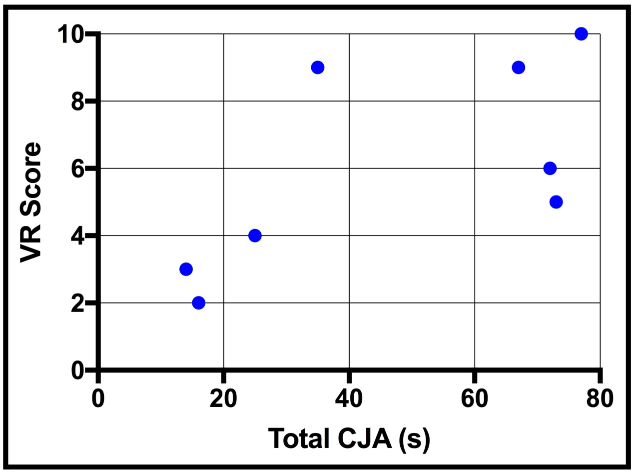
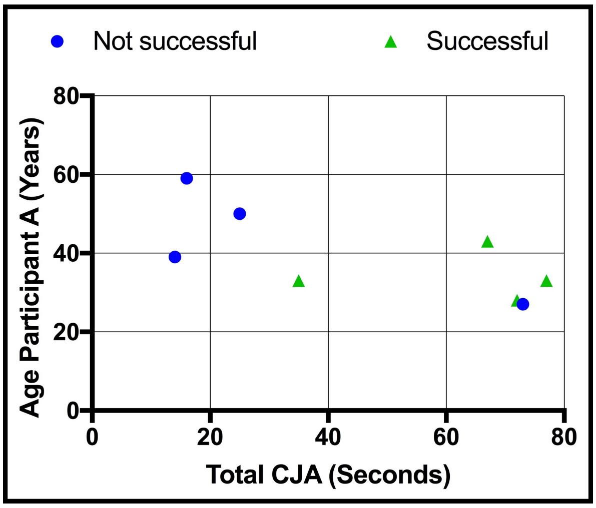


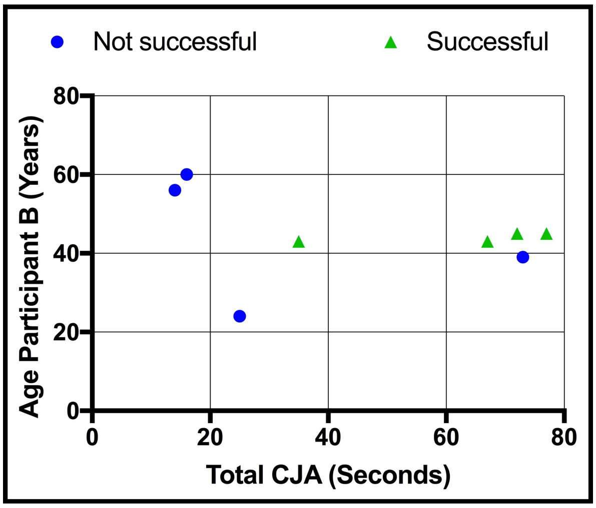
Figure 7: The total CJA in seconds is compared to the CVR. A higher CVR correlates with a longer lasting CJA.

5.2 SUS Interpretation

The SUS is an effective, reliable tool for measuring the usability of a wide variety of products and services [1]. The average of



(a) Participant A; teacher (explainer)



(b) Participant B; learner (listener)

Figure 8: The total CJA in seconds is compared to the age of the participants.

69.8 is considered above average, while scores below 68 are below average. The score was interpreted by putting it on a curved grade scale ranging from F (absolutely unsatisfactory) to A+ (absolutely satisfactory) [31]. These grades should be interpreted with caution [1]; this thesis regards them as a way to communicate the SUS outcome with non-human factors professionals. The grades can be found in Table 2. The SUS score that was found in this research can serve as a benchmark for the further development of the planning tool.

5.3 Comparative Experiment

The time to measure the distance between the logistics rack and the wall (Distance 1 as part of task 1) was distilled from the video material. At times the manual syncing of the audio and video tracks

Table 2: Total CJA and successfulness of the collaboration per pair of participants.

Group	SUS	Grade
Layout Experts	66.9	C
Complementary Experts	72.8	B-
Total Average	69.8	C

left room for ambiguities, as there was a delay in the steamed and recorded audio compared to the locally recorded audio. To measure correctly, the screen recording of the LayCAD expert user was considered determinative.

Performing the task in VR took on average 29.75 seconds (SD=13.9) whereas the duration for this in LayCAD was 14.25 seconds (SD=5.6). A paired sample t-test did not reveal a significant difference between the conditions (p = .053). See Appendix A for the full SPSS output.

5.4 Indirect Heuristics

At the instances the collaborative interaction broke, Nielsen's Heuristics are used to categorize UI problems. This can be called indirect heuristics. The mix between inspection evaluation and evaluation through observation interprets heuristics as an annotation scheme for video analysis. This can help to find affordances and limitation of a VR layout planning system.

An important heuristic describes the match between a system and the real world. The VR environment was representative of a real-world factory, but interaction-wise the system was not always similar. To allow for user- control and freedom, functionalities like Undo and Redo should be supported. The VR system gave the possibility to fully reset space, but Undo and Redo were lacking and missed. Also, a volume control was missing, resulting in some users having trouble hearing the other user.

Users should not have to wonder whether different actions mean the same thing, thus it is advised to pay attention to platform conventions. Conventions like teleportation and depth- perception were implemented in the system, as well as for example an iconic speech bubble when a participant was talking. But inside (VIVE-products) and across (all mixed reality applications) VR systems leak a common user-interface. It is VIVE platform convention that *grabbing* an item is done using the side buttons, but many participants were not able to reach those, even after 30 minutes of using the system intensively.

It is believed in the field of HCI that each UI element or piece of information in an interface uses the user's attention and cognitive capabilities, thus a minimalist design aesthetic is advised. The avatars used in the implementation are quite complex and distracting. Also, they overpromise and underdeliver; their big eyes are always staring past the gaze of the beholder.

During all sessions, there were many problems related to physical space constraints. The physical room was smaller than the virtual room and many times users wandered off to the point they had to be protected by an evaluator in order to not bump into a wall. Using the collaborative VR system on a wide scale inside Scania would imply creating several VR rooms or designing and implementing a system that is optimised for using whilst sitting behind an office desk.

Glitchiness should be minimised, but during the evaluation sessions, there were glitches. Once participant A (Teacher) was not able to see participant B (Learner), immediately there was no opportunity to collaborate. Also, the Yellow Rack that should be moved during task 2, was prone to tip over, leading to the chaos that would require an Undo function. The problem occurred all times and only one user was able to solve it.

6 DISCUSSION

This section contains a reflection on the conducted study and a discussion on the relation between the research, the related work and real-world practices.

Every study that involved sessions with people has been exposed to practical problems that require improvisation and problem solving, even though these stories do not usually make it to the final paper. On the morning of the day that half of the user evaluation sessions were planned, the power in the whole town of Södertälje was out. This was problematic since the participants of the user studies were all experts users with limited time in their agendas. The project got somewhat delayed.

The SUS scores turned out to be above average, but still relatively low if we use the curved grade scale comparison [31]. It is a known effect for the satisfaction scores of newer users to be significantly lower than the scores of more (VR) experienced people [3]. The users got less than one hour to experience the VR system, so in hindsight, it would have been possibly more informative to have longer sessions. A limitation was the finite availability of expert users. Half of the participants indicated that they preferred their computer system language to be set to Swedish, but the current VR system is only available in English.

Most scatterplots showed low R-squared values in a regression analysis, still, low R squared variables are not intrinsically problematic. In some fields, it is suspected that R-squared values are found to be low. For example, any discipline that attempts to predict human behaviour, such as psychology, typically has R-squared values lower than 50% . People are simply harder to predict than physical processes[4].

The sessions were conducted in office spaces and not in specialized VR rooms. Therefore there were a lot of problems related to physical space constraints. The physical room was smaller than the virtual room and many times users wandered off to the point they had to be protected by an evaluator in order to not bump into a wall. Using the collaborative VR system on a wide scale inside Scania would imply creating several VR rooms or designing and implementing a system that is optimised for using whilst sitting behind an office desk.

In some sessions, the learning was very passive and the teacher did not get a lot of feedback; this could be due to the experiment design. The teacher was explaining using the digital measurement tool, and the learner had no other task then to listen. The teachers tried to give the same instructions and feedback to all participants, but in some (crucial) instances the participants pointed most of their questions towards the researcher instead of communicating with each other.

Six pairs of participants were male-male (MM) and two pairs were female-female (FF). Even though the amounts of subjects in these two conditions was small, the successfulness of the MM pair had a very different relation to the total CJA duration than the FF pairs. Independent of this, the study evaluates a learner-teacher relation and these are relations with cultural-historical gender inequality. If the research would be replicated with the purpose of gathering more qualitative results, it would be advised to make all possible pairs (MM, FF, MF, FM) to account for and rule out any biases associated with gender roles.

The coding done in the video analysis is done by one researcher due to limited time and manpower. To evaluate the quality of the coding scheme and to rule out inconsistencies in coding a higher amount of coders would be advised. With more coders an inter-rater agreement (Kappa) can be calculated.

7 CONCLUSION

To conclude, this research investigated the main research question: *What are the affordances and limitations of a multi-user VR system for supporting layout experts, project managers, and project members with expertise in logistics, human factors and maintenance in their collaborative task of planning and evaluating factory layouts of vehicle manufacturers as measured through usability inspection and collaborative joint attention (CJA)?*

An interactive, multi-user VR system for collaborative factory planning was created following the requirements that arose from Scania's use case. The system was evaluated using mixed methods using quantitative measures (SUS) and qualitative measures (video analysis through coding). The qualitative measure was concerning the collaborative capabilities of the system; a pair of participants was asked to perform a measuring task, whilst both being inside the virtual environment. The field of view of the participants was video recorded and analysed using a coding scheme. The coding scheme contained the following leading definition of CJA: "*Participant B is actively involved with the measuring-tape (and process) that participant A manipulates in such a way as to explain something to participant B (e.g., Participant B asks an in-depth question that shows understanding).*" The moments in the interaction where CJA was interrupted were analysed using a selected set of heuristics [25, 27, 28].

The system's score on the SUS appeared to be above average, but participants with higher experience in factory planning gave higher scores. Also as legibility check; a higher VR score (previous experience in VR minus the chance of motion sickness) correlated positively with a higher SUS score.

The observation through video- and audio recording and the accompanying indirect heuristic evaluation lead to major insights in the problems that users the system could face. Many users were off to a bad start due to problems in the physical world. The VIVE controller uses four main inputs, all utilised in the system: tracked position, trackpad, trigger and side buttons. The side buttons were used to *grab and hold* an object, but not all participants were able to reach the side buttons, even after 30 minutes of being in the experience.

[4] Regression Analysis interpretation - via Minitab blog - Accessed July 6, 2018 http://blog.minitab.com/blog/adventures-in-statistics-2/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit

There were numerous problems related to the physical room being smaller than the virtual room. Users wandered off to the point they had to be protected by an evaluator in order to not bump into a wall. This is a safety concern that should be taken care of before implementing this system throughout Scania; for example by building big rooms with soft walls or by researching VR systems that are optimised for use whilst sitting behind a desk.

The findings have implications for builders and evaluators of multiparty VR systems that allow for collaboration. The evaluators need to consider including CJA as one of their dependent variables. Many findings lead to questions that could be addressed in future work, these will be discussed in the next section (7.1). The current system is an early version of the system that is now benchmarked and ready to be compared with future versions of the system. The usability problems found can be input for requirements for future development.

7.1 Future Work

The research is considering the limitations and affordances of using a VR system for factory layout planning. The process that is evaluated is a collaborative process that is measured by CJA. The CJA was captured using video analysis of screen recordings of the field of view of the VR participants. The accuracy of the joint attention can be increased by using an eye-tracking system. The qualitative and quantitative measurements and insights can be expended by placing third-perspective cameras in the virtual world.

The methodological problem of the learner being somewhat passive can be overcome by giving important tasks to the learner. If the learner would measure the distances, instructed by the teacher through the VR system, this would amount to an interesting experiment setup. This can also solve the challenge of the participants pointing their attention towards the evaluator too much.

Further hardware evaluation seems in order; the participants reacted uniquely to the systems physical controllers. There is a variety of hardware existing on the market. Also, gloves with trackers could be evaluated. A collaborative planning system that is optimised for the user being seated behind in an office chair behind a desk is an interesting subject for future work.

The current system can be evaluated using sessions with bigger groups of collaborators, possibly over bigger distances, possibly internationally over the internet. The latter would bring forth questions of cultural diversity. There could be sessions where different types of avatars are compared since the current avatars proved to be distracting at times.

Some research points toward Mixed Reality being the future of computing, so it would be useful to explore. When the future factory floor still only exists on paper, there is nothing to augment any digital representation on yet, but the collaborative capabilities might be improved.

During the literature review that accompanied the requirements phase of the system, longer distance navigation inside VR was explored. A common teleportation system was implemented, but there is an opportunity for future research and development in this area. The shortest distance between two points is often not a straight line. Because *zooming out* is logarithmic, it is always shorter to *fly* to a birds-eye perspective before *zooming in* at your desired destination [12]. Factory floors can be several acres large, although the current VR environment only displayed 2 large classroom-sized rooms.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to Dr Mario Romero (Associate Professor, EECS, KTH Royal Institute of Technology) for sharing his knowledge with me during the course of this thesis project, mostly through feedback sessions. Secondly, I would like to thank Lars Hanson, Franz A. Waker and Andreas Lind from Scania CV AB for providing me with this use case and assisting me during the project. Thirdly, I thank my examinor Tino Weinkauf and opponent Raghu Nayyar. Lastly, I would like to thank my friends and family for their aid and assistance.

REFERENCES

[1] A. Bangor, P. Kortum, and J. Miller. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

[2] L. P. Berg and J. M. Vance. Industry use of virtual reality in product design and manufacturing: a survey. *Virtual Reality*, 21(1):1–17, Mar. 2017.

[3] S. Borsci, S. Federici, S. Bacci, M. Gnaldi, and F. Bartolucci. Assessing User Satisfaction in the Era of User Experience: Comparison of the SUS, UMUX, and UMUX-LITE as a Function of Product Experience. *International Journal of HumanfiComputer Interaction*, 31(8):484–495, Aug. 2015.

[4] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[5] J. E. Cutting. How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments, & Computers*, 29(1):27–36, Mar. 1997.

[6] A. de Giorgio, M. Romero, M. Onori, and L. Wang. Human-machine Collaboration in Virtual Reality for Adaptive Production Engineering. *Procedia Manufacturing*, 11:1279–1287, Jan. 2017.

[7] J. Du, Y. Shi, Z. Zou, and D. Zhao. CoVR: Cloud-Based Multiuser Virtual Reality Headset System for Project Communication of Remote Users. *Journal of Construction Engineering and Management*, 144(2):04017109, 2017.

[8] F. Dupont, T. Duval, C. Fleury, J. Forest, V. Gouranton, P. Lando, T. Laurent, G. Lavou, and A. Schmutz. Collaborative Scientific Visualization: The COLLAVIZ Framework. In *JVRC 2010 (2010 Joint Virtual Reality Conference of EuroVR - EGVE - VEC)*, Fellbach, Germany, Oct. 2010.

[9] G. Elwyn, A. Edwards, P. Kinnersley, and R. Grol. Shared decision making and the concept of equipoise: the competences of involving patients in healthcare choices. *Br J Gen Pract*, 50(460):892–899, 2000.

[10] R. Fleck, Y. Rogers, N. Yuill, P. Marshall, A. Carr, J. Rick, and V. Bonnett. Actions Speak Loudly with Words: Unpacking Collaboration Around the Table. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, pages 189–196, New York, NY, USA, 2009. ACM.

[11] E. Friesen and P. Ekman. Facial action coding system: a technique for the measurement of facial movement. *Palo Alto*, 1978.

[12] G. W. Furnas and B. B. Bederson. Space-scale Diagrams: Understanding Multiscale Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, pages 234–241, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[13] P. Galambos, . Csap, P. Zentay, I. M. Flp, T. Haidegger, P. Baranyi, and I. J. Rudas. Design, programming and orchestration of heterogeneous manufacturing systems through VR-powered remote collaboration. *Robotics and Computer-Integrated Manufacturing*, 33:68–77, June 2015.

[14] S. W. Greenwald, W. Corning, and P. Maes. Multi-User Framework for Collaboration and Co-Creation in Virtual Reality. 12th International Conference on Computer Supported Collaborative Learning (CSCL), 2017.

[15] S. W. Greenwald, M. Funk, L. Loreti, D. Mayo, and P. Maes. EVA: Exploratory Learning with Virtual Companions Sharing Attention and Context. In *2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT)*, pages 26–30, July 2016.

[16] J. Grbel, T. Thrash, C. Hlscher, and V. R. Schinazi. Evaluation of a conceptual framework for predicting navigation performance in virtual reality. *PloS one*, 12(9):e0184682, 2017.

[17] S. G. Hart. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage Publications Sage CA: Los Angeles, CA, 2006.

[18] M. Hermann, T. Pentek, and B. Otto. Design Principles for Industrie 4.0 Scenarios. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 3928–3937, Jan. 2016.

[19] K. Kumpulainen and A. Mikkola. Boundary crossing of discourses in pupils' chat interaction during computer-mediated collaboration. *Learning, Culture and Social Interaction*, 3(1):43–53, Mar. 2014.

[20] J. J. LaViola. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, 32(1):47–56, Jan. 2000.

[21] P. Leinonen, S. Jrvel, and P. Hkkinen. Conceptualizing the awareness of collaboration: A qualitative study of a global virtual team. *Computer Supported Cooperative Work (CSCW)*, 14(4):301–322, 2005.

[22] A. L. Lewy and G. Dawson. Social stimulation and joint attention in young autistic children. *Journal of Abnormal Child Psychology*, 20(6):555–566, Dec. 1992.

[23] E. Matsas and G.-C. Vosniakos. Design of a virtual reality training system for humanfi?!robot collaboration in manufacturing tasks. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 11(2):139–153, May 2017.

[24] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[25] R. Murtza, S. Monroe, and R. J. Youmans. Heuristic Evaluation for Virtual Reality Systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 61, pages 2067–2071. SAGE Publications Sage CA: Los Angeles, CA, 2017.

[26] B. A. Nardi. *Context and Consciousness: Activity Theory and Human-computer Interaction.* MIT Press, 1996. Chapter 9, page 199.

[27] J. Nielsen. Usability inspection methods. In *Conference companion on Human factors in computing systems*, pages 413–414. ACM, 1994.

[28] J. Nielsen and R. Molich. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256. ACM, 1990.

[29] T. Piumsomboon, Y. Lee, G. Lee, and M. Billinghurst. Covar: a collaborative virtual and augmented reality system for remote collaboration. In *SIGGRAPH Asia 2017 Emerging Technologies*, page 3. ACM, 2017.

[30] M. Sacco, G. Dal Maso, F. Milella, P. Pedrazzoli, D. Rovere, and W. Terkaj. Virtual factory manager. In *International Conference on Virtual and Mixed Reality*, pages 397–406. Springer, 2011.

[31] J. Sauro and J. R. Lewis. *Quantifying the user experience: practical statistics for user research.* Morgan Kaufmann, Amsterdam, Netherlands, second edition.. edition, 2016.

[32] M. Schenk, S. Stra\s sburger, and H. Kissner. Combining virtual reality and assembly simulation for production planning and worker qualification. In *Proc. of International Conference on Changeable, Agile, Reconfigurable and Virtual Production*, 2005.

[33] J. D. Stets, Y. Sun, W. Corning, and S. Greenwald. Visualization and Labeling of Point Clouds in Virtual Reality. *arXiv:1804.04111 [cs]*, pages 1–2, 2017. arXiv: 1804.04111.

[34] J. C. Tang. Findings from observational studies of collaborative work. *International Journal of Man-Machine Studies*, 34(2):143–160, Feb. 1991.

[35] T. S. Tullis and J. N. Stetson. A comparison of questionnaires for assessing website usability. In *Usability professional association conference*, pages 1–12, 2004.

[36] R. Vertegaal. The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 294–301. ACM, 1999.

[37] T.-H. Wu, F. Wu, C.-J. Liang, Y.-F. Li, C.-M. Tseng, and S.-C. Kang. A virtual reality tool for training in global engineering collaboration. *Universal Access in the Information Society*, pages 1–13, 2017.

[38] X. Yang, R. C. Malak, C. Lauer, C. Weidig, H. Hagen, B. Hamann, J. C. Aurich, and O. Kreylos. Manufacturing system design with virtual factory tools. *International Journal of Computer Integrated Manufacturing*, 28(1):25–40, Jan. 2015.

[39] D. Zeltzer. Autonomy, interaction, and presence. *Presence: Teleoperators & Virtual Environments*, 1(1):127–132, 1992.

Appendices

A PAIRED SAMPLE T-TEST

Paired Samples Correlations

	N	Correlation	Sig.
VR_s - LayCAD_s	6	.047	.930

Paired Samples Test

	Paired Differences				
				95% Confidence Interval of the Difference	
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper
VR_s - LayCAD_s	-15.83333	15.35470	6.26853	-31.94710	.28043

Paired Samples Test

	t	df	Sig. (2-tailed)
VR_s - LayCAD_s	-2.526	5	.053

Figure 9: Paired sample T-test comparing time to completion of users measuring Distance 1 in VR and in LayCAD.

B CALCULATION TABLE SYSTEM USABILITY SCALE

SYSTEM USABILITY SCALE (SUS)

VR system for Factory Layout Planning											
PARTICIPANT	#	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
11	User's Scale Position	4	3	3	2	4	2	4	4	2	5
12	User's Scale Position	4	4	4	3	3	3	5	2	4	4
21	User's Scale Position	3	2	3	4	3	3	4	2	4	4
22	User's Scale Position	4	2	4	4	3	3	3	3	3	4
31	User's Scale Position	4	2	4	3	4	2	2	1	4	2
32	User's Scale Position	3	2	3	4	2	3	4	2	4	2
41	User's Scale Position	5	3	3	1	3	3	5	1	4	1
42	User's Scale Position	5	2	4	4	5	2	4	2	4	2
51	User's Scale Position	3	2	4	4	4	3	4	2	3	4
52	User's Scale Position	4	2	4	5	3	3	5	2	5	2
61	User's Scale Position	5	1	4	2	5	1	5	1	5	1
62	User's Scale Position	5	3	4	3	3	2	4	2	4	2
71	User's Scale Position	4	1	4	1	4	2	5	2	5	1
72	User's Scale Position	4	2	5	1	4	2	3	2	4	2
81	User's Scale Position	4	1	5	1	2	1	4	1	5	1
82	User's Scale Position	3	1	4	2	4	2	4	2	3	1

PARTICIPANT	#	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Individual SUS Score
11	SUS Contribution	3	2	2	3	3	3	3	1	1	0	52.5
12	SUS Contribution	3	1	3	2	2	2	4	3	3	1	60.0
21	SUS Contribution	2	3	2	1	2	2	3	3	3	1	55.0
22	SUS Contribution	3	3	3	1	2	2	2	2	2	1	52.5
31	SUS Contribution	3	3	3	2	3	3	1	4	3	3	70.0
32	SUS Contribution	2	3	2	1	1	2	3	3	3	3	57.5
41	SUS Contribution	4	2	2	4	2	2	4	4	3	4	77.5
42	SUS Contribution	4	3	3	1	4	3	3	3	3	3	75.0
51	SUS Contribution	2	3	3	1	3	2	3	3	2	1	57.5
52	SUS Contribution	3	3	3	0	2	2	4	3	4	3	67.5
61	SUS Contribution	4	4	3	3	4	4	4	4	4	4	95.0
62	SUS Contribution	4	2	3	2	2	3	3	3	3	3	70.0
71	SUS Contribution	3	4	3	4	3	3	4	3	4	4	87.5
72	SUS Contribution	3	3	4	4	3	3	2	3	3	3	77.5
81	SUS Contribution	3	4	4	4	1	4	3	4	4	4	87.5
82	SUS Contribution	2	4	3	3	3	3	3	3	2	4	75.0

AVG	69.8
SD	13.3

Figure 10: Table calculating SUS from the user's answers to the questions.

C PRE-SESSION QUESTIONNAIRE

Age Integer
Gender ☐Female ☐Male ☐Prefer not to say ☐Other (short text answer)
Country of birth Short text answer
Working title Short text answer
How often do you use a computer? ☐Daily ☐Weekly ☐Monthly ☐Occasionally ☐Never
Preferred language in computer systems ☐English ☐Swedish ☐Other (short text answer)
How often do you physically meet up with colleges or suppliers, discussing around a printed 2D map/drawing?
☐Daily ☐Weekly ☐Monthly ☐Occasionally ☐Never (This question concerns the transition from physical to digital tools)
Did you ever use a Virtual Reality headset before? ☐Yes ☐No ☐Maybe
How often do you work with Virtual Reality? ☐Daily ☐Weekly ☐Monthly ☐Occasionally ☐Never
Are you likely is it that you get motion sickness? I do not feel motion sickness **1 - 2 - 3 - 4 - 5** I am sensitive to motion sickness
(If you have tried VR applications, you can think did you feel discomfort easily when using them. If you have not experience in VR, think of feeling sick in a car for example.)
Did you follow a course in LayCAD? ☐Yes ☐No
How many years of experience do you have in using LayCAD? Integer
How often do you work with LayCAD? ☐Daily ☐Weekly ☐Monthly ☐Occasionally ☐Never

D POST-SESSION QUESTIONNAIRE[5]

1. I think that I would like to use this system frequently Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
2. I found the system unnecessarily complex Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
3. I thought the system was easy to use Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
4. I think that I would need the support of a technical person to be able to use this system Strong. D. **1-2-3-4-5** Strong. A.
5. I found the various functions in this system were well integrated Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
6. I thought there was too much inconsistency in this system Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
7. I would imagine that most people would learn to use this system very quickly Strongly Disagree **1-2-3-4-5** Strongly Agree
8. I found the system very cumbersome (=taggigt/besvärlig) to use Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
9. I felt very confident using the system Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
10. I needed to learn a lot of things before I could get going with this system Strongly Disagree **1-2-3-4-5** Strongly Agree
***11.* I felt that I could see and understand the other participant** Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
***12.* I felt that the other participant could see and understand me** Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
***13.* Moving around the virtual environment was easy** Strongly Disagree **1 - 2 - 3 - 4 - 5** Strongly Agree
***14.* I was able to anticipate what would happen next in response to the actions that I performed** S.D. **1 - 2 - 3 - 4 - 5** S.A.

[5] *QUESTION 11 TO 14 ARE NOT PART OF THE SUS, BUT ARE CONCERNING COLLABORATION*