Just Undo It: Exploring Undo Mechanics in Multi-User Virtual Reality

Julian Rasch
LMU Munich
Munich, Germany
julian.rasch@ifi.lmu.de

Florian Perzl
LMU Munich
Munich, Germany
florian.perzl@campus.lmu.de

Yannick Weiss
LMU Munich
Munich, Germany
yannick.weiss@ifi.lmu.de

Florian Müller
LMU Munich
Munich, Germany
florian.mueller@ifi.lmu.de

Figure 1: In this paper we explore the influence of different implementations of undo functionalities for multi-user Virtual Reality in varying modes of collaboration on the users' effectiveness, efficiency and social connection.

ABSTRACT

With the proliferation of VR and a metaverse on the horizon, many multi-user activities are migrating to the VR world, calling for effective collaboration support. As one key feature, traditional collaborative systems provide users with undo mechanics to reverse errors and other unwanted changes. While undo has been extensively researched in this domain and is now considered industry standard, it is strikingly absent for VR systems in research and industry. This work addresses this research gap by exploring different undo techniques for basic object manipulation in different collaboration modes in VR. We conducted a study involving 32 participants organized in teams of two. Here, we studied users' performance and preferences in a tower stacking task, varying the available undo techniques and their mode of collaboration. The results suggest that users desire and use undo in VR and that the choice of the undo technique impacts users' performance and social connection.

CCS CONCEPTS

• **Human-centered computing** → *Collaborative and social computing*; **Virtual reality**; *HCI theory, concepts and models*; *Interaction techniques.*

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05…$15.00
https://doi.org/10.1145/3613904.3642864

KEYWORDS

SocialVR, Undo, CSCW, Multi-User, Connectedness, Virtual Reality

ACM Reference Format:
Julian Rasch, Florian Perzl, Yannick Weiss, and Florian Müller. 2024. Just Undo It: Exploring Undo Mechanics in Multi-User Virtual Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3613904.3642864

1 INTRODUCTION

With increasing numbers of social Virtual Reality (VR) applications, both in the personal entertainment domain [50] as well as the professional training and qualification domain [2, 43], Computer-Supported Cooperative Work (CSCW) in VR increases as well. In these shared VR environments users can co-exist or collaborate with others. As a consequence, individual as well as collaborative tasks intertwine in many scenarios in the shared VR and can affect other users of the environment. Like in the physical world, small errors caused by one user or the technical system (e.g. Tracking Errors) can destroy the progress of others and lead to undesired changes in the VR environment.

In traditional computer systems, one established solution to this challenge are undo actions [3, 48], reverting the system or parts of the system to a previous state. While originally this is a single-user feature, subsequent work from CSCW extends this *personal or individual undo* to be suitable for collaborative work as well, by extending the range of effect beyond the individual actions. While many other established mechanics from traditional user interface (UI)s, like pointing [36], selecting [4], and manipulating [39], are well researched and established in collaborative VR [25, 37], the usage of undo functions remains niche in VR applications. Besides only a few recent exceptions [33, 54] addressing specific use cases, undo features in VR are not investigated by researchers yet.

In this paper, we revisit questions from CSCW to identify how to design undo mechanics for multi-user VR and how users will utilize them. Based on a thorough review of related work we compare three undo techniques for VR object manipulation, namely IndividualUndo, SelectiveUndo, WorldUndo and a NoUndo baseline, each with different effect range in two different modes of collaboration. In our user study (n=32) 16 pairs of participants work Divided and Collaborative on a tower construction task with one UndoTechnique or the NoUndo baseline available. Our aim is to determine users' approval of the proposed undo concepts in the different CollaborationModes and their willingness to use these features in the future. Additionally, we seek to understand how the user experience of the UndoTechniques changes based on the type of cooperation, considering factors such as efficiency, recoverability, users' willingness to take risks, intuitiveness, user expectations, and enjoyment.

The contribution of our paper is two-fold. First, we contribute the results of a controlled experiment exploring the strengths and weaknesses of the proposed UndoTechniques for different CollaborationModes. Second, based on our results, we derive design recommendations for the implementation of multi-user UndoTechniques in VR. Our results indicate, that like in traditional computer systems users appreciate the availability of UndoTechniques, as it helps them recover from mistakes and revert unwanted changes in the VR environment. UndoTechniques like the WorldUndo, which allow to undo all changes in the VR environment independent of the source of the change, increases the connection between users in the same VR environments, but also leads to higher reciprocal disturbance of the users.

2 RELATED WORK

To situate our work, we give an overview of different undo mechanics established in traditional computer systems and summarize the state of collaboration in VR environments.

2.1 Undo Actions in Collaborative Systems

Today, undo mechanics are one of the standard features of many computer systems and are recommended for well-designed UIs of all kinds [46]. The feature is present in traditional desktop computers, same as in mobile devices [28, 45]. Teitelman [48] introduced one of the first instances of an undo mechanic as a feature for the *programmer's assistant* more than 50 years ago. 17 years after this, Yang [53] states that "*Most sophisticated interface systems should be provided with an undo support*" and identifies the undo mechanic as one of three core support features for user interfaces to recover from errors and unwanted situations, besides stops and escapes. Further, Myers [32] work on visual programming paradigms points out the relevance of undo mechanisms in graphical user interfaces (GUIs).

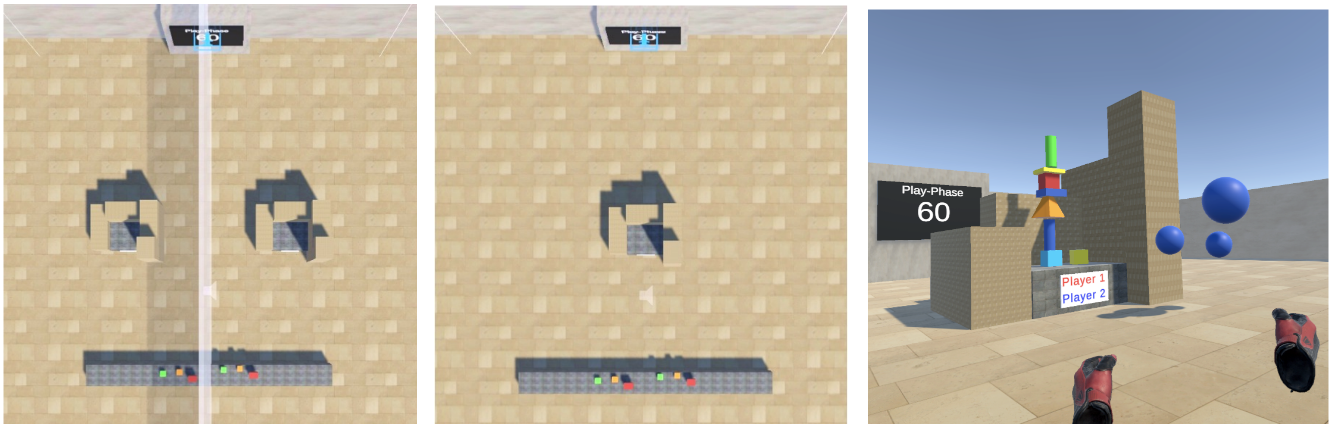
For multi-user systems Abowd and Dix [1] emphasize the need for undo support in the context of synchronous group editors, where multiple streams of activity can easily induce errors. In this context Abowd and Dix [1] discuss emerging problems regarding different roles and ownership with respect to objects in the system and identify two models of undo for multi-user systems. A *local undo* only affects changes made by the users themselves, and a *global undo* affects all changes made to the system. Extending on this, Prakash and Knister [40] propose a *selective undo*, allowing to apply undo only to certain objects in the system. The associated selection can be based on any attribute of the objects, e.g., the identity of the previous user, the manipulation time, or the region of the object. In their subsequent work Prakash and Knister [41] propose the *region undo* besides the per-user history undo and a multiple-operation undo as three practical examples for their *selective undo*. Also, more recent work, e.g., by Cass et al. [10] advocates for the use of selective undo mechanisms to allow users to also undo actions in a non-linear fashion and maintain dependencies. In the context of large interactive surfaces, Seifried et al. [44] investigate different undo techniques for co-located collaborative workspaces on interactive screens. Here, as well, the authors resort to the three established undo concepts *global*, *personal*, and *selective or regional undo*, each with a different range of effect. These examples underline that undo mechanics are well established and researched in many computer systems, however, their applicability and benefit for VR systems remains unexplored.

2.2 Collaboration in VR

In recent years, we have seen increasing use of VR as a foundation for novel collaborative multi-user experiences in research and industry. This resulted in a large variety of collaborative VR applications from entertainment [31, 50] to professional usage [19] and groupware [14, 20]. Designers and Engineers use collaborative VR applications to iterate over designs in Computer Aided Design (CAD) application [26, 29, 47] and Collaborative Learning Environments [23, 35] offer new ways to engage with learning content. While some applications focus more on transferring the workflow from 2D screens to VR, others make use of the new possibilities, e.g., by changing the size of the collaborating users [52]. Other research explores the use of multi-user VR in the context of psychotherapy [30, 34, 49] or for collaborative immersive visual analysis of multidimensional data [9]. Lastly, there are multiple examples for multi-user VR-trainings from the industrial [2] and medical domain [27, 43]. As a reoccurring pattern, here we can observe two main collaboration types known from traditional computer systems again: divided and collaborative work. These two extremes on an actual continuous spectrum are used, e.g., by Xia et al. [52] and Pinho et al. [37] as a representative distinction for different modes of VR collaboration.

Many of the interaction techniques to support collaboration in traditional multi-user collaboration tools, such as synchronized pointing, selecting, and manipulating, made the leap into the VR domain already [37]. While these techniques have been thoroughly studied and are now part of many collaborative VR systems in the industry, undo mechanics in these systems are notably absent to a large part. This is in stark contrast to participants in VR user studies expressing their desire for undo techniques if not present [35] and researchers encouraging future work to investigate undo mechanics [16] for VR as well.

Recently, research started exploring aspects of undo mechanics in VR, in order to revisit questions from CSCW and the adaptability to the domain of VR. Zhang et al. [54] propose a git-like version control system for collaborative content creation in VR to identify similarities and differences to desktop-based systems. Friedman



(a) Top view for DIVIDED task (b) Top view for COLLABORATIVE task (c) Participants view for COLLABORATIVE task

Figure 2: The task environment of our user study. (a) Top view of the task environment for the DIVIDED task. A transparent wall separates the game area to prevent interference between two players. Each player has their own stacking area as well as spawn area of the building blocks. (b) Top view of the task environment for the COLLABORATIVE task. The players share one stacking and block spawning area. (c) Participants perspective during the COLLABORATIVE task. The center of the image shows the stacking area, on which participants have to place the building blocks. The stairs around the stacking area can be used by participants to reach the top of the tower also for increased heights. The avatar of the other player is represented by one sphere for the head and two smaller spheres for the hands. In the back a "screen" displays the remaining time to construct the tower. Not in the picture is the area to pick up the building blocks. In this case, participants already stacked 7 blocks successfully.

et al. [18] implement a VR time travel experience to study users' responses to the illusion of time travel in the context of a virtual trolley problem [17]. And Müller et al. [33] investigate the effects of an undo mechanic for point&teleport-steps on the user behavior in an explorative navigation task. Further, individual single-user VR applications like, e.g., *Tilt Brush*[1] provide undo features.

Beyond the specific focus on VR interfaces, previous work also points out the need for undo actions in Augmented Reality (AR) [12]. Here, Piumsomboon et al. [38] discuss potential gestures for AR interfaces for 40 common tasks in AR including the undo function and Kaufmann [24] maps undo functions to a hand-held tracked panel in an AR learning environment.

While these works are exciting and deliver important contributions to the community, they do not offer insights into the transferability of established undo mechanics to VR applications. In particular, to the best of our knowledge, there is no prior work on undo mechanics for multi-user VR applications. This is supported by Ens et al. [15], who investigated the status of Mixed Reality (MR) groupware and did not find the term *undo* a single time.

With the increasing number of collaborative multi-user applications for VR on the one hand and the established effectiveness of undo mechanics as a support feature for traditional (collaborative) computing systems on the other hand, we see the need for a systematic understanding of the transferability of undo actions to the domain of collaborative VR. To assess which characteristics of traditional undo mechanics can or cannot be adopted to these use cases is a complex and nontrivial endeavor as it requires taking into account the intricacies of multi-user VR applications. As one key element unique to VR, we see the blending of users' actions to the computer system and changes in the VR environment itself, resulting in a strong intertwining of the two. Consequently, undo actions of users result in a change in the users' surroundings. With this work, we aim to contribute to a better systematic understanding of characteristics of undo actions in VR.

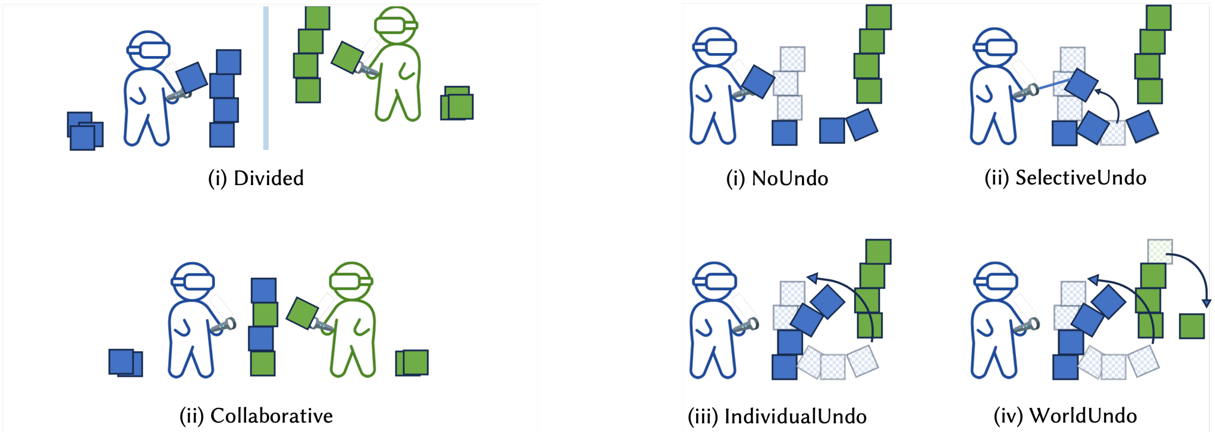
[1]https://www.tiltbrush.com

3 METHODOLOGY

The analysis of previous work (see Section 2) revealed that undo techniques are an established and well-researched means in traditional computer systems to support individual and collaborative activities. On the other hand, we found a surprising lack of systematic insights into the use and best practices of undo techniques for VR, especially for collaborative applications. Based on the findings from related work and current practices in VR, we formulate the following research hypotheses:

H1: The availability of undo in VR increases users' performance.
H2: The availability of undo in VR increases the user experience.
H3: Undo techniques that affect all users increase the social connectedness of users in VR.
H4: Undo techniques that affect all users increase the interference of users during non-collaborative work in VR.

Many collaborative interaction techniques from traditional computer systems, like synchronized pointing, selecting, and manipulating, are adopted to collaborative VR already [37]. With H1 and H2, we investigate if the implementation of undo mechanics as a support feature for collaborative VR applications can assist users similar to traditional (collaborative) computer systems [46]. With H3 and H4, we study the effects different established forms of undo [1, 40] have on the relation between two users.



(a) CollaborationMode (b) UndoTechnique

Figure 3: (a) The 2 levels of the independent variable CollaborationModes with the (i) Divided case on top and the (ii) Collaborative case bottom. (b) The 4 levels of the independent variable UndoTechnique with the baseline condition (i) NoUndo as well as the 3 different UndoTechniques (ii) SelectiveUndo, (iii) IndividualUndo and (iv) WorldUndo.

To test our hypotheses, we explore the influence of 1) the mode of collaboration between users to account for different task scenarios common to co-located collaborative work [37, 44, 52] and 2) the range of effect of the undo technique as established in CSCW [1, 40].

3.1 Design and Task

For our user study, we designed a construction task for two participants in VR. As the canvas for the undo techniques, we selected three-dimensional primitive shapes as suitable artifacts representative for the three-dimensionality of the VR domain. These afford all interactions inherent to VR and shape the environment itself. We instructed the participants to build a tower with blocks of different basic 3D geometries while being in the same virtual space as another participant, as shown in Figure 2. To account for different scenarios, participants worked on the task either together or alone, depending on the condition. When working alone, we adjusted the scene as shown in Figure 2a, to provide a stacking area for each participant and placed a transparent wall separating the game space of the virtual world to avoid unintended interference between the participants. In both cases, the goal is to build a tower as high as possible within a task time of 4 minutes. We chose this time frame as it allowed participants enough time to reach a critical stability of the tower while limiting the total study time to a maximum of 90 minutes. During the task participants moved the blocks from a spawning location to a target location, where they stacked them. Once they placed a block on the tower, a new block appeared in the spawning location, ensuring sufficient blocks throughout the condition. With increasing height of the tower, the towers' stability decreases naturally, resulting in an eventual collapse. We informed participants that the height of the tower at the end of the task time is relevant to ensure participants approach the task carefully. We chose a minimalistic avatar design to visualize the users to reduce distractions, and encourage participants to focus on the task. We opted for a controller-based input device for the input modality as the current de facto standard for VR interactions. We chose this task for the following reasons. First, it uses the standard VR manipulation techniques (translate, rotate, grab, release) and requires locomotion. Second, the task can be solved individually as well as collaboratively. Third, it changes the environment of the participant. And fourth, it allows us to measure the progress and success of participants. The study design was approved by our institution's ethics committee.

3.2 Independent Variables

As described at the beginning of this section, we investigate the effects of the CollaborationMode and available UndoTechnique to discover the influence on the dependent variables. We vary these two independent variables with the following levels:

CollaborationMode: To attribute for different modes of collaboration, we chose two levels as shown in Figure 3a:
- ***Divided*** Players build their individual tower in their separate game space.
- ***Collaborative*** Players build a tower together in a joint game space.

UndoTechnique: Based on the concept described before, we investigate three different levels of undo functions and a baseline. The UndoTechniques are shown in Figure 3b (i)-(iv):
- ***NoUndo*** This level represents the current de facto standard VR interactions without any undo function.
- ***SelectiveUndo*** Players can undo object manipulations of the object selected by the ray interactor.
- ***IndividualUndo*** Players can undo their own object manipulations in a linear way.
- ***WorldUndo*** Players can undo all object manipulations in the scene in a linear way, also affecting the actions of the other player.

Based on the presented levels of the two independent variables, participants experience 8 different conditions throughout the study. To prevent learning effects and reduce order-effects and carry-over-effects between the conditions, we counterbalance the conditions using the *balanced latin square design*.

We chose not to include the respective other user's avatar in any range of effect for the following reasons. First, we want not to objectify users. Second, we want to maintain individuals' freedom of movement as users do not like to be moved by others' [42]. Third, as mentioned before, the aspect of locomotion is covered already [33], although only in a single-user context. This results in a 2 x 4 space shown in Figure 3, which we investigate in the context of our controlled experiment, described in more detail in the following.

3.3 Dependent Variables

To assess the influence of the CollaborationMode and UndoTechnique, we survey the influence on the following measures:

IOS Score The Inclusion of Other in the Self (IOS) is a single-item questionnaire proposed by Aron et al. [5] and measures the perceived psychological closeness to other persons or groups. With this measure, we aim to study H3.

GEQ Score The Behavioural Involvement Component of the Social Presence Module in the Game Experience Questionnaire (GEQ) by IJsselsteijn et al. [22] measures how much attention participants pay to their counterpart. The mentioned component uses six of the 17 items of the Social Presence Module of the GEQ and uses a five-point Likert scale for answers. We include this measure to study H3 and H4.

RTLX The Raw Nasa-TLX (RTLX) [21] measures how demanding each condition is. The six items of the RTLX use a 0-100 scale for answers to study H1 and H2.

Custom Questionnaire The questionnaire on an established five-point Likert scale holds eight items regarding the feeling of DisturbingOthers, FeltDisturbed, Control, Frustration, Success, Recover, Enjoyment and DesiredFutureUsage. The results section presents the complete statement for each item. We chose this custom questionnaire in addition to the established questionnaires to allow for specific insights for H1, H2, and H4 for the study on hand.

Besides the survey data, we also log the users' interactions throughout each condition to assess the influence of the CollaborationMode and UndoTechnique. During the experiment, we logged the following data:

NumberOfGrabs The total number of grab interactions a participant performed during a condition as an efficiency and engagement measure. With this measure, we want to understand if the availability of an undo feature affects the frequency of the conventional grab interaction and study H1.

NumberOfUndos The total number of undo operations a participant performed during a condition as a measure of participants' acceptance and usage. With this measure, we want to understand if participants use an undo feature if available and study H1, H2, and H4.

TowerHeight The final height of the constructed tower in meters at the end of a condition as a measure of the performance. With this measure, we aim to detect changes in participants' performance on the task for different undo features and study H1 and H4.

We excluded time as an efficiency measure, as we needed a fixed time per condition to study the effects on participants' willingness for last-minute changes and create time pressure towards the end of each condition.

3.4 Apparatus

We implemented the virtual environment and interaction techniques as detailed above as a multi-player VR application and deployed it on two connected computers with two head-mounted display (HMD)s using a shared tracking space. We developed the study design and its functionalities using the Unity Version 2021.3.14f. We used two HTC Vive Pro as VR devices, each consisting of a HMD, two HTC motion controllers, and two base stations for tracking. We executed the application on two VR computers with the following specifications: Intel Core i7-10700, 16 GB RAM, NVIDIA GeForce RTX 2060 SUPER. For connecting the individual users' applications, we used the multi-player networking framework Photon Pun[2] for Unity, which supports room creation, matchmaking, and event-based communication for real-time scenarios. To connect to the internet and minimize networking complications, both computers were connected to the same network via LAN cable.

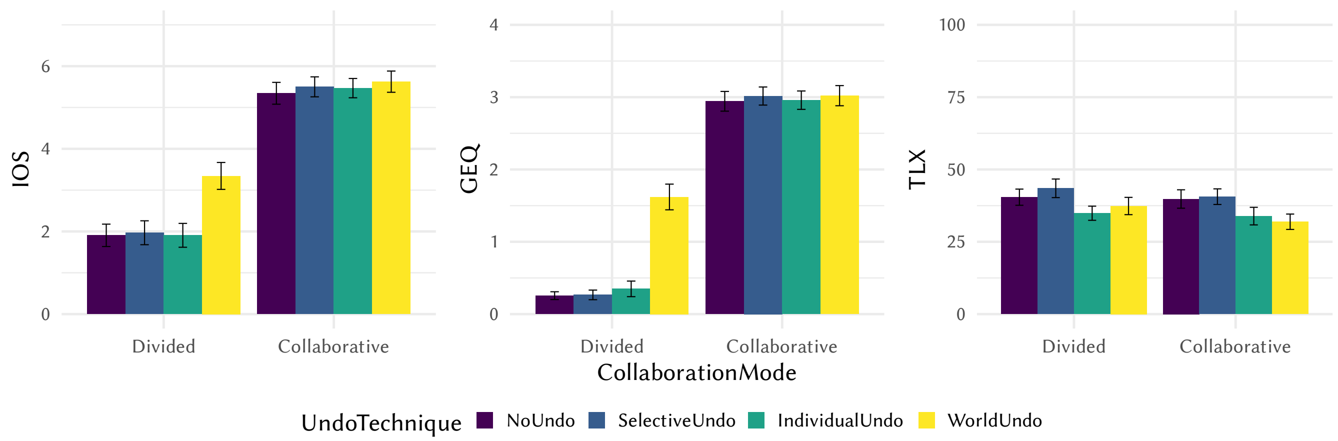
To implement the UndoTechniques, we store the position and rotation of each object together with a time-stamp and the manipulating agent in a data frame throughout the runtime of the VR application. When a participant uses the undo feature, we access the previous states of objects and assign the position accordingly. For the SelectiveUndo, we use a ray interactor for selection. Depending on the current condition, we access all objects, the selected object, or all objects manipulated by one player from the data frame.

We used a total area of 3.2 x 5.5 meters as physical space and set up two individual VR spaces with active Steam Guard to prevent collision with walls or the other participant.

3.5 Procedure

After welcoming the participants, we provide an overview of the study's objectives and their tasks. Participants are then asked to sign a consent form to authorize the collection and processing of their data. They then complete a pre-survey covering demographics, experience with virtual reality, and familiarity with the other participants. Next, we demonstrate fundamental VR controls for teleportation and object interaction, allowing participants a few minutes to practice before the study begins. We then explain their primary task in VR, namely constructing a tower using building blocks within a four-minute time frame, aiming to maximize its height. We emphasize that building blocks should remain stationary when time expires to contribute to their score. We then let participants enter a tutorial scene to allow them to familiarize themselves with basic controls, ensuring they understand the task and are able to perform it. Additionally, before each test condition, participants have one minute to familiarize themselves with the provided undo function of the condition. Throughout the conditions, participants can talk with each other. After each of the eight conditions, they are

[2]https://www.photonengine.com/en-US/Photon



(a) IOS score **(b) GEQ score** **(c) RTLX score**

Figure 4: The mean results for (a) IOS, (b) GEQ, and (c) RTLX as a bar chart plot. The error bars indicate the standard error.

asked to complete a questionnaire on a PC. After finishing all conditions, participants take a short break before filling out a post-study survey to evaluate and provide feedback on the undo functions relative to each other. After this, we collect further qualitative feedback from the participants. Finally, we provide compensation for their participation. Throughout the study of 90 minutes participants experience 8 different conditions for 4 minutes each in a counterbalanced order.

3.6 Participants

We recruited 32 participants (18 female, 14 male) with a mean age of 26.56, ranging from 19 to 58. Of the participants, 18 knew the other participant, while 14 did not know the other participant. 10 participants stated to be a *experienced VR user*, 10 stated to be a *sporadic VR user* and 12 had *no VR Experience* at all.

3.7 Analysis

For analyzing the non-parametric data, we applied the Aligned Rank Transform (ART) proposed by Wobbrock et al. [51]. For significant results, we follow the ART-C procedure suggested by Elkin et al. [13]. We report the generalized ETA squared η_G^2 as a measure of the effect and classify it in alignment with Bakeman [6]. Here, we use suggestions by Cohen [11] for small (> .0099), medium (> .0588), or large (> .1379) effect size.

For analyzing the parametric data, we tested the data with Shapiro-Wilk's and Mauchly's tests for normality of the residuals and sphericity assumptions. When the assumption of sphericity was violated, we used the Greenhouse-Geisser method to correct the tests. We used two-way repeated-measures ANOVAs to identify significant effects and applied Bonferroni-corrected t-tests for post-hoc analysis. If normality was violated, we performed a non-parametric analysis as described before.

To analyze the count data from the log files, we fitted Poisson regression models and applied Type III Wald chi-square tests for significance testing. Here, we used the Tukey method for p-value adjustments.

4 RESULTS

This section presents the results of our controlled experiment.

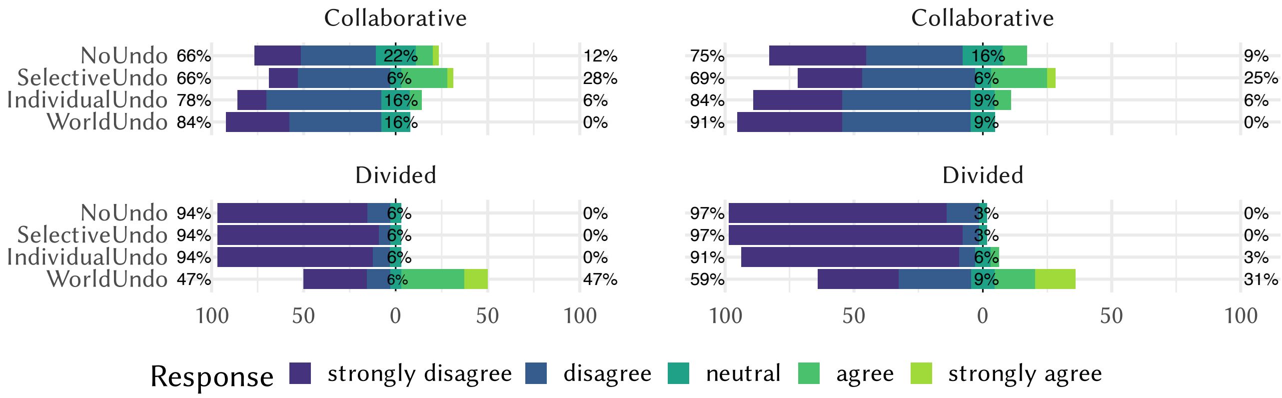
4.1 IOS Score

For the single-item IOS questionnaire, we took the participants' ratings and compared them directly as proposed by Aron et al. [5]. We analyzed the IOS score as a measure of social connectedness between participants. Here, we found mean values ranging from $M = 5.62, SD = 1.45$ (COLLABORATIVE, WORLDUNDO) to M= 1.91, SD = 1.63 (DIVIDED, INDIVIDUALUNDO) shown in Figure 4. The ART ANOVA showed a significant ($F_{1,31} = 146.70$, $p < .001$) main effect for the COLLABORATIONMODE on the IOS scale with a large (η_G^2 = 0.82) effect size. Post-hoc tests revealed significantly ($p < .001$) higher ratings for COLLABORATIVE ($M = 5.48, SD = 1.40$) compared to DIVIDED ($M = 2.28, SD = 1.76$).

We also found a significant ($F_{3,93} = 12.52$, $p < .001$) main effect for the UNDOTECHNIQUE on the IOS scale with a large (η_G^2 = 0.28) effect size. Post-hoc tests revealed significantly higher ratings for WORLDUNDO ($M = 4.48, SD = 2.01$) compared to all other levels (NOUNDO: $M = 3.62, SD = 2.29, p < .001$, SELECTIVEUNDO: $M = 3.73, SD = 2.32, p < .001$, INDIVIDUALUNDO: $M = 3.69, SD = 2.32, p < .01$).

Further, we found interaction effects ($F_{3,93} = 5.56$, $p < .01$) with a large (η_G^2 = 0.15) effect size. While we could not find differences between the UNDOTECHNIQUES in the COLLABORATIVE case (all $p > .05$), WORLDUNDO received significantly higher ($p < .001$) ratings compared to all other levels in the DIVIDED case.

These results support H3, showing that UNDOTECHNIQUES with a global range of effect like WORLDUNDO can increase the social connectedness of the users.



(a) DisturbingOthers

(b) FeltDisturbed

Figure 5: Participants' responses regarding (a) DisturbingOthers and (b) FeltDisturbed on a 5-point Likert scale. The percentage number indicates the proportion of the answers for negative, neutral, and positive responses.

4.2 GEQ Score

For the GEQ, we evaluated the Behavioral Involvement Component as an average of its items, according to the scoring guideline [22]. Evaluating the Behavioral Involvement Component of the GEQ Social Presence Module, the analysis yielded values ranging from $M = 3.02, SD = 0.78$ (Collaborative, WorldUndo) to $M = 0.25, SD = 0.3$ (Divided, NoUndo).

Using the ART ANOVA, we found a significant ($F_{1,31} = 358.836$, $p < .001$) main effect for the CollaborationMode on the GEQ rating with a large ($\eta^2_G = 0.92$) effect size. Post-hoc tests revealed significantly ($p < .001$) higher ratings for Collaborative ($M = 2.98, SD = 0.74$) compared to Divided ($M = 0.62, SD = 0.85$). We also found a significant ($F_{3,93} = 18.09$, $p < .001$) main effect for the UndoTechnique on the GEQ rating with a large ($\eta^2_G = 0.36$) effect size. Here, post-hoc tests revealed significantly ($p < .001$) higher ratings for WorldUndo ($M = 2.32, SD = 1.14$) compared to all other levels. (NoUndo: $M = 1.60, SD = 1.47$, SelectiveUndo: $M = 1.64, SD = 1.50$, IndividualUndo: $M = 1.65, SD = 1.47$).

Again, we found a significant ($F_{3,93} = 17.83$, $p < .001$) interaction effect with a large ($\eta^2_G = 0.36$) effect size. While we could not find differences between the UndoTechniques in the Collaborative case ($p > .05$), WorldUndo received significantly ($p < .001$) higher ratings compared to all other levels in the Divided case.

These results support H3 and H4, again showing that UndoTechniques with a global range of effect like WorldUndo can increase users' social connectedness and mutual interference.

4.3 Raw TLX Score

We calculated the RTLX score as proposed by Hart [21]. Evaluating the RTLX, we found values ranging from $M = 43.5, SD = 18.2$ (Divided, SelectiveUndo) to $M = 31.9, SD = 15.0$ (Collaborative, WorldUndo). Our ART ANOVA showed a significant ($F_{3,93} = 9.34$, $p < .001$) main effect for the UndoTechnique on the RTLX score with a large ($\eta^2_G = 0.23$) effect size. Here, post-hoc tests revealed significantly higher ratings for NoUndo ($M = 40.1, SD = 16.8$) compared to IndividualUndo ($M = 34.4, SD = 15.5$) and WorldUndo ($p < .01$) ($M = 34.6, SD = 16.1$) as well as significantly higher ratings for SelectiveUndo ($M = 42.0, SD = 16.7$) compared to IndividualUndo ($p < .001$) and WorldUndo ($p < .01$). We could not find a significant main effect ($F_{1,31} = 3.91$, $p > .05$) for the CollaborationMode nor interaction effects ($F_{3,93} = .41$, $p > .05$) between the two independent variables.

These results only partially support H1 and H2 since not all UndoTechniques positively affected the performance and user experience. We discuss this in more detail in Section 5.

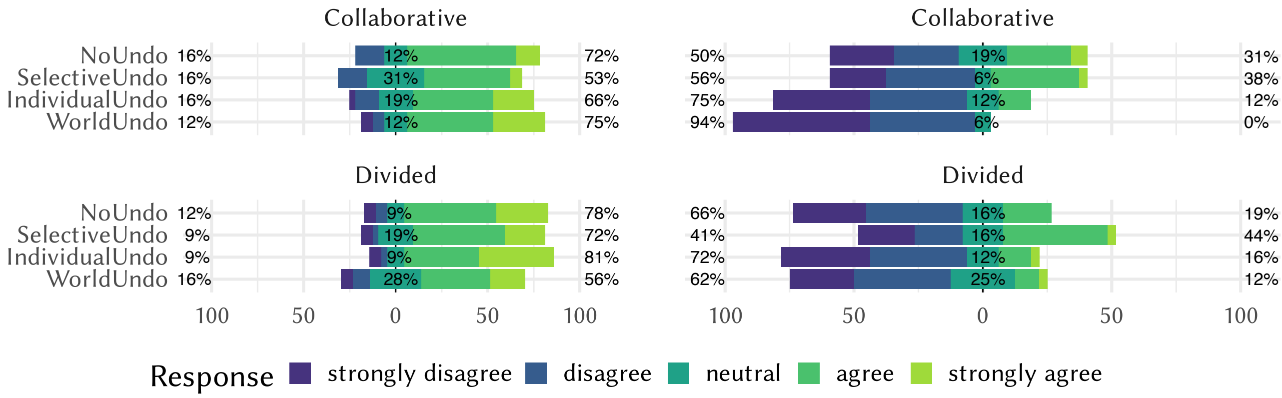
4.4 Custom Questionnaire

In the following, we present the results of the ART ANOVA of our custom questionnaire.

4.4.1 DisturbingOthers "I felt that I disturbed the other player". We found a significant ($F_{1,31} = 33.86$, $p < .001$) main effect for the CollaborationMode on the DisturbingOthers rating with a large ($\eta^2_G = 0.52$) effect size. Post-hoc tests revealed significantly higher ratings for Collaborative compared to Divided ($p < .001$). We also found a significant ($F_{3,93} = 5.02$, $p < .01$) main effect for the UndoTechnique on the DisturbingOthers rating with a medium ($\eta^2_G = 0.13$) effect size. Post-hoc tests revealed significantly higher ratings for WorldUndo compared to NoUndo ($p < .05$) and IndividualUndo ($p < .01$).

Further, we found an interaction effect ($F_{3,93} = 24.07$, $p < .001$) with a large ($\eta^2_G = 0.43$) effect size. While we could not find differences between the UndoTechniques in the Collaborative case ($p > .05$), WorldUndo received significantly ($p < .001$) higher ratings compared to all other levels in the Divided case.

These results support H4, showing UndoTechniques with a global range of effect like WorldUndo can increase mutual interference of the users.



(a) Control (b) Frustration

Figure 6: Participants' responses regarding (a) Control and (b) Frustration on a 5-point Likert scale.

4.4.2 FeltDisturbed *"I felt that the other player disturbed me".*

We found a significant ($F_{1,31} = 12.86$, $p < .01$) main effect for the CollaborationMode on the FeltDisturbed rating with a large ($\eta^2_G = 0.29$) effect size. Post-hoc tests revealed significantly higher ratings for Collaborative compared to Divided ($p < .001$). We also found a significant ($F_{3,93} = 7.42$, $p < .001$) main effect for the UndoTechnique on the FeltDisturbed rating with a large ($\eta^2_G = 0.19$) effect size. Post-hoc tests revealed significantly higher ratings for WorldUndo compared to NoUndo ($p < .01$) and IndividualUndo ($p < .001$).

Again, we found interaction effects ($F_{3,93} = 12.46$, $p < .001$) with a large ($\eta^2_G = 0.28$) effect size. While we could not find differences between the UndoTechniques in the Collaborative case ($p > .05$), WorldUndo received significantly ($p < .001$) higher ratings compared to all other levels in the Divided case.

These results support H4, showing UndoTechniques with a global range of effect like WorldUndo can increase mutual interference of the users.

4.4.3 Control *"I felt in control over my surroundings".*

We found no significant ($p > .05$) main effect for CollaborationMode and UndoTechnique on the Control rating. However, we found significant ($F_{3,93} = 4.97$, $p < .01$) interaction effects with a medium ($\eta^2_G = 0.13$) effect size. However, post-hoc tests did not confirm ($p > .05$) significant differences between the groups, and therefore neither contradict nor support H1, H2, and H4.

4.4.4 Frustration *"I felt frustrated".*

We found a significant ($F_{3,93} = 7.09$, $p < .001$) main effect for the UndoTechnique on the Frustration rating with a large ($\eta^2_G = 0.18$) effect size. Post-hoc tests revealed significantly higher ratings for SelectiveUndo compared to IndividualUndo ($p < .01$) and WorldUndo ($p < .001$). We could not find a significant main effect ($F_{1,31} = .21$, $p > .05$) for the CollaborationMode nor interaction effects ($F_{3,93} = 1.55$, $p > .05$) between the two independent variables.

These results neither contradict nor support H1, H2, and H4.

4.4.5 Success *"I felt like I solved the task successfully".*

We found a significant ($F_{3,93} = 7.32$, $p < .001$) main effect for the UndoTechnique on the Success rating with a large ($\eta^2_G = 0.19$) effect size. Post-hoc tests revealed significantly higher ratings for IndividualUndo and WorldUndo compared to SelectiveUndo ($p < .001$). Again, we could not find a significant main effect ($F_{1,31} = .00$, $p > .05$) for CollaborationMode nor interaction effects ($F_{3,93} = 1.67$, $p > .05$) between the two independent variables.

These results only partially support H2, as not all UndoTechniques show a positive effect on the user experience. We discuss this in more detail in Section 5.

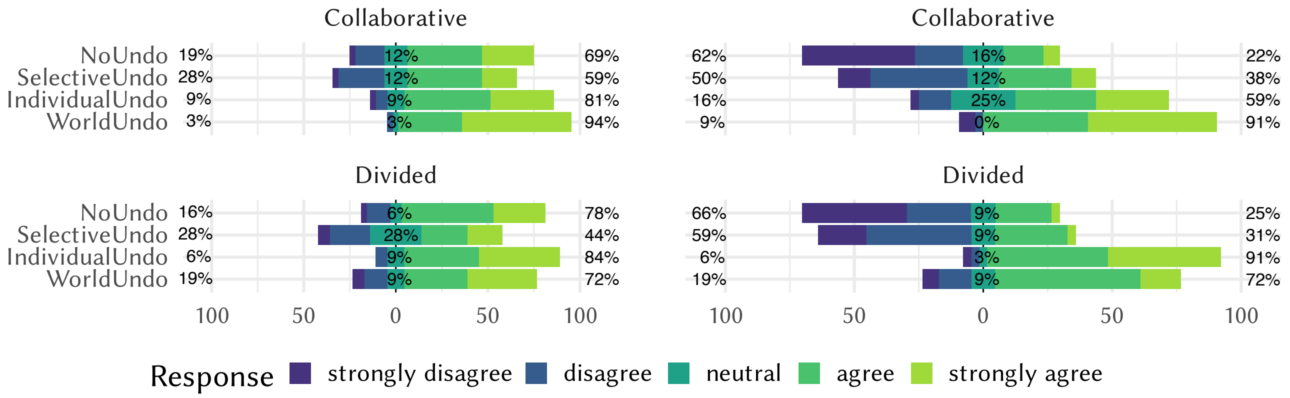
4.4.6 Recover *"I felt that I could easily recover from my mistakes".*

We found a significant ($F_{3,93} = 32.12$, $p < .001$) main effect for the UndoTechnique on the Recover rating with a large ($\eta^2_G = 0.50$) effect size. Post-hoc tests revealed significantly higher ratings for IndividualUndo and WorldUndo compared to SelectiveUndo ($p < .001$) as well as for IndividualUndo and WorldUndo compared to NoUndo ($p < .001$). We could not find a significant main effect ($F_{1,31} = 1.63$, $p > .05$) for CollaborationMode, but found a significant ($F_{3,93} = 5.46$, $p < .01$) interaction effect with a large ($\eta^2_G = 0.14$) effect size. While in the Collaborative case, the WorldUndo received higher ratings compared to IndividualUndo, in the Divided case, this inverts, resulting in higher ratings for IndividualUndo compared to the WorldUndo. However, these differences are not significant ($p > .05$).

These results partially support H1 and H2, as again, not all UndoTechniques show a positive effect on the performance and user experience. We discuss these results in more detail in Section 5.

4.4.7 Enjoyment *"I enjoyed using the undo feature".*

We found a significant ($F_{3,93} = 48.05$, $p < .001$) main effect for the UndoTechnique on the Enjoyment rating with a large ($\eta^2_G = 0.60$) effect size. Post-hoc tests revealed significantly higher ratings for IndividualUndo and WorldUndo compared to NoUndo and SelectiveUndo ($p < .001$). We could not find a significant main effect



(a) Success

(b) Recover

Figure 7: Participants' responses regarding (a) Success and (b) Recover on a 5-point Likert scale.

($F_{1,31} = 1.81$, $p > .05$) for CollaborationMode but found significant ($F_{3,93} = 8.67$, $p < .001$) interaction effects with a small ($\eta^2_G = 0.05$) effect size. While in the Collaborative case WorldUndo received non-significant higher ratings compared to IndividualUndo, again in the Divided case, this inverts, resulting in significantly ($p < .05$) higher ratings for IndividualUndo compared to WorldUndo.

These results partially support H2, as they are inconsistent across the UndoTechniques. We discuss this in more detail in Section 5.

4.4.8 DesiredFutureUsage "I want to use the undo feature in the future". Again, we found a significant ($F_{3,93} = 35.27$, $p < .001$) main effect for the UndoTechnique on the DesiredFutureUsage rating with a large ($\eta^2_G = 0.53$) effect size. Post-hoc tests revealed significantly higher ratings for IndividualUndo and WorldUndo compared to NoUndo and SelectiveUndo ($p < .001$). We could not find a significant main effect ($F_{1,31} = 3.47$, $p > .05$) for CollaborationMode. Again, we found significant ($F_{3,93} = 4.10$, $p < .01$) interaction effects with a medium ($\eta^2_G = 0.11$) effect size. While in the Collaborative case the WorldUndo received non-significant higher ratings compared to IndividualUndo, again in the Divided case, this inverts, resulting in significantly higher ($p < .05$) ratings for IndividualUndo compared to WorldUndo.

These results partially support H2, as only IndividualUndo and WorldUndo increased the user experience, while SelectiveUndo did not.

4.5 Number of Grab Actions

We fitted a Poisson regression model to analyze the NumberOfGrabs as an efficiency measure. Here, we found values ranging from $M = 21.4, SD = 8.42$ (Collaborative, WorldUndo) to $M = 30.8, SD = 11.2$ (Divided, NoUndo). Our analysis shows a significant ($\chi^2(1) = 27.01$, $p < .001$) main effect for the CollaborationMode. Post-hoc tests revealed significantly ($p < .001$) higher values for Divided ($M = 27.7, SD = 10.4$) compared to Collaborative ($M = 22.6, SD = 9.19$). We also found a significant ($\chi^2(3) = 21.25$, $p < .001$) main effect for the UndoTechnique. Here, post-hoc tests revealed significantly lower values for WorldUndo ($p < .001$) ($M = 23.4, SD = 9.50$) and IndividualUndo ($p < .01$) ($M = 24.2, SD = 9.97$) compared to NoUndo ($M = 27.4, SD = 11.3$). We could not find a significant ($\chi^2(3) = 2.29$, $p > .05$) interaction effect.

These results partially support H1, as only IndividualUndo and WorldUndo increased the performance, while SelectiveUndo did not.

4.6 Number of Undo Actions

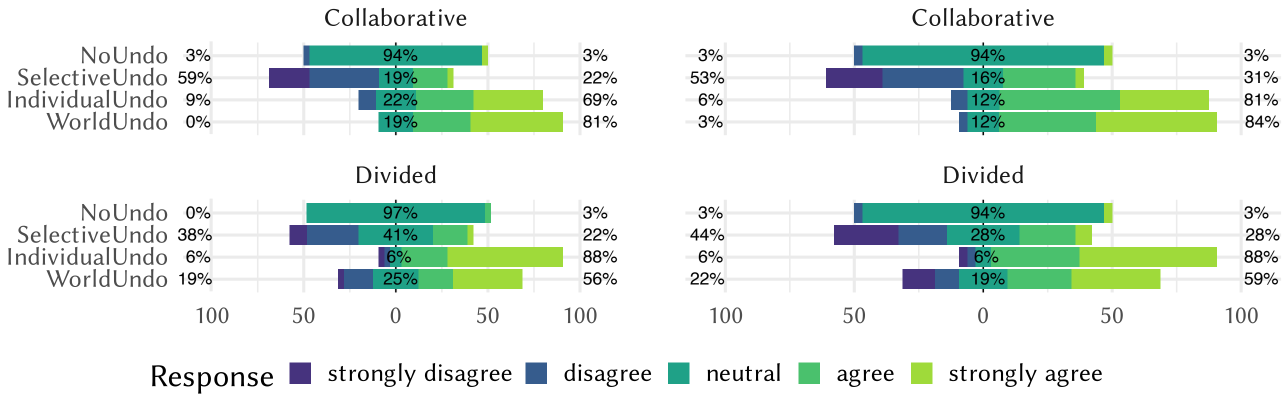
To analyze the undo actions, we excluded the trials for the NoUndo conditions. For the remaining conditions, we fitted a Poisson regression model and found values ranging from from $M = 12.1, SD = 13.5$ (Divided, WorldUndo) to $M = 4.69, SD = 7.70$ (Collaborative, WorldUndo).

We found a significant ($\chi^2(1) = 23.59$, $p < .001$) main effect for the CollaborationMode on the counted NumberOfUndos. Post-hoc tests revealed significantly ($p < .001$) higher values for Divided ($M = 8.42, SD = 11.8$) compared to Collaborative ($M = 4.38, SD = 7.10$). We also found a significant ($\chi^2(2) = 6.77$, $p < .05$) main effect for the UndoTechnique. However, post-hoc tests did not confirm this observation ($p > .05$). We found a significant ($\chi^2(2) = 15.80$, $p < .001$) interaction effect. While the WorldUndo received significantly ($p < .05$) lower ratings compared to SelectiveUndo in the Collaborative case, this inverts in the Divided case, where WorldUndo received non-significantly ($p > .05$) higher ratings compared to all other levels.

Since these results do not allow for a simple conclusion for H1, H2, and H4, we discuss them in more detail in Section 5.

4.7 Final Tower Height

As a performance measure, we analyzed the final TowerHeight. Performing a two-way RM ANOVA we found values ranging from



(a) Enjoyment **(b) DesiredFutureUsage**

Figure 8: Participants' responses regarding (a) Enjoyment and (b) DesiredFutureUsage on a 5-point Likert scale.

$M = 4.05, SD = 1.05$ (Collaborative, WorldUndo) to $M = 2.51, SD = 1.03$ (Divided, SelectiveUndo). Our RM ANOVA shows a significant ($F_{1,31} = 29.22$, $p < .001$) main effect for the CollaborationMode on the TowerHeight with a medium ($\eta^2_G = 0.11$) effect size. Post-hoc tests revealed significantly ($p < .001$) higher ratings for Collaborative ($M = 3.56, SD = 1.15$) compared to Divided ($M = 2.75, SD = 1.13$). We also found a significant ($F_{2.73,84.78} = 2.99$, $p < .05$) main effect for the UndoTechnique on the TowerHeight rating with a large ($\eta^2_G = 0.34$) effect size. Post-hoc tests revealed significantly ($p < .05$) higher ratings for WorldUndo ($M = 3.38, SD = 1.33$) compared to SelectiveUndo ($M = 2.82, SD = 1.08$). We could not find significant ($F_{2.90,89.83} = 2.5$, $p > .05$) interaction effects.

Again, these results only support H1 and H4 partially, as not all UndoTechniques show a positive effect on the performance and mutual interference. We discuss this in more detail in Section 5.

4.8 Qualitative Feedback

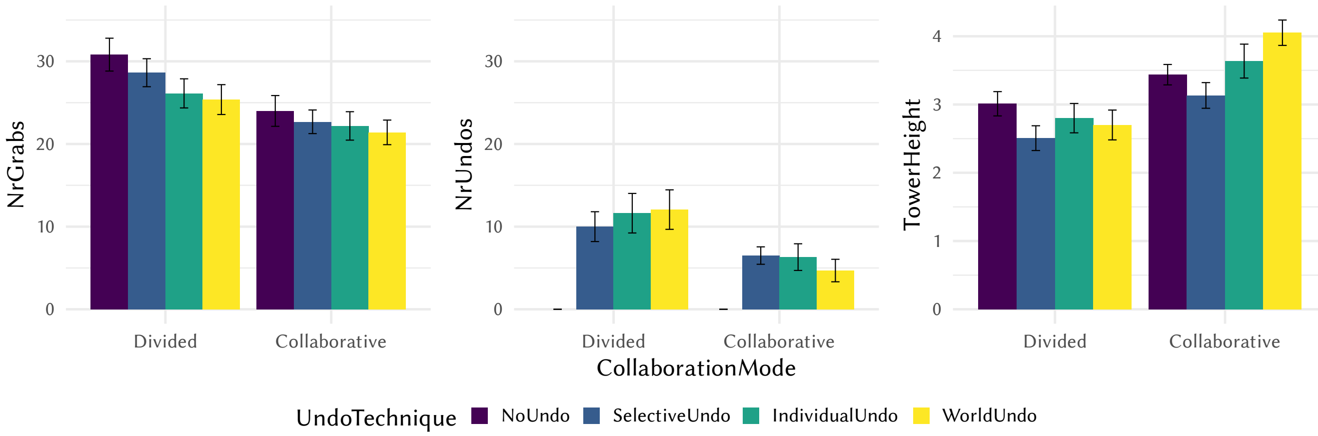
Besides the Questionnaires and Logging, we also collected qualitative feedback from participants through positive and negative comments after each condition. We used a thematic analysis to identify themes within these comments following the process by Blandford et al. [8]. To do so, three researchers individually coded a sample of 15% of the comments before discussing and agreeing on the final codes together. As a last step, one researcher reviewed the remaining material, coding them with the agreed codes. In this section, we briefly highlight participants' statements using italic typesets for the themes and quotes for participants' statements. In general, participants commented more about the available UndoTechniques than their CollaborationMode and their comments aligned with the quantitative results.

When NoUndo was available, participants commented on their *worry of mistakes* and bad *recoverability* and stated there was "*no room for errors*" (P4) as "*mistakes are devastating*" (P9). Participants also "*felt the pressure of not having the undo feature*" (P13). As a consequence, participants commented both positively and negatively about their more *thoughtful actions* in these conditions. Another theme within the positive comments was participants' *independence*, and participants liked that they were "*not disturbed by other player*" (P3) and "*not disturbing other player*" (P11).

For the SelectiveUndo, the dominant theme for both positive and negative comments was the *granularity of control*. On the positive side, this was reflected by comments like "*nice to undo one selected object only*" (P23) while on the negative side by comments like "*undoing one cube at a time didn't help or made it worse*" (P1). Another theme was the need for *memorization*, and participants disliked that they had to "*remember the correct order to undo objects*" (P26). This links to the next theme *efficiency*, which was mostly present in the negative comments by statements like "*Undoing something takes a lot of time*" (P12).

After using the IndividualUndo, participants positively stated they were "*not afraid to make mistakes*" (P9) and had the "*confidence to take risks*" (P27), resulting in *no worry of mistakes*. Other themes among the positive comments were the *efficiency* as well as *recoverability* reflected by comments like "*The individual undo works well and lets you recover from big mistakes fast*" (P26).

For WorldUndo, participants again positively commented they had *no worry of mistakes* and liked that they "*didn't have to think about my/our mistakes*" (P16). While participants commented on this mainly in the Collaborative case, in the Divided case, the dominant theme among the negative comments was participants *dependency*. This was, on the one hand, expressed by comments regarding the own progress "*I could not accomplish anything since the other player was constantly undermining my efforts with their undo*" (P9), but on the other hand, regarding empathy with the other player "*I didn't want to interrupt the other persons flow, so i felt uncomfy in certain situation, especially if I made a mistake and need to use the undo function*" (P4).



(a) NumberOfGrabs **(b) NumberOfUndos** **(c) TowerHeight**

Figure 9: The mean results of our Logged Data of the user study. (a) Shows the NumberOfGrabs, (b) the NumberOfUndos on the same scale, and (c) the TowerHeight. The error bars show the standard error of the data.

5 DISCUSSION

The results presented in Section 4 indicate that our participants overall liked and utilized UndoTechniques when available in our study environment. Our qualitative and quantitative data allows insights into participants' usage preference and their acceptance of the different UndoTechniques in different CollaborationModes. In this section we discuss the results with regard to our hypotheses, compare the strengths and shortcomings of the different UndoTechniques, and give recommendations on how to best use them.

5.1 Undo Increases Perceived Success, but Not Necessarily Actual Success

Our results indicate that participants used the undo feature whenever available. In particular, the NumberOfUndos proves that participants used all UndoTechniques if they can. However, we observe differences between the techniques. Independent of the collaboration type, users embraced the IndividualUndo and WorldUndo as it increased their perceived Success and helped them recover from mistakes. It also reduced participants' Frustration and task load during the task as evident from the RTLX-score, supporting H2. Further, participants reported that they enjoyed using those techniques and want to use them in the future. Contrary to how participants perceived their performance, interestingly, they did not actually perform better throughout by using the available techniques. In the Divided case, the TowerHeight did not change significantly when utilizing the UndoTechniques. On the other hand, when working Collaborative, the usage of WorldUndo helped to achieve the highest TowerHeight. Consequently, this does not universally support or contradict our H1 as we can observe differences in between the different UndoTechniques and CollaborationModes regarding the actual performance.

We attribute the generally positive participants' response towards the undo features to participants' familiarity with the concept in traditional computer systems and their need to undo mistakes in the context of our study. By providing a mechanism to revert mistakes, we satisfy this need and encourage participants to take more risk in their construction style. Our results showing divergence of actual and perceived success are in line with findings from previous studies [33]. As Archer et al. [3] observe "*The nature of errors may change with experience, but their occurrence does not.*" hinting at the fact, that an undo not only aims at increasing performance. The comments received by participants underline this aspect relevant to undo, namely that they were "*not afraid to make mistakes*" and were willing to "*take risks*".

Reflecting on our results, we suggest including undo operations in future VR applications to reduce frustration and increase the feeling of success. In VR applications for collaborative design processes, such undo mechanics could, for example, motivate users to explore new variations, and users of virtual learning environments could profit from reduced fear of mistakes.

5.2 World Undo Connects Users, Individual Undo Prevents Mutual Interference

Besides the task performance, the choice of the UndoTechnique also influenced the perceived social connectedness between the participants. This is evident from the ratings on the IOS and GEQ scores that indicate higher social connectedness for WorldUndo compared to IndividualUndo, supporting H3. Our data shows that this is especially true for the Divided case, where the social connectedness was lower otherwise. While WorldUndo increased social connectedness, on the other hand, it also caused an increase in reciprocal disturbances, which backs H4.

We speculate that in the Collaborative case, the actual collaboration and common task goal outweigh the influence of the UndoTechniques. For Divided work, however, WorldUndo proved to be crucial to fostering social connectedness. We assume that when working separately, the occasional consequence of the other user's use of the WorldUndo is sufficient to remind users of their existence and actions. While the WorldUndo works as a means to provide awareness of the other user, it can also influence and disturb the other player, as undo actions by one user directly influence the progress of the other user. As expected, this is particularly problematic in the divided case, where users are focused on their own tasks.

Considering these results, we suggest selecting an undo technique based on weighing the goals and the mode of collaboration between the users. If users work Collaborative, the WorldUndo does not impose negative effects, and participants favored this technique above all others. If users work Divided and it is relevant not to disturb them, the IndividualUndo should be selected, resulting in a lower social connection to other users however. If users work Divided but a high social connection is the design goal, the WorldUndo should still be selected, even though this will result in participants' mutual disturbance.

5.3 Selective Undo Is Not Suitable for Situations Where Large Changes to the Scene Need to Be Reverted

In contrast to the IndividualUndo and WorldUndo, the SelectiveUndo was not received well for the task at hand. This UndoTechnique increased the task load and Frustration and helped significantly less in Recovering from mistakes compared to the WorldUndo and IndividualUndo. When utilizing this technique, participants also felt less Success and Enjoyment.

We attribute these results to the inappropriateness of the chosen form of the SelectiveUndo technique for the chosen task, not as a general problem of this type of undo. In the task, undoing only single building blocks was rarely beneficial, as the tower more often collapsed as a connected bigger structure. In these cases our implementation of a SelectiveUndo required similar effort by participants as rebuilding a new tower, potentially with better stability. Further, the task featured strong dependencies of the manipulated artifacts, and undoing a single mistake did not recover the resulting consequences, like e.g., a more linear undo like IndividualUndo or WorldUndo. While related work identified the isolated manipulation and maintaining of subsequent manipulation steps as one strength of the selective undo [7], in our study, this proved not to be useful, as mirrored by participants commenting this technique being "*not helpful*".

This inadequacy of a technique is mirrored in the use of IndividualUndo in the Collaborative case. The alternating stacking of elements caused a time dependency, limiting users to the last undo step with the IndividualUndo.

Following our results, we recommend a thorough assessment of the use cases before selecting an undo technique, in particular, a SelectiveUndo. While SelectiveUndo was generally rejected in the quantitative and qualitative data in our controlled experiment, we nevertheless consider this SelectiveUndo or another possible implementation [40] suitable for other task types and also advantageous over WorldUndo and IndividualUndo. Future work is necessary to provide a more in-depth analysis of different implementations of selective undo for VR.

6 LIMITATIONS AND FUTURE WORK

Our data as well as comments from participants prove the need for undo actions in the VR domain as well. In the context of our study, we could identify strengths and shortcomings of the various UndoTechniques through questionnaires, qualitative comments, and logged data. As evident from our discussion, this is only a first step in transferring findings from traditional CSCW into the VR domain and we are confident that this work will serve as a base for future work to further address these challenges. Throughout our study and the consecutive analysis, we identified several limitations imposed by our study design as well as directions for future work, which we discuss in this section.

6.1 External Validity and Real-World Applicability

In this paper, we contributed the results of an experiment that explored UndoTechniques for VR in a highly artificial task and environment, while enforcing strict CollaborationModes. As the first work investigating undo mechanics for multi-user systems in VR, we adopted such a highly controlled approach to provide a solid foundation for future work. Therefore, we selected representative extremes of collaborative situations on the continuum between joint work on a common goal to independent work with individual goals. Further, we opted for a task that was suitable for both of these collaborative situations and that allowed us to quantify the performance of the participants. In realistic scenarios, collaboration situations may not fall within these extremes and may be subject to constant change. Further, external influences such as objects that are not part of the actual task or actions that do not result from user interactions may have an impact. We acknowledge that such changes could yield other results and thus further work is needed in this area.

6.2 Scaleability

In our experiment, we decided on a dyadic task, as the minimum to study collaborative work and potential usage of UndoTechniques. We expect, that with increasing numbers of users, an UndoTechnique like WorldUndo can easily become very unfavorable due to the potential for high disturbance, as found already in our dyadic case. At the same time, as discussed before, IndividualUndo techniques might be useless in complex scenarios with highly intertwined actions of multiple users. In this context, we see the potential for the exploration of variations of SelectiveUndo techniques, balancing between individual understandable yet ineffective and global effective yet chaotic actions.

6.3 Ownership of Objects in the Shared Virtual Space

Implementing an IndividualUndo allows users to undo their actions. This brings up the question, of what these "own actions" are. In our experiment we only considered the objects manipulated directly by a user, as "their" objects. One can ask, if a manipulation resulting from an object manipulated by one player, should also be considered this player's action. The consequent question is which margin of manipulation should one consider here. Which limits for the translation or rotation of an affected object to choose? Could the mere contact with an object be considered manipulation already? With regard to our task, we opted against this implicit definition of manipulation as this implementation would blur the borders of the individual and world undo. However, these are highly relevant questions and future work should address the effect of different implementations of these.

6.4 Continuous versus Discrete Steps

During the development of our controlled experiment, we had to make certain design decisions. In contrast to most undos known from traditional computer systems, we chose to implement continuous undo functions. This obviously is one relevant variable and as shown in related work [33] has a strong influence on how users perceive and understand their environment. We chose the continuous implementation, as we derived from related work, that discrete undo steps will impose problems in understanding the mutual undo actions. Future work should investigate these effects of different implementations of a discrete and continuous undo again in the context of multi-user undo. Here, we speculate to find differences concerning performance, and especially mutual understanding of undone actions. We chose to exclude the time passing between the manipulation of an object and the execution of an undo feature from the undo timeline. In doing so, the users experience an effect right after starting the undo, independent of how long the last action has passed. We also chose to exclude the other user's avatar from the range of effect of any undo feature, as explained in Section 3. All of these design decisions could be chosen as independent variables. However, studying the influence of all of them was beyond the scope of our work and should be addressed in future work.

7 CONCLUSION

In this paper, we investigated the transferability of established undo techniques from the traditional CSCW domain to their use in multi-user collaborative VR and derived usage guidelines based on our findings. Our results clearly show that we should provide users with one form of undo in comparable VR tasks, as this increases users' perceived success, helps them to recover from mistakes, and reduces participants' frustration and task load. Which form of undo is best suited highly depends on the social constellation as well as the task goal of the VR scene. We are confident our work provides essential findings relevant to both follow-up studies as well as designers of VR applications. The proposed techniques originating from related work can enrich future collaborative systems by providing users with undo mechanics and therefore improving their experience. As one key consideration when designing UndoTechniques for multi-user VR, one must consider how important social connection between users is and how the mode of collaboration of the users will be. Here, one must weigh between higher social connectedness and potential disturbance between users.

REFERENCES

[1] Gregory D Abowd and Alan J Dix. 1992. Giving undo attention. *Interacting with Computers* 4, 3 (Dec. 1992), 317–342. https://doi.org/10.1016/0953-5438(92)90021-7

[2] Víctor H. Andaluz, Jorge S. Sánchez, Carlos R. Sánchez, Washington X. Quevedo, José Varela, José L. Morales, and Giovanny Cuzco. 2018. Multi-user Industrial Training and Education Environment. In *Augmented Reality, Virtual Reality, and Computer Graphics*, Lucio Tommaso De Paolis and Patrick Bourdot (Eds.). Vol. 10851. Springer International Publishing, Cham, 533–546. https://doi.org/10.1007/978-3-319-95282-6_38 Series Title: Lecture Notes in Computer Science.

[3] James E. Archer, Richard Conway, and Fred B. Schneider. 1984. User Recovery and Reversal in Interactive Systems. *ACM Transactions on Programming Languages and Systems* 6, 1 (Jan. 1984), 1–19. https://doi.org/10.1145/357233.357234

[4] Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (May 2013), 121–136. https://doi.org/10.1016/j.cag.2012.12.003

[5] Arthur Aron, Elaine N. Aron, and Danny Smollan. 1992. Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology* 63, 4 (Oct. 1992), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

[6] Roger Bakeman. 2005. Recommended effect size statistics for repeated measures designs. *Behavior Research Methods* 37, 3 (Aug. 2005), 379–384. https://doi.org/10.3758/BF03192707

[7] Thomas Berlage. 1994. A selective undo mechanism for graphical user interfaces based on command objects. *ACM Transactions on Computer-Human Interaction* 1, 3 (Sept. 1994), 269–294. https://doi.org/10.1145/196699.196721

[8] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. *Qualitative HCI Research: Going Behind the Scenes.* Springer International Publishing, Cham. https://doi.org/10.1007/978-3-031-02217-3

[9] Simon Butscher, Sebastian Hubenschmid, Jens Müller, Johannes Fuchs, and Harald Reiterer. 2018. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. https://doi.org/10.1145/3173574.3173664

[10] Aaron G. Cass, Chris S. T. Fernandes, and Andrew Polidore. 2006. An empirical evaluation of undo mechanisms. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*. ACM, Oslo Norway, 19–27. https://doi.org/10.1145/1182475.1182478

[11] Jacob Cohen. 2013. *Statistical Power Analysis for the Behavioral Sciences* (0 ed.). Routledge. https://doi.org/10.4324/9780203771587

[12] Marcelo De Paiva Guimaraes and Valeria Farinazzo Martins. 2014. A Checklist to Evaluate Augmented Reality Applications. In *2014 XVI Symposium on Virtual and Augmented Reality*. IEEE, Piata Salvador, 45–52. https://doi.org/10.1109/SVR.2014.17

[13] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. ACM, Virtual Event USA, 754–768. https://doi.org/10.1145/3472749.3474784

[14] Clarence Ellis and Jacques Wainer. 1999. Groupware and Computer Supported Cooperative Work. In *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, 425–457.

[15] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. 2019. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies* 131 (Nov. 2019), 81–98. https://doi.org/10.1016/j.ijhcs.2019.05.011

[16] Seth M. Feeman, Landon B. Wright, and John L. Salmon. 2018. Exploration and evaluation of CAD modeling in virtual reality. *Computer-Aided Design and Applications* 15, 6 (Nov. 2018), 892–904. https://doi.org/10.1080/16864360.2018.1462570

[17] Philippa Foot. 1967. The Problem of Abortion and the Doctrine of the Double Effect. (1967).

[18] Doron Friedman, Rodrigo Pizarro, Keren Or-Berkers, SolÃ¨ne Neyret, Xueni Pan, and Mel Slater. 2014. A method for generating an illusion of backwards time travel using immersive virtual realityâ€"an exploratory study. *Frontiers in Psychology* 5 (Sept. 2014). https://doi.org/10.3389/fpsyg.2014.00943

[19] Gernot Goebbels and Vali Lalioti. 2001. Co-presence and co-working in distributed collaborative virtual environments. In *Proceedings of the 1st international conference on Computer graphics, virtual reality and visualisation*. ACM, Camps Bay, Cape Town South Africa, 109–114. https://doi.org/10.1145/513867.513891

[20] Simon N. B. Gunkel, Hans M. Stokking, Martin J. Prins, Nanda Van Der Stap, Frank B. Ter Haar, and Omar A. Niamut. 2018. Virtual reality conferencing: multi-user immersive VR experiences on the web. In *Proceedings of the 9th ACM Multimedia Systems Conference*. ACM, Amsterdam Netherlands, 498–501. https://doi.org/10.1145/3204949.3208115

[21] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (Oct. 2006), 904–908. https://doi.org/10.1177/154193120605000909

[22] Wijnand A IJsselsteijn, Yvonne AW De Kort, and Karolien Poels. 2013. *The game experience questionnaire.* Technical Report. Technische Universiteit Eindhoven.

[23] Randolph L. Jackson and Eileen Fagan. 2000. Collaboration and learning within immersive virtual reality. In *Proceedings of the third international conference on Collaborative virtual environments*. ACM, San Francisco California USA, 83–92. https://doi.org/10.1145/351006.351018

[24] Hannes Kaufmann. 2003. Collaborative Augmented Reality in Education.

[25] Joseph J. LaViola, Ernst Kruijff, Ryan P. McMahan, Doug A. Bowman, and Ivan Poupyrev. 2017. *3D user interfaces: theory and practice* (second edition ed.). Addison-Wesley, Boston. OCLC: ocn935986831.

[26] V.D. Lehner and T.A. DeFanti. 1997. Distributed virtual reality: supporting remote collaboration in vehicle design. *IEEE Computer Graphics and Applications* 17, 2 (April 1997), 13–17. https://doi.org/10.1109/38.574654

[27] Dieter Lerner, Stefan Mohr, Jonas Schild, Martin Göring, and Thomas Luiz. 2020. An Immersive Multi-User Virtual Reality for Emergency Simulation Training: Usability Study. *JMIR Serious Games* 8, 3 (July 2020), e18822. https://doi.org/10.2196/18822

[28] Marco Loregian. 2008. Undo for mobile phones: does your mobile phone need an undo key? do you?. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*. ACM, Lund Sweden, 274–282. https://doi.org/10.1145/1463160.1463190

[29] Morad Mahdjoub, Davy Monticolo, Samuel Gomes, and Jean-Claude Sagot. 2010. A collaborative Design for Usability approach supported by Virtual Reality and a Multi-Agent System embedded in a PLM environment. *Computer-Aided Design* 42, 5 (May 2010), 402–413. https://doi.org/10.1016/j.cad.2009.02.009

[30] Maria Matsangidou, Boris Otkhmezuri, Chee Siang Ang, Marios Avraamides, Giuseppe Riva, Andrea Gaggioli, Despoina Iosif, and Maria Karekla. 2022. "Now i can see me" designing a multi-user virtual reality remote psychotherapy for body weight and shape concerns. *Human–Computer Interaction* 37, 4 (July 2022), 314–340. https://doi.org/10.1080/07370024.2020.1788945

[31] Joshua McVeigh-Schultz, Anya Kolesnichenko, and Katherine Isbister. 2019. Shaping Pro-Social Interaction in VR: An Emerging Design Framework. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–12. https://doi.org/10.1145/3290605.3300794

[32] Brad A. Myers. 1990. Taxonomies of visual programming and program visualization. *Journal of Visual Languages & Computing* 1, 1 (March 1990), 97–123. https://doi.org/10.1016/S1045-926X(05)80036-9

[33] Florian Müller, Arantxa Ye, Dominik Schön, and Julian Rasch. 2023. UndoPort: Exploring the Influence of Undo-Actions for Locomotion in Virtual Reality on the Efficiency, Spatial Understanding and User Experience. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–15. https://doi.org/10.1145/3544548.3581557

[34] Christian Paping and Willem-Paul Brinkman. 2010. An Explorative Study into a Tele-delivered Multi Patient Virtual Reality Exposure Therapy System. (2010).

[35] Zhenhui Peng, Chuhan Shi, and Xiaojuan Ma. 2021. Exploring User Experience and Design Opportunities of Desktop Social Virtual Reality for Group Learning Activities. In *The Ninth International Symposium of Chinese CHI*. ACM, Online Hong Kong, 105–111. https://doi.org/10.1145/3490355.3490367

[36] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*. ACM, Brighton United Kingdom, 99–108. https://doi.org/10.1145/3131277.3132180

[37] Marcio S. Pinho, Doug A. Bowman, and Carla M. Dal Sasso Freitas. 2008. Cooperative object manipulation in collaborative virtual environments. *Journal of the Brazilian Computer Society* 14, 2 (June 2008), 53–67. https://doi.org/10.1007/BF03192559

[38] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-Defined Gestures for Augmented Reality. In *Human-Computer Interaction – INTERACT 2013*, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, Paula Kotzé, Gary Marsden, Gitte Lindgaard, Janet Wesson, and Marco Winckler (Eds.). Vol. 8118. Springer Berlin Heidelberg, Berlin, Heidelberg, 282–299. https://doi.org/10.1007/978-3-642-40480-1_18 Series Title: Lecture Notes in Computer Science.

[39] I. Poupyrev, T. Ichikawa, S. Weghorst, and M. Billinghurst. 1998. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. *Computer Graphics Forum* 17, 3 (Aug. 1998), 41–52. https://doi.org/10.1111/1467-8659.00252

[40] Atul Prakash and Michael J. Knister. 1992. Undoing actions in collaborative work. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work - CSCW '92*. ACM Press, Toronto, Ontario, Canada, 273–280. https://doi.org/10.1145/143457.143527

[41] Atul Prakash and Michael J. Knister. 1994. A framework for undoing actions in collaborative systems. *ACM Transactions on Computer-Human Interaction* 1, 4 (Dec. 1994), 295–330. https://doi.org/10.1145/198425.198427

[42] Julian Rasch, Vladislav Dmitrievic Rusakov, Martin Schmitz, and Florian Müller. 2023. Going, Going, Gone: Exploring Intention Communication for Multi-User Locomotion in Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–13. https://doi.org/10.1145/3544548.3581259

[43] Jonas Schild, Sebastian Misztal, Beniamin Roth, Leonard Flock, Thomas Luiz, Dieter Lerner, Markus Herkersdorf, Konstantin Weaner, Markus Neuberaer, Andreas Franke, Claus Kemp, Johannes Pranqhofer, Sven Seele, Helmut Buhler, and Rainer Herpers. 2018. Applying Multi-User Virtual Reality to Collaborative Medical Training. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Tuebingen/Reutlingen, Germany, 775–776. https://doi.org/10.1109/VR.2018.8446160

[44] Thomas Seifried, Christian Rendl, Michael Haller, and Stacey Scott. 2012. Regional undo/redo techniques for large interactive surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Austin Texas USA, 2855–2864. https://doi.org/10.1145/2207676.2208690

[45] Keigo Shima, Ryosuke Takada, Kazusa Onishi, Takuya Adachi, Buntarou Shizuki, and Jiro Tanaka. 2015. AirFlip-Undo: Quick Undo using a Double Crossing In-Air Gesture in Hover Zone. In *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, Daegu Kyungpook Republic of Korea, 97–98. https://doi.org/10.1145/2815585.2815737

[46] Ben Shneiderman and Catherine Plaisant. 2005. *Designing the user interface: strategies for effective human-computer interaction* (4. ed ed.). Pearson/Addison-Wesley, Boston, Mass. Munich.

[47] R. Stark, J.H. Israel, and T. Wöhler. 2010. Towards hybrid modelling environments—Merging desktop-CAD and virtual reality-technologies. *CIRP Annals* 59, 1 (2010), 179–182. https://doi.org/10.1016/j.cirp.2010.03.102

[48] Warren Teitelman. 1972. Automated programmering: the programmer's assistant. In *Proceedings of the December 5-7, 1972, fall joint computer conference, part II on - AFIPS '72 (Fall, part II)*. ACM Press, Anaheim, California, 917. https://doi.org/10.1145/1480083.1480119

[49] Kristen M Triandafilou, Daria Tsoupikova, Alexander J Barry, Kelly N Thielbar, Nikolay Stoykov, and Derek G Kamper. 2018. Development of a 3D, networked multi-user virtual reality environment for home therapy after stroke. *Journal of NeuroEngineering and Rehabilitation* 15, 1 (Dec. 2018), 88. https://doi.org/10.1186/s12984-018-0429-0

[50] C. Wienrich, K. Schindler, N. Dollinger, S. Kock, and O. Traupe. 2018. Social Presence and Cooperation in Large-Scale Multi-User Virtual Reality - The Relevance of Social Interdependence for Location-Based Environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Reutlingen, 207–214. https://doi.org/10.1109/VR.2018.8446575

[51] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 143–146. https://doi.org/10.1145/1978942.1978963

[52] Haijun Xia, Sebastian Herscher, Ken Perlin, and Daniel Wigdor. 2018. Spacetime: Enabling Fluid Individual and Collaborative Editing in Virtual Reality. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. ACM, Berlin Germany, 853–866. https://doi.org/10.1145/3242587.3242597

[53] Yiya Yang. 1988. Undo support models. *International Journal of Man-Machine Studies* 28, 5 (May 1988), 457–481. https://doi.org/10.1016/S0020-7373(88)80056-7

[54] Lei Zhang, Ashutosh Agrawal, Steve Oney, and Anhong Guo. 2023. VRGit: A Version Control System for Collaborative Content Creation in Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–14. https://doi.org/10.1145/3544548.3581136