DOI: 10.1111/cgf.14550
Eurographics Conference on Visualization (EuroVis) 2022
R. Borgo, G. E. Marai, and T. Schreck
(Guest Editors)

Volume 41 (2022), Number 3

Hybrid Touch/Tangible Spatial Selection in Augmented Reality

Mickael Sereno[1], Stéphane Gosset[1], Lonni Besançon[2], and Tobias Isenberg[1]

[1]Université Paris-Saclay, CNRS, Inria, LISN, France [2]Linköping University, Sweden

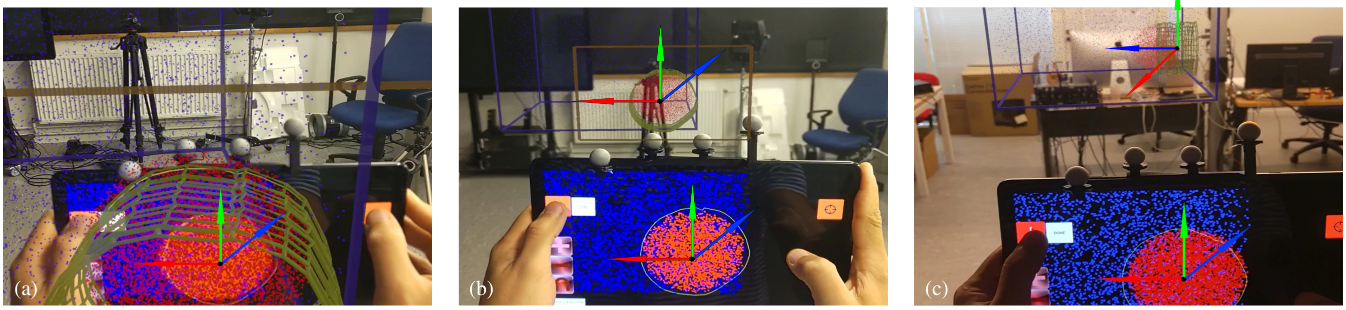


Figure 1: *Our three AR mappings. In NA (a), the user extrudes the drawn 2D shape in a direct approach. The position of the virtual tablet overlaps the position of the physical tablet. In RA (b), the orientations of the virtual and physical tablets with respect to the physical space is the same during an extrusion. However, the user controls the virtual plane with relative motions. In this example, the user stepped back before starting remotely an extrusion. Finally, RF (c) creates two coordinate systems. Movements from the one defined by the physical tablet are mapped to the coordinate system defined by the virtual table. In this example, the user, after having placed the tablet in front of the dataset, stepped back and rotated around the dataset which the user now is facing sideway. As the user moves along the normal of the physical tablet, its virtual counterpart also moves along its own normal axis (the blue arrows).*

Abstract

We study tangible touch tablets combined with Augmented Reality Head-Mounted Displays (AR-HMDs) to perform spatial 3D selections. We are primarily interested in the exploration of 3D unstructured datasets such as cloud points or volumetric datasets. AR-HMDs immerse users by showing datasets stereoscopically, and tablets provide a set of 2D exploration tools. Because AR-HMDs merge the visualization, interaction, and the users' physical spaces, users can also use the tablets as tangible objects in their 3D space. Nonetheless, the tablets' touch displays provide their own visualization and interaction spaces, separated from those of the AR-HMD. This raises several research questions compared to traditional setups. In this paper, we theorize, discuss, and study different available mappings for manual spatial selections using a tangible tablet within an AR-HMD space. We then study the use of this tablet within a 3D AR environment, compared to its use with a 2D external screen.

CCS Concepts

• ***Human-centered computing*** $\rightarrow$ ***Mixed / augmented reality; Touch screens; Scientific visualization;***

1. Introduction

Many scientists work to understand data that is inherently three-dimensional, such as in physics and biology. They explore such data traditionally as projections on 2D screens, with an inherently limited depth perception. As the availability of immersive devices increases, however, hybrid or fully stereoscopic data exploration and analysis become possible. We explore the use of augmented reality head-mounted displays (AR-HMDs) as the basis of such an approach, and ask about how best to interact in this space. We notably focus on the interactive selection of data points based on their 3D spatial properties. Such spatial selection—a fundamental exploratory tool [Tuk77, KI13]—is essential for users to specify regions of interest (ROIs) in unstructured datasets as they often cannot be isolated by filtering some properties. While such interaction has been studied in the past, most research focused on 2D rendering as opposed to immersive means [BYK*21]. We investigate the use of AR-HMDs as they integrate interaction, visualization, and the users' physical 3D spaces, and as their immersive displays are suitable for 3D data visualization [FCL09, KMLM16, HRD*19]. Standalone AR-HMDs (e. g., Microsoft's HoloLens) are now also powerful enough for real-time 3D rendering. This environment being new, many research questions remain regarding how to explore large scientific datasets and how AR-based exploration tools differ from past setups.

An essential point to consider is that AR-HMDs cannot rely on the same input as PCs. Researchers have proposed to explore speech input, eye tracking, and mid-air gestures for AR-HMDs, yet this input is often limited. Compared to single-action commands (speech, some gestures), continuous interactions suffer from low precision and fatigue. Moreover, it is difficult to distinguish spatial input from casual motion (eye tracking, mid-air gestures). One option is to rely on hybrid setups that use both an AR-HMD and a PC together (e. g., [BBK*06, WBR*20]), yet often one also wants to be independent of a workstation and benefits from an AR-HMD's mobility. Hence, we explore the use of a portable tablet as a secondary device, whose combination with AR-HMDs also provides flexible input—yet without being physically stationary.

We study spatial selection techniques in which users fully control what they can select, i. e., they have high Degrees of Freedom (DoF) [BSA*19]. While mid-air 3D input suffers from instability [AA13, AKA*17] and induces fatigue [HRGMI14], tablets can provide beneficial input for tasks within immersive systems [SGHV19] such as spatial selection [MMNK*20]. We based our spatial selection on Besançon et al.'s [BSA*19] Tangible Brush, as it is flexible and uses a tangible tablet. The users draw a 2D lasso on the tablet's touch screen and then extrude it in 3D by moving the tablet. While this manual interaction showed benefits compared to a semi-automatic approach that relies on computer-based decisions, Tangible Brush decouples the user's output and input spaces: the output is rendered only as a projection on an external 2D screen and the input is entered in the user's physical space. This decoupling may explain the high reported mental effort for participants. We thus adapted Tangible Brush by replacing the external 2D view with an AR stereoscopic one. This new configuration poses numerous research questions, as now the tablet's position has a meaning in both the users' physical and visualization spaces. Compared to the original, our AR setup may lower the users' mental demand as, for the extrusion, both the input and the output spaces coincide with the users' physical space.

Our contribution is threefold. First, we examine multiple position mappings of a tangible tablet used inside the AR 3D space for manual spatial selections. Second, we study these mappings to find the most suitable one for AR-HMDs. Third, we compare our AR-based spatial selection technique using the selected mapping to the original, to better understand the implications of merging the user's input space within the user's physical and visualization 3D spaces.

2. Related Work

In our work we rely on past research on hybrid systems that include immersive devices, AR-related input techniques, and manual spatial selection like the traditional Tangible Brush as we summarize next.

2.1. Hybrid Immersive Environment

Past research showed the benefits of immersive setups for user cognition and perception [WF96, WM08, HRD*19]. To support scientists in their work, researchers investigated stereoscopic displays coupled with traditional 2D interfaces to provide precise 2D input, enter texts, and manipulate widgets as in traditional exploratory tools [BBK*06, RD19, WBR*20]. For instance, Bornik et al. [BBK*06] combined a stereoscopic 3D display with a PC to manipulate volumetric medical data, while Kim et al. [KJK*15] added tactile interaction to a CAVE. Wang et al. [WBR*20] studied the use of an AR-HMD as a complementary screen to a workstation for particle physicists, the latter for access to traditional analytics tools. They found that, while scientists prefer a hybrid environment to a standalone AR-HMD or workstation, they need to walk around the data and use dedicated 3D input. Yet Bornik et al. [BBK*06] and Wang et al. [WBR*20] do not provide such a functionality as their interfaces were in 2D and stationary. Bach et al. [BSB*18] stated that their participants, in the AR-HMD condition, preferred to stand up—although half of the tasks did not require 3D input. A tablet—as we use it—may solve this mobility issue, while still offering 2D input. Surale et al. [SGHV19] explored its use in a Virtual Reality (VR) environment for CAD software. Depending on the usual speed/accuracy trade-off, users preferred to manipulate their views and data either on the tablet or via the 3D VR interfaces. In our system, users perform 3D spatial selections by extruding a 2D shape in its physical space. We use a tablet to input the 2D shape and rely on the physical user space for the extrusion.

We do not envision AR-HMDs plus tablets to fully replace workstations, which have higher resolutions, computing power, more precise 2D input strategies, and more powerful software [WBG*19]. Such a system, however, can be used for dedicated tasks such as collaborative work where most of the users' time would concern discussions, comparisons, and agreements, which require a set of fundamental yet less complex exploratory tools [SWB*22, Ser21]. We thus consider the design of a complete application and focus, in this paper, on spatial selections as one of the fundamental 3D exploration techniques [KI13, Tuk77]. For this interaction, we use the tablet to provide mobile input and, at the same time, serve as a tangible device, whose interaction space is closely related to the user's physical space and the current display space [CBL*17].

2.2. Mid-Air Gestures and Tangible Interactions

AR-HMDs tightly connect the physical world with the display and interaction spaces for 3D physical interactions [CBL*17], e. g., some mid-air input and tangible interactions. This close connection may reduce the users' cognitive load [CBL*17] and be suitable for tasks that require input with many DoF [HvDG94, BSB*18] and tangible interaction. Bach et al. [BSB*18] showed that the tight connection may explain why AR-HMDs were fastest, with similar error rates as other forms of input for the placement of cutting planes. In addition to its function as a 2D interface, Surale et al. [SGHV19] also used a tablet for 3D tangible input. They found that the choice of touch or tangible tablet input depends on the task and the needed precision.

Passive haptic feedback and planar constraints of tablets can improve the comfort and accuracy of 2D and 3D input [HvDG94, dHKP02, AKA*17, MMNK*20]. Montano-Murillo et al. [MMNK*20] and Arora et al. [AKA*17] showed that using a real tablet compared to a virtual 2D plane in a VR setup improves the user's performance. Reipschläger and Dachselt [RD19] augmented a tabletop with an AR-HMD for CAD sketching, with users drawing shapes and extrude them on the table and the AR-HMD rendering the

3D result. These examples show that, compared to view-occluding VR-HMDs, AR-HMDs allow users to better interact with additional, physical devices and we rely on an AR-HMD coupled with a tablet for volumetric selection for the same reason.

2.3. Volumetric/Spatial Selection

Many 3D selection techniques exist. We separate the selection of explicit objects from the spatial selection of subvolumes. The first group relies on predefined shapes or objects in the 3D space that users explicitly mark, where most techniques rely on raycasting [GB06, RN10, RCK*17, BPC19] in 3D environments or tap and lassos in 2D ones [XFAT12, YEII16, CZY*20]. The second allows users to specify 3D subspaces [AA13, BSA*19, BYK*21]. We focus here on unstructured datasets which do not include explicit objects but instead sample physical properties of, e. g., particles or volume cells, which one can filter to make selections. Jackson et al. [JLS*13] filter vector fields based on a tangible cylinder's orientation. Hurter et al. [HRD*19] brush paths using two VR controllers and select paths that connect both targeted points. Akers et al. [ASM*04] filter brain white matter pathways using 3D boxes associated to Boolean operations. Yu et al. [YEII16] select regions based on the density of the field and 2D lasso input. Montano-Murillo et al.'s [MMNK*20] hybrid VR-HMD plus tablet setup allows one to select points in cloud point datasets. The user first frames the ROI using the touch tablet and then uses the tablet as a magical lens to target at with a VR controller and raycasting. Transfer functions can also extract ROIs [KIL*03, GXY12]. Wiebel et al. [WVFH12] selects a ROI based on the 3D color and opacity field as defined by the transfer function, the virtual camera transformation, and the user's 2D input. While transfer function manipulations and Wiebel et al.'s algorithm [WVFH12] can be implemented on tablets, they often require time to find the suitable function, and it may not fully capture the user's intended ROI. Moreover, not all datasets possess suitable physical properties. We are thus interested in general techniques that work for every type of datasets (i. e., point clouds [Ste12, CZY*20], volumes [WVFH12], and vector fields [JLS*13]). We also cannot only rely on predefined geometries, e. g., boxes [UZW*07], as we want a flexible, data-independent way to specify ROIs. Moreover, we are interested in better understanding the users' mental models in AR environments, compared to non-immersive ones, for 6 DoF interaction techniques that make use of a tangible tablet.

We thus rely on Tangible Brush [BSA*19], on whose tablet interface a user draws a lasso in an orthographic view and extrudes it in virtual 3D space by moving the tablet. The orthographic view may lead to lower cognitive loads and higher accuracy as both its interactive and its visualization spaces are in 2D [BSB*18, BIRW19], while perspective views would not allow users to understand how the lasso shape relates to the objects in the scene. Because the tablet then does not provide depth cues, Besançon et al. [BSA*19] provided a second stationary 2D screen which shows the position of the virtual camera in the virtual space. The coordinate system of the interaction was reset each time, based on the orientation of the tablet as there were no absolute 3D coordinates. They compared Tangible Brush with SpaceCast [YEII16], a semi-automatic technique that infers the selection based on minimal user's input. Tangible Brush was slower but was more accurate and versatile. The authors recorded a

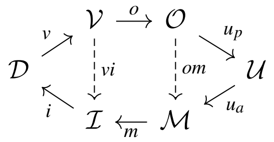


Figure 2: *Bruckner et al.'s model [BIRW19] of spatial directness. Spaces $\mathcal{D}$—data; $\mathcal{V}$—visualization; $\mathcal{O}$—output; $\mathcal{U}$—user; $\mathcal{M}$—manipulation; $\mathcal{I}$—interaction. Arrows represent transitions between spaces.*

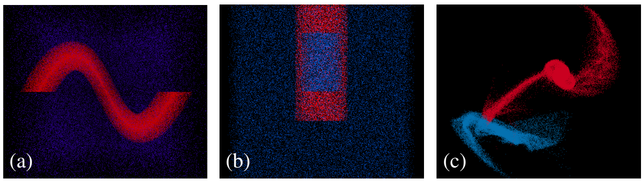


Figure 3: *The datasets we studied. (a) helical spring shape (see also Fig. 4(a) and Fig. 5). It represents non-linear structures (e. g., blood vessels), used to encourage unconstrained operations; (b) outer shell of a cylinder; (c) collision of two galaxies as a real use-case scenario. All datasets measure 50 × 50 × 50 cm³ in the AR space. In our study, valid points are in red, and invalid points are in blue. Figures (b) and (c) are from Besançon et al.'s paper [BSA*19] (© Wiley, used with permission).*

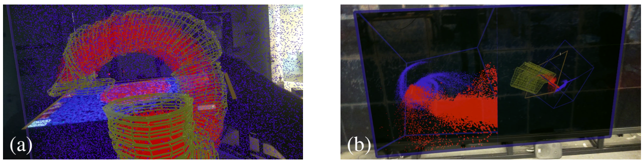


Figure 4: *Screenshots of the HoloLens' interface during extrusions for (a) the AR conditions and (b) the 2D condition (see also [BSA*19]). In the AR condition using NA, the user drew a circle on the tablet, which the user extruded by physically following, with the tablet, the Spring the AR-HMD renders. It results in the yellow wire-framed mesh that shows the current selection to (in)validate.*

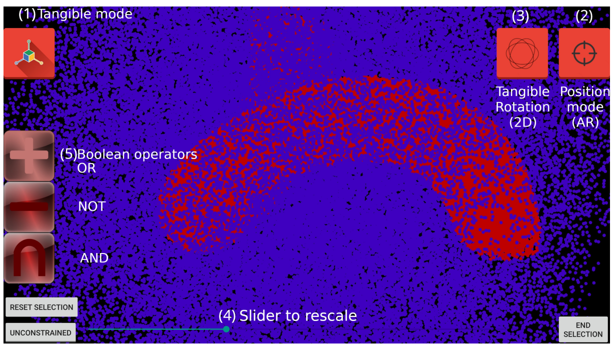


Figure 5: *Tablet interface while performing spatial selection.*

high mental load, however, which may be due to the decoupling between Tangible Brush's input (3D tangible interactions) and output (2D external screen). In our work we modified Tangible Brush and replaced the stationary screen with an AR-HMD as we explain next.

3. AR Concept and Differences to Tangible Brush

In our AR version we replace the external 2D screen with an AR-HMD stereoscopic view. Users can then use their physical tablet superimposed with its virtual counterpart. We define three mappings for the actual realization of our AR Tangible Brush (see Fig. 1). In all three cases users need to set the virtual camera of the tablet (still using an orthographic projection), before actually drawing the lasso. This allows users to relate their lasso-to-be-extruded with the dataset. We freeze the 2D orthographic view on the tablet for a given position, allowing users to draw their lasso on a static view for precise input [LODI16]. We describe next our three AR mappings.

Naïve Approach (NA): As a ***Naïve Approach*** (NA), we map the position, orientation, and size of the tablet from the physical space to their virtual counterparts. While it is the spatially most direct approach, NA has limitations. First, the size of the tablet (10"–12" diagonal for many models) is a strong limitation, as most 3D visualizations would have to be scaled down to "fit inside the tablet" and the global size of the data representation constantly has to be adapted to the intended selection. Second, users sometimes need small targeted selections. As visualizations can be large in size, a user may lose the overall spatial context as they need to be physically close to the intended selection. This may lead to parts of the data not fitting in the user's field-of-view anymore. Finally, interacting inside the 3D visual representation can feel uncomfortable due to the vergence-accommodation conflict [CKC*10, RD19]. The tablet also possesses an active screen which is then hidden by the AR-HMD overlay, leading to conflicts. These reasons led us to create a virtual selection tablet, different in size, position, and orientation to its physical counterpart, and to use relative motions. We propose two types of relative motions as we discuss next.

Relative-Aligned (RA) and Relative-Full (RF): For RA and RF, the user first initializes the virtual tablet within the 3D visualization space using NA. They then move the physical tablet and observe the motions of the virtual tablet in the visualization space, to either start an extrusion or to replace the virtual tablet from afar. Those mappings are then suitable for remote interactions on datasets that require large overviews. For RF, similar to the original implementation, we use a clutched interaction so that the tablet's actual orientation may differ from its virtual alignment. In contrast, the physical and the virtual tablets have the same orientation in RA. RA is thus useful for users focusing on the AR space, while RF with its clutched interaction is suitable for users focusing on the tablet.

In Appendix A we show how Bruckner et al.'s model [BIRW19] (Fig. 2) can be used to classify each technique depending on their directness. In the remainder of the paper, we reuse the following two spaces that Bruckner et al. introduced: The output space $\mathcal{O}$ which is the screen on which the user focuses (the screen of the tablet, the AR view, or the 2D external screen), and the manipulation space $\mathcal{M}$, which is the tablet the user manipulates.

To summarize the analysis of directness, the original technique uses indirect interactions, in particular due to the dimension differences between spaces $\mathcal{O}$ (2D screens) and $\mathcal{M}$ (3D tangible tablet). Moreover, as the external screen can be placed anywhere in the room, 3D movements ($\mathcal{M}$) of the tablet may not be aligned with what users perceive on the external screen ($\mathcal{O}$). This analysis may explain why Besançon et al.'s [BSA*19] participants experienced a high mental load, as directness is usually linked to the users' mental model, especially for novel interaction paradigms [BIRW19].

For AR mappings, NA is the spatially most direct method: it merges, with respect to the AR-HMD, the interaction/manipulation and the visualization/output spaces within the users' physical space. NA, however, has some limitations as we described. We consider NA and RA to be "direct" forms of interaction when compared to the original technique, even if one relies on absolute positions (NA) and the other on relative motions (RA). This difference, however, makes NA more direct than RA. RF with its clutched interaction is the least direct mapping when users focus on the AR space. Due to this property, however, RF is the spatially most direct approach when users focus on the tablet. To better understand the implications of these differences, we need to empirically study the different interaction mappings. Such an experiment may also lead to a better understanding of the user space $\mathcal{U}$ (i. e., the user's mental model) during interaction, which may depend on the screen on which the user focuses ($\mathcal{O}$), due to its influence on the level of spatial directness.

4. Experimental Verification

We conducted two controlled experiments. In the first we determined the "best" AR mapping from the three we introduced in Sect. A, and in the second we compare that to the original Tangible Brush. We make all gathered data available online along their respective pre-registrations (details below). We did not pre-register the criteria for selecting the AR mapping of the second experiment because they depended on the results from the first. We recruited 18 participants per experiment. Each participant performed only one of the two experiments to avoid learning effects. We reused the Galaxies and the Cylinder datasets from Besançon et al.'s study [BSA*19] (Fig. 3(b)–(c)). As all their datasets rely on extruded shapes along an axis, we added a helical Spring as a third dataset (Fig. 3(a)).

4.1. Implementation

Our system (`github.com/MickaelSereno/SciVis_Server/tree/CHI2020`) comprises four devices: an AR-HMD Microsoft HoloLens 2 (Fig. 4), a Samsung Galaxy Tab S4 tablet (Fig. 5) which we track with a VICON, and a server. Users can move and rotate the datasets, which we placed at a fixed position, on the tablet via FI3D [YSI*10]. While Tangible Brush works for any 3D data, we used point clouds in our study to measure the accuracy of the techniques we compare. We give more details of our implementation in Appendix B as, in this paper, we focus on our two experiments.

When users start a selection, we show a yellow outline to represent the virtual tablet in the AR view, and an orthographic projection of the data as seen from its virtual position on the tablet. The near-clipping plane is placed at $z = 0$ within the tablet camera's space. Normally the tablet's physical motion does not affect its virtual counterpart, but users can update the virtual tablet's size and the position on demand. The "position" mode repositions the virtual tablet in the 3D space using NA, and the "tangible" mode applies the current mapping to either re-adjust the position of the virtual tablet or to extrude a drawn 2D lasso. Users can adjust the size of the virtual tablet with a slider. As in the original work, we provide Boolean operations to adjust the selection (AND, OR, and NOT) and a constrained extrusion mode that moves the virtual tablet only along its normal. By removing hand shaking and tracking noise for the x- and y-axes, this mode can improve the accuracy of the selection when users extrude along the tablet's depth axis. During extrusions, the AR-HMD renders the volume outline the user created.

To not influence our study with the differences of the screen quality of AR-HMDs compared to large 2D screens, we simulate the original technique inside the AR-HMD using a 1.22 m × 0.68 m virtual screen. We adapted the original technique as follows to make it more comparable with our AR implementations, and to accommodate the AR-HMD setup: (i) we added the FI3D widgets to move/rotate datasets, (ii) we set the bird-of-view position to ease users with translations which do not rely on tangible movements anymore, (iii) we added the 3D wireframe volumes for users to (in)validate, (iv) we added tangible rotation to adjust the 3D scene at any step of the extrusion as rotations are strong depth cues, and (v) we allowed users to scale the tablet's view.

4.2. Tasks

Both experiments rely on the same tasks. For each mapping, participants had three training trials with simple shapes to ensure that they were well-trained. The first training session was the longest (up to 30 minutes) as participants also had to learn the tablet's interface and the tasks. Then, after a break, they performed the selections on the three datasets (Fig. 3), and we measured the quantitative metrics. For each dataset, participants performed two consecutive trials, resulting in six trials per mapping per participant. We asked participants to select the red and avoid the blue dots, and to be as accurate as possible without noticeably slowing down. To avoid overly long trials, we explained that it was nearly impossible to be perfectly accurate. We rendered selected dots with a lower saturation (i. e., tending to white) than non-selected dots, to allow participants to understand both selection status and dot character (red vs. blue).

5. Experiment 1: Best AR Mapping

We recruited 18 participants (ages 20–45, AVG=26, SD=5.8) to study the three AR mappings in our first experiment, pre-registered [BPSS*21, Har22] at `osf.io/pwauq`. We documented differences with the pre-registration on OSF (see `Errata.docx`). Participants performed all the trials for a mapping, before moving to the next one. For a given participant, we presented the datasets in the same order for each mapping. We counter-balanced the dataset order using a cyclic Latin square on $(P_{ID}/3) \bmod 3$ (integer division), and the technique order on $P_{ID} \bmod 3$, where ID is unique and ranges from 0 to 17. After each mapping, participants answered the corresponding part of the questionnaire (see additional materials). At the end of the experiment, participants answered the remaining general questions. Participants were unfamiliar with AR (1–5 on a 5-point Likert Scale, AVG=2.2, SD=1.5) and VR headsets (1–5 on a 5-point Likert Scale, AVG=2.1, SD=1.1), except for the 5 AR or VR researcher participants, but rated the VICON tracking as accurate (3–7 on a 7-point Likert Scale, AVG=5.8, SD=1.0). We hypothesize that:

- **H1** NA will be the fastest approach, as it is the most direct one.
- **H2** NA will be the most physically tiring mapping where users need to hold the tablet high. Indeed, visualizations are expected to be aligned with the user's eyes or chest to be comfortable, similar to Bach et al.'s [BSB*18] study.
- **H3** When focusing on the AR space, RA will be more precise and require less cognition than RF, and vice-versa otherwise. The accuracy and cognitive load of RF and RA will depend on which "screen" a participant focuses, as both mappings have different degrees of directness but rely on the same input modalities.
- **H4** We expect users to mainly focus on the AR space as they will need some depth perception to correctly move the tablet in the 3D space, where the AR-HMD outperforms the tablet.
- **H5** Participants will prefer RA. The selection often requires to scale up the virtual tablet as users cannot, otherwise, select what they intend to because the frame will be too small. NA is inappropriate to get the overview of the 3D scene (see Sect. 3). Participants will also focus on the AR view (**H4**), making RA more suitable than RF (**H3**) due to the alignment of spaces $\mathcal{O}$ and $\mathcal{M}$.

5.1. Results

We gathered $participants \times conditions \times datasets \times trials = 18 \times 3 \times 3 \times 2 = 324$ data points. We report our results using estimation techniques based on 95% confidence intervals (CIs) instead of p-values [Dra16, AGM19, BD19]. We refrain from dichotomously interpreting our results to follow current best practices [BD19, CDBG20, HHC*21]. Except where stated otherwise, we computed CIs from bootstrap resampling and computed all within-subjects pairwise comparisons by bootstrapping geometric means on ratios to emphasize effect sizes. We removed one trial $(P_{ID}; Mapping; Data; Trial) = (P_1; RA; Cylinder; 1)$ as the participant selected only a very small portion of the data, surely due to miss-clicks with the buttons "reset selection" or "end-task."

Completion Time. We started the timer at the beginning of each trial, which we paused during the computations of volumetric selections. To correct for positive skewness, we analyzed log-transformed speed metrics [SL10] and present the CIs of their anti-logged results (Fig. 6), which is standard for such data [JDF13, BIAI17]. We did not find evidence of one mapping being faster than another, not supporting **H1**. We experienced a learning effect (Fig. 7) as participants tried to understand and find a strategy for each dataset to solve the problems during their first respective mapping, before sticking to it for the others. Especially for the Galaxies and the Spring, the possible solutions were not obvious to find or to apply, when compared to the Cylinder where participants understood directly that they (1) can select the outer part by extruding a cylinder and (2) remove its inner part with a smaller cylinder. To a lesser extent, users got used to the devices as all techniques use the same input $\mathcal{M}$ and output $\mathcal{O}$ modalities, strengthening again this learning effect.

We found strong evidence that the Spring dataset was the longest to complete, which participants solved using two strategies. Some defined it mathematically using constrained Boolean operators. The number of operations is here higher than in the other datasets. Others followed the Spring using 6 DoF unions, which require slow movements as unconstrained operations need more control for accuracy.

Workload. We used the NASA-TLX [Har06] questionnaire to measure the participants' workload. We performed pairwise comparisons as a within-subject design to calibrate participants (Fig. 8(b)). Fig. 8 shows strong evidence that NA felt the least performant, some evidence that RA felt the most performant, and some evidence that

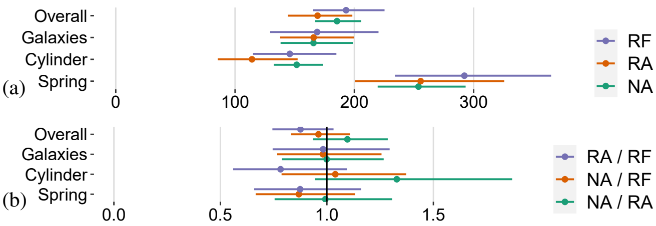


Figure 6: *Raw data (a) and pairwise within-subjects comparisons (b) of the Task-Completion Time, 1st experiment.*

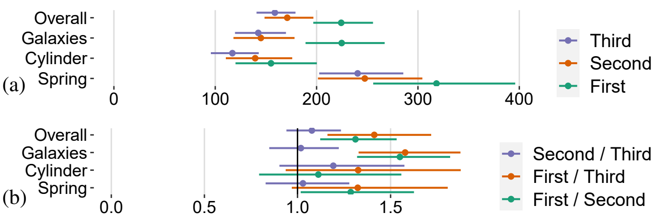


Figure 7: *Raw data (a) and pairwise within-subjects comparisons (b) of the Task-Completion Time following each participant's technique order, 1st experiment. This analysis was not pre-registered.*

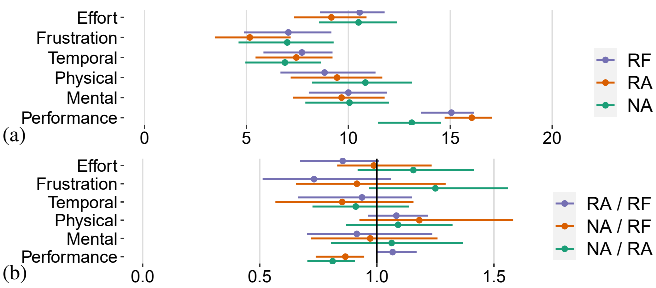


Figure 8: *Raw data (a) and pairwise within-subjects comparisons (b) for TLX workload, 1st experiment. We inverted the Performance metric compared to the original NASA-TLX questionnaire.*

RA required less effort than RF to participants. Finally, we found no evidence of a difference between NA, RA, and RF w.r.t. mental or physical workloads, not supporting **H2**.

Accuracy. To measure accuracy, we computed the F1 and MCC [CJ20] metrics, similar to Besançon et al. [BSA*19] (see their paper for the definitions). As both metrics yielded similar conclusions, we present here only the MCC score and report the F1 score in Appendix C. All mappings have a similar accuracy (Fig. 9(a)), with some evidence that RF is the most accurate one (Fig. 9(b)). There is weak evidence that RA is more accurate than NA for the Cylinder and the Galaxies, but with a small effect size ($< 5\%$).

Focus. Both screens were useful during extrusions (Fig. 10). Still, participants mainly relied on the AR view except for RF with the Galaxies, partially supporting **H4**. We cannot, however, support **H3** as we found evidence that RF is the most accurate mapping for participants who mainly relied on the AR view. This is especially true for NA (Spring and Cylinder), which we explain by its level of directness with respect to the AR view. As the Galaxies dataset contains much empty spaces compared to the others, participants could

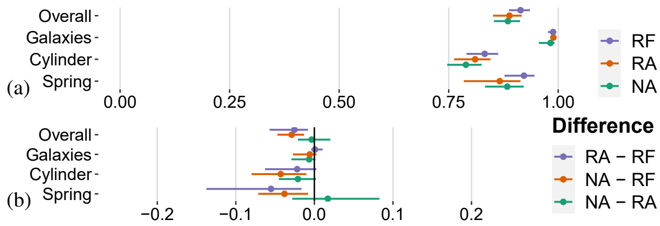
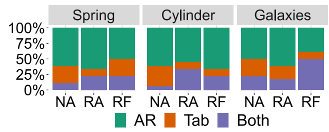


Figure 9: *Raw data (a) and pairwise within-subjects comparisons (b) of the MCC accuracy score, 1st experiment.*

Figure 10: *Proportion of which screen participants focused on during the extrusion steps, 1st experiment.*



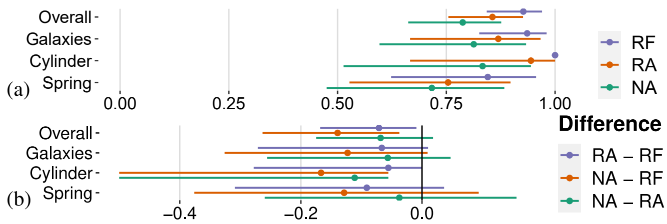
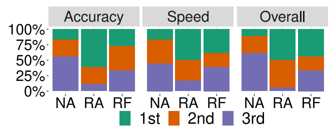


Figure 11: *Raw data (a) and pairwise within-subjects comparisons (b) of the proportion of constraint operations that contributed to the final results over unconstrained ones, 1st experiment.*

Figure 12: *Results of participants' rankings of the three mappings, 1st experiment; 1st is best.*



see the screen of the tablet as they interacted within the dataset. This may explain why they focused more on the tablet in the Galaxies dataset compared to the others using NA. Participants marked both answers in the questionnaire if they focused similarly on both views.

Constrained. The level of directness may explain why NA and RA might invite more unconstrained operations than RF (Fig. 11). We were surprised to see that participants solved the Spring primarily using constrained operations, while we designed it to encourage unconstrained ones. P_0, P_9, and P_{17} stated that it was difficult to perform unconstrained operations for this dataset, for which they spent a lot of time. P_{17} stated that he used only constrained operations using the relative mappings, just as he does with CAD software.

Preferences. Fig. 12 shows evidence that RA and RF were preferred to NA. Participants rated RA as being more accurate than RF, but the data (Fig. 9) suggests the opposite. We found strong evidence that participants preferred RA over RF overall. As participants mainly focused on the AR view during extrusion, we support **H5**.

Observations. P_{0-1}, P_{4-5}, and P_{10-11} stated that they cannot correctly see both the 3D view and the tablet when interacting inside the data with NA. P_1, however, said that she felt more in control with NA after some trials, maybe due to its higher directness.

Opinions and strategies differ among participants w.r.t. RF and RA. P_5 stated that RA was hard to use when she did not rely on constrained operations. P_0 found RA and RF to be similar in most cases where she relied on constrained operations. In our video recordings we saw participants P_{0-1}, P_{3-5}, P_{7-10}, and P_{12-16}, in some trials with RA and RF, operating the 2D virtual plane in a sideways fashion. This gave them a better perspective (and thus accuracy; **H3**) compared to the most direct mapping (NA). It is also useful, e. g., for the Cylinder whose inner part relies on two parallel planes easily visible from the side. In particular in RF, P_{3-6}, P_{10}, and P_{12} aligned the virtual tablet with the cylinder's circle side in the Cylinder dataset. They then physically moved around it to face the virtual plane at an angle of 90°. They then moved the tablet during the extrusion toward the cylinder's rectangular side, i. e., the tablet's normal. The virtual tablet (and the extrusion by extension) then moves along the depth axis of the cylinder, i. e., the virtual tablet's normal. P_0 manipulated the physical tablet from the other side of the Cylinder, facing it at 180°. These points may explain RF's higher accuracy compared to the other modes and why participants using RF focused mainly on the AR-HMD. The users' performance may depend on the possible rotation values (e. g., 45°, 90°, or 180°), which should be investigated as future work. However, we saw some participants hesitating about how they should physically place/orient their body in the 3D space before extruding the lasso in RF, which may be due to the disconnect between $\mathcal{O}$ and $\mathcal{M}$, explaining our workload results. Especially, P_1 had spatial issues to move the tablet in RF.

In the Galaxies data, users can select the red galaxy by extruding along the normal of the almost-planar blue galaxy, where users stop to maximize their accuracy. P_3, P_{9-10}, and P_{17} rotated the dataset before starting the selection. P_{17}, in addition, lowered the dataset to put the red galaxy at a comfortable position. He then selected it from top to bottom. P_0, P_3, and P_4 placed the virtual plane near the blue galaxy, before selecting the red one which was on top. They kneeled or lowered themselves to place the virtual plane and check its position. Finally, P_{11} selected the red galaxy using the unconstrained mode. In RA, he focused on the AR view sideways while moving the tablet around.

5.2. Discussion

Our participants mainly performed constrained operations, which correspond to target selection tasks. Reflecting on studies about Fitt's law [Bal04, MAK*88] and P_9's comments about her screen focus, we hypothesize that users focus on the AR view during their initial coarse but fast movements (ballistic phase [GH88]), and later on the tablet for precise movements. Indeed, the HMD facilitates a good 3D understanding of the dataset and its spatial relationship with the virtual plane. These facilitate fast movements towards the target, while the tablet, with its near-clipping plane, renders a precise position of where the virtual plane is within its neighborhood. We considered this hypothesis in the design of our second experiment.

The NA condition was not the preferred one (**H5**) mostly because of a lack of scene overview and because the AR hologram is rendered on top of the tablet, making both hardly readable. P_{10} used the relative mappings to place the tablet before the extrusion instead of using the "Position" button (which uses NA), to avoid this conflict. While we expected RA to be the most accurate (**H3**) with participants mainly focusing on AR (**H4**), the data shows that RF was the most accurate. However, our participants perceived RA to be more accurate than RF. Yet, the directness difference of RA and RF, when AR is the focus, may have biased the participants. Directness is then not always an indication for accuracy and can even bias users. Still, all mappings had a good accuracy. The main difference between RA and RF is the user's mental model. Results show that participants preferred RA over RF, possibly due to the virtual and the real tablet sharing the same orientation in AR space, on which most participants focused during extrusion.

Based on our results, we suggest to use relative mappings instead of NA. Users can still perform absolute operations by applying relative mappings with a relative distance of 0. Using the "Position" mode, users can replace the virtual tablet once they have drawn the lasso before starting an extrusion. In the next experiment, we compared the original setup with this AR one. As the AR mapping, we selected RA as it was the most preferred one. One may have chosen to compare a clutched interaction for both the original and the AR setups due to better accuracy. Should we find strong evidence that RA is more accurate than the original setup in this second study, we can also be confident that RF as well will be more accurate than the original setup, hence our choice. Moreover, as all mappings reached similar accuracy, we preferred to focus this experiment on studying the users' mental model rather than their level of performance.

6. Experiment 2: AR vs. Original

We recruited 18 participants, as pre-registered at `osf.io/rvpuc`. We had to replace one participant with another as this person did not respect our experiment protocol, but we include their results on OSF. The second experiment used the same protocol, tasks, and datasets as the first. We counter-balanced techniques (our *AR* technique and the original *2D* one) using cyclic Latin squares on P_{ID} mod 2 and datasets on $(P_{ID}/2)$ mod 3. In the 2D condition, participants could use the complete set of translations provided by the FI3D widgets. Participants (ages 21–36, AVG=25, SD=4.1) were unfamiliar with AR (1–2 on a 5-point Likert Scale, AVG=1.4, SD=0.50) and VR headsets (1–3 on a 5-point Likert Scale, AVG=1.7, SD=0.69), but rated the VICON tracking as accurate (3–7 on a 7-point Likert Scale, AVG=5.4, SD=1.1). We hypothesize that:

- **H6** *AR* will outperform *2D* both in cognitive load and participant preferences as participants now have better depth perception through the 3D stereoscopic view, and that *AR* is more direct than *2D* when focusing on the AR view.
- **H7** *AR* will be more accurate than *2D* for the same reasons, and because the input modalities are similar.
- **H8** *AR* will be faster than *2D* as *AR* allows participants to place the virtual tablet in a 1:1 mapping prior to extrusions, compared to *2D* which needs to either place the tablet using relative translations, or to move the dataset using the FI3D widgets.
- **H9** During extrusions, participants will use the AR view for coarse movements and use the tablet's view for refinements.

6.1. Results

We use the same reporting and analysis as for our first experiment.

Completion Time. We did not find evidence that one technique would be faster (Fig. 13), not supporting **H8**. We again observed a learning effect across all datasets between the first and the second

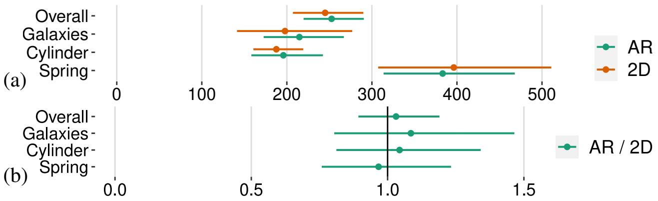


Figure 13: *Raw data (a) and pairwise within-subjects comparisons (b) for Task-Completion Time, 2^{nd} experiment.*

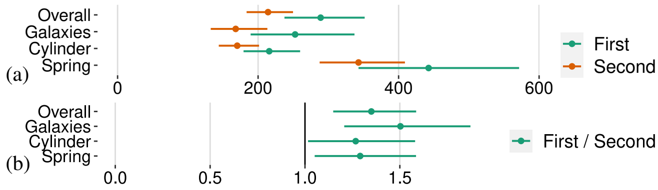


Figure 14: *Raw data (a) and pairwise within-subjects comparisons (b) of the Task-Completion Time following each participant's technique order, 2^{nd} experiment. This analysis was not pre-registered.*

techniques as ordered per participant (Fig. 14). Finally, we again see strong evidence that the Spring dataset was the slowest to solve.

Workload. We found strong evidence that users felt more performant, were less frustrated, needed to work less hard, and needed less mental effort at solving tasks using *AR* compared to *2D* (Fig. 15). We also note that the difference is large in most cases (Fig. 15(b)). This may be due to the AR view offering strong spatial cues about the spatial relationship between virtual tablet and 3D visualization and helps users to understand the 3D visualization, and due to the difference of the directness of the techniques. We thus confirm **H6**. The lack of strong evidence for differences in physical demand may be due to both techniques relying on the same input modalities.

Accuracy. Both techniques are similarly accurate (Fig. 16). *AR* seems more accurate than *2D* for the Galaxies ($< 1\%$) and the Cylinder ($\approx 3.6\%$), but with a small effect size (Fig. 16(b)). As users can observe more closely in AR compared to a 2D screen that does not offer zooming, we think that participants better fine-tuned their selections by altering small groups of dots in *AR* compared to *2D*. We found weak evidence that *2D* is more accurate than *AR* for the Spring ($\approx 1.5\%$). Overall, we cannot conclude that one technique is more accurate than another, not supporting **H7**. We thus cannot conclude about the RF mapping accuracy in AR compared to the original interaction. These diverging results may come from the participants' strategies concerning the use of constrained operations.

Constrained. When examining the ratio of useful constrained operations over total number of useful operations (Fig. 17), it seems that *AR* calls for more unconstrained operations compared to *2D* (Fig. 17(b)). This may be due to users having a better 3D understanding of the dataset in AR, especially for the Spring, and thus making them willing to work with 6 DoF. Moreover, the level of directness of RA (*AR*) over *2D* may invite more 6 DoF operations. However, unconstrained operations for the Spring are less accurate than constrained ones. This may explain the difference in accuracy.

Participants performed unconstrained operations in a *Naïve Approach* in *AR*: they placed the tablet near the Spring, drew a circle, and again replaced the tablet near the Spring before extruding. P_0 tried twice to solve the Spring in unconstrained mode in *AR*. He then tried and succeed at solving the Spring using the constrained mode. P_9 used constrained operations for the Spring in *2D* (his first condition) and used unconstrained operations in *AR*, where the 3D geometry was better understood. P_6 and P_8 used unconstrained operations in *AR* for their first trial of the Spring, but reverted to constrained operations for all the three other trials involving the Spring.

Focus. Fig. 18 shows strong evidence that users focus more on the External Screen in *AR* (the 3D AR view) than in *2D* (the large 2D

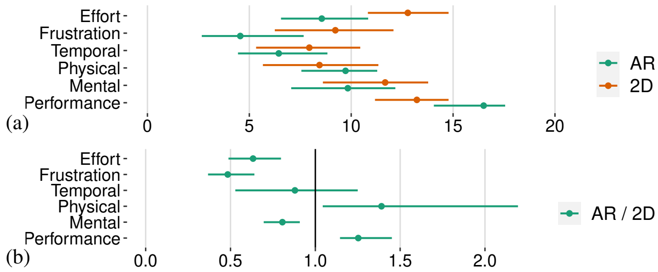


Figure 15: *Raw data (a) and pairwise within-subjects comparisons (b), of the TLX workload results, 2^{nd} experiment.*

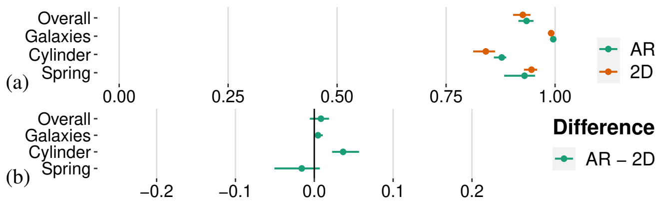


Figure 16: *Raw data (a) and pairwise within-subjects comparisons (b) of the MCC accuracy score, 2^{nd} experiment.*

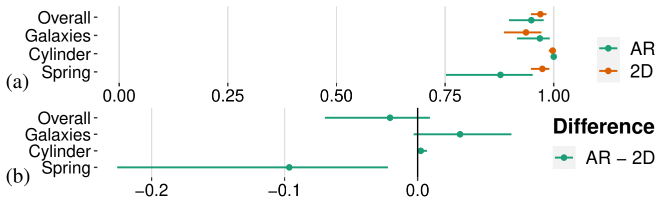
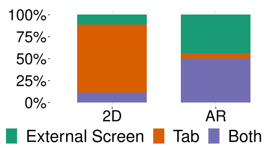


Figure 17: *Raw data (a) and pairwise within-subjects comparisons (b) of the proportion of constraint operations that contributed to the final results over unconstrained ones, 2^{nd} experiment.*

Figure 18: *Proportion of which screen participants focused on while placing the tablet with respect to the visualization prior extrusions, 2^{nd} experiment.*



screen) when placing the virtual tablet before starting an extrusion. Fig. 19 also shows the AR view supported our participants better than the 2D view. In *AR*, the 3D view gives strong spatial cues about where the virtual tablet resides w.r.t. the dataset, compared to the 2D screen. In *2D*, however, the camera associated with the FI3D widgets on the tablet matches its virtual position, which helps users to orient the dataset with respect to the virtual tablet.

Preferences. Fig. 20 shows strong evidence that users prefer *AR* over *2D*. Participants could rate each technique as being equal,

Figure 19: *Raw data (a) and pairwise within-subjects comparisons (b) on support of external screen (AR view or external 2D screen) to participants on a 7-point Likert Scale, 2nd experiment.*

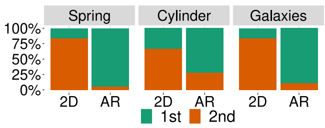


Figure 20: *Participants' technique ranking for each dataset, 2nd experiment. 1 is best.*

which we encoded with a value of "1st position" for both techniques. The effect size, however, is weaker for the Cylinder dataset. As *2D* allowed participants to manipulate the 3D visualization with respect to the static virtual tablet, some first oriented the Cylinder to face the virtual tablet along its extrusion axis and then moved it along the depth axis. Using the constrained mode, they only needed to move the tablet forward to first add the red dots, and then to remove the inner part of the Cylinder, which is a simple strategy.

Observations. *Extrusion in AR:* P_{0-6}, P_{8-14}, and P_{16} primarily focused on the AR view when extruding. P_0 and P_{10} said that they looked at the tablet only to remove the inner part of the Cylinder, which helped them with its near-clipping plane. Interestingly, P_8 used the AR view to check when to stop and if she was doing as expected, and used the tablet to know where to start using its near-clipping plane. P_1 used the AR view to understand the relationship between the virtual tablet and the dataset. She said that she rarely looked at the tablet compared to *2D*, except when she needed accuracy. P_5 used the tablet only to check if he had reached the end of the extrusion. P_{13} stated that constrained selections are like sliding movements, and he did not look much at both screens, except to check the start and the end of the movement on the AR view when required, e. g., with the inner part of the Cylinder. P_2, P_6, and P_9 focused on the tablet only during unconstrained operations (for the Spring), to check that their lasso was aligned with the circle the orthographic view of the tablet renders. P_2 also commented on the conflict between the AR and tablet views when both are merged.

In addition, some participants used the AR view during the whole process or to refine their movements. P_2 looked at it to check when the extrusion should end. P_3, P_4, P_{11-12}, P_{14}, and P_{16-17} used the AR view during the whole process. Finally, P_7 and P_{15} focused on the tablet most of the time and used the AR view to check their selections. P_{15} said that he especially looked at the tablet at the beginning and the end of his selections. Overall, we found conflicting evidence regarding **H9**, depending on participants' preferences and strategies. We therefore cannot conclude anything about it.

Extrusion in 2D: Before starting the extrusion, some participants sometimes placed the dataset (using the FI3D widgets) with respect to the fixed virtual tablet, instead of placing the virtual tablet with respect to the fixed dataset. They first oriented the dataset correctly, before translating it on the depth axis of the virtual plane, resulting in few but accurate operations for the placement before starting an extrusion. P_{13} commented that he used the 2D screen to adapt the depth position (with respect to the virtual tablet) of the dataset, before starting an extrusion. Once the dataset and the virtual tablet were aligned, P_0 used "Tangible Rotation" of the dataset in the 2D condition to arrange the plane of the virtual tablet to be parallel to the view shown on the external screen, for Cylinder and Spring (where he used only constrained movements). This allowed him to reduce the problem along one dimension to remove, e. g., the inner part of the Cylinder. P_7 used the same strategy all over her 2D condition trials. This "Tangible Rotation" can thus reduce the users' workload and can improve the original technique.

In *2D*, compared to *AR*, most participants (P_{1-2}, P_{4-5}, P_{7-8}, P_{10-14}, P_{11}, and P_{16-17}) looked primarily at the tablet during extrusions, maybe due to the mapping being defined by the screen of the tablet, which aligns spaces $\mathcal{O}$ and $\mathcal{M}$. P_{12} said that he better understood the mapping in *2D* than in *AR* when he focused on the tablet. P_5, P_8, and P_{12} used the external screen only to verify their extrusions. P_{10} was not facing the virtual screen in some trials but used it sometimes when she was lost in the 3D space. P_6 and P_{10} said that the perspective view as defined by the virtual tablet (left part of the external screen) was not useful at all. P_{11} never used the external screen during extrusions. P_{17}, curiously, used the external screen when she wanted to be accurate. She did not comment, however, what part of the external screen she was looking at.

Some participants, however, used both screens about equally. P_3 used the external 2D screen to understand the 3D geometry and extrusion axis, while he used the tablet's near-clipping plane to check when to stop. He also commented that he had issues with understanding the 3D scene in projected *2D* compared to *AR*. P_{15} looked at the tablet only at the beginning and the end of the extrusions, and relied on the external 2D screen when making coarse movements. Finally, P_6 and P_9 primarily used the right part of the external screen for their extrusions. However, P_9 commented that he focused more on the tablet during the last trials when he understood how the near-clipping plane of the tablet can support him.

6.2. Discussion

While we did not find evidence that *AR* would improve the users' accuracy and speed, our results answer our main research question about the users' workload (**H6**). Users largely preferred and felt more comfortable using a 3D AR view over a 2D screen. By providing them with an AR-HMD, we shifted the users' attention from the tablet to the AR view. This is a fundamental difference compared to the original Tangible Brush where most users primarily relied on the view of the tablet for almost everything. Our evidence shows that the AR-HMD and the tablet can be used jointly with different available strategies. Some users used the near-clipping plane of the tablet to be accurate, and others relied on the AR-HMD for most interactions. We saw three strategies to place the virtual tablet with respect to the 3D dataset. The first one relies on the FI3D widgets. Designers may consider what parameters to give to the FI3D camera, which should depend on the users' tasks (e. g., placing the dataset in the 3D space or with respect to the virtual tablet). While Besançon et al.'s [BSA*19] participants seemed to have troubles with rotating their datasets using tangible rotations, we did not experience such an issue with the FI3D widgets. The second strategy is to place the virtual tablet in a 1:1 mapping using the "Position" mode. The last one is to remotely control the virtual tablet using the "Tangible" mode to apply the current mapping. More research is needed to understand which strategies work best and under what scenarios.

7. Overall Discussion

We now discuss the implications of our findings for future research and the limitations of our experimental design.

Full 6-DoF Manipulations Are Not Always Preferred. We found evidence that users in immersive contexts might use full 6-DoF manipulations more than in non-immersive ones. This echoes previous research [BCD*18, DMI*18] on the benefits of 3D rendering environments to offer more intuitive 3D object manipulations. Nonetheless, we note that our participants used constrained manipulations very frequently. Our results add then to the discussion of integrating or separating DoF when performing 3D manipulations. Past research found conflicting evidence on what can be preferred: while DoF integration may lead to faster manipulations [WBAI19], DoF separation can lead to more precise and less frustrating manipulations [NBBW09, VCB09, MCG10, SW11, BIAI17, WBAI19]. Moreover, past research focusing on docking tasks for touch displays show that users generally decompose the task into a translation task and a rotation task, instead of doing everything simultaneously [MCG12, BBC13]. This probably explains why some of our participants used the FI3D widgets instead of tangible operations to align the virtual tablet and the dataset correctly in the 2D condition, as the FI3D widgets facilitate such a decomposition.

While most research so far highlighted that DoF integration and separation depended on the input technology [LJKM*17], our findings suggest that the degree of immersion of the output modalities also matters, as we saw that 3D AR views seem to invite more DoF integration. These results are particularly important when considering how interaction paradigms such as tangible input and mid-air gestures are praised for their "naturalness" [IU97]. Our results show evidence that such unconstrained operations also may not be preferred to be constantly used, even in a fully immersive setup.

All those points remind us of discussions and our experiences that there may be a limit of our human biomechanical or mental abilities that leads users to avoid manipulating more than 4 DoF simultaneously [MCG12, BBC13]. Yet, as this observation was not one of our primary research outcomes, further investigations on the impact of immersion on preferences for DoF integration is needed.

Spatial Directness Affects Mental Models. Using Bruckner et al.'s model [BIRW19] (Sect. 3), we showed that not all our AR techniques have the same directness, even though they rely on the same input modalities. Still, while our techniques are more direct than the original Tangible Brush, we did not perceive that directness would strongly influence performance. However, the users' mental model, strategies, and perceptions are different. We especially see that relying on a 3D display can shift the user attention from what they manipulate (the tangible tablet) to the explicitly shown data on the AR-HMD, which gives scene overviews and spatial cues.

Tablet Control for AR-HMDs. Several research projects combine AR-HMDs with workstations or tablets [BBK*06, RD19, WBR*20]. Our work adds to this discussion that rendering holograms through the AR-HMD around the secondary device engenders conflicts and make users uncomfortable. For instance, P_2 of the second experiment said that he had hard time to follow the Spring using unconstrained operations as he could not check on his tablet if the lasso was aligned with the "circular shape" of the spring. Designers should thus consider the density of the used visual representations in AR. At the same time, however, they should also distinguish visualizations where users need large overviews to those that need small overviews because spatially large representations combined with a relatively small mobile device would require constant focus switches. Except for the Galaxies, all our datasets are dense and require large overviews. A future study would shed light on different interaction designs for sparse datasets that require small overviews.

Limitations. Despite our careful design, our experiments have limitations. Our measurement of task completion time is a major one. Through our observations of the videos, it seems that participants spent a lot of time in their first trial with a dataset attempting to understand it and trying different strategies, before sticking to a suitable one. This may impact our results of the speed metric and of the participants' strategies. We share this limitation with previous experiments in the literature using similar experimental designs [YEII16, BSA*19]. Moreover, we did not associate a dialog box with the "Reset Selection" action which was sometimes pressed inadvertently, impacting our time measurements.

The accuracy of the tracking is another limitation. As the server merges the VICON tracking and the HoloLens one, fast head rotations can lead to a lag between both devices, which leads to synchronization conflicts on the server. Moreover, the HoloLens has difficulties to relocate itself during fast movements, i. e., the origin of its coordinate system is noisy, which strengthen this synchronization issue. Similar to raycasting paradigms, rotation noise can lead to strong displacements in, e. g., the orthographic projection of the tablet. However, this is true only when users rotate their head unnaturally fast. While all our participants rated the tracking as being accurate, we could solve this issue with an integrated tracking on the tablet, such as what current VR-HMDs do with their controllers.

8. Conclusion

Using Bruckner et al.'s model [BIRW19] (Sect. 3), we showed that our three adaptations of Tangible Brush within AR-HMD are more direct than the original technique. Directness in AR, however, is not always synonymous to performance, preference, strategies, and users' workload as we showed in our first experiment (Sect. 5). Specifically, we found strong evidence that users generally prefer to rely on remote interactions over interacting inside the visualization where the tangible display conflicts with the AR view, conflicting with traditional thoughts that users prefer direct manipulations in 3D spaces. As future work we envision to study the implications of the tangible tablet for spatial selections in VR environments as, compared to AR-HMD optical see-through ones, users cannot move as comfortably and cannot see the tablet with the same fidelity. However, VR environments benefit from a better management of visual conflicts between the virtual datasets and the physical tablet.

We then compared one remote adaptation in AR to the original Tangible Brush setup in a second experiment (Sect. 6). We found strong evidence that AR views, for 3D volumetric selections, lower significantly the users' workload compared to the 2D condition, while having similar accuracy and speed. Our results also show that users behave differently in AR environments compared to 2D ones, which should be further investigated in future work.

References

[AA13] Argelaguet F., Andujar C.: A survey of 3D object selection techniques for virtual environments. *Comput Graph 37*, 3 (May 2013), 121–136. doi:10.1016/j.cag.2012.12.003.

[ABB*01] Azuma R., Baillot Y., Behringer R., Feiner S., Julier S., MacIntyre B.: Recent advances in augmented reality. *IEEE Comput Graph Appl 21*, 6 (Nov. 2001), 34–47. doi:10.1109/38.963459.

[AGM19] Amrhein V., Greenland S., McShane B.: Scientists rise up against statistical significance. *Nature 567*, 7748 (2019), 305–307. doi:10.1038/d41586-019-00857-9.

[AKA*17] Arora R., Kazi R. H., Anderson F., Grossman T., Singh K., Fitzmaurice G.: Experimental evaluation of sketching on surfaces in VR. In *Proc. CHI* (2017), ACM, New York, pp. 5643–5654. doi:10.1145/3025453.3025474.

[ASM*04] Akers D., Sherbondy A., Mackenzie R., Dougherty R., Wandell B.: Exploration of the brain's white matter pathways with dynamic queries. In *Proc. VIS* (2004), IEEE CS, Los Alamitos, pp. 377–384. doi:10.1109/VISUAL.2004.30.

[Azu97] Azuma R. T.: A survey of augmented reality. *Presence (Camb) 6*, 4 (Aug. 1997), 355–385. doi:10.1162/pres.1997.6.4.355.

[Bal04] Balakrishnan R.: "Beating" Fitts' law: Virtual enhancements for pointing facilitation. *Int J Hum Comput Stud 61*, 6 (Dec. 2004), 857–874. doi:10.1016/j.ijhcs.2004.09.002.

[BBC13] Brouet R., Blanch R., Cani M.-P.: Understanding hand degrees of freedom and natural gestures for 3D interaction on tabletop. In *Proc. INTERACT* (2013), Springer, Berlin, pp. 297–314. doi:10.1007/978-3-642-40483-2_20.

[BBK*06] Bornik A., Beichel R., Kruijff E., Reitinger B., Schmalstieg D.: A hybrid user interface for manipulation of volumetric medical data. In *Proc. 3DUI* (2006), IEEE CS, Los Alamitos, pp. 29–36. doi:10.1109/VR.2006.8.

[BCD*18] Büschel W., Chen J., Dachselt R., Drucker S., Dwyer T., Görg C., Isenberg T., Kerren A., North C., Stuerzlinger W.: Interaction for immersive analytics. In *Immersive Analytics*. Springer Intern., Cham, 2018, ch. 4, pp. 95–138. doi:10.1007/978-3-030-01388-2_4.

[BD19] Besançon L., Dragicevic P.: The continued prevalence of dichotomous inferences at CHI. In *Proc. CHI* (May 2019), ACM, New York, pp. 14:1–14:11. doi:10.1145/3290607.3310432.

[BIAI17] Besançon L., Issartel P., Ammi M., Isenberg T.: Mouse, tactile, and tangible input for 3D manipulation. In *Proc. CHI* (2017), ACM, New York, pp. 4727–4740. doi:10.1145/3025453.3025863.

[BIRW19] Bruckner S., Isenberg T., Ropinski T., Wiebel A.: A model of spatial directness in interactive visualization. *IEEE Trans Vis Comput Graph 25*, 8 (Aug. 2019), 2514–2528. doi:10.1109/TVCG.2018.2848906.

[BPC19] Baloup M., Pietrzak T., Casiez G.: RayCursor: A 3D pointing facilitation technique based on raycasting. In *Proc. CHI* (2019), ACM, New York, pp. 101:1–101:12. doi:10.1145/3290605.3300331.

[BPSS*21] Besançon L., Peiffer-Smadja N., Segalas C., Jiang H., Masuzzo P., Smout C., Billy E., Deforet M., Leyrat C.: Open science saves lives: lessons from the COVID-19 pandemic. *BMC Medical Research Methodology 21*, 117 (June 2021), 1–18. doi:https://doi.org/10.1186/s12874-021-01304-y.

[BSA*19] Besançon L., Sereno M., Ammi M., Yu L., Isenberg T.: Hybrid touch/tangible spatial 3D data selection. *Comput Graph Forum 38*, 3 (June 2019), 553–567. doi:10.1111/cgf.13710.

[BSB*18] Bach B., Sicat R., Beyer J., Cordeil M., Pfister H.: The hologram in my hand: How effective is interactive exploration of 3D visualizations in immersive tangible augmented reality? *IEEE Trans Vis Comput Graph 24*, 1 (Jan. 2018), 457–467. doi:10.1109/TVCG.2017.2745941.

[BYK*21] Besançon L., Ynnerman A., Keefe D. F., Yu L., Isenberg T.: The state of the art of spatial interfaces for 3D visualization. *Comput Graph Forum 40*, 1 (Feb. 2021), 293–326. doi:10.1111/cgf.14189.

[CBL*17] Cordeil M., Bach B., Li Y., Wilson E., Dwyer T.: Design space for spatio-data coordination: Tangible interaction devices for immersive information visualisation. In *Proc. PacificVis* (2017), IEEE CS, Los Alamitos, pp. 46–50. doi:10.1109/PACIFICVIS.2017.8031578.

[CDBG20] Cockburn A., Dragicevic P., Besançon L., Gutwin C.: Threats of a replication crisis in empirical computer science. *Commun ACM 63*, 8 (July 2020), 70–79. doi:10.1145/3360311.

[CJ20] Chicco D., Jurman G.: The advantages of the matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics 21*, 6 (2020), 1–13. doi:10.1186/s12864-019-6413-7.

[CKC*10] Chan L.-W., Kao H.-S., Chen M. Y., Lee M.-S., Hsu J., Hung Y.-P.: Touching the void: Direct-touch interaction for intangible displays. In *Proc. CHI* (2010), ACM, New York, pp. 2625–2634. doi:10.1145/1753326.1753725.

[CZY*20] Chen Z., Zeng W., Yang Z., Yu L., Fu C.-W., Qu H.: LassoNet: Deep lasso-selection of 3D point clouds. *IEEE Trans Vis Comput Graph 26*, 1 (Jan. 2020), 195–204. doi:10.1109/TVCG.2019.2934332.

[dHKP02] de Haan G., Koutek M., Post F. H.: Towards intuitive exploration tools for data visualization in VR. In *Proc. VRST* (2002), ACM, New York, pp. 105–112. doi:10.1145/585740.585758.

[DMI*18] Dwyer T., Marriott K., Isenberg T., Klein K., Riche N., Schreiber F., Stuerzlinger W., Thomas B.: Immersive analytics: An introduction. In *Immersive Analytics*. Springer Intern., Cham, 2018, ch. 1, pp. 1–23. doi:10.1007/978-3-030-01388-2_1.

[Dra16] Dragicevic P.: Fair statistical communication in HCI. In *Modern Statistical Methods for HCI*. Springer Intern., Cham, 2016, ch. 13, pp. 291–330. doi:10.1007/978-3-319-26633-6_13.

[FCL09] Forsberg A., Chen J., Laidlaw D.: Comparing 3D vector field visualization methods: A user study. *IEEE Trans Vis Comput Graph 15*, 6 (Nov. 2009), 1219–1226. doi:10.1109/TVCG.2009.126.

[GB06] Grossman T., Balakrishnan R.: The design and evaluation of selection techniques for 3D volumetric displays. In *Proc. UIST* (2006), ACM, New York, pp. 3–12. doi:10.1145/1166253.1166257.

[GH88] Gan K.-C., Hoffmann E. R.: Geometrical conditions for ballistic and visually controlled movements. *Ergonomics 31*, 5 (May 1988), 829–839. doi:10.1080/00140138808966724.

[GXY12] Guo H., Xiao H., Yuan X.: Scalable multivariate volume visualization and analysis based on dimension projection and parallel coordinates. *IEEE Trans Vis Comput Graph 18*, 9 (2012), 1397–1410. doi:10.1109/TVCG.2012.80.

[Har06] Hart S. G.: Nasa-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 50*, 9 (2006), 904–908. doi:10.1177/154193120605000909.

[Har22] Haroz S.: Comparison of preregistration platforms. doi:10.31222/osf.io/zry2u.

[HHC*21] HELSKE J., HELSKE S., COOPER M., YNNERMAN A., BESANCON L.: Can visualization alleviate dichotomous thinking effects of visual representations on the cliff effect. *IEEE Trans Vis Comput Graph 63*, 8 (Apr. 2021), 70–79. doi:10.1109/TVCG.2021.3073466.

[HRD*19] HURTER C., RICHE N. H., DRUCKER S. M., CORDEIL M., ALLIGIER R., VUILLEMOT R.: FiberClay: Sculpting three dimensional trajectories to reveal structural insights. *IEEE Trans Vis Comput Graph 25*, 1 (Jan 2019), 704–714. doi:10.1109/TVCG.2018.2865191.

[HRGMI14] HINCAPIÉ-RAMOS J. D., GUO X., MOGHADASIAN P., IRANI P.: Consumed endurance: A metric to quantify arm fatigue of mid-air interactions. In *Proc. CHI* (2014), ACM, New York, pp. 1063–1072. doi:10.1145/2556288.2557130.

[HvDG94] HERNDON K. P., VAN DAM A., GLEICHER M.: The challenges of 3D interaction: A CHI '94 workshop. *SIGCHI Bull 26*, 4 (Oct. 1994), 36–43. doi:10.1145/191642.191652.

[IU97] ISHII H., ULLMER B.: Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proc. CHI* (1997), ACM, New York, pp. 234–241. doi:10.1145/258549.258715.

[JDF13] JANSEN Y., DRAGICEVIC P., FEKETE J.-D.: Evaluating the efficiency of physical visualizations. In *Proc. CHI* (2013), ACM, New York, pp. 2593–2602. doi:10.1145/2470654.2481359.

[JLS*13] JACKSON B., LAU T. Y., SCHROEDER D., TOUSSAINT K. C., KEEFE D. F.: A lightweight tangible 3D interface for interactive visualization of thin fiber structures. *IEEE Trans Vis Comput Graph 19*, 12 (Dec. 2013), 2802–2809. doi:10.1109/TVCG.2013.121.

[KI13] KEEFE D. F., ISENBERG T.: Reimagining the scientific visualization interaction paradigm. *Computer 46*, 5 (May 2013), 51–57. doi:10.1109/MC.2013.178.

[KIL*03] KNISS J., IKITS M., LEFOHN A., HANSEN C., PRAUN E., ET AL.: Gaussian transfer functions for multi-field volume visualization. In *Proc. VIS* (2003), IEEE CS, Los Alamitos, pp. 497–504. doi:10.1109/VISUAL.2003.1250412.

[KJK*15] KIM K., JACKSON B., KARAMOUZAS I., ADEAGBO M., GUY S. J., GRAFF R., KEEFE D. F.: Bema: A multimodal interface for expert experiential analysis of political assemblies at the Pnyx in ancient Greece. In *Proc. 3DUI* (2015), IEEE CS, Los Alamitos, pp. 19–26. doi:10.1109/3DUI.2015.7131720.

[KMLM16] KWON O.-H., MUELDER C., LEE K., MA K.-L.: A study of layout, rendering, and interaction methods for immersive graph visualization. *IEEE Trans Vis Comput Graph 22*, 7 (July 2016), 1802–1815. doi:10.1109/TVCG.2016.2520921.

[LJKM*17] LAVIOLA JR J. J., KRUIJFF E., MCMAHAN R. P., BOWMAN D., POUPYREV I. P.: *3D User Interfaces: Theory and Practice*. Addison-Wesley, Boston, 2017.

[LODI16] LÓPEZ D., OEHLBERG L., DOGER C., ISENBERG T.: Towards an understanding of mobile touch navigation in a stereoscopic viewing environment for 3D data exploration. *IEEE Trans Vis Comput Graph 22*, 5 (May 2016), 1616–1629. doi:10.1109/TVCG.2015.2440233.

[MAK*88] MEYER D. E., ABRAMS R. A., KORNBLUM S., WRIGHT C. E., SMITH J. E.: Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychol Rev 95*, 3 (July 1988), 340–370. doi:10.1037/0033-295X.95.3.340.

[MCG10] MARTINET A., CASIEZ G., GRISONI L.: The effect of DOF separation in 3D manipulation tasks with multi-touch displays. In *Proc. VRST* (2010), ACM, New York, pp. 111–118. doi:10.1145/1889863.1889888.

[MCG12] MARTINET A., CASIEZ G., GRISONI L.: Integrality and separability of multi-touch interaction techniques in 3D manipulation tasks. *IEEE Trans Vis Comput Graph 18*, 3 (Jan. 2012), 369–380. doi:10.1109/TVCG.2011.129.

[MDB*17] MERRIAUX P., DUPUIS Y., BOUTTEAU R., VASSEUR P., SAVATIER X.: A study of vicon system positioning performance. *Sensors 17*, 7 (2017), 1591:1–1591:18. doi:10.3390/s17071591.

[MMNK*20] MONTANO-MURILLO R. A., NGUYEN C., KAZI R. H., SUBRAMANIAN S., DIVERDI S., MARTINEZ-PLASENCIA D.: Slicing-volume: Hybrid 3D/2D multi-target selection technique for dense virtual environments. In *Proc. VR* (2020), IEEE CS, Los Alamitos, pp. 53–62. doi:10.1109/VR46266.2020.00023.

[NBBW09] NACENTA M. A., BAUDISCH P., BENKO H., WILSON A.: Separability of spatial manipulations in multi-touch interfaces. In *Proc. GI* (2009), CHCCS, Toronto, pp. 175–182. URL: https://graphicsinterface.org/proceedings/gi2009/gi2009-23/.

[RCK*17] RO H., CHAE S., KIM I., BYUN J., YANG Y., PARK Y., HAN T.: A dynamic depth-variable ray-casting interface for object manipulation in AR environments. In *Proc. SMC* (2017), IEEE CS, Los Alamitos, pp. 2873–2878. doi:10.1109/SMC.2017.8123063.

[RD19] REIPSCHLÄGER P., DACHSELT R.: DesignAR: Immersive 3D-modeling combining augmented reality with interactive displays. In *Proc. ISS* (2019), ACM, New York, pp. 29–41. doi:10.1145/3343055.3359718.

[RN10] ROSA D. A. W., NAGEL H. H.: Selection techniques for dense and occluded virtual 3D environments, supported by depth feedback: Double, bound and depth bubble cursors. In *Proc. SCCC* (2010), IEEE CS, Los Alamitos, pp. 218–225. doi:10.1109/SCCC.2010.51.

[Ser21] SERENO M.: *Collaborative Data Exploration and Discussion Supported by Augmented Reality*. Theses, Université Paris-Saclay, Dec. 2021. URL: https://tel.archives-ouvertes.fr/tel-03550513/document.

[SGHV19] SURALE H. B., GUPTA A., HANCOCK M., VOGEL D.: TabletInVR: Exploring the design space for using a multi-touch tablet in virtual reality. In *Proc. CHI* (2019), ACM, New York, pp. 13:1–13:13. doi:10.1145/3290605.3300243.

[SL10] SAURO J., LEWIS J. R.: Average task times in usability tests: What to report? In *Proc. CHI* (2010), ACM, New York, pp. 2347–2350. doi:10.1145/1753326.1753679.

[Ste12] STENHOLT R.: Efficient selection of multiple objects on a large scale. In *Proc. VRST* (2012), ACM, New York, pp. 105—112. doi:10.1145/2407336.2407357.

[SW11] STUERZLINGER W., WINGRAVE C. A.: The value of constraints for 3D user interfaces. In *Virtual Realities*. Springer, Vienna, 2011, ch. 11, pp. 203–223. doi:10.1007/978-3-211-99178-7_11.

[SWB*22] SERENO M., WANG X., BESANÇON L., MCGUFFIN M. J., ISENBERG T.: Collaborative work in augmented reality: A survey. *IEEE Trans Vis Comput Graph* (2022). To appear. doi:10.1109/TVCG.2020.3032761.

[Tuk77] TUKEY J. W.: *Exploratory Data Analysis*, vol. 2. Pearson, London, 1977.

[UZW*07] ULINSKI A., ZANBAKA C., WARTELL Z., GOOLKASIAN P., HODGES L. F.: Two handed selection techniques for volumetric data. In *Proc. 3DUI* (2007), IEEE CS, Los Alamitos, pp. 107–114. doi:10.1109/3DUI.2007.340782.

[VCB09] VEIT M., CAPOBIANCO A., BECHMANN D.: Influence of degrees of freedom's manipulation on performances during orientation tasks in virtual reality environments. In *Proc. VRST* (2009), ACM, New York, pp. 51–58. doi:10.1145/1643928.1643942.

[WA04] WARE C., ARSENAULT R.: Frames of reference in virtual object rotation. In *Proc. APGV* (2004), ACM, New York, pp. 135–141. doi:10.1145/1012551.1012576.

[WBAI19] WANG X., BESANÇON L., AMMI M., ISENBERG T.: Augmenting tactile 3D data navigation with pressure sensing. *Comput Graph Forum 38*, 3 (June 2019), 635–647. doi:10.1111/cgf.13716.

[WBG*19] WANG X., BESANÇON L., GUÉNIAT F., SERENO M., AMMI M., ISENBERG T.: A vision of bringing immersive visualization to scientific workflows. In *Workshop CHI-IA* (2019).

[WBR*20] WANG X., BESANÇON L., ROUSSEAU D., SERENO M., AMMI M., ISENBERG T.: Towards an understanding of augmented reality extensions for existing 3D data analysis tools. In *Proc. CHI* (2020), ACM, New York, pp. 1–13. doi:10.1145/3313831.3376657.

[WF96] WARE C., FRANCK G.: Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Trans Graph 15*, 2 (Apr. 1996), 121–140. doi:10.1145/234972.234975.

[WM08] WARE C., MITCHELL P.: Visualizing graphs in three dimensions. *ACM Trans Appl Percept 5*, 1 (Jan. 2008), 2:1–2:15. doi:10.1145/1279640.1279642.

[WVFH12] WIEBEL A., VOS F. M., FOERSTER D., HEGE H.-C.: WYSIWYP: What you see is what you pick. *IEEE Trans Vis Comput Graph 18*, 12 (Dec. 2012), 2236–2244. doi:10.1109/TVCG.2012.292.

[XFAT12] XU P., FU H., AU O. K.-C., TAI C.-L.: Lazy selection: A scribble-based tool for smart shape elements selection. *ACM Trans Graph 31*, 6 (Nov. 2012), 142:1–142:9. doi:10.1145/2366145.2366161.

[YEII16] YU L., EFSTATHIOU K., ISENBERG P., ISENBERG T.: CAST: Effective and efficient user interaction for context-aware selection in 3D particle clouds. *IEEE Trans Vis Comput Graph 22*, 1 (Jan. 2016), 886–895. doi:10.1109/TVCG.2015.2467202.

[YSI*10] YU L., SVETACHOV P., ISENBERG P., EVERTS M. H., ISENBERG T.: FI3D: Direct-touch interaction for the exploration of 3D scientific visualization spaces. *IEEE Trans Vis Comput Graph 16*, 6 (Nov./Dec. 2010), 1613–1622. doi:10.1109/TVCG.2010.157.