Authors' version. Related publications at: https://hci.uni-konstanz.de/publications/

Jonathan Wieland, Samar Abed, Anke V. Reinschluessel, Johannes Zagermann, Harald Reiterer, Tiare Feuchtner. Extending the Fishing Reel – Improving Multiple Object Selection in VR Using Transparency and a Resizable Pointer. *AlpCHI 2026*, Ascona, Switzerland, 2026, doi: 10.1145/3780045.3780046.

Extending the Fishing Reel – Improving Multiple Object Selection in VR Using Transparency and a Resizable Pointer

Jonathan Wieland
Human Computer Interaction Group
University of Konstanz
Konstanz, Germany
jonathan.wieland@uni-konstanz.de

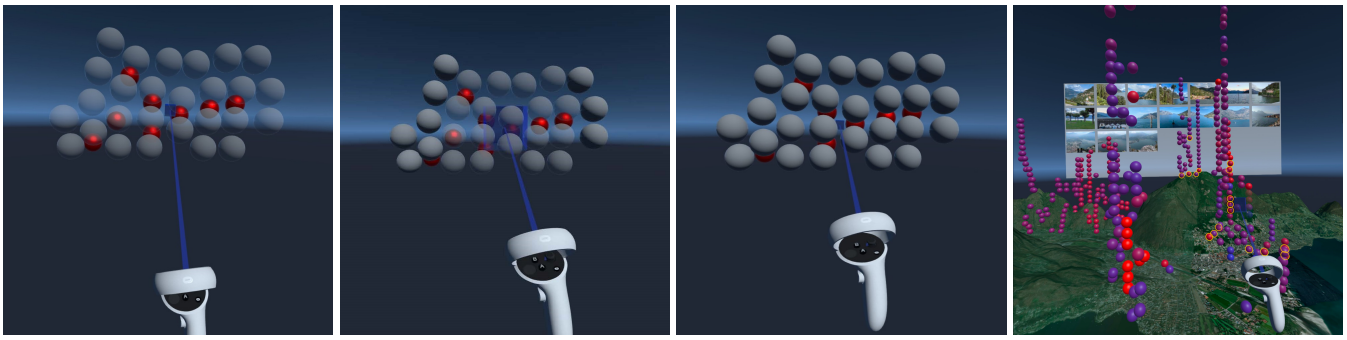
Samar Abed
Human Computer Interaction Group
University of Konstanz
Konstanz, Germany
samar.abed@uni-konstanz.de

Anke V. Reinschluessel
Human Computer Interaction Group
University of Konstanz
Konstanz, Germany
anke.reinschluessel@uni-konstanz.de

Johannes Zagermann
Human Computer Interaction Group
University of Konstanz
Konstanz, Germany
johannes.zagermann@uni-konstanz.de

Harald Reiterer
Human Computer Interaction Group
University of Konstanz
Konstanz, Germany
harald.reiterer@uni-konstanz.de

Tiare Feuchtner
Human Computer Interaction Group
University of Konstanz
Konstanz, Germany
tiare.feuchtner@uni-konstanz.de



(a) Transparency Layer (TL) (b) Transparency Flashlight (TF) (c) No Transparency (NT) (d) Demo App 'VR Photo Explorer'.

Figure 1: We introduce a novel VR multi-selection technique that addresses the problem of occlusion with a transparency layer. We compare a non-transparent baseline (c) to two transparency variants: (a) Transparency Layer (TL) applies transparency to all spheres between the selection box and the user; (b) Transparency Flashlight (TF) renders only spheres around the box semi-transparent. (d) We illustrate our technique in the demo app 'VR Photo Explorer'. Map Data: © Mapbox, © OpenStreetMap.

Abstract

One of the key challenges in distant object selection in virtual environments (VE) is 'target occlusion.' This paper presents a novel approach to overcome this issue when selecting multiple (occluded) objects at a distance in VR. We extend the well-known 'fishing reel' metaphor, adding transparency within the selection area and allowing manipulation of its size. In a user study (n = 30, within-subject design), we investigate two transparency concepts ('layer' vs. 'flashlight') and compare their efficiency against no transparency. Our results indicate that our transparency approach reduces physical demand and the need for adjusting one's position in space. They also show that density of objects in the VE influences performance, indicating that our method provides a greater benefit in dense settings. We demonstrate our technique by visualizing geo-tagged Flickr photos on a 3D alpine map, where users can select individual images from dense clusters to view in a floating gallery.

CCS Concepts

• **Human-centered computing** → **User studies**; Walkthrough evaluations; **Virtual reality**; **Laboratory experiments**; **Pointing**.

Keywords

virtual reality, multiple object selection, lab study, user study

ACM Reference Format:
Jonathan Wieland, Samar Abed, Anke V. Reinschluessel, Johannes Zagermann, Harald Reiterer, and Tiare Feuchtner. 2026. Extending the Fishing Reel – Improving Multiple Object Selection in VR Using Transparency and a Resizable Pointer. In *1st International Conference on Human-Computer Interaction in the Alps (AlpCHI 2026), March 01–05, 2026, Ascona, Switzerland.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3780045.3780046

ACM ISBN 979-8-4007-1948-6/2026/03
https://doi.org/10.1145/3780045.3780046

1 Introduction

Selection is one of the four basic 3D manipulation tasks identified by LaViola Jr. et al. [13]. Nearly all kinds of interaction in virtual environments (VE) involve some form of selection, e. g., choosing menu items. In VEs with many objects, for example, in immersive analytics, the user might not only want to select single but multiple objects. This can be tiresome when the selection method is 'one by one', i. e., sequential. One potential solution is providing a so-called multiple object selection (MOS) technique. There are a few approaches to achieve multiple selections with more ease, such as brushes, lassos, or 'magic wands' [21, 24]. In addition, users might also face the problem of occlusion when selecting multiple objects. In more crowded VEs (cf. Fig. 1), target objects may not be completely visible, and selecting these with ease poses another challenge to the selection method. While there are various methods to disambiguate a selection (e. g., [1, 2]), they generally target single object selection processes.

To address the issue of selecting multiple occluded objects, we combined a ray-casting method with (1) a transparency layer to overcome the problem of occlusion and (2) a resizable selection cube for easy multiple object selection. Instead of a simple depth indicator on the ray (cf. [2, 7]), we add a resizable box that allows direct selection of every object within the box. Similar to the interaction proposed as the 'fishing reel metaphor' [3], the user can manipulate the distance of the box on the ray using the controls on the controller. By controlling the distance of the selection box, the user further controls a transparency layer that renders any objects between the user and the box transparently to increase the visibility of the objects at a distance. To evaluate our approach, we did a design study with an early prototype to uncover usability issues and then ran a user study with 30 participants, comparing two variants of transparency visualization to a standard ray-casting method. All three variations included our multiple object selection box. Our results show that by using our transparency approach, participants' walking distance was significantly lower, and the task completion time was lower, especially for densely arranged maps. Furthermore, the physical demand was significantly lower for the transparency conditions. These results indicate that by adding transparency, we can provide an easy way to overcome occlusion – which could also be applied to single object selection without the need for complex weighting algorithms as proposed by related works [5, 17].

In addition to the controlled evaluation, we implemented a demo application that visualizes geo-tagged Flickr photos on a 3D alpine terrain map (see Fig. 1(d)). The spheres can be color-coded by attributes such as popularity, season, or license, and users can employ our multi-selection technique to explore these and view selected images in a floating gallery.

2 Related Work

In this paper, we focus on pointing as a selection method for *distant* objects. Therefore, we only report on pointing techniques for (multiple) distant objects. For more comprehensive overviews, including a classification of selection techniques, we refer to Argelaguet and Andujar [1] and Yu et al. [26].

The earliest method to select objects at a distance was proposed by Mine [15] and was essentially a ray casting without explicitly stating it. To improve manipulation, Bowman and Hodges [3] extended it by allowing to change the distance to the selected object and coined it as using "a 'fishing reel' metaphor". Variations of the ray casting method are the 'Depth Ray', 'Lock Ray', and 'Flower Ray' [7]. The Depth Ray is similar to the fishing reel; a depth marker on the ray allows for adjusting the selection depth and selecting the nearest object. For the Lock Ray, the procedure is similar, but only after 'locking the ray' does the depth marker appear. The Flower Ray also first locks all objects along the ray and then shows them at the front, arranged like a flower, to select the intended object in a second step. The selection technique 'IntenSelect' modifies the ray to 'bend' towards objects nearby by a calculated score to make selection easier [5]. A similar approach was followed by Steinicke et al. [20], who also bent a ray towards the closest selectable object. The work of Schmidt et al. [17] also tries to use weighting schemes to determine which object(s) the user wanted to select based on the frustum along the pointing direction. The 'SQUAD technique' [12] is similar to the Flower Ray, but the target area is a ray cast, creating a sphere-shaped selection area at the destination, and its content is then split up to refine the selection in multiple steps. Cashion et al. [4] expand the SQUAD technique by adding the possibility to zoom, calling it 'Expand'. As most previous systems allow selecting multiple objects by selecting them one by one (sequential approach), others allow automatically selecting multiple objects in local proximity to the target using a 'magic wand' [21] or a variable cone [19]. Ro et al. [16] targeted the issue of selection ambiguity in the depth and created a depth-sensitive ray with a manual fine-tuning of the depth. A variation of the ray cast was presented by Baloup et al. [2], introducing the 'RayCursor', which allows manipulation of the depth of the selection. Wang et al. [22] present an approach to allow for targeted selection of occluded densely placed objects by maximizing "the visible footprint" in the center of the selection area.

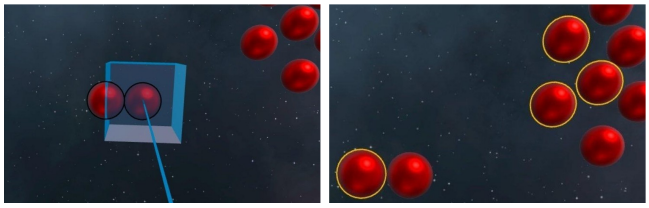


Figure 2: Objects with a black outline intersect the selection box (left), and selected items have a yellow outline (right).

Many approaches in research focus on selecting objects at a distance, but only a small subset addresses the interaction complexity of occluded objects at a distance – part of the identified key challenges for VR object selection in a recent survey [26]. Therefore, in this paper, we present an approach that allows detailed (multiple) object selection at a distance by making the hidden objects visible.

3 Design of the MOS Techniques

Extending on related works, we designed a MOS technique with three different ways for handling occlusion. Our final design, as

described in the following, resulted from a design review involving seven VR experts who tested an initial prototype of our technique [10], as a fast and effective test of our concept [14].

All three variants allow volume-based selection (cf. [25]) of multiple objects at an adjustable distance with a resizable selection box on a ray. Objects intersecting with the box can be selected or deselected and provide visual feedback (black outline when (de-)selectable, yellow when selected; see Fig. 2).

We used the thumbstick on the controller to adjust the distance and the size of the selection box (see Fig. 3). We used a natural mapping (e. g., also used by [16]) as a forward movement increases the distance and backward movement decreases the distance. Moving the thumbstick to the left reduces the box size to a minimum of 0.1 m edge length (0.001 m^3), and moving it to the right increases the size up to 1.5 m edge length (3.38 m^3). The intensity of moving the thumbstick is mapped to increase/decrease the speed. Additionally, it is possible to change both aspects (size and distance) at once by a diagonal movement of the thumbstick, i. e., moving it to the bottom-right corner decreases the distance while increasing the selection box's size.

Selection of objects within the box happens by pressing the trigger button on a virtual reality (VR) controller while deselection is pressing the hand trigger (cf. Fig. 3). Multiple objects can be (de-)selected by continuously pressing either trigger button and moving it through multiple objects in a brush-like fashion [24] (also called 'sequential object selection').

Figure 3: The thumbstick controls the ray length and the size of the selection box (left). The trigger button is used for selection and the hand trigger for deselection (right).

To solve the occlusion problem, we designed two ways to make distant occluded objects more visible: *Transparency Layer (TL)* and *Transparency Flashlight (TF)*. For both methods, the distance of the selection box defines the opacity behavior of surrounding objects, with objects behind the box rendered opaque and objects between the user and the box rendered transparent. As a control condition, we also designed a version without transparency (*No Transparency (NT)* – see Fig. 1(c)).

The design of TL incorporates a view-filling layer at the distance of the selection box that renders everything between the selection box and the user semi-transparent (see Fig. 1(a)). This allows seeing the objects clearer to improve the selection.

The variant TF also renders objects between the selection box and the user (semi-)transparent but not all objects. Inspired by a flashlight highlighting only part of the environment, within an area double the size of the selection box, everything is (semi-)transparent (see Fig. 1(b)). Therefore, the size of the area depends on the size of the selection box – the smaller the box, the fewer objects in front are transparent. This method was integrated to investigate the effect of whether it influences the user experience when all closer objects are transparent.

4 Demo Application 'VR Photo Explorer'

To illustrate the applicability of our techniques beyond abstract tasks, we implemented a demo application that visualizes geotagged Flickr photos on a 3D terrain map for any location, such as an alpine landscape (see Fig. 1(d)). The terrain is rendered with Mapbox[1] elevation and imagery data, and photos are fetched from the Flickr API[2]. Based on the available metadata of each photo, it is represented as a sphere at its geographic coordinates, and colored by attributes such as popularity, season, or license. To reduce overlap in dense areas, spheres are vertically offset with a repulsion scheme, creating a bar-chart-like appearance while preserving map positions. Photos selected by the user are displayed in a floating gallery above the map. This allows users to explore photo clusters, for example along a planned hiking route. With our multi-selection technique, they can browse the most-viewed images at huts or summits along the route, or filter by season to compare the landscape looks in different conditions. Users could then adjust their route, e. g., to include particularly scenic viewpoints in the gallery.

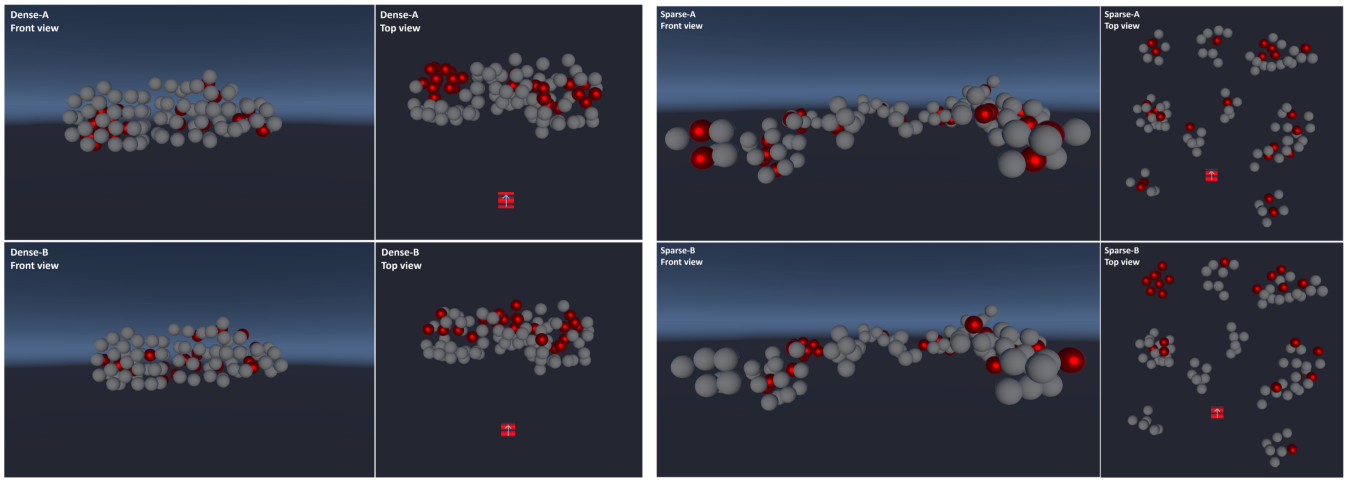
5 User Study

To evaluate our techniques in a controlled environment, we conducted a within-subjects user study with our three techniques representing the conditions (TL, TF, and NT). As selection in dense environments is more challenging, we included four VEs with two different object densities (sparse and dense, see Fig. 4). We chose two different arrangements for both types of density maps to ensure that any difference is not solely based on one specific arrangement. For the dense maps, all spheres are in one 'cloud', and for the sparse maps, they are distributed in nine smaller clouds over the whole space. Each object was a sphere with a diameter of 20 cm, and we ensured no spheres overlapped.

We are interested in the movement behavior, performance, user experience, and perceived task load. We measured movement behavior in the form of walking distance, and performance was evaluated by task completion time and error rate. To investigate user experience and task load, we used the established questionnaires User Experience Questionnaire (UEQ) [18] and NASA Task Load Index (NASA TLX) [8, 9] as well as a semi-structured interview.

We used the Meta Quest 2[3] as apparatus, and participants held one accompanying controller in their dominant hand to perform the study task. The VE was developed using Unity 2020.3.36[4] using the XR Interaction Toolkit (version 1.0.0-pre.8)[5] and the Oculus XR Plugin (version 1.10.1)[6]. We configured our lab to provide a walkable area of 5.0 × 5.0 meters, and participants were allowed to

[1]Data from *Mapbox* https://www.mapbox.com/about/maps and *OpenStreetMap* http://www.openstreetmap.org/copyright – last accessed: 12th Dec. 2025
[2]https://www.flickrhelp.com/hc/en-us/articles/4404070036884-Flickr-API – last accessed: 12th Dec. 2025
[3]https://www.meta.com/quest/products/quest-2/ – last accessed: 9th Dec. 2025
[4]https://unity.com/ – last accessed: 9th Dec. 2025
[5]https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@1.0/manual/index.html – last accessed: 9th Dec. 2025
[6]https://docs.unity3d.com/2021.1/Documentation/Manual/com.unity.xr.oculus.html – last accessed: 9th Dec. 2025



(a) Dense Maps A and B

(b) Sparse Maps A and B

Figure 4: We explore our technique for selecting targets (red spheres) in two dense maps (Dense-A and Dense-B) and two sparse maps (Sparse-A and Sparse-B). Here, we visualize both the front view and top view out of the VR scene for better visibility.

walk around. Our prototype featured a virtual guardian boundary that was activated when participants approached within 40 cm of the designated area's perimeter, thereby preventing collisions with walls or exiting the allocated space. All participants started at the same starting position close to one edge of the area, facing toward its center. Feedback on the number of selected spheres was displayed next to the controller without an indication of whether these were correctly selected or not.

5.1 Participants

We recruited 30 participants (17 females, 12 males, one non-binary) between 19 and 33 years old ($M = 23.7$, $SD = 3.66$) at our local university who had diverse backgrounds (e. g., biological sciences, psychology, or mathematics). Thirteen participants had prior experiences in immersive VR, and nine had experience with 3D modeling software. All participants were right-handed; 18 had visual impairments corrected by glasses ($n = 12$) or contact lenses ($n = 6$), and two were affected by dyschromatopsia.

5.2 Task

Each of the maps in Fig. 4 included 100 spheres, 25 red spheres as the target objects, and 75 grey spheres as clutter. Therefore, the task was to select all 25 red spheres and none of the grey ones. Participants started the task by clicking on a start button, and the participants decided by themselves when they were done by pressing 'A'. For each condition (TL, TF, NT), the participants had to select the 25 spheres on all four density maps in a counterbalanced order.

5.3 Procedure

We welcomed participants and informed them about the purpose and procedure of our study. Afterward, participants signed a consent form and completed a demographic questionnaire. The examiner then introduced participants to the HMD and offered support for properly fitting the device and picking up the controller with their dominant hand. Participants performed three sets in total, one for each condition in a counterbalanced order. Each set consisted of a step-by-step tutorial explaining the general use of the specific technique, a training phase, the actual study tasks (see above), and filling out the NASA TLX and UEQ. Participants were asked to complete the tasks as fast and accurately as possible. Sessions were concluded with a semi-structured interview and took approximately one hour. Participants were compensated for their time.

6 Results

In the following, we present the main findings of our lab study. We use subscript $_{TL}$ for indicating *Transparency Layer*, subscript $_{TF}$ for *Transparency Flashlight*, and subscript $_{NT}$ for *No Transparency*; likewise we use subscript $_{D}$ for the dense environments and subscript $_{S}$ for the sparse environments. We used IBM SPSS[7] to analyze the data, assumed an alpha level of .05 for statistical significance, and applied the Bonferroni correction to adjust for multiple comparisons.

6.1 Movement Behaviour

The average walking distances for the dense and sparse maps and all three conditions are visualized in Fig. 5(a). As Shapiro-Wilk tests indicated significant deviations from normality, we used a non-parametric approach for statistical analysis. Friedman's ANOVA revealed statistically significant differences between conditions for the average walking distances in the dense maps ($\chi^2(2) = 21.067$, $p < .001$). Post-hoc pairwise comparisons showed that participants, on average, walked shorter distances in the TL condition ($Mdn_{TL} = 6.19$, $z = -4.389$, $p < .001$) and the TF condition ($Mdn_{TF} = 7.76$, $z = -3.357$, $p = .002$) compared to the NT condition ($Mdn_{NT} = 10.66$). Also, for the sparse maps, Friedman's ANOVA revealed statistically significant effects between conditions ($\chi^2(2) = 6.2$,

[7] https://www.ibm.com/products/spss-statistics – last accessed: 9th Dec. 2025

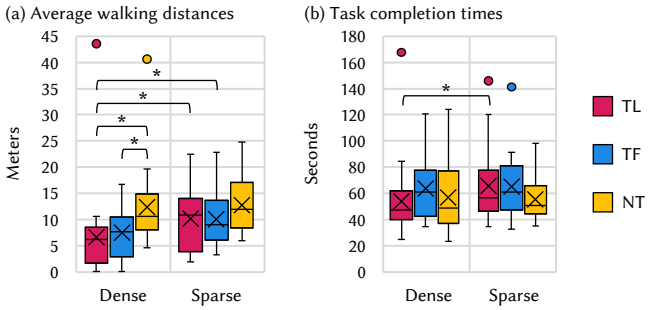


Figure 5: (a) Average walking distances (in meters) for dense and sparse maps. (b) Boxplots for average task completion times (in seconds) for dense and sparse maps.

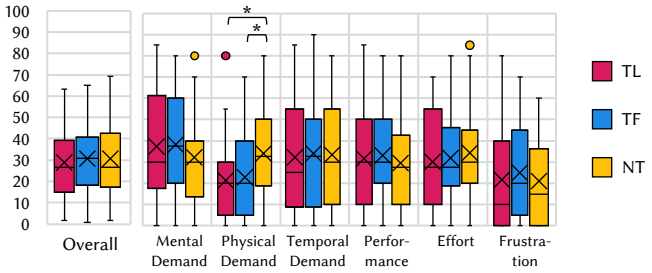


Figure 6: Boxplots for NASA TLX scores.

$p = .045$). However, post-hoc analysis did not show any statistically significant effects. When comparing the two density maps with each other, Wilcoxon signed rank tests showed that average walking distances were lower for the dense maps with the TL ($Mdn_D = 6.19$ m, $Mdn_S = 10.92$ m, $z = 3.240, p = .001$) and the TF condition ($Mdn_D = 7.76$ m, $Mdn_S = 9.02$ m, $z = 2.828, p = .005$).

6.2 Performance

We measured average task completion times (see Fig. 5(b)) and error rates for both density maps to assess the performance of our techniques. As for each of these data sets, Shapiro-Wilk tests indicated significant deviations from normality, we used a non-parametric approach for statistical analysis. In terms of task completion times, Friedman's ANOVA revealed no statistically significant differences between conditions for both density maps. However, comparing the two density maps with each other, Wilcoxon signed rank tests revealed that the arrangement of targets has an influence on the condition TL: as could be expected, task completion times were significantly lower for the dense maps ($Mdn_D = 47.6$ s) compared to the sparse maps ($Mdn_S = 56.83$ s, $z = 2.931, p = .003$). Analyzing the error rate, which is defined as the number of 'clutter spheres' that were wrongly selected by participants, we found no significant differences – neither between conditions nor between maps.

6.3 Workload

The workload was assessed using the NASA TLX, and the results are shown in Fig. 6. Regarding the overall scores, analysis with Friedman's ANOVA revealed no significant differences. An analysis of the subscales revealed a significant difference for the subscale *Physical Demand* ($\chi^2(2) = 16.253, p < .001$). Post-hoc comparisons showed that participants rated both the TL ($Mdn_{TL} = 20, z = -3.227, p = .004$) and the TF condition ($Mdn_{TF} = 20, z = -2.969, p < .009$) significantly better than the NT condition ($Mdn_{NT} = 32.5$). The mean reported physical demand for TL is 21.33 ($SD_{TL} = 18.80$), for TF, it resulted in a mean of 22.83 ($SD_{TF} = 20.12$), while the NT condition had an average of 34 ($SD_{NT} = 20.10$).

6.4 User Experience

User experience was assessed using the UEQ (see Fig. 7). While Friedman's ANOVA revealed no significant differences between the conditions, all conditions achieved high values (i. e., positive evaluation > 0.8, scale $[-3; +3]$). In direct comparison, the TL condition received the highest ratings for attractiveness ($1.91_{TL} > 1.77_{NT} > 1.57_{TF}$), pragmatic (efficiency, perspicuity, dependability) ($1.82_{TL} > 1.70_{NT} > 1.59_{TF}$), and hedonic (stimulation, novelty) qualities ($1.63_{TL} > 1.48_{TF} > 1.34_{NT}$).

In terms of user preference, participants provided a fractional raking at the end of the study (see Fig. 8(a)). We resolved ties by using ranks (1.5 for tie on rank 1, 2.5 for tie on rank 2). Friedman's ANOVA showed no statistically significant differences between subjective rankings of conditions. While 11 participants placed TL on rank 1, another 10 ranked it on rank 3 (rank 2: $n = 5$, tie: $n = 4$). A similar unclear picture was painted for the TF condition: eight participants ranked it 1, 10 ranked it 2, and 9 ranked it 3 (ties: $n = 3$). Although NT had the lowest mean rank ($M_{NT} = 2.05$), with $n = 6$ for rank 1, 8 placed it on the second rank and 9 on third place, with $n = 7$ ties with the other conditions.

When asked which variation they would use for dense or sparse maps, participants chose TL for the dense maps 14 times and for sparse maps 15 times. The TF method was chosen similarly often, 13 times for either map, while NT was only chosen 5 times for dense maps and 8 times for sparse maps (see Fig. 8(b)).

During the interview, the participants mentioned that the TL condition is 'easy to learn and use', 'convenient', and that it is easy to find the targets. Furthermore, they liked that they did not need to move to see everything. On the other hand, they also criticized the transparency as 'irritating' and 'confusing'. The condition TF received similar feedback, such as being described as 'easy' and 'clear', and also that it is less physically demanding. However, they also found it 'confusing' and 'challenging' to use. Participants also mentioned that it was fun to have transparency and be able to walk around. Condition NT was also mentioned as being fun and also as more 'natural', but participants also perceived it as 'stressful' and 'annoying', as well as that they were not sure if they picked the correct spheres. They did not like that they *had to* walk.

7 Discussion

In this study, we investigate the effect of using two variants of transparency to target the occlusion problem when selecting objects in VR. Based on the movement behavior, which we measured in walking distance during the tasks, we can conclude that using transparency helped with occlusion, as we observed significantly less walking in space to select the target objects. We observed this behavior for both types of maps, with objects arranged more densely

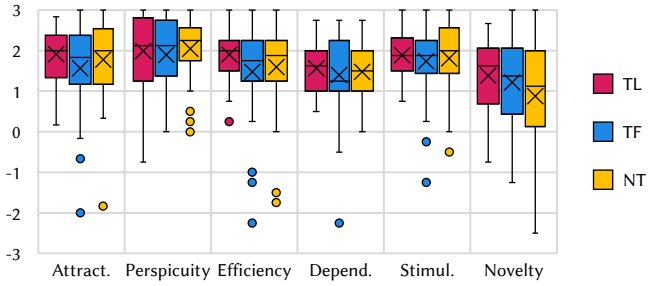


Figure 7: Boxplots for the UEQ scores.

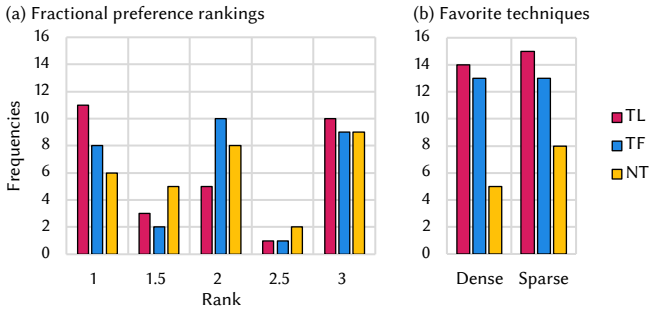


Figure 8: Bar charts for the (a) factional preference rankings and the (b) favorite techniques per density map.

and those arranged more sparsely. This result is also supported by the results of the NASA TLX, where the physical demand was significantly lower for both transparency conditions. This is especially important, as ray casting methods often suffer from fatigue [6, 23]. Also, when examining large data collections, e. g., in immersive analytics, the users spend more time in the setup, and longer walking distances can add to the experience fatigue. The effect is especially prominent on dense maps, where the average walking distance was the lowest overall and especially compared to the NT condition (6.66 m (TL) & 7.58 m (TF) vs. 12.37 m (NT)). Participants also positively commented on the fact that this visualization reduced their need to walk around the space, and the UEQ subscale 'Efficiency' received the highest scores for TL. One participant (P19) even said that "this is going to be more than annoying" when starting on the NT condition. Even though both map structures (dense and sparse) resulted in shorter walking distances for the transparency conditions, the effect was more prominent for dense maps, which might be rooted in the fact that the objects were more concentrated in one place and required less area exploration. This was also visible in the task completion times, as for the TL condition, people were significantly faster on dense maps compared to sparse maps.

While the walking distance and related measures showed a significant difference between the conditions, the majority of the other measures were not significant. We also did not see a difference between the two transparency conditions, and we assume that the users focused on the objects to select and by pointing there, in both transparency conditions, the area of interest was sufficiently better visible, making any differences negligible. Further research should investigate this in greater depth, to see if there are spatial arrangements of data, where either method provides benefits. Performance measures might have been impacted by low contrast, as participants reported that "it was difficult to differentiate between the colors" when the objects were semitransparent (P20). This might have had a negative impact on both task completion time and user experience. Even though the transparency was helpful in terms of reducing the walking distance, and participants also reported that "transparency helps a lot" (P17) and that it is easy to "see through" the clutter (P14), future work has to explore the issue of reduced contrast. Another factor potentially influencing task completion time was that we observed participants walking around to make sure they collected the correct spheres before ending the trial, even though the instruction was 'as fast as possible'.

Despite not being significantly different, all UEQ results indicate good user experience, with values > 0.8 (scale: [-3,3]), e. g., pragmatic qualities: $TL = 1.82; NT = 1.70; TF = 1.59$ (cf. sec. 6.4, Fig. 7). This aligns with the qualitative results from the interviews, as participants found the NT condition was more natural and "intuitive" (P25), and the overall interaction concept of all conditions was perceived as "fun."

One interesting aspect of the interviews was that the participants reported that the controls for the depth and size of the box were (too) sensitive, as the feedback during the design review was that it was too slow. To achieve a grounded understanding of what is a comfortable speed, it might be worthwhile to investigate the speed separately and also take training effects into account, as many of our participants did not have prior experience with VR.

In this work, we demonstrated (1) that our approach of an adjustable selection box – adjustable in size and distance – works for selecting multiple objects at a distance and (2) that transparency provides a measurable advantage. While we used an artificial study task for the controlled evaluation of our techniques, we also implemented the 'VR Photo Explorer' (cf. Sec. 4). This highlights how the method can be applied to existing datasets with dense and overlapping content, where reducing physical navigation and effort is more relevant than selection accuracy. Future work could further investigate our techniques in other real-world settings, e. g., in cooperation with experts from immersive analytics or geographic visualization, where large datasets naturally suffer from occlusion.

8 Conclusion

This paper presents a novel approach for multiple object selection in VR that addresses the issue of occlusion when interacting with distant and densely placed virtual objects or data points. Our approach is an extension of ray casting that integrates transparency and adjustable selection areas, inspired by Bowman and Hodges's 'fishing reel' metaphor [3]. In a user study ($n = 30$), we compared two approaches of integrating transparency (layer vs. flashlight) against no transparency. Results indicate that transparency reduces physical demand and the need for repositioning, with stronger benefits in dense object arrangements. Further, we demonstrate the applicability of our techniques in an exemplary 'VR Photo Explorer' application, in which geo-tagged Flickr photos are displayed on a 3D terrain map. A further potential application domain involves the visualization of large amounts of data, such as in immersive analytics (e. g., [11]).

Acknowledgments

This research was funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) – Project-ID 251654672 – TRR 161.

References

[1] Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (2013), 121–136.

[2] Marc Baloup, Thomas Pietrzak, and Géry Casiez. 2019. Raycursor: A 3d pointing facilitation technique based on raycasting. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[3] Doug A. Bowman and Larry F. Hodges. 1997. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics* (Providence, Rhode Island, USA) *(I3D '97)*. Association for Computing Machinery, New York, NY, USA, 35–ff. https://doi.org/10.1145/253284.253301

[4] Jeffrey Cashion, Chadwick Wingrave, and Joseph J. LaViola Jr. 2012. Dense and Dynamic 3D Selection for Game-Based Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics* 18, 4 (2012), 634–642. https://doi.org/10.1109/TVCG.2012.40

[5] Gerwin de Haan, Michal Koutek, and Frits H. Post. 2005. IntenSelect: Using Dynamic Object Rating for Assisting 3D Object Selection. In *Proceedings of the 11th Eurographics Conference on Virtual Environments* (Aalborg, Denmark) *(EGVE'05)*. Eurographics Association, Goslar, DEU, 201–209.

[6] João Marcelo Evangelista Belo, Anna Maria Feit, Tiare Feuchtner, and Kaj Grønbæk. 2021. XRgonomics: Facilitating the Creation of Ergonomic 3D Interfaces. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 290, 11 pages. https://doi.org/10.1145/3411764.3445349

[7] Tovi Grossman and Ravin Balakrishnan. 2006. The Design and Evaluation of Selection Techniques for 3D Volumetric Displays. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology* (Montreux, Switzerland) *(UIST '06)*. Association for Computing Machinery, New York, NY, USA, 3–12. https://doi.org/10.1145/1166253.1166257

[8] 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.

[9] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[10] Rex Hartson and Pardha Pyla. 2019. Chapter 25 - Analytic UX Evaluation: Data Collection Methods and Techniques. In *The UX Book (Second Edition)* (second edition ed.), Rex Hartson and Pardha Pyla (Eds.). Morgan Kaufmann, Boston, 545–560. https://doi.org/10.1016/B978-0-12-805342-3.00025-4

[11] Sebastian Hubenschmid, Jonathan Wieland, Daniel Immanuel Fink, Andrea Batch, Johannes Zagermann, Niklas Elmqvist, and Harald Reiterer. 2022. ReLive: Bridging In-Situ and Ex-Situ Visual Analytics for Analyzing Mixed Reality User Studies. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 24, 20 pages. https://doi.org/10.1145/3491102.3517550

[12] Regis Kopper, Felipe Bacim, and Doug A. Bowman. 2011. Rapid and Accurate 3D Selection by Progressive Refinement. In *Proceedings of the 2011 IEEE Symposium on 3D User Interfaces (3DUI '11)*. IEEE Computer Society, USA, 67–74.

[13] Joseph J LaViola Jr., Ernst Kruijff, Ryan P McMahan, Doug Bowman, and Ivan P Poupyrev. 2017. *3D user interfaces: theory and practice*. Addison-Wesley Professional.

[14] Thomas Memmel, Fredrik Gundelsweiler, and Harald Reiterer. 2007. Agile human-centered software engineering. In *BCS-HCI'07: 21st British HCI Group Annual Conference on People and Computers*. 167–175.

[15] Mark R Mine. 1995. Virtual environment interaction techniques. *UNC Chapel Hill CS Dept* (1995).

[16] Hyocheol Ro, Seungho Chae, Inhwan Kim, Junghyun Byun, Yoonsik Yang, Yoonjung Park, and Tackdon Han. 2017. A dynamic depth-variable ray-casting interface for object manipulation in ar environments. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2873–2878.

[17] Greg Schmidt, Yohan Baillot, Dennis G. Brown, Erik B. Tomlin, and J. Edward II Swan. 2006. Toward Disambiguating Multiple Selections for Frustum-Based Pointing. In *Proceedings of the IEEE Conference on Virtual Reality (VR '06)*. IEEE Computer Society, USA, 129. https://doi.org/10.1109/VR.2006.133

[18] Martin Schrepp, Jörg Thomaschewski, and Andreas Hinderks. 2017. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 6 (12/2017 2017), 103–108. https://doi.org/10.9781/ijimai.2017.09.001

[19] Rongkai Shi, Yushi Wei, Xuning Hu, Yu Liu, Yong Yue, Lingyun Yu, and Hai-Ning Liang. 2024. Experimental Analysis of Freehand Multi-object Selection Techniques in Virtual Reality Head-Mounted Displays. *Proc. ACM Hum.-Comput. Interact.* 8, ISS, Article 529 (Oct. 2024), 19 pages. https://doi.org/10.1145/3698129

[20] Frank Steinicke, Timo Ropinski, and Klaus Hinrichs. 2006. Object selection in virtual environments using an improved virtual pointer metaphor. In *Computer Vision and Graphics: International Conference, ICCVG 2004, Warsaw, Poland, September 2004, Proceedings*. Springer, 320–326.

[21] Rasmus Stenholt. 2012. Efficient Selection of Multiple Objects on a Large Scale. In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology* (Toronto, Ontario, Canada) *(VRST '12)*. Association for Computing Machinery, New York, NY, USA, 105–112. https://doi.org/10.1145/2407336.2407357

[22] L. Wang, J. Chen, Q. Ma, and V. Popescu. 2021. Disocclusion Headlight for Selection Assistance in VR. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE Computer Society, Los Alamitos, CA, USA, 216–225. https://doi.org/10.1109/VR50410.2021.00043

[23] Colin Ware and Leonard Slipp. 1991. Using velocity control to navigate 3D graphical environments: A comarison of three interfaces. In *Proceedings of the Human Factors Society Annual Meeting*, Vol. 35. Sage Publications Sage CA: Los Angeles, CA, 300–304.

[24] Graham J Wills. 1996. Selection: 524,288 ways to say" this is interesting". In *Proceedings IEEE Symposium on Information Visualization'96*. IEEE, 54–60.

[25] Zhiqing Wu, Difeng Yu, and Jorge Goncalves. 2023. Point-and volume-based multi-object acquisition in vr. In *IFIP Conference on Human-Computer Interaction*. Springer, 20–42.

[26] Difeng Yu, Tilman Dingler, Eduardo Velloso, and Jorge Goncalves. 2024. Object Selection and Manipulation in VR Headsets: Research Challenges, Solutions, and Success Measurements. *ACM Comput. Surv.* 57, 4, Article 98 (Dec. 2024), 34 pages. https://doi.org/10.1145/3706417