3D Tabletop AR: A Comparison of Mid-Air, Touch and Touch+Mid-Air Interaction

Carole Plasson, Dominique Cunin, Yann Laurillau, Laurence Nigay

3D Tabletop AR: A Comparison of Mid-Air, Touch and Touch+Mid-Air Interaction

Carole Plasson[1,2]
[1]Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
F-38000 Grenoble, France
carole.plasson@univ-grenoble-alpes.fr

Dominique Cunin[2]
[2]Ecole Superieure d'Art et Design Grenoble-Valence
F-26000 Valence, France
dominique.cunin@esad-gv.fr

Yann Laurillau[1]
[1]Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
F-38000 Grenoble, France
yann.laurillau@univ-grenoble-alpes.fr

Laurence Nigay[1]
[1]Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG
F-38000 Grenoble, France
laurence.nigay@univ-grenoble-alpes.fr

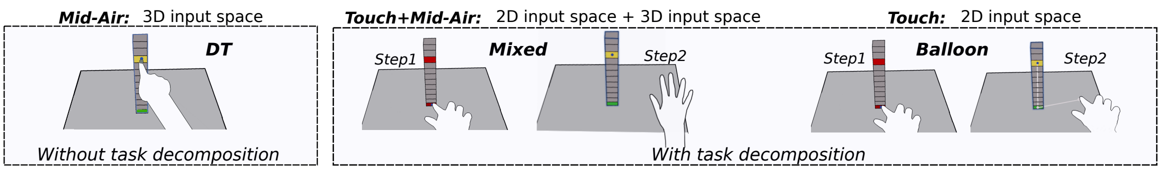


Figure 1: The 3 compared techniques classified according to their input spaces and whether they involve task decomposition

ABSTRACT

This paper contributes a first comparative study of three techniques for selecting 3D objects anchored to the table in tabletop Augmented Reality (AR). The impetus for this study is that touch interaction makes more sense when the targeted objects are anchored to the table. We experimentally compare touch and a mixed (touch+mid-air) techniques with the common direct mid-air technique. The touch and mixed techniques involve a decomposition of the 3D task into a 2D task by touch on the table followed by a 1D task by touch or mid-air interaction. Results show that: (1) The touch and mixed techniques present completion times similar to the mid-air technique and are more accurate than the mid-air technique; (2) The mixed technique defines a good compromise between accuracy of touch interaction and speed of mid-air interaction.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Mixed / augmented reality**.

KEYWORDS

Augmented Reality; HMD; Tabletop; 3D interaction; Touch; Mid-air

ACM Reference Format:
Carole Plasson, Dominique Cunin, Yann Laurillau, and Laurence Nigay. 2020. 3D Tabletop AR: A Comparison of Mid-Air, Touch and Touch+Mid-Air Interaction. In *International Conference on Advanced Visual Interfaces (AVI '20), September 28-October 2, 2020, Salerno, Italy*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3399715.3399836

AVI '20, September 28-October 2, 2020, Salerno, Italy

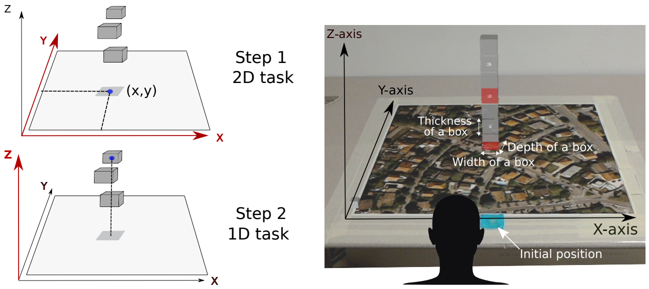
ACM ISBN 978-1-4503-7535-1/20/09…$15.00
https://doi.org/10.1145/3399715.3399836

1 INTRODUCTION

Tabletop Augmented Reality (AR) systems combine a tangible surface and virtual objects. They support a variety of applications such as architecture, urban design/visualization systems [3], games and 3D modelling [13]. In this paper, we study interaction techniques for selecting 3D objects linked with/anchored to the table in tabletop AR. Anchored objects are frequent in various applications as illustrated in the following scenario: *Presentation of flats for sale. A 3D virtual model of a district is placed on the table and allows buyers to visualize the buildings and the roads. The property developer selects flats in buildings in order to present a virtual tour of them.*

Beyond 3D objects directly placed on the table as in the above scenarios, 2D projections of floating objects (e.g., shadows) could be displayed on the table for selection. As 3D anchored objects are omnipresent in tabletop AR, it is important to understand the most effective interaction for selecting them. While 3D AR selections mainly rely on mid-air input space ("air-tap" gesture with the Microsoft HoloLens [10]), tabletop AR systems provide a physical support (i.e. the table) to the users and thus an additional 2D input space allowing touch interaction [1, 4, 7, 15]. Touch-only techniques use task decomposition to select 3D objects. Previous studies compared mid-air and touch interaction [1, 9, 16] but primarily focus on selecting 3D objects without a visual link with the table. These studies found that touch interaction reduces fatigue in comparison with mid-air interaction [1, 16], but their results about the task decomposition are contradictory. In addition to techniques involving 2D or 3D input space, mixed techniques integrate and unify these two input spaces "to bridge this divide" [2]: these mixed selection techniques seamlessly combine touch and mid-air interaction [2, 20] and involve a third input space (2D + 3D space).

This paper contributes a comparative study of three techniques (see Figure 1) that goes beyond previous work (1) by considering objects anchored to the table and (2) by comparing 3D mid-air



(a) 3D selection task decomposition **(b) The 3 dimensions of the boxes**

Figure 2: Selection of 3D objects anchored to the table

interaction and 2D touch interaction with mixed touch+mid-air interaction. Moreover we provide recommendations on input spaces that enrich existing design insights for tabletop AR as in [13].

2 3D OBJECT SELECTION TECHNIQUES

We designed 3 techniques for selecting a 3D box in a stack of 3D boxes. Figure 1 classifies them according to their input interaction space and whether they involve task decomposition.

1) DirectTouch - DT is a 3D direct mid-air technique considered as a baseline (see Figure 1, left). Users directly touch the targeted box to select it. A blue sphere follows the users' fingertip. A selection occurs when this sphere is inside a box.

The two following techniques decompose the 3D selection task into a 2D task (Step 1) and a 1D task (Step 2) as shown in Figure 2a.

2) Mixed combines touch and mid-air interaction (see Figure 1, center). Step 1, users select the base of the stack by touching the table: the blue sphere which follows their fingertip should be totally inside the base. The sphere can then only move along the Z-axis. Step 2, users move their hand in mid-air, the height of the blue sphere follows the height of their finger.

3) Balloon is a touch-only technique largely inspired by *Balloon Selection* [1] which uses a metaphor of the manipulation of a helium balloon attached to a string (see Figure 1, right). The string is held by two control points placed on the table and the balloon moves up or down by reducing or increasing the distance between the two points. Step 1, users select the base of the stack, and thus place a first control point at the center of the base. Step 2, users place the same finger on the table (wherever they want) in order to place a second point thus making the virtual string appear. The balloon is the same blue sphere as in *DT* and *Mixed*. Users move their finger on the table to move the sphere up or down.

With the 3 techniques, users move a virtual blue sphere used as a 3D cursor, and a selection occurs when this sphere is entirely inside a box for 1 second (dwell time mechanism) [11, 18]. When the timer starts, a gauge is displayed as a feedback; if the sphere moves outside the target the timer is reset. The targeted box is red, all others are grey. All boxes are semi-transparent to make the blue sphere visible inside them and they turn from grey to yellow when they are selected. When the users touch the base of a stack on the table during step 1 (*Mixed*, *Balloon*), this base turns from red to green and the stack is highlighted in blue. When using *DT*, the stack is always highlighted and the base is green (see Figure 1, left).

The 3 techniques are implemented using a HoloLens (two HD see-through displays $FoV = 30x17.5$ degrees and frame rate = 60Hz) to visualize the 3D virtual scene. We use Optitrack cameras and a reflective marker to track the users' fingertip (100 fps). An "Image Target" that Vuforia Engine can detect and track is put on the non-interactive table. Once the image is detected, we create and place a world anchor (using Unity for HoloLens) near the image and stop Vuforia. This anchor is a link between the real world (the table, the marker) and the virtual world.

In the two following experiments, we study the effect of the input spaces and of the task decomposition by comparing the 3 techniques for the selection of (1) visible objects (Section 3) and (2) occluded objects in a dense environment (Section 5).

3 EXPERIMENT 1: VISIBLE OBJECTS

We compare the performance of the techniques for the selection of visible boxes by varying their shape *SHAPE* and their position *HEIGHT* in one stack (close to or far from the table).

Hypotheses. - *H1.* Task decomposition does not impact performance of *Balloon* and *Mixed*: they perform better (time, accuracy) than *DT*. - *H2.* *DT* is less accurate than the others. Interacting in mid-air is not free from difficulties [8] due to hand tremor, fatigue [1, 16] and perception issues such as the distance estimation [6, 14, 17].

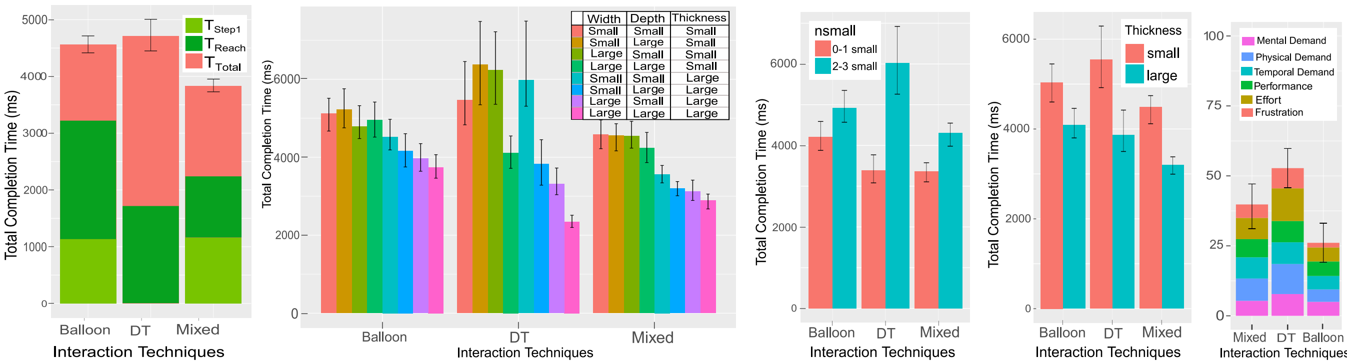
Design and Participants. The experiment is designed as a within-subjects user study, with 3 independent variables: *SHAPE* (8 shapes), *HEIGHT* (Height 1, 2 & 3, see Figure 2b) and the technique *TECH* (*Balloon*, *Mixed*, *DT*). The shapes result from the combination of small (1.2cm) and large (3.6cm) values of depth, width and thickness. The diameter of the blue sphere (3D cursor) is 0.8 cm. We recruited 15 unpaid volunteers (10 males, 5 females), ranging from ages 20 to 39 ($m = 27.67$, $sd = 4.97$). None of them was an expert in AR.

Task and Procedure. A stack of boxes is displayed on the table, at the center of the "Image Target" (see Figure 2b). The task consists of selecting the red box. Participants are asked to stand in front of the table during the experiment. They place their finger in the blue square (i.e. initial position) before starting each task. After a training phase with a set of 11 trials, participants perform 72 selections per technique: 3 repetitions (to study learning and tiredness effect) of pseudo-random order of 24 *SHAPE* × *HEIGHT* combinations. The order of the techniques is counterbalanced with a Latin square. At the end of each technique, participants fill a RTLX (according to [5], NASA-TLX and RTLX perform equally well). We conclude by an interview. The experiment lasts approximately 1 hour.

Measures. For performance, we measure 1) the completion time to select the base of the stack (for *Balloon* & *Mixed*), 2) the completion time to reach the target for the first time (including the time to select the base for *Balloon* & *Mixed*); 3) the total completion time to select the target (time to reach the target + validation phase with potential corrective movements when entering/ leaving the box). We remove 1s (dwell time) from the total completion time. For accuracy, we record the number of times the timer is reset during the validation phase. We call this metric the number of errors.

4 RESULTS OF EXPERIMENT 1

Completion times follow a normal distribution (Shapiro-Wilk test), so we use ANOVAs and t-tests with Bonferroni adjustment for pairwise comparisons. For the non-parametric analysis of accuracy, we use ART [19] with a Bonferroni correction.



(a) Total time T_{Total}, time to reach the target T_{Reach} and time to select the base T_{Step1}

(b) Total time per shape

(c) Total time per shape according to the number of small dimensions

(d) Total time per shape according to thickness

(e) Total workload between 0 and 100

Figure 3: Total Completion Time and workload per technique. We report means and 95% confidence intervals

4.1 Objective Measures: Completion Time

In comparing users' performance during the 3 repetitions per technique, we do not find learning or tiredness effects.

4.1.1 Selecting a box. For the total completion time (see Figure 3a), we find a main effect for *TECH* ($F_{2,28} = 9.11$, $p = 0.0009$), *SHAPE* ($F_{7,98} = 131$, $p < 0.0001$) and *HEIGHT* ($F_{2,28} = 94$, $p < 0.0001$). The average total completion times are 3.84s for *Mixed*, 4.56s for *Balloon* and 4.72s for *DT*. Pairwise comparisons show that it takes statistically less time to select a box with *Mixed* than with *DT* ($p = 0.03$) and with *Balloon* ($p = 0.041$). We do not find a significant difference between *DT* and *Balloon*. The average time to select the base is 1.1s: 25% of total time for *Balloon*, 30% of total time for *Mixed*.

4.1.2 Reaching a box. For the time to reach the target for the first time (see Figure 3a), we find a main effect for *TECH* ($F_{2,28} = 62.1$), *SHAPE* ($F_{7,98} = 46$), and *HEIGHT* ($F_{2,28} = 101$) with $p < 0.0001$. This average completion time is 1.9s with *DT*, 2.2s with *Mixed* and 3.2s with *Balloon*. Post-hoc analysis shows that it takes statistically more time to reach the target with *Balloon* than with *Mixed* ($p = 0.0007$) and *DT* ($p < 0.0001$). Even if *DT* seems to be faster than *Mixed*, we do not find a significant difference between them.

4.1.3 Interaction effect between TECH and SHAPE. For the total completion time, we find that *TECH* has significant interaction effects with *SHAPE* ($F_{14,196} = 26.5$, $p < 0.0001$, see Figure 3b). For *DT*, results show large variations and a striking difference between two groups of shapes: the 4 shapes with 0 or 1 small dimension (T_{Total} for *DT* is 3.4s) and the 4 shapes with 2 or 3 small dimensions (T_{Total} for *DT* is 5.99s). *Mixed* and *DT* are significantly slower with shapes having at least 2 small dimensions ($p < 0.008$), see Figure 3c. We do not find a significant difference for *Balloon*. For shapes with at most 1 small dimension, *Mixed* and *DT* are statistically faster than *Balloon* ($p < 0.03$). We do not find a significant difference between *Mixed* and *DT*. For shapes with at least 2 small dimensions, *Mixed* is significantly faster ($p = 0.0004$) than *DT*. No significant difference is found between *Balloon* and mid-air techniques.
Aggregating the shapes by thickness (see Figure 3d) reveals that only the techniques using mid-air are statistically faster for shapes with a large thickness ($p < 0.0001$). Aggregating the shapes by width or by depth reveals a significant effect of these 2 dimensions for *DT* only ($p = 0.001$ for width, $p = 0.022$ for depth).

4.1.4 Interaction effect between TECH and HEIGHT. For the total completion time, we find that *TECH* has significant interaction effects with *HEIGHT* ($F_{4,56} = 3.26$, $p = 0.018$). The time to reach any of the boxes with *DT* is almost the same (regardless of the height). Although the 2 other techniques seem to be faster to reach the box at *Height 1* than other boxes, the only significant difference is found for *Balloon* between *Height 1* and *Height 3* ($p < 0.0001$). For the total completion time, pairwise comparisons show that *Height 1* is selected faster than *Height 2* ($p < 0.05$) for all techniques, and *Height 3* ($p \leq 0.003$) for the techniques using decomposition. We do not find significant differences between *Height 2* and *Height 3*.

4.2 Objective Measures: Accuracy

4.2.1 Average number of errors. We find a main effect for *TECH* ($F_{2,994} = 308.2977$), *SHAPE* ($F_{7,994} = 149.9787$) and *HEIGHT* ($F_{2,994} = 122.24$) with $p < 0.0001$. With only 0.5 error, *Balloon* is significantly more accurate than *Mixed* (0.8 error, $p = 0.003$) and than *DT* (2.2 errors, $p = 0.0002$).

4.2.2 Interaction effect between TECH and SHAPE. We find that *TECH* has significant interaction effects with *SHAPE* ($F_{14,994} = 31.1439$) with $p < 0.0001$. We observe a difference for *DT* between 2 groups of shapes: the 4 shapes with at most 1 small dimension (1.19 error) and the 4 other shapes (3.27 errors). The 2 other techniques are only impacted by the thickness of the boxes: they perform less than 0.4 error on average for all shapes having a large thickness, and 1 error on average for all shapes having a small thickness.

4.2.3 Interaction effect between TECH and HEIGHT. We find that *TECH* has significant interaction effects with *HEIGHT* ($F_{4,994} = 35.5997$) with $p < 0.0001$. We do not find a significant difference between heights for *Balloon*. Pairwise comparisons show that *DT* and *Mixed* are statistically more accurate for *HEIGHT 1* than upper heights ($p < 0.03$). To explain this, we observed that participants' hand was resting on the table to select boxes close to the table (*HEIGHT1*). We do not find significant difference between others.

4.3 Subjective Measures

Balloon and *DT* require respectively the lowest and the highest workload (Figure 3e). *Balloon* is rated as the least mentally demanding (8/11 participants), the least physically demanding (14/15 participants), the most successful (13/15 participants), the most accurate (all participants) and the fastest technique (9/15 participants). *Balloon* is the most preferred technique for all the participants, and *DT* is the least preferred technique for 11/15 participants. A majority of them prefers *Mixed* and *Balloon*. Even if *DT* is found intuitive, they report that task decomposition significantly reduces the mental and physical fatigue and especially frustration. Participants would choose *Balloon* instead of *Mixed* due to its perceived accuracy providing a sentiment of control, in contrast to *Mixed* which is found difficult to use for boxes with a small thickness. Although the results show that *Mixed* is the fastest technique on average, a majority of participants think that *Balloon* is the fastest technique by far. Concerning distance perception, 2 participants report having issues with all techniques and 9 participants report having issues with *DT* only. To select small bases with *Balloon* and *Mixed*, most of the participants exploit the table as they first put their finger on the table near the base and then drag it until they reach the stack.

5 EXPERIMENT 2: OCCLUSION & DENSITY

In this experiment, we study the impact of density and occlusion on users' preferences and observe how they use the 3 techniques. We hypothesize that density has a negative impact on *Balloon* (*H3*): passing through virtual objects during step 2 might disturb users. In contrast with the first experiment, a wall of high opaque distractors is inserted in front of the targeted stack, occluding it. Users are forced to lean forward to see the target. Several other distractors are placed on the table. To reduce the number of variables, targets are always at height 2 avoiding participants from putting their hand on the table during the validation phase. The first experiment uncovered that performance is impacted by the number of small dimensions of the target. Therefore, we choose 3 shapes only: one with 3 large dimensions, one with only a small thickness to equally influence the 3 techniques and finally one with a small thickness and a small depth.

We recruited 9 volunteers (5 males, 4 females), ranging from ages 24 to 37 ($m = 29$, $sd = 3.57$). All of them participated in the first experiment, with a delay of 6 months between the experiments. They perform 36 tasks per technique in random order. The order of the techniques is counterbalanced with a Latin square. At the end of each technique, participants fill a RTLX. After using the 3 techniques, the participants rank them in order of preference. During a final qualitative session, participants use the same techniques but without occlusion of the target (i.e. the foremost distractor of the wall is removed). Then, users are interviewed about 1) the perceived impact of density when the target is visible, 2) their new preference order and 3) their use of the techniques depending on visibility conditions. The experiment lasts about 40 minutes.

As in the first experiment, results show that *Balloon* and *DT* are the techniques requiring respectively the lowest and the highest workload. *Balloon* is the most preferred technique for 7/9 participants, the 2 others choose *Mixed*. *DT* is the least preferred technique for 7/9 participants. For the 3 techniques, participants have to lean forward and change their view point to see the stack. No consensus emerges on the technique being the least impacted and the most impacted by occlusion. According to them, occlusion has a small impact on the difficulty of the task (for *Mixed* and *Balloon*, we measure that occlusion slightly increases the time to select the base by 0.12 second). Except for the view point, they report using all the techniques the same way as when selecting visible targets. For instance, they directly go through obstacles to reach the boxes instead of moving above as in [12]. With visible targets, 7/9 participants report no or a very small impact of density on how to use the techniques. Only one participant reported being disturbed by distractors for touch interaction with *Balloon* because of the objects hiding her/his hand (for step 2). For a large majority of the participants, the techniques involving task decomposition are preferred, regardless of visibility and density conditions.

	Benefits	*Limitations*
2D input space Touch (Balloon)	• most accurate ✔ [1] • average completion time similar to *DT* ✔ [1] • fast 2D task on the table, regardless of the visibility condition • low sensitivity to small boxes • most preferred • impression of control • constant physical support • low perceived workload • low fatigue ✔ [1, 16]	• slower than *Mixed* to select a box • slowest to reach a box • slower for high boxes
Mixed input space Touch+Mid-air (Mixed)	• fastest interaction on average, ✘ contradicting [9, 16] on task decomposition • faster than *Balloon* to select large boxes • faster than *DT* to select small boxes • more accurate than *DT* • fast 2D task on the table, regardless of the visibility condition • accurate & fast for lowest boxes (hand resting on the table)	• less accurate than *Balloon* • sensitive to the thickness of the box
3D input space Mid-air (DT)	• fast to reach a box regardless of its height • faster than *Balloon* to select large boxes • intuitive • accurate & fast for lowest boxes (hand on the table)	• least accurate ✔ • slower than *Mixed* to select boxes on average • sensitive to the 3 box dimensions • highest perceived workload • physical fatigue ✔ [1, 16] • frustration

Table 1: Benefits and limitations of the techniques. We indicate whether a result confirms ✔ or contradicts ✘ those of other studies.

6 DISCUSSION AND CONCLUSION

Table 1 summarizes the main results. Overall our studies highlight the usefulness of a tangible support and the efficiency of task decomposition. The 2D selection by touch on the table is easy, meaningful and thus fast, regardless of the size of the stack base, the visibility of the targets and the density of the environment. Our results partially confirm *H1*: *Mixed* is on average the fastest technique and even though *Balloon* is slower than *DT* to reach a box, its accuracy allows users to select it with an average total time similar to *DT*. The low accuracy of *DT* confirms *H2*. Also, the touch-based techniques are still efficient and appreciated in dense environments whereas they require users to interact through obstacles (*H3* rejected). For the second step of the decomposition, we observe a fast height adjustment in mid-air for *Mixed*, which is slightly impacted by the height of the boxes (in contrast to *Balloon*). Thus, we suggest unifying the 2D and 3D input spaces and using techniques like *Mixed* which fruitfully combines the accuracy of touch-based interaction and the speed of mid-air interaction. Moreover *Mixed* can be extended to select out-of-reach objects: the first step is performed by indirect relative touch inputs on the table to control a cursor, while the second step, performed by mid-air interaction is unchanged. Such an extension could be compared with a ray-casting technique.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the AP2 project ANR-15-CE23-0001.

REFERENCES

[1] H. Benko and S. Feiner. 2007. Balloon Selection: A Multi-Finger Technique for Accurate Low-Fatigue 3D Selection. In *2007 IEEE Symposium on 3D User Interfaces*. https://doi.org/10.1109/3DUI.2007.340778

[2] H. Benko, E. W. Ishak, and S. Feiner. 2005. Cross-dimensional gestural interaction techniques for hybrid immersive environments. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* 209–216. https://doi.org/10.1109/VR.2005.1492776

[3] Nicolas J. Dedual, Ohan Oda, and Steven K. Feiner. 2011. Creating Hybrid User Interfaces with a 2D Multi-touch Tabletop and a 3D See-through Head-worn Display. In *Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR '11)*. IEEE Computer Society, Washington, DC, USA, 231–232. https://doi.org/10.1109/ISMAR.2011.6092391

[4] Martin Hachet, Benoit Bossavit, Aurélie Cohé, and Jean-Baptiste de la Rivière. 2011. Toucheo: Multitouch and Stereo Combined in a Seamless Workspace. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. ACM, New York, NY, USA, 587–592. https://doi.org/10.1145/2047196.2047273

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

[6] Ernst Kruijff, J. Edward Swan, and Steven K. Feiner. 2010. Perceptual issues in augmented reality revisited. *2010 IEEE International Symposium on Mixed and Augmented Reality* (2010), 3–12.

[7] A. Martinet, G. Casiez, and L. Grisoni. 2010. The design and evaluation of 3D positioning techniques for multi-touch displays. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*. 115–118. https://doi.org/10.1109/3DUI.2010.5444709

[8] Sven Mayer, Valentin Schwind, Robin Schweigert, and Niels Henze. 2018. The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 653, 13 pages. https://doi.org/10.1145/3173574.3174227

[9] Daniel Mendes, Fernando Fonseca, Bruno Araujo, Alfredo Ferreira, and Joaquim Jorge. 2014. Mid-Air Interactions Above Stereoscopic Interactive Tables. *IEEE Symposium on 3D User Interfaces 2014, 3DUI 2014 - Proceedings.* https://doi.org/10.1109/3DUI.2014.6798833

[10] Microsoft. 2017. Microsoft HoloLens. https://www.microsoft.com/en-us/hololens

[11] Choon Hong Peck. 2001. Useful Parameters for the Design of Laser Pointer Interaction Techniques. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (Seattle, Washington) *(CHI EA '01)*. Association for Computing Machinery, New York, NY, USA, 461–462. https://doi.org/10.1145/634067.634333

[12] Carole Plasson, Dominique Cunin, Yann Laurillau, and Laurence Nigay. 2019. Tabletop AR with HMD and Tablet: A Comparative Study for 3D Selection. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) *(ISS '19)*. Association for Computing Machinery, New York, NY, USA, 409–414. https://doi.org/10.1145/3343055.3360760

[13] Patrick Reipschläger and Raimund Dachselt. 2019. DesignAR: Immersive 3D-Modeling Combining Augmented Reality with Interactive Displays. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) *(ISS '19)*. Association for Computing Machinery, New York, NY, USA, 29–41. https://doi.org/10.1145/3343055.3359718

[14] Rebekka S. Renner, Boris M. Velichkovsky, and Jens R. Helmert. 2013. The Perception of Egocentric Distances in Virtual Environments - A Review. *ACM Comput. Surv.* 46, 2, Article 23 (Dec. 2013), 40 pages. https://doi.org/10.1145/2543581.2543590

[15] Sven Strothoff, Dimitar Valkov, and Klaus Hinrichs. 2011. Triangle Cursor: Interactions with Objects Above the Tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (Kobe, Japan) *(ITS '11)*. ACM, New York, NY, USA, 111–119. https://doi.org/10.1145/2076354.2076377

[16] M. Veit, A. Capobianco, and D. Bechmann. 2011. An experimental analysis of the impact of Touch Screen Interaction techniques for 3-D positioning tasks. In *2011 IEEE Virtual Reality Conference*. 75–82. https://doi.org/10.1109/VR.2011.5759440

[17] Peter Willemsen, Amy A. Gooch, William B. Thompson, and Sarah H. Creem-Regehr. 2008. Effects of Stereo Viewing Conditions on Distance Perception in Virtual Environments. *Presence: Teleoper. Virtual Environ.* 17, 1 (Feb. 2008), 91–101. https://doi.org/10.1162/pres.17.1.91

[18] Andrew Wilson and Steven Shafer. 2003. XWand: UI for Intelligent Spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 545–552. https://doi.org/10.1145/642611.642706

[19] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. ACM, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[20] R. Xiao, J. Schwarz, N. Throm, A. D. Wilson, and H. Benko. 2018. MRTouch: Adding Touch Input to Head-Mounted Mixed Reality. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1653–1660. https://doi.org/10.1109/TVCG.2018.2794222