Precise Tool to Target Positioning Widgets (TOTTA) in Spatial Environments: A Systematic Review

Mine Dastan 🕑, Michele Fiorentino ២ and Antonio E. Uva ២

Fig. 1. The Tool to Target (TOTTA) widget designs from the 24 papers analyzed chronologically. For each frame, the Tool widget (TO) is depicted at the top left, and the Target widget (TA) at the bottom right.

Abstract—TOTTA outlines the spatial position and rotation guidance of a real/virtual tool (TO) towards a real/virtual target (TA), which is a key task in Mixed reality applications. The task error can have critical consequences regarding safety, performance, and quality, such as surgical implantology or industrial maintenance scenarios. The TOTTA problem lacks a dedicated study and it is scattered in different domains with isolated designs. This work contributes to a systematic review of the TOTTA visual widgets, studying 70 unique designs from 24 papers. TOTTA is commonly guided by the visual overlap –an intuitive, pre-attentive "collimation" feedback– of simple shaped widgets: Box, 3D Axes, 3D Model, 2D Crosshair, Globe, Tetrahedron, Line, Plane. Our research discovers that TO and TA are often represented with the same shape. They are distinguished by topological elements (e.g. edges/vertices/faces), colors, transparency levels, and added. shapes, widget quantity, and size. Meanwhile some designs provide continuous "during manipulation feedback" relative to the distance between TO and TA by text, dynamic color, sonification, and amplified graphical visualization. Some approaches trigger discrete "TA reached feedback" such as color alteration, added sound, TA shape change, and added text. We found the lack of golden standards, including in testing procedures, as current ones are limited to partial sets with different and incomparable setups (different target configurations, avatar, background, etc.). We also found a bias in participants: right-handed, young male, non-color impaired.

Index Terms— Virtual environments, 3D user interface, tool to target manipulation, widgets, 3D positioning.

1 INTRODUCTION

Mixed reality (MR) and relative spatial environments are increasing in popularity thanks to the steady evolution of technology and the reduced costs of head-mounted displays, improved tracking, software ecosystems like SteamVR, and development tools such as Unity3D [1, 2]. However, MR spatial environment management is rather limited to gaming and entertainment. Among the key reasons for the lack of adoption in "serious" applications is the inadequate performance of spatial interactions [3, 4, 5]. Many industrial applications (surgery, maintenance, welding, etc.) require the user to handle a tool and perform spatial dexterity tasks with 6 degrees of freedom (6 DOF,

• Mine Dastan, Antonio E. Uva and Michele Fiorentino are with the Polytechnic University of Bari. E-mail: mine.dastan@poliba.it

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org.

Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxx/.

three positional and three rotational), with a low margin of error to avoid consequences in terms of cost, time, and even human safety [6], 7, 8, 9]. MR has been demonstrated to be beneficial in training and guiding spatial operations using virtual reality (VR) and executing in the real world using augmented reality (AR) [10, 11, 12]. These spatial operations are challenging since they demand precision manipulation of virtual/real tools [13, 14, 15].

Spatial XR interactions used in wide application fields require precise tool-to-target (TOTTA) tasks (position/rotation). In medical surgical procedures, where implementing haptic feedback is complex and costly [16, 17, 18, 19] and the surgeon must place the tool accurately on the preoperative target. Moreover, in industrial scenarios, tasks such as quality checks of diagnostic sensors to verify carbon fiber structures [20, 21, 22] requires a certain precision.

Although a 3D tool to target positioning tasks is important, there is not a specific definition of this problem that differentiates it from general 2D/3D UI research like "drag-and-drop," "aiming/aligning," and "docking" [23]. Therefore, we define it as the TOTTA (TO-to-TA), a spatial positioning task. TOTTA is usually guided by visual elements -the widgets [24, 25]- to be aligned by spatial skills in 3D space. These widgets help manipulate the tool (physical or virtual TO) towards a spatial target (physical or virtual TA). The user manipulates a tool towards a target, supported by tool widgets (TO) and target widgets (TA), to achieve higher precision and performance. Table 1: Summary of the selected papers' used technology, interaction, visualization, and tracking device.

No.	Ref.	Authors	Technology	Interaction	Visualization Device	Tracking Device	Field
1	[26]	Boritz et al.	Desktop VR	Object-Mouse	CrystalEyesT	Fastrak	Near
2	[27]	Fiorentino et al.	Projected VR	Controller	Vertical screen,	ART	Near
			HMD		Projectors		
3	[28]	Veit et al.	VR HMD	Gesture-	CrystalEyes CE-2	ART	Far
				Touch			
4	[29]	Ragan et al.	VR HMD	Controller	Virtual Research V8	Optitrack	Near
					HMD		Far
5	[30]	Raj et al.	Desktop VR	Object	Computer	Kinect v2 InterSense	Far
						inertia cube3	
6	[31]	Ha et al.	AR HMD	Gesture	ACCUPIX my bud	PTAMM camera tracking	Near
7	[32]	Vuibert et al	Desktop VR	Gesture-	Vision RF shutter	Optitrack Flex: V100	Far
				Object	glass	Motion capture	
8	[33]	Wang et al.	VR HMD	Object-Touch	eMagin z800 HMD	PhaseSpace motion	Near
						capture	
9	[34]	Feng et al.	Fish tank VR	Object-Touch	Nvidia 3D Vision	Polhemus Fastrak	Near
					glasses		
10	[35]	Mendes et al.	VR HMD	Object-Touch	Gear VR, Samsung s6	Kinect v2	Near
11	[36]	Krichenbauer et	VST-HMD	Object	Oculus Rift, OVR	Leap Motion	Near
		al.			vision stereo		Far
12	[37]	Ro et al.	AR HMD	Gesture-	HoloLens	Kinect v2	Far
	5203			Touch			
13	[38]	Kım et al.	VR HMD	Controller	Oculus Rift Consumer	Positional	Near
14	[39]	Schlunsen et al.	VR HMD	Controller	HTC Vive pro	Leap Motion	Far
15	[40]	Heinrich et al.	Projected AR	Object	Projector	Fusion track 500	Near
16	[41]	Sun et al.	Web VR	Mouse	Computer	NA	Far
17	[42]	Andersen et al.	AR HMD	Gesture	HoloLens 1	6-DoF V	Far
18	[43]	Liu et al.	VR HMD	Gaze	HTC Vive Pro-Eye	6Dof VR, eye tracking	Far
19	[44]	Weiß et al.	AR HMD/VR/	Object	HTC Vive, HoloLens,	Vuforia image target,	Near
	54.53		Projected AR	1	projector		
20	[45]	Fuvattanasılp et	Handheld AR	Touch	Apple 1Pad Pro (2017)	AR marker	Far
	54.63	al.	THE THE CO	a . "	TIMO TY		
21	[46]	Lee et al.	VR HMD	Controller	HTC Vive	6-DoF VR	Near
- 22	[47]	V	VD UMD	C	Dise No. 2 E	(D-EVD tra-alc	Far
22	[4/]	r u et al.		Gaze	Pico Neo 2 Eye	o-DoF VK, eye tracking	Far
23	[48]	Dastan et al.	VKHMD	Controller	Uculus Quest 2		Near
24	[49]	Ganias et al.	VK HMD	Controller	HIC Vive Pro Eye	I wo base stations	Near

The literature on the design of TOTTA is very varied, and the impact of widgets and UI is crucial yet underestimated, especially considering that they can improve user performance, usability, comfort, and ease of use [50, 51, 52]. Poor design can also have side effects, like distraction from the primary task and increased mental and physical demands [53, 54, 55].

In the past, 3D tracking technology was the main TOTTA limiting factor due to low precision, accuracy, and latency [56], [57]. However, MR technology is evolving rapidly, and more precise and cost-effective devices are expected [58]. Even assuming ideal tracking, the basic knowledge, methods, and tools for TOTTA guidance interface design (UI) are still open research [59, 60]. TOTTA literature is scattered into different application domains, different research objectives (hardware validation, widget validation), and MR setups and lacks well-established guidelines.

To our knowledge, no previous literature review addressed TOTTA positioning widgets. The main contributions of this work are:

- C1: Define the TOTTA problem and its importance,
- C2: Provide a systematic review of TOTTA visual widgets,
- C3: Classify and compare the approaches,
- C4: Discuss open issues and trends,
- C5: Guides for further research for optimal design TOTTA,

C6: Interest in MR developers in many domains to save time and resources and create effective, easy-to-use spatial interfaces and applications.

The research questions that motivate the literature research on the precise TOTTA spatial visual guidance in MR are: "What are the existing approaches in literature?", "How are the TOTTA widgets designed?" and "How are TOTTA widgets evaluated?".

2 METHODOLOGY SYSTEMATIC REVIEW

Being a multidisciplinary and cross-domain problem, the query definition is divided into four categories: "spatiality," "task," "technology," and "objective" (Fig. 2). We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for conducting a systematic review [61].







Fig. 3. PRISMA flow chart approved by [61].

This method is often used in conducting literature reviews [62, 63, 64]. We launched the search on the Scopus, ACM, and IEEE databases in August 2023. The PRISMA-based data collection stage is summarized in Fig. 3. From the initial 206 papers; we eliminated the papers that do not present or describe the widgets, the ones that focus on only hardware devices without user interfaces, and tasks that do not include a tool or target widgets. These aspects were essential in our systematic review and pointed out the key connection of papers, that they all present a spatial manipulation task of a visual real/virtual tool to a real/virtual target using TOTTA visual widgets. These criteria enabled detailed manual reading of the final selection of 24 papers in Table 1.

2.1 **Bibliometric analysis**

We visualized the keywords of the selected papers using word clusters [39] (Fig. 4). This graph shows that "3D user interfaces" and "3D interaction" are the key topics, and "computer-assisted surgery" and "object manipulation" are common applications. The colors demonstrate how recent studies (yellow) VR first, then AR, later in green to "user studies "and visualization, as the effect of natural technology evolution and optimization.

Fig. 5 supports the topic's novelty and demonstrates a positive trend. More than 50 percent of the found literature papers have been published in the last five years, and the increased number of citations and the majority of the papers are conference proceedings (16/24 papers, 67%).

The application domain is general in most of the cases (19/24 papers, 80%), with some verticalizations in medical (2) and industrial (3). This can be explained by the fact that TOTTA is a very common task, and these widgets can be useful in many domains.



Fig. 4. Bibliographic keywords connection of the TOTTA papers.



Fig. 5. The 24 TOTTA paper's publication year and citations.

2.2 **TOTTA** widgets in literature

This section presents the key aspects of the TOTTA widget design in chronological order, along with the main research drivers and findings.

Boritz et al. [26] propose a direct midair TOTTA interface using a new tetrahedron-shaped physical input device. The TOTTA widgets, which mimic the input device, are two identical tetrahedrons (1.73cm height and 0.87cm width), with one face perpendicular to the base and with a checkerboard texture (Fig. 6). The user must align TO on TA geometric shapes. When the distance is less than a threshold (0.5 cm), a red box appears at the tip of the TA. They found higher positional error along z (front direction of depth) with the monoscopic display vs stereo.



Fig. 6. Boritz et al. [26] use Tetrahedron, No: 1.

Fiorentino et al. [27] implement TO with a 3D Model - a simplified representation of the physical pen held by the user-; for TA, a green Box with colored 3D Axes (Fig. 7). With only a 3D positioning test, they find that difficulty varied with the target position: the targets in front of the user and above the head lead to greater error.



Fig. 7: Fiorentino et al. use Model TO, Box TA and Axis TA.

Veit et al. [28] investigate touch screen interactions with DOF separation (height axis vs. depth axis) using a monochromatic (green) 3D Axes TO (15 cm length) and a semi-transparent blue globe TA (7.5 cm radius). At collimation, the TO 3D Axes flash (Fig. 8). Results indicate that isolating the depth axis manipulation increases precision, while haptic cues do not improve user precision.



Fig. 8. Veit et al. [28] use 3D Axes TO and Globe TA, No:2.

Ragan et al. [29] propose a multi-touch input device for TOTTA. They use two Boxes of different sizes: TO is opaque blue, and the TA is semi-transparent purple, with eight small gray spheres in the vertexes. As additional feedback, the spheres turned red when only a vertex was aligned and green when all eight were aligned (Fig. 9). They found that touch-based interfaces increased the task completion time compared to wand or joystick interactions.



Fig. 9. Ragan et al. [29] use Box No: 3.

Raj et al. [30] utilize two different colored 3D Axes: TO, purple, and TA, gray. Both 3DAxes had unique-colored spheres at the endpoints to match the orientation (Fig. 10). The results differed between participants' gender and video game experience. This paper shows how the user's gender, avatar representation, and experience can influence performance. The self-avatar visualization resulted in a slightly faster rotation time than a sphere visualization of an avatar.



Fig. 10. Raj et al. [30] use 3D Axes, No:4.

Ha et al. [31] implemented a bare-hand user interface with a green Box for TO and a blue Box for TA combined with semitransparent grey guidelines and shadows where the lines intersect in the AR environment (Fig. 11). The TA Box turns red as the target achieved feedback. This design is interesting as it claims that anteroposterior depth visual feedback by shadows and guidelines enables precise manipulation.



Vuibert et al. [32] compared the tetrahedrons versus two chair models in desktop VR; they use a pair of same-sized Tetrahedrons rendered as a colored wireframe (Fig. 12). TO has an opaque sphere in the center, and TA has a larger semi-transparent sphere. This research found that virtual 3D models can perform better (time and precision) than tetrahedrons, probably by leveraging natural human skills from the real world. However, the 3D Model must allow unique positional and angular collimation.



Wang et al. [33] developed an Object Impersonation metaphor that enables switching the DRIVE (avatar view on a tablet and tetrahedron's view on HMD) and VIEW methods (avatar view on HMD and tetrahedron's view on a tablet). They use a pair of samesize Tetrahedrons with small colored spheres in the vertexes and additional Crosshair (Fig. 13). This widget comprises two coplanar, perpendicular lines (forming a 2D cross) surrounded by concentric circles. TO is semi-transparent blue with gray edges and orange crosshair, and TA is non-transparent turquoise with nontransparent crosshair. The object impersonation method gave better orientational precision but required higher cognitive demand.



Fig. 13. Wang et al. [33], Tetrahedron & Crosshair, No: 7.

Feng et al. [34] evaluate one-hand free vs novel two-handed input devices for 7 DOF manipulation techniques. The input device corresponds to virtual spherical cursors (blue-left, pink-right). They use Box TOTTA with colored faces, but different in size and frame color: red (TO) and white (TA) (Fig. 14). The virtual cursors are linked with an orange-colored cylinder "spindle" with a small red sphere mid-point. The technique's result is equivalent when the TOTTA size is the same and faster with their input device when the TOTTA size is different.



Fig. 14. Feng et al. [34] use Box, No:8.

Mendes et al. [35] present an opaque 3D Model with colored 3D Axes with sphere endpoints for TO, and a transparent 3D Model for TA (Fig. 15). Interactive secondary feedback is provided by the Model TO color gradually turning green with the distance from TA. An interesting aspect is the 3D axes allowed to control a single DOF. The PRISM technique dynamically adjusts between hand and object motion ratio. Experimentation demonstrated how DOF separation brings benefits but at the cost of task completion time.



Fig. 15. Mendes et al. [35],3D Axes, No: 9.

Krichenbauer et al. [36] compare AR vs VR, TO as a composition of an opaque Box with a single red face textured with the number two, colored 3D Axes, and a wireframe-colored Globe. (Fig. 16). TA is a box like TO with different sizes, transparency, and dashed edges. The results claim that AR resulted in faster completion time than VR when using a 3D input device and mouse.



Fig. 16. Krichenbauer et al. [36], 3D Axes/Globe/Box; No:10.

Ro et al. [37] present a novel physical input device using the Laser pointer metaphor AR pointer. The user collimates the Box TO (green, blue, red) remotely with a touch on the mobile device to a different size semitransparent TA (Fig. 17). The AR pointer performs better in task completion time than the direct free hand 3D manipulation metaphor.



Fig. 17. Ro et al. [37] use Box No: 11.

Kim et al. [38] compared "DOF *separation* (1DOF, only axishandles)", "*without* DOF *separation* (3DOF, only center-handle)," and "switchable DOF (1-DOF/2-DOF/3DOF, center, and axishandles)" for mid-air manipulation. They utilize TO as a combination of a 3D Model (a teapot), colored 3D Planes, and colored 3D Axes with spheres in the endpoint (Fig. 18). TA is represented only as a semi-transparent 3D model. During manipulation, the constrained axis becomes yellow along lines or planes. The switchable DOF outperformed others in terms of time and precision efficiency.



Fig. 18. Kim et al. [38] use 3D Axes/Plane/Model, No: 12.

Schlunsen et al. [39] evaluate free hand vs widget-based manipulation techniques and different multimodal cues for 3D manipulation of system control tasks. Green framed gray Box TO with 3D Axes TO (translation) with spheres in the endpoints and gray framed Globe TO (rotation) is used (Fig. 19). TA is a brown framed yellow box. The free-hand manipulation resulted in faster and most preferred by the participants. They claim that multimodal feedback (audio) improved the user experience.



Fig. 19. Schlunsen et al. [39], 3D Axes/Globe/Box, No: 13.

Heinrich et al. [40] compare three visual widgets (circle, the crosshair, and the arrow concept) for AR-supported medical needle insertion (Fig. 20). They use different-sized 2D Crosshair TOTTA, small crosshair TO (color change red-orange-green), and bigger crosshair (transparent- yellow-green- red for depth feedback). Each concept has distinct color mapping and indicator scaling. The Crosshair outperformed in orientation and depth parameters. The results for the color and indicator scaling factors are less consistent.



Fig. 20. Heinrich et al. Center TOTTA, No: 14.

Sun et al. [41] compare DOF manipulation modes in WebVR to explore user workload and task performance effects. TO use colored 3D Axes (translation) with cones in the endpoints and Globe TO (rotation) with three wireframes with different colored rings with spheres of interaction points (Fig. 21). The manipulated axis appeared yellow while others disappeared (1DOF). Multiple DOFs provide less perceived workload and higher presence. The results indicated that users feel less workload or more presence and tend to spend less time completing tasks on WebVR.



Fig. 21. Sun et al. [41] use Sphere/Axes/Plane, No: 15.

Andersen et al. [42] elaborate on three semi-transparent widget designs for mid-air interaction. The 3D Axes TOTTA, Crosshair TOTTA (white and red), and triangular pyramid TOTTA (Fig. 22). The crosshair and triangular TOTTA have the shortest alignment time. In contrast, 3D Axes TOTTA performs best in translation and rotation errors. A novel aspect is that visual elements' size affects how far the user extends the arm, influencing torque forces.



Fig. 22. Andersen et al. [42] use Crosshair, No: 16.

Liu et al. [43] evaluate which type of gaze-based manipulations (eye vs. head) performs best when combined with OrthoGaze. OrthoGaze allows the user to manipulate gray Box TO using the orthogonal Planes TO. The TA is represented as the green-colored Box TA (Fig. 23). During TOTTA, the gaze-selected Plane activates, and the user adjusts the 2-DoF position by looking at the target location and confirming placement through a gaze dwell. The eye gaze results are more accurate than the head gaze for continuous aiming.



Fig. 23. Liu et al. [43] use Box and Plane, No: 17.

Weib et al. [44] investigate DIY tasks such as woodworking (drill, saw, and screw) using various levels of guidance: 2D video instructions, VR, and AR. They use distinct-size Crosshairs for TO and TA (Fig. 24). The TOTTA gives dynamic feedback to the user to avoid user error during drilling. The results indicate that context-aware situated visualizations are less likely to rely on empirical methods.



Fig. 24. Weib et al. [44] use Crosshair, No: 18.

Fuvattanasilp et al. [45] implemented SlidAR+, a novel handheld AR device (HAR) with an interaction method. They use a redcolored Line TO and a translucent, green-colored arrow with a thin blue line TA (Fig. 25). The user matches the TO (red arrow) with the TA (green arrow with the base of the virtual pillar). To align perfectly, a red line appears from the arrow TO tip to the TA, and the user slides TO using the Line. SlidAR+ resulted in faster task completion and is preferred by users.



Fig. 25. Fuvattanasilp et al. [45] use Line, No: 19.

Lee et al. [46] propose a novel near-field interaction metaphor for distant object manipulations. They compare widget-based metaphor with unimanual metaphor (one hand & scaled replicated model) and bimanual metaphor (both hand & scaled replicated model). The Box TO has eight pair-colored small spheres in the vertex points (Fig. 26). For translation, a colored 3D Axes TO is used with a small yellow cube visually indicating the manipulated axis and red text for numerical feedback. For rotation, a colored wireframe Globe TO with a sphere is used. TA is a white wireframe with colored spheres on vertex points. The unimanual metaphor has the highest efficiency, the widget-based metaphor has the slowest, and the bimanual metaphor, with a scaled replica, grants the lower movement. Interestingly, subjective impressions are most favorable with the bimanual metaphor.



Fig. 26. Lee et al. Sphere/3D Axes/Plane TOTTA, No: 20.

Yu et al. [47] compare four gaze-supported interaction techniques: gaze grab, remote hand, 3D Magic gaze, and implicit gaze. They used a 3D Model for TO (a rabbit) and a transparent replica (Fig. 27). During the manipulation, the TO has a blue outline, and when the target is achieved, it turns red. The results indicate that gaze does not influence performance when the TO is in front of the user, but it can be useful for distant targets and larger spaces. The gaze input reduces the arms fatigue issue and potentially allows future TOTTA manipulation.



Dastan et al. [48] present a 5DOF guidance applied to dental implantology as the rotation along the drill axis is not influent. TO comprises three triangle pairs - colored differently for each direction- and two semi-circle pairs - colored differently for each rotation (Fig. 28). The pairs visualize in real time and amplify the position and rotation distance values. TA is a static green line with a concentrical cylinder. This approach leverages human reification from Gestalt theory [34], seeking a quick, pre-attentive reaction. Their method performed better in angular (with major effects) and positional precision and accuracy, with less mental demand and frustration than the literature. However, this gain is obtained with a significant increase in task time and physical demand.





Ganias et al. [49] compare grasping visualizations, auto-pose (realistic grasp), single pose (hands do not penetrate with object), and disappearing hand (hand disappears). They used a colored Box for TO (solid) and a transparent replica for TA (Fig. 29). At the end of each positioning, the visual cue of the changed color was used (the yellow returned green). The results indicated no significant difference in user performance in any visualizations. The auto pose is a user preference and provides a stronger perceived sense.



Fig. 29. Ganias et al. Box TOTTA, No:24.

3 WIDGET DESIGN ANALYSIS

A key aspect of TOTTA widgets is the distance of interactions. Near-field interactions allow the ability to direct manipulation of tool and target in proximity and are advantageous for the precision of small-size objects. On the other hand, the far-field interactions rely on distant object manipulations beyond the user's arm reach, which is advantageous for large tools and target size and distance flexibility. There is no prevalence between the near-field (12/24 papers, 50%) or far-field interaction (9/24 papers, 37%), and few studies (3/24 papers, 13%) employ both (Tab. 1).

The TOTTA may require various levels of DOF, so we analyze that it requires superimposing or aiming. The most common widget guidance mechanic is the visual 3D superimposition of TO over TA (20/24 papers, 83%) (Fig. 30).

A limited number also includes the scaling (8/24 papers, 33%) with the superimposing task, which consists of changing the size of the TO widget to match the TA. Visual scaling has no direct meaning for TOTTA collimation, even if it may have the potential to set some tool parameters (e.g., drill speed value). However, this usage is not envisioned in the selected papers.

Few papers prefer the aiming task (4/24 papers, 17%), which provides a reference point for the user to aim and align the TO used for 5DOF tasks (e.g., needle insertion, drill positioning).





For a precise TOTTA, users receive multimodal guidance feedback associated with TOTTA widgets. At first glance, TOTTA visual widgets use basic shapes. However, each TOTTA representation is unique in detail. Therefore, the whole design gathered in a TOTTA widget cluster, resulting in 70 different graphical designs (Fig. 1).

In the next subsections, we analyzed the widget guidance mechanics through their feedback to the user. We divided the TOTTA widget design as feedback factors, "Collimation Feedback," "End of task feedback," and lastly, "During manipulation feedback."

3.1 Collimation Feedback

The TO and TA shapes must provide positional and orientation graphical/geometrical clues for unique collimation configuration. This aspect is not trivial and is key in widget design. We found



Fig. 31. Eight different visual designs were used for the TO and TA widgets.

eight diverse types of shapes: Box, 3D Axes, 3D Model, Crosshair, Globe, Tetrahedron, Lines, and Planes (Fig. 31).

The box is the most preferred geometry (21/70 design, 30%, 10 TO/11 TA). It is simple, has predictable orthogonal angles, and is easy to implement in all graphical engines. The Box design is preferred by 10/24 papers [28, 30, 33, 36, 37, 38, 40, 42, 45, 49] (Fig. 32). It's also curious how boxes are used in physical child toys in learning motor skills. Box widgets with different dimensions can provide unique positional reference and partial angular. For angular, unique primary Box is supported by additional shapes or different colored/textured faces, edge styles, or sizes.

3D Axes are also common (11/70 design, 16%, 9TO/2TA), with a clear and familiar design for CAD users and gamers. The 3D Axis design is preferred by 9/24 papers [27, 34, 35, 38, 40, 41, 45, 48, 65] (Fig. 32). The presence of a center supports positional placement, and the orthogonal lines facilitate angular arrangements. TO and TA are differentiated by colors or additional geometries, like spheres in the vertices.

3D Models (9/70 design, 13%, 5TO/4TA) (e.g., rabbit, teapot, arrow) instead of basic geometries. This design is preferred by 5/24 papers [34, 36, 37, 44, 46] (Fig. 32). TO and TA 3D Models are commonly differentiated by color and transparency. However, the choice of a specific model is limited as some 3D models may be inefficient in positional and orientational guidance. Therefore, they are substandard or rarely used by the end user.

Crosshair (8/70 design, 11%, 4TO/4TA) is common in aviation, military, and healthcare interfaces. The Crosshair design was chosen from 4/24 papers [32, 39, 41, 43] (Fig. 32). Planar or 3D crosshairs are used in spatial interactions with reduced DOF, like needle insertion and drilling.

Globe (6/70 design, 9%, 5TO/1TA) is a familiar widget design for desktop applications and games. Interestingly, 6/24 papers (Fig. 32) preferred Globe and they are often preferred singularly only TO or only TA without having them together in the task [27, 35, 38, 40, 45, 48]. Globe is often represented as wireframe rings or semi-transparent to avoid visual occlusion.

The tetrahedron (6/70 design, 9%, 3TO/3TA) is geometrically the simplest (having minimum entities) shape for the TOTTA widget. 3/24 papers use it [21, 27, 28] (Fig. 32). The corners help with orientation but are less familiar compared to the other shapes and have no orthogonal angles.

Line (6/70 design, 9%, 4TO/2TA)) can provide single DOF guidance with reference to mid/end points by distinctive styles (dashed, solid, transparent, or color). However requires the combination of elements to be functional in the 3D space. The Line widget is preferred by 4/24 papers [33, 36, 41, 46] (Fig. 32).

Planes (4/70 design, 6%, 2TO/1TA) are usually represented in three perpendicular surfaces. Their intersection can also generate 3D axes. However, Planes are less intuitive and more prone to visual occlusion than others. The Planes is preferred by 4/24 papers [38, 41, 43, 46] (Fig. 32). Some other visual representation methods are used in TOTTA to support collimation feedback (Fig. 33).



Fig. 32. Single basic shape type used by literature.

As theorized by perception, the same or similar shapes for TO and TA are used principally (22/24 papers, 92%).

To further distinguish the tools from the targets, colored minor parts of widgets (such as frames/vertices/faces) (18/24 papers, 75%) or entirely colored widgets (17/24 papers, 71%) are used frequently.

Transparency (16/24 paper, 67%) is also used to reduce visual cluttering and enhance depth perception during collimation.

Following, TO\TA pairs are differentiated by containing additional/different geometrical elements like small spheres and cubes (14/24 papers, 59%).

Some preferred (13/24 papers, 54%) more than one design of widgets, we defined them as "Mixed widgets", [38, 39, 46]. They have used the composition of shapes (from Fig 31), requiring more cues for guidance—even redundant in some cases—since they can result in more complexity.

Finally, the different widget sizes (11/24 papers, 46%) are used to distinguish TO from TA, especially if the task requires resizing.



Fig. 33. Collimation feedback mechanics: graphical clues.

3.2 During Manipulation Feedback

This feedback supports the user during the TOTTA task through continuous guidance feedback (Fig. 34). Only 14/24 literature papers use during manipulation continuous feedback; they are interactive and not just signals. The majority of these TOTTA designs use continuous color change (e.g., red-orange-green) to convey guidance (9/24 papers, 13%).

Text feedback (2/24 papers, 13%) is also preferred to indicate the real-time error values. Some use sonification (2/24 papers, 8%), which is a simultaneously generated sound (e.g., drums with variable rhythm) to guide the user interactively [32, 39]. Text feedback can provide precision control during TOTTA. However, since there is a continuous change of text in the field of view, it may frustrate the user or cause a high task load.

As dynamic novel approach visual error visualization (1/24 papers, 4%) is used during the manipulation. This approach promises intuitive feedback; it allows dynamic visualization of the target distance by widgets' forms of distance [48]. Another unique aspect is that the widgets disappear at the target threshold, reducing clutter. The widgets' more dynamic and complex behavior is demonstrated to improve user performance but at the cost of physical and cognitive demand. This aspect of widget design has an undisclosed potential to guide the user along the interaction in addition to collimation.



Fig. 34. During manipulation, continuous feedback.

3.3 TA Reached Feedback

As the TO and TA shapes begin to superimpose, visual collimation feedback becomes increasingly inefficient. Therefore, "supplementary" feedback is often provided concurrently with the collimation (Fig. 35). This feedback conveys the end-of-task information as signals and they are triggered at the TOTTA positional and rotational distance threshold, probably to substitute haptic feedback (e.g. vibration) [31, 38, 39, 43].

Commonly instant color change is used after the task completion (17/24 papers, 71%) as it is intuitive for the user and is easy to implement. Further, at the collimation event, playing a completion sounds (9/24 papers, 38%) followed by a new object/target appearance (6/24 papers, 25%) [33] and text flash (1/24 papers, 4%) "Right There!".



Fig. 35. TA reached Feedback at the end of the task.

4 EVALUATION METHODS

TOTTA widgets perform differently and can be physically and mentally demanding for the user. Not much research has been done to highlight how the literature compares and evaluates different TOTTA designs. Therefore, in this section, we analyzed the literature evaluation methods in four subsections: research questions, Participants, Procedure, and Metrics.

4.1 Research Questions

Common literature research questions are manipulation methods (10/24 papers, 42%): DOF separation, multi-level DOF, learning (knowledge retention, skill acquisition, and transferability), or user experience evaluation (presence and engagement), Fig. 36.



Fig. 36. Research Questions of the TOTTA papers.

A secondary research question concerns interaction devices (5/24 papers, 21%), such as hand-held controllers, finger-tracking, and eye-tracking, and their effect on user experience regarding usability, user satisfaction, and performance. A third common research question concerns visualization techniques (5/24 papers,

21%) and the effects of AR/VR user experience on presence and engagement.

Our research discovers that the research questions on widget visual design are limited (3/24 papers, 13%) despite its significant impact on the user experience and performance.

Other research questions are specific: gender video game experience affects self-avatar representation (1/24 papers, 4%) on the 3D manipulation task.

4.1 Participants

Within-subject (20/24 papers, 83%) is the most utilized methodology among the literature papers. The average number of participants is 21 +—SD 12.9, with a minimum age of 18 and a maximum age of 50. Most are right-handed (92%), young (SD 3.1, mean 25.8), and male (65%) (Fig. 37). The majority are student participants (90%), some unpaid volunteers (4 /24 papers), and some paid (4/ 24 papers), while the rest is not specified.



Fig. 37. Participants' demographic information.

4.2 Procedure

TOTTA validations are more commonly tested in VR (19/24 papers, 79%) than in AR. Twelve papers use VR HMD, five desktop VR, three AR HMD, two Projected AR, one Fish Tank VR, one VST-HMD, and one hand-held AR, as shown in Table 1.

Often the participants are provided with a hand-held device (17/24 papers, 71%) since it simulates the feeling of a TO being picked up or provides physical feedback when interacting with an input device or controller, Table 1. A custom-made object input device is the most preferred (9/24 papers, %37), followed by controller (7/29 papers, %29) and mouse (2/24 papers, %8) input devices. Moreover, 5/24 papers (%21) involved a free-hand interaction, and 2/24 papers (%8) included a gaze interaction as well. In some cases, the participants were also provided with virtual avatars and input device representation (8/24 papers, 33%), which influences self-perception and the sense of presence in mixed realities.

4.3 Metrics

The literature papers measure performance, acceptance, and preference using similar metrics. The most inquired metric is task completion time (24/24 papers, 100%), followed by subjective data (20/24 papers, 83%), positional error (15/24 papers, 63%), angular error (11/24 papers, 46%), scale error (4/24 papers, 79%), and count (e.g., attempts, click) (4/24 papers, 17%) (Fig. 38).

We analyzed the Subjective data since it is essential for comprehensively understanding the user experience, needs, emotions, and behaviors collected during the experiment. The most investigated metric is user preference (14/24 papers, 58%), followed by the perceived performance (9/24 papers, 38%), NASA-TLX (7/24 papers, 29%), difficulty rate (6/24 papers, 25%), comfort rate (6/24 papers, 25%), SUS (4/24 papers, 17%), intuitiveness (4/24 papers, 17%), PQ-presence questionnaire (3/24, 13%), UEQ (2/24 papers, Fig. 39. The least used metrics are AttrakDiff (1/24 papers, 5%), ARI (1/24 papers, 4%), and HARUS (1/24 papers, 4%).

5 DISCUSSION

This systematic review showed that TOTTA widget designs are non-standard and differ in design, feedback, and interaction. In all studies, the TOTTA guidance is supported by the Gestalt theory's



Fig. 38: Quantitative TOTTA collected metrics.

visual overlap or cognitive psychology as a "collimation feedback," intuitive and pre-attentive feedback.

The visual overlap of similar basic shapes is supported by reification, a pre-attentive human capability of interpreting visual information as theorized by the Gestalt laws, such as proximity, closure, similarity, and continuation.

This aspect is not always supported from a theoretical cognitive point of view in the studies. Box, 3D Axis, and 3D Models are the most preferred ones. However the geometries that are missing a center (e.g., wireframe Box) can have problems in precise positioning.

The alignment of the same basic-shaped TO and TA presents the challenge of differentiating them. Some evaluated approaches are frames/vertices/faces color or style (75%), different-colored TO/TA (71%), transparency (67%), additional shapes (59%), multiple widgets (54%), and size (46%). Using colored, partial elements and transparency, TO and TA are differentiated. Transparency can also be beneficial for reducing visual occlusion during the TOTTA task.

Although some patterns are visible and probably lead to effective solutions, few studies provide scientific ground for visual and interaction design. We argue that a deeper comprehension of the perception of shapes (e.g., Gestalt theories) can lead to better results in terms of performance and usability.

Another aspect is the common design of the continuous "during manipulation feedback." Sonification and animations can bring large margins of improvement. The latest papers are evolving from the one-time signal at goal reach to a potentially more effective, dynamic, and responsive guidance method.

Another key finding is the lack of a well-established golden standard and direct comparisons of the present widgets or a partial set. The hardware systems used in the selected papers vary in terms of immersion, tracking, and visual quality, deeply impacting the resulting experience, user performances, and acceptance.

The TOTTA validation methods found in this systematic review use quite different -thus not comparable-experiment designs (target configuration, avatar, background, etc.), and we spot some bias (VR is more tested than AR, right-handed, and male participants). Another aspect to highlight is that the accessibility issues are not investigated or mitigated (e.g., color code and color blindness), and none of the studies reviewed addressed this topic.

5.1 Limitations and Future Works

This study acknowledges its limitations because no previous research has been done in the literature to analyze the TOTTA widget design and evaluation methods. We examined papers that required a specific tool to align targets using visual widgets. Tasks such as assembly were excluded since the user searches for tools and targets in the environment and constructs a whole. This limitation allows us to analyze each method deeply.

For future works, it is important to evaluate the possible tradeoff between the widget complexity and the cognitive overload as experienced in some experiments [33, 48]. We think that to achieve better performance, as requested by the industry, widget behavior will increase in complexity, and the future challenge is to balance this with users' cognitive overload.



Fig. 39. The subjective TOTTA collected metrics.

Also, DOF separation improved user precision so that it may be investigated better as a design solution in future studies. On the grounds of this study, we can draw the future TOTTA research:

- Define and implement a standard experiment framework.
- Equally compare the existing TOTTA widgets.
- Improve widget design by perception theory.
- Improve widgets with continuous guidance.
- Enforce diversity in user tests.

5.2 Key findings

We extracted some main research outcomes and key findings approved by TOTTA papers, considering that they are valid in the specific context, widget design, and experimental conditions.

- TOTTA stereoscopic performs better than monoscopic view [26].
- 6DOF direct hand manipulations [29, 39], DOF separation [28, 35], switchable DOF [38], increased precision, completion time, and qualitative results.
- Self-avatar visualizations reduce task time [30], and participants preferred the semi-transparent hand [31].
- Object impersonation (user embodiment in TO) provides better orientational error but increases cognitive demand than DRIVE and VIEW metaphors [33].
- If TO vs. TA sizes differ (the user must also apply to scale), the physical input device influences the task precision and time, and bimanual interactions are better suited than unimanual [34].
- The participants perform faster in AR than in VR with a 3D input device [36].
- Free-hand manipulation is faster than widget-based manipulation and is preferred by the participants, and multimodal cues improve the user experience [39].
- The 2D crosshair performs better than the 2D arrow for translation and rotation [40].
- 3D Model TOTTA performs better than 3D Axes/Globe [46].
- The tetrahedron shape has better orientational precision, but the task time increases [32] compared to the model shape.
- Gaze-based manipulation causes more fatigue than controllerbased manipulation [43].

6 CONCLUSION

This work defined the TOTTA guidance problem and its importance in several applications. We provided a systematic review, starting from 206 papers, and we selected 24 papers that deeply analyzed dedicated widgets. We described the existing approaches, studied TOTTA widget design characteristics and evaluation methods, and presented key findings of the selected papers. We are encouraged by this study's findings, which identify the gaps and provide context for future research. This research provides interesting perspectives and guides for the many research applications and industries where the precise tool to target object manipulation is required, such as medicine, aviation, industry, retail, etc. This study concludes with acknowledgment of the need to address TOTTA problems and the necessity of a standardized but flexible evaluation methodology due to the different declinations of MR systems in the scientific community.

ACKNOWLEDGMENTS

Co-founded by MICS (Made in Italy - Circular and Sustainable) -National Recovery and Resilience Plan (PNRR), 4-2-1.3 - No. 341, Italian Ministry of University and Research, European Union – NextGenerationEU PE00000004, CUP D93C22000920001. Cofunding by Emme Evolution S.r.l. FESR 2014 – 2020 Objectivo Convergenza. Project development organizations: with Polytechnic University of Bari, CUP B95H22000810007.

REFERENCES

- M.-U. Erdenebat, Y.-T. Lim, K.-C. Kwon, N. Darkhanbaatar, and N. Kim, "Waveguide-Type Head-Mounted Display System for AR Application," in *State of the Art Virtual Reality and Augmented Reality Knowhow*, InTech, 2018. doi: 10.5772/intechopen.75172.
- [2] R. Anthony J. de Belen, H. Nguyen, D. Filonik, D. Del Favero, and T. Bednarz, "A systematic review of the current state of collaborative mixed reality technologies: 2013–2018," *AIMS Electronics and Electrical Engineering*, vol. 3, no. 2, pp. 181–223, 2019, doi: 10.3934/ElectrEng.2019.2.181.
- [3] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *IEEE Comput Graph Appl*, vol. 21, no. 6, pp. 34–47, 2001, doi: 10.1109/38.963459.
- [4] C. Hand, "A Survey of 3D Interaction Techniques," Computer Graphics Forum, vol. 16, no. 5, pp. 269–281, Dec. 1997, doi: 10.1111/1467-8659.00194.
- [5] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev, "An Introduction to 3-D User Interface Design," *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 1, pp. 96– 108, Feb. 2001, doi: 10.1162/105474601750182342.
- [6] X. Xu, E. Mangina, and A. G. Campbell, "HMD-Based Virtual and Augmented Reality in Medical Education: A Systematic Review," *Front Virtual Real*, vol. 2, Jul. 2021, doi: 10.3389/frvir.2021.692103.
- [7] A. Martin-Gomez, U. Eck, and N. Navab, "Visualization Techniques for Precise Alignment in VR: A Comparative Study," in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, Mar. 2019, pp. 735–741. doi: 10.1109/VR.2019.8798135.
- [8] E. Laviola, M. Gattullo, V. M. Manghisi, M. Fiorentino, and A. E. Uva, "Minimal AR: visual asset optimization for the authoring of augmented reality work instructions in manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 119, no. 3–4, pp. 1769–1784, Mar. 2022, doi: 10.1007/s00170-021-08449-6.
- [9] L. Rodriguez, F. Quint, D. Gorecky, D. Romero, and H. R. Siller, "Developing a Mixed Reality Assistance System Based on Projection Mapping Technology for Manual Operations at Assembly Workstations," *Procedia Comput Sci*, vol. 75, pp. 327– 333, 2015, doi: 10.1016/j.procs.2015.12.254.
- [10] H. J. Kang, J. Shin, and K. Ponto, "A Comparative Analysis of 3D User Interaction: How to Move Virtual Objects in Mixed Reality," in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, Mar. 2020, pp. 275–284. doi: 10.1109/VR46266.2020.00047.
- [11] S. Schmidt, O. J. A. Nunez, and F. Steinicke, "Blended Agents: Manipulation of Physical Objects within Mixed Reality Environments and Beyond," in *Symposium on Spatial User Interaction*, New York, NY, USA: ACM, Oct. 2019, pp. 1–10. doi: 10.1145/3357251.3357591.
- [12] J. Müller, R. Rädle, and H. Reiterer, "Virtual Objects as Spatial Cues in Collaborative Mixed Reality Environments," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, May 2016, pp. 1245–1249. doi: 10.1145/2858036.2858043.
- [13] P. M. Kuber and E. Rashedi, "Alterations in Physical Demands During Virtual/Augmented Reality-Based Tasks: A Systematic Review," Ann Biomed Eng, vol. 51, no. 9, pp. 1910–1932, Sep. 2023, doi: 10.1007/s10439-023-03292-0.
- [14] A. Çöltekin *et al.*, "Extended Reality in Spatial Sciences: A Review of Research Challenges and Future Directions," *ISPRS Int J Geoinf*, vol. 9, no. 7, p. 439, Jul. 2020, doi: 10.3390/ijgi9070439.

- [15] J. Buchner, K. Buntins, and M. Kerres, "The impact of augmented reality on cognitive load and performance: A systematic review," *J Comput Assist Learn*, vol. 38, no. 1, pp. 285–303, Feb. 2022, doi: 10.1111/jcal.12617.
- [16] G. Frediani and F. Carpi, "Tactile display of softness on fingertip," *Sci Rep*, vol. 10, no. 1, p. 20491, Nov. 2020, doi: 10.1038/s41598-020-77591-0.
- [17] M. Zhou, S. Tse, A. Derevianko, D. B. Jones, S. D. Schwaitzberg, and C. G. L. Cao, "Effect of haptic feedback in laparoscopic surgery skill acquisition," *Surg Endosc*, vol. 26, no. 4, pp. 1128– 1134, Apr. 2012, doi: 10.1007/s00464-011-2011-8.
- [18] I. El Rassi and J.-M. El Rassi, "A review of haptic feedback in tele-operated robotic surgery," *J Med Eng Technol*, vol. 44, no. 5, pp. 247–254, Jul. 2020, doi: 10.1080/03091902.2020.1772391.
- [19] J. T. Verhey, J. M. Haglin, E. M. Verhey, and D. E. Hartigan, "Virtual, augmented, and mixed reality applications in orthopedic surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 16, no. 2, Apr. 2020, doi: 10.1002/rcs.2067.
- [20] H. Eschen, T. Kötter, R. Rodeck, M. Harnisch, and T. Schüppstuhl, "Augmented and Virtual Reality for Inspection and Maintenance Processes in the Aviation Industry," *Procedia Manuf*, vol. 19, pp. 156–163, 2018, doi: 10.1016/j.promfg.2018.01.022.
- [21] A. Muñoz, X. Mahiques, J. E. Solanes, A. Martí, L. Gracia, and J. Tornero, "Mixed reality-based user interface for quality control inspection of car body surfaces," *J Manuf Syst*, vol. 53, pp. 75–92, Oct. 2019, doi: 10.1016/j.jmsy.2019.08.004.
- [22] J. Rehbein, S.-J. Lorenz, J. Holtmannspötter, and B. Valeske, "3D-Visualization of Ultrasonic NDT Data Using Mixed Reality," J Nondestr Eval, vol. 41, no. 1, p. 26, Mar. 2022, doi: 10.1007/s10921-022-00860-7.
- [23] S. Zhai, P. Milgram, and W. Buxton, "The influence of muscle groups on performance of multiple degree-of-freedom input," in *Proceedings of the SIGCHI conference on Human factors in computing systems common ground - CHI '96*, New York, New York, USA: ACM Press, 1996, pp. 308–315. doi: 10.1145/238386.238534.
- [24] B. D. Conner, S. S. Snibbe, K. P. Herndon, D. C. Robbins, R. C. Zeleznik, and A. van Dam, "Three-dimensional widgets," in *Proceedings of the 1992 symposium on Interactive 3D graphics SI3D '92*, New York, New York, USA: ACM Press, 1992, pp. 183–188. doi: 10.1145/147156.147199.
- [25] G. Ramos, M. Boulos, and R. Balakrishnan, "Pressure widgets," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA: ACM, Apr. 2004, pp. 487–494. doi: 10.1145/985692.985754.
- [26] J. Boritz and K. S. Booth, "A study of interactive 6 DOF docking in a computerised virtual environment," in *Proceedings. IEEE* 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180), IEEE Comput. Soc, pp. 139–146. doi: 10.1109/VRAIS.1998.658476.
- [27] M. Fiorentino, G. Monno, P. A. Renzulli, and A. E. Uva, "3d pointing in virtual reality: experimental study," 2003.
- [28] M. Veit, A. Capobianco, and D. Bechmann, "An experimental analysis of the impact of Touch Screen Interaction techniques for 3-D positioning tasks," in 2011 IEEE Virtual Reality Conference, IEEE, Mar. 2011, pp. 75–82. doi: 10.1109/VR.2011.5759440.
- [29] E. Ragan, C. Wilkes, D. A. Bowman, and T. Hollerer, "Simulation of Augmented Reality Systems in Purely Virtual Environments," in 2009 IEEE Virtual Reality Conference, IEEE, Mar. 2009, pp. 287–288. doi: 10.1109/VR.2009.4811058.
- [30] M. Raj, S. H. Creem-Regehr, K. M. Rand, J. K. Stefanucci, and W. B. Thompson, "Kinect based 3D object manipulation on a desktop display," in *Proceedings of the ACM Symposium on Applied Perception*, New York, NY, USA: ACM, Aug. 2012, pp. 99–102. doi: 10.1145/2338676.2338697.
- [31] T. Ha, S. Feiner, and W. Woo, "WeARHand: Head-worn, RGB-D camera-based, bare-hand user interface with visually enhanced depth perception," in 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, Sep. 2014, pp. 219–228. doi: 10.1109/ISMAR.2014.6948431.
- [32] V. Vuibert, W. Stuerzlinger, and J. R. Cooperstock, "Evaluation of Docking Task Performance Using Mid-air Interaction Techniques," in *Proceedings of the 3rd ACM Symposium on*

Spatial User Interaction, New York, NY, USA: ACM, Aug. 2015, pp. 44–52. doi: 10.1145/2788940.2788950.

- [33] J. Wang and R. W. Lindeman, "Object impersonation: Towards effective interaction in tablet- and HMD-based hybrid virtual environments," in 2015 IEEE Virtual Reality (VR), IEEE, Mar. 2015, pp. 111–118. doi: 10.1109/VR.2015.7223332.
- [34] J. Feng, I. Cho, and Z. Wartell, "Comparison of Device-Based, One and Two-Handed 7DOF Manipulation Techniques," in *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, New York, NY, USA: ACM, Aug. 2015, pp. 2–9. doi: 10.1145/2788940.2788942.
- [35] D. Mendes, F. Relvas, A. Ferreira, and J. Jorge, "The benefits of DOF separation in mid-air 3D object manipulation," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, New York, NY, USA: ACM, Nov. 2016, pp. 261–268. doi: 10.1145/2993369.2993396.
- [36] M. Krichenbauer, G. Yamamoto, T. Taketom, C. Sandor, and H. Kato, "Augmented Reality versus Virtual Reality for 3D Object Manipulation," *IEEE Trans Vis Comput Graph*, vol. 24, no. 2, pp. 1038–1048, Feb. 2018, doi: 10.1109/TVCG.2017.2658570.
- [37] H. Ro, J.-H. Byun, Y. J. Park, N. K. Lee, and T.-D. Han, "AR Pointer: Advanced Ray-Casting Interface Using Laser Pointer Metaphor for Object Manipulation in 3D Augmented Reality Environment," *Applied Sciences*, vol. 9, no. 15, p. 3078, Jul. 2019, doi: 10.3390/app9153078.
- [38] M. Kim and J. Han, "Effects of Switchable DOF for Mid-Air Manipulation in Immersive Virtual Environments," *Int J Hum Comput Interact*, vol. 35, no. 13, pp. 1147–1159, Aug. 2019, doi: 10.1080/10447318.2018.1514163.
- [39] R. Schlünsen, O. Ariza, and F. Steinicke, "A VR Study on Freehand vs. Widgets for 3D Manipulation Tasks," in *Proceedings* of Mensch und Computer 2019, New York, NY, USA: ACM, Sep. 2019, pp. 223–233. doi: 10.1145/3340764.3340791.
- [40] F. Heinrich, F. Joeres, K. Lawonn, and C. Hansen, "Comparison of Projective Augmented Reality Concepts to Support Medical Needle Insertion," *IEEE Trans Vis Comput Graph*, vol. 25, no. 6, pp. 2157–2167, Jun. 2019, doi: 10.1109/TVCG.2019.2903942.
- [41] W. Sun et al., "Workload, Presence and Task Performance of Virtual Object Manipulation on WebVR," in 2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), IEEE, Dec. 2020, pp. 358–361. doi: 10.1109/AIVR50618.2020.00073.
- [42] D. Andersen and V. Popescu, "AR Interfaces for Mid-Air 6-DoF Alignment: Ergonomics-Aware Design and Evaluation," in 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, Nov. 2020, pp. 289–300. doi: 10.1109/ISMAR50242.2020.00055.
- [43] C. Liu, A. Plopski, and J. Orlosky, "OrthoGaze: Gaze-based threedimensional object manipulation using orthogonal planes," *Comput Graph*, vol. 89, pp. 1–10, Jun. 2020, doi: 10.1016/j.cag.2020.04.005.
- [44] M. Weib, K. Angerbauer, A. Voit, M. Schwarzl, M. Sedlmair, and S. Mayer, "Revisited: Comparison of Empirical Methods to Evaluate Visualizations Supporting Crafting and Assembly Purposes," *IEEE Trans Vis Comput Graph*, vol. 27, no. 2, pp. 1204–1213, Feb. 2021, doi: 10.1109/TVCG.2020.3030400.
- [45] V. Fuvattanasilp *et al.*, "SlidAR+: Gravity-aware 3D object manipulation for handheld augmented reality," *Comput Graph*, vol. 95, pp. 23–35, Apr. 2021, doi: 10.1016/j.cag.2021.01.005.
- [46] C.-Y. Lee, W.-A. Hsieh, D. Brickler, S. V. Babu, and J.-H. Chuang, "Design and Empirical Evaluation of a Novel Near-field Interaction Metaphor on Distant Object Manipulation in VR," in *Symposium on Spatial User Interaction*, New York, NY, USA: ACM, Nov. 2021, pp. 1–11. doi: 10.1145/3485279.3485296.
- [47] D. Yu et al., "Gaze-Supported 3D Object Manipulation in Virtual Reality," in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, New York, NY, USA: ACM, May 2021, pp. 1–13. doi: 10.1145/3411764.3445343.
- [48] M. Dastan, A. E. Uva, and M. Fiorentino, "Gestalt Driven Augmented Collimator Widget for Precise 5 DOF Dental Drill Tool Positioning in 3D Space," in 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, Oct. 2022, pp. 187–195. doi: 10.1109/ISMAR55827.2022.00033.
- [49] G. Ganias, C. Lougiakis, A. Katifori, M. Roussou, Y. Ioannidis, and Ioannis P. Ioannidis, "Comparing Different Grasping

Visualizations for Object Manipulation in VR using Controllers," *IEEE Trans Vis Comput Graph*, vol. 29, no. 5, pp. 2369–2378, May 2023, doi: 10.1109/TVCG.2023.3247039.

- [50] S. Brewster, "The design of sonically-enhanced widgets," *Interact Comput*, vol. 11, no. 2, pp. 211–235, Dec. 1998, doi: 10.1016/S0953-5438(98)00028-9.
- [51] RobertC. Zeleznik et al., "An Interactive 3D Toolkit for Constructing 3D Widgets," in Proceedings of SIGGRAPH 93, 1993, pp. 81–84.
- [52] E. H. Korkut and E. Surer, "Visualization in virtual reality: a systematic review," *Virtual Real*, vol. 27, no. 2, pp. 1447–1480, Jun. 2023, doi: 10.1007/s10055-023-00753-8.
- [53] T. Galais, A. Delmas, and R. Alonso, "Natural interaction in virtual reality," in *Adjunct Proceedings of the 31st Conference on l'Interaction Homme-Machine*, New York, NY, USA: ACM, Dec. 2019, pp. 1–9. doi: 10.1145/3366551.3370342.
- [54] W. Sun et al., "Workload, Presence and Task Performance of Virtual Object Manipulation on WebVR," in 2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), IEEE, Dec. 2020, pp. 358–361. doi: 10.1109/AIVR50618.2020.00073.
- [55] S. Kim, M. A. Nussbaum, and J. L. Gabbard, "Influences of augmented reality head-worn display type and user interface design on performance and usability in simulated warehouse order picking," *Appl Ergon*, vol. 74, pp. 186–193, Jan. 2019, doi: 10.1016/j.apergo.2018.08.026.
- [56] Y. Park, V. Lepetit, and Woontack Woo, "Multiple 3D Object tracking for augmented reality," in 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE, Sep. 2008, pp. 117–120. doi: 10.1109/ISMAR.2008.4637336.
- [57] S. Rokhsaritalemi, A. Sadeghi-Niaraki, and S.-M. Choi, "A Review on Mixed Reality: Current Trends, Challenges and Prospects," *Applied Sciences*, vol. 10, no. 2, p. 636, Jan. 2020, doi: 10.3390/app10020636.
- [58] M. Walker, T. Phung, T. Chakraborti, T. Williams, and D. Szafir, "Virtual, Augmented, and Mixed Reality for Human-robot Interaction: A Survey and Virtual Design Element Taxonomy," *ACM Trans Hum Robot Interact*, vol. 12, no. 4, pp. 1–39, Dec. 2023, doi: 10.1145/3597623.
- [59] J. Bergström, T.-S. Dalsgaard, J. Alexander, and K. Hornbæk, "How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, May 2021, pp. 1–20. doi: 10.1145/3411764.3445193.
- [60] A. Sadeghi Milani, A. Cecil-Xavier, A. Gupta, J. Cecil, and S. Kennison, "A Systematic Review of Human–Computer Interaction (HCI) Research in Medical and Other Engineering Fields," *Int J Hum Comput Interact*, vol. 40, no. 3, pp. 515–536, Feb. 2024, doi: 10.1080/10447318.2022.2116530.
- [61] A. Liberati *et al.*, "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration," *J Clin Epidemiol*, vol. 62, no. 10, pp. e1–e34, Oct. 2009, doi: 10.1016/j.jclinepi.2009.06.006.
- [62] K. Nesenbergs, V. Abolins, J. Ormanis, and A. Mednis, "Use of Augmented and Virtual Reality in Remote Higher Education: A Systematic Umbrella Review," *Educ Sci (Basel)*, vol. 11, no. 1, p. 8, Dec. 2020, doi: 10.3390/educsci11010008.
- [63] H. Sumdani, P. Aguilar-Salinas, M. J. Avila, S. R. Barber, and T. Dumont, "Utility of Augmented Reality and Virtual Reality in Spine Surgery: A Systematic Review of the Literature," World Neurosurg, vol. 161, pp. e8–e17, May 2022, doi: 10.1016/j.wneu.2021.08.002.
- [64] M. L. Rethlefsen *et al.*, "PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews," *Syst Rev*, vol. 10, no. 1, p. 39, Jan. 2021, doi: 10.1186/s13643-020-01542-z.
- [65] E. Ragan, C. Wilkes, D. A. Bowman, and T. Hollerer, "Simulation of Augmented Reality Systems in Purely Virtual Environments," in 2009 IEEE Virtual Reality Conference, IEEE, Mar. 2009, pp. 287–288. doi: 10.1109/VR.2009.4811058.