Evaluation of AR visualization approaches for catheter insertion into the ventricle cavity

Benmahdjoub, Mohamed; Thabit, Abdullah; Van Veelen, Marie Lise C.; Niessen, Wiro J.; Wolvius, Eppo B.; Van Walsum, Theo

DOI
10.1109/TVCG.2023.3247042

Publication date
2023

Document Version
Final published version

Published in
IEEE Transactions on Visualization and Computer Graphics

Citation (APA)
Benmahdjoub, M., Thabit, A., Van Veelen, M. L. C., Niessen, W. J., Wolvius, E. B., & Van Walsum, T. (2023). Evaluation of AR visualization approaches for catheter insertion into the ventricle cavity. *IEEE Transactions on Visualization and Computer Graphics*, *29*(5), 2434-2445. https://doi.org/10.1109/TVCG.2023.3247042

Evaluation of AR visualization approaches for catheter insertion into the ventricle cavity

Mohamed Benmahdjoub*, Abdullah Thabit*, Marie-Lise C. van Veelen, Wiro J. Niessen *Member, IEEE*,
Eppo B. Wolvius, and Theo van Walsum *Member, IEEE*

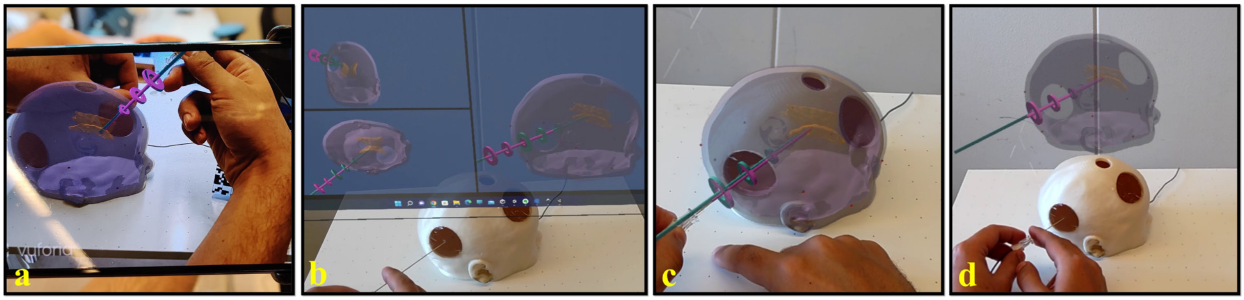


Fig. 1: Evaluated visualization approaches for the insertion task. a: smartphone-based AR (Smart2), b: see-through 2D window (Win2), c: fully aligned model (Align3), d: rotationally aligned model (Off3).

Abstract—Augmented reality (AR) has shown potential in computer-aided surgery. It allows for the visualization of hidden anatomical structures as well as assists in navigating and locating surgical instruments at the surgical site. Various modalities (devices and/or visualizations) have been used in the literature, but few studies investigated the adequacy/superiority of one modality over the other. For instance, the use of optical see-through (OST) HMDs has not always been scientifically justified. Our goal is to compare various visualization modalities for catheter insertion in external ventricular drain and ventricular shunt procedures. We investigate two AR approaches: (1) 2D approaches consisting of a smartphone and a 2D window visualized through an OST (Microsoft HoloLens 2), and (2) 3D approaches consisting of a fully aligned patient model and a model that is adjacent to the patient and is rotationally aligned using an OST. 32 participants joined this study. For each visualization approach, participants were asked to perform five insertions after which they filled NASA-TLX and SUS forms. Moreover, the position and orientation of the needle with respect to the planning during the insertion task were collected. The results show that participants achieved a better insertion performance significantly under 3D visualizations, and the NASA-TLX and SUS forms reflected the preference of participants for these approaches compared to 2D approaches.

Index Terms—Computer-assisted surgery, Surgical navigation systems, Augmented reality, Augmented reality visualization, Needle guidance, External ventricular drain, Ventricular shunt, User study

✦

1 INTRODUCTION

Hydrocephalus is one of the most common medical conditions in pediatric neurosurgery. It occurs when there is an accumulation of cerebrospinal fluid (CSF) in the brain, which is typically associated with enlargement of the ventricles and elevated intracranial pressure [32, 68]. Placement of an external ventricular drain (EVD) or ventricular shunt (VS) are common procedures that are used to treat hydrocephalus. In both procedures a catheter is inserted into the ventricle cavity to remove CSF buildup [32]. This is often preformed blindly using a free hand approach, where surgeons place the catheter based on anatomical surface landmarks. Free hand catheter placement has been reported to have high occurrence of misplacement, requiring multiple insertions of the catheter which may cause unnecessary brain damage [4, 13].

Computer assisted image-guidance in neurosurgery is widely used and has shown to improve surgical procedures [61]. Surgical navigation systems allow for locating surgical instruments with respect to the patient's position. This is done by tracking the patient and instruments in the operative field and visualizing the patient-specific data, such as CT or MRI scans, and surgical plan on a 2D monitor. This technology generally makes use of optical/electromagnetic tracking systems (OTS/EMTS) to track sensors attached to instruments or anatomical regions of interest [14]. For instance, EM-based navigation has been proposed to guide catheter placement in EVD and ventricular shunt procedures [4, 33, 44]. However, the use of these conventional navigation systems comes with few challenges: depth perception, hand-eye coordination, and the switch of focus between the surgical site and the screen. These challenges could possibly be addressed by the use of extended reality (XR)-based surgical navigation [6, 8].

There are two areas of extended reality: one where a user gets immersed in virtual environments, which is commonly known as virtual

---

- ** The first two authors contributed equally to this work.*
- *Mohamed Benmahdjoub and Abdullah Thabit are with the Biomedical Imaging Group Rotterdam, Department of Radiology & Nuclear medicine and the Department of Oral and Maxillofacial Surgery, Erasmus MC, 3015 GE Rotterdam, The Netherlands.*
- *Marie-Lise C. van Veelen is with the Department of Neurosurgery, Erasmus MC, 3015 GE Rotterdam, The Netherlands.*
- *Wiro J. Niessen is with the Biomedical Imaging Group Rotterdam, Department of Radiology & Nuclear medicine, Erasmus MC, 3015 GE Rotterdam, The Netherlands, and with the Department of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands.*
- *Eppo B. Wolvius is with the Department of Oral and Maxillofacial Surgery, Erasmus MC, 3015 GE Rotterdam, The Netherlands.*
- *Theo van Walsum is with the Biomedical Imaging Group Rotterdam, Department of Radiology & Nuclear medicine, Erasmus MC, 3015 GE Rotterdam, The Netherlands.*

Manuscript received 14 October 2022; revised 13 January 2023; accepted 30 January 2023. Date of publication 22 February 2023; date of current version 29 March 2023. Digital Object Identifier no. 10.1109/TVCG.2023.3247042

reality (VR), and one where virtual objects are brought into the real world, which is commonly known as augmented reality (AR).

VR has been introduced in education [55], training skills [38, 74], and medical tasks for various applications [21]: from mental health and physical rehabilitation [48, 63] to diagnosis and planning of surgical interventions [58, 69]. AR has also been used for training [12, 28, 73], tele-mentoring [41, 59], and entertainment [16, 35]. In addition, AR has the quality of including real-time user assistance. In industry for instance, Radkowski et al. [56] built an AR assembly system of mechanical axial piston motors, making use of textual and animation information on 2D images. Similarly, AR has been investigated as a navigation tool during surgery such as for dental, craniomaxillofacial or spinal procedures [9, 25, 37]. In neuro-navigation, AR has improved surgical procedures and provided more intuitive view of the planning [15, 47, 65]. Rae et al. [57] used AR to mark neurosurgical burr hole placements, Fatih et al. [27] provided a proof of concept for locating brain tumors using the Microsoft HoloLens, whereas Thabit et al. [67] used a head-mounted display (HMD) to locate cranial sutures in craniosynostosis surgery.

For EVD placement, several AR-based systems have been suggested in the literature. Sun et al. [62] focused on catheter tracking using HL. This allowed the visualization of the catheter when it's conventionally invisible as the insertion starts. Li et al. [40] proposed a Microsoft HoloLens (HL) guidance system where a manually registered skull, the insertion axis, and the ventricles were visualized. Azimi et al. and Van Gestel et al. [5, 70] conducted a phantom study, using the HL and an automatic landmark-based alignment approach, to visualize the ventricles and the insertion path. Moreover, Palumbo et al. [49] investigated EVD placement using HL2 and a surface-based registration allowing for a fast workflow for emergency cases. Note that in all these studies, an optical see-through (OST) device was used (HL/HL2). However, various technologies for visualizing AR virtual models have been reported. These technologies can be categorized into optical-see through (OST), video-see through (VST), and projection-based devices [60]. In OST devices, the virtual models are added to the real world using additive screens such as the HL. In VST devices, the virtual plan is added to the video feed from a camera such as using a smartphone, a tablet or screen-based HMDs [6]. In projection-based AR, the virtual models are projected on the physical surface of the target structure [20].

The way of presenting the virtual model and the planning may play a vital role in identifying target structures and locating surgical instruments. Different visualization techniques have been presented in the literature, such as smartphone/tablet-based AR [76], window-screen AR [52], and 3D overlays either directly aligned with the patient [18] or adjacent to the patient [42]. Little is known on how these visualization approaches perform in comparison to each other and which of these visualization approaches is more suitable for guidance in surgical procedures, such as VS or EVD catheter insertion.

Our study, therefore, investigates needle insertion placement under various AR visualization approaches. The main contribution is a within-subject study which evaluates user's performance in an insertion task when using an AR OST or VST device. In this study, participants performed multiple needle/catheter insertion tasks under various visualization conditions relying only on perceptual matching [64].

The remainder of this manuscript is structured as follows: Section 2 discusses AR applications and AR visualization approaches for surgical guidance. Section 3 contains the experimental setup, design, and protocols for the user evaluation. Next, in Section 4, the results, including a statistical analysis performed on the experiment data, are presented. The manuscript is concluded by a discussion on the study outcomes, recommendations, limitations, future perspectives, and conclusions.

2 RELATED WORK

In our work, we focus on four different visualization approaches that have been reported in many AR surgical navigation systems. These include Smartphone-based AR, 2D window/screen-based AR, and 3D based AR with direct and adjacent overlays. Garcia-Mato et al. [19] used a smartphone mounted on a mechanical arm to show the virtual model overlaid in craniosynostosis surgery. In their study, the authors reported that surgeons would prefer using an HMD or a smartphone to show the AR visualization over external screens. However, the HMD visualization was not tested. Therefore, it was not clear whether it would have been preferred over using the smartphone.

Pellegrino et al. [52] used the Microsoft HoloLens for dynamic navigation in dental implant placement surgery via a 2D virtual window. Their approach allowed to visualize the planning directly within the operative field next to the patient and was shown to be useful. In another work, Iqbal et al. [29] replicated the user interface from the screen of a surgical robot in orthopedic assisted robotic surgery. They found no changes to the clinical metrics or operating time compared to the screen, whereas the post-operative survey indicated satisfaction of the usefulness of the AR visualization. However, in both applications, despite bringing the display closer to the patient, the problem of 2D-3D mapping is not solved and surgeons were required to perform this mapping mentally.

A recent review article on OST head-mounted displays in surgery, showed that virtual content is most commonly presented as directly overlaid on the physical target (i.e., fully aligned) [18]. This visualization approach has the advantage of superimposing a 3D virtual model to match the real target, making it intuitive and informative. For instance, Jiang et al. [30] conducted an animal study assessing the feasibility of drilling trajectories by projecting the trajectories on the real animal.

Another approach that has been used for guidance during surgery is the projection of an adjacent preoperative 3D reconstructed model of the patient. Lui et al. and Tang et al. [42, 66] implemented this approach using HL and HL2 respectively. The purpose is to eliminate the 3D mental reconstruction intraoperatively, allowing for a more intuitive visualization without looking at a distant screen. However, adjacent 3D overlay that is static and does not follow the patient's rotation requires either manual adjustment of the model's rotation or 3D mental-rotation mapping by the surgeon so that the model's rotation would correspond to that of the patient.

Although the use of OST HMDs has proven to provide better performance in achieving manual tasks [70], only few studies addressed the differences between AR devices and/or visualizations for a given task, especially for surgical interventions. Qian et al. [54] compared five factors while using three different OST devices: text readability, contrast perception, mental load, frame rate and system lag. This study compares only OST devices. Additionally, the compared factors were unrelated to any given task: for surgical interventions requiring target reaching, such as needle insertion, the usage or effectiveness of OST devices cannot be concluded. Long et al. [43] compared two AR approaches, a smartphone and a HL against CBCT-guided fluoroscopy navigation. The authors concluded that AR reduced the placement time significantly, while no significant difference was noticed for the needle displacement. However, the study included few participants (n = 6) which might impact the statistical significance when comparing the three groups (Smartphone, HL, CBCT). Park et al. [50] investigated the use of HL2 for needle insertion to reach out-of-plane lesions comparing it to traditionally CT-guided insertion. The results confirmed an improvement when using AR in terms of placement accuracy, time, and the number of passes performed. Heinrich et al. [23, 24] proposed GlyphVis as a visualization approach which consists of an adapted crosshair providing depth and orientation information based on the target's color. This visualization was implemented for four different devices: stationary tablet, OST HMD (HL), a spatial AR projector-camera-system, and 2D screen. Afterwards, the four devices were compared through a user study concluding that the time-to-completion and angular reduction were in favor of projection-based approaches. In a second step, GlyphVis was compared, under HL and projector-based AR, to PathVis, which consists of showing the 3D model of the needle already placed at the right location, and to the SeeThrough visualization where a cutout area consists of a grid showing a red extrapolated virtual needle providing a perception of depth insertion. The results showed no signficant difference between the projector-based AR and the OST HMD.

To the best of our knowledge, studies investigating the different de-

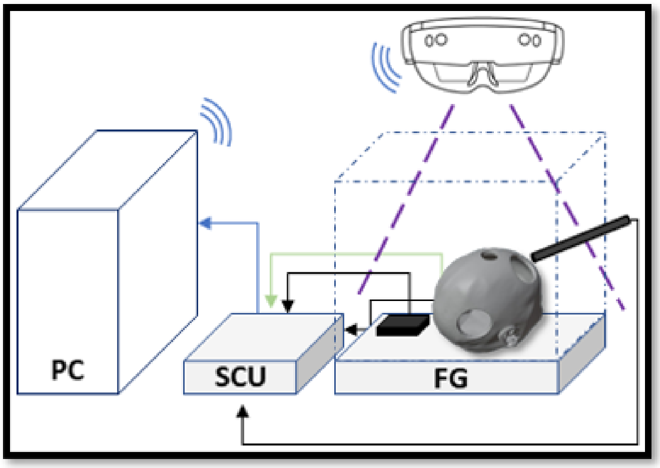


Fig. 2: The OST-HMD-based AR system setup. SCU: System Control Unit of the EMTS. FG: Field Generator. Black arrows: EM coils communication wires. SCU-PC arrow: serial connection. Dashed cube: EMTS's tracking volume. Black cube: QR-EM marker. Black cylinder: needle. The spatial information is sent from the PC to the AR device through Wi-Fi.

vices and/or visualization in needle insertion/drilling tasks are lacking. Therefore, our study aims to contribute to this research area by: (1) proposing a new visualization approach which consists of an adjacent model overlay that is rotationally aligned (i.e., follows the patient's rotation), (2) comparing state-of-the-art AR approaches that have not been assessed previously in the same experiment, and (3) specifically, providing a quantitative comparison between 3D and 2D approaches when performing spatial tasks in AR. In this way, we hope to suggest implicit recommendations, based on objective and subjective evidences, through a user study experiment, which hopefully would help in better design of AR-guided needle/drilling systems.

More specifically, this paper assesses the impact of different devices and visualizations on the accuracy of needle insertion tasks. It compares the use of smartphone-based AR and OST HMD-based AR visualized as a 2D window, a fully aligned virtual plan, and a rotationally aligned virtual plan for EVD/VS catheter placement.

3 MATERIALS AND METHODS

3.1 Visualization techniques

In this within-subject study, the AR visualization approach was considered as an independent variable. The following approaches are included in our study:

- *Smart2:* this 2D AR approach is smartphone based (Fig. 1 - a); it is common in many studies in the literature [17,34,72,75,76].
- *Win2:* 2D adjacent window: this AR approach consists of projecting a 2D navigation window containing multiple viewpoints of the 3D world through a see-through HMD (Fig. 1 - b).
- *Align3:* a classical 3D AR navigation approach where the preoperative model is aligned with the patient (Fig. 1 - c).
- *Off3:* a 3D AR approach where the 3D model is rotationally aligned, but spatially displaced (i.e., displayed next to the model) (Fig. 1 - d); the projected model follows the rotations of the tracked patient with a translation offset.

3.2 Materials

3.2.1 AR system

The AR system is composed of an EMTS, NDI Aurora (v2), which is capable of tracking EM coils located in its tracking volume (see Fig. 2). The EMTS is linked to the AR devices (HL2 or smartphone) using a multimodal QR-EM marker [8]. The marker is composed of an EM coil attached rigidly to a QR-code (Vuforia image target) that is trackable by the AR device's camera using Vuforia Engine [1] . The QR-code and the EM coil have been calibrated to allow for locating the AR device with respect to the EMTS's coordinate system [8]. The calibration of the QR-EM marker was done once and the same calibration matrix was used for both AR devices (HL2 and smartphone).

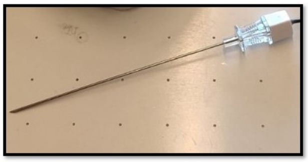


Fig. 3: The needle used for the insertion task.

A tracked pointer is first used to perform a point-based registration of the planning model to the patient model, and then an EM-tracked needle is used for the insertion task. The needle has an EMTS sensor providing its translation and rotation in the EMTS's coordinate system (see Fig. 3).

3.2.2 Implementation details

In this section, we provide more details about how the AR systems that were used in the experiment function. All of the following systems made use of the same tracking technology (EM):

- Smartphone (Smart2): a smartphone was attached to a flexible arm. It was positioned along the line of sight of the user so that its camera can track the QR-EM marker. Subsequently, the virtual plan was visualized in the smartphone screen for the user to operate the insertion task. For this study a Xiaomi 11T smartphone with a screen dimension of 6.81 inch was used (see Fig. 1 - a).
- 2D window (Win2): a navigation app in Unity was created with three views: a main camera view from the perspective of the user with a lateral view of the phantom (see Fig. 1 - b), and two additional camera views from the other axes presenting a top and frontal view of the phantom and planning. The navigation window was created to mimic the 2D visualization in traditional navigation systems, where the user has to look at the three views and mentally map them into a 3D perspective to navigate surgical instruments. The screen of the navigation app running on the PC was mirrored in the HL2 using Mirage [36]. The 2D window was placed next to the phantom and users were allowed to adjust its position based on preference.
- Fully aligned (Align3): a visualization showing the virtual model and the planning in 3D, directly overlaid on the physical 3D printed phantom. The user is allowed to rotate or move the phantom in all directions during the experiment (see Fig. 1 - c).
- Rotationally aligned (Off3): similar to Align3, the visualization shows the virtual model and the planning in 3D and follows the physical phantom orientation but it has been offsetted (translated) in the upward direction and has no overlap with the physical phantom (see Fig. 1 - d).

We categorize these approaches in two classes: (1) 2D approaches: *(Smart2)* and *(Win2)*, and (2) 3D approaches: *(Align3)* and *(Off3)*.

3.2.3 Phantom and 3D planning

A CT image (anonymized) of a head was used. The surface of the head was segmented using MevisLab [3]. Several divots (for registration) and three holes were added to the resulting 3D model using Blender [2] (see Fig. 4). The three holes represent regions where an insertion could be made: one at the back side of the head to simulate a catheter insertion for VS, and one at the front to simulate a catheter insertion for EVD. The third hole at the top was used for additional virtual

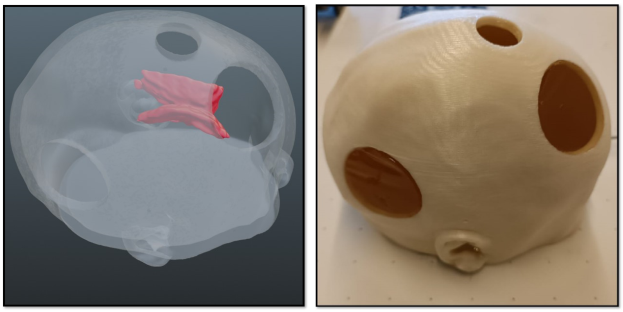


Fig. 4: The 3D model and the printed head phantom used for the experiment.

insertions to acquire more data points at various depths inside the head. Subsequently, the 3D model was printed (5mm thickness) and was filled with gelatin-water mixture (8%) to simulate the properties of soft tissue. Each hole was covered with a thin piece of self-adhesive tape to mimic the skin and to permit positioning the needle without direct penetration of the gelatin. The tape was replaced after every visualization approach. Finally, an EM coil was rigidly attached to the head, and an image-to-patient registration procedure was performed using a point-based approach [26] by locating divots visible on the CT image and also on the printed model.

To prevent participants to 'reuse' needle tracks from previous participants, the planned trajectories were slightly different for subsequent participants. In addition, after every experimental day, the holes on the phantom were covered, and the model was put inside the microwave to liquefy the gelatin, and then was placed in the fridge to regain its gel properties.

3.2.4 Needle and trajectory visualization

The approach used to visualize the needle and the pre-planned trajectory is important in the context of AR. In this study, we opted for a virtual extension (VE) visualization suggested by Benmahdjoub et al. [7], which confirmed the study of Peillard et al. [51], highlighting the importance of virtual-to-virtual matching in achieving a good performance in an alignment task when using see-through headsets. This visualization approach suggests the addition of virtual elements to the needle model that are not necessarily existing on the real/physical instrument or the planned trajectory. In this study, the extensions were represented by three rings which are parallel and in a descending size order the closer they get to the tip of the needle. The needle was visualized in pink, whereas the pre-planned trajectory was visualized in green (see Fig. 1).

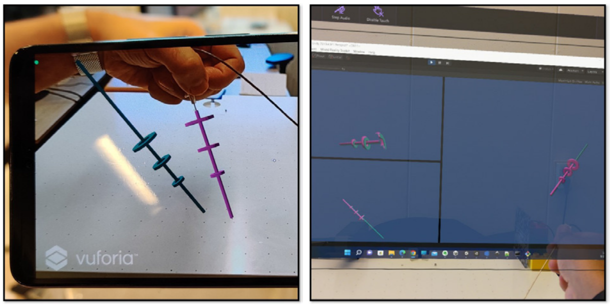
3.3 Experimental setup

The goal of the experiments is to assess subjective and objective measures of user performance under various AR approaches.

3.3.1 Task

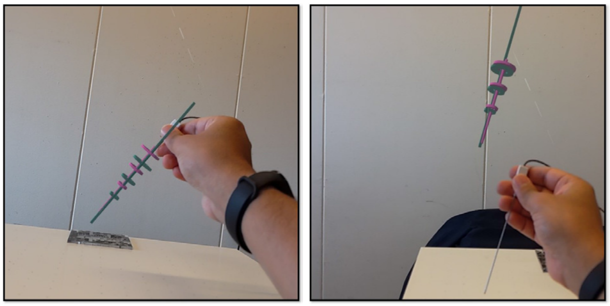
The task is the alignment and the subsequent insertion of a needle with a planned trajectory (see video attached in the supplementary material). Participants were instructed to locate the entry point on the surface and align the needle with the planning's orientation, and then insert the needle until the target was reached (when the augmentation of the needle matches that of the planned trajectory). The participants were not rushed into finishing the task, and were asked to perform their best insertion regardless of the time taken. Furthermore, in order to reduce the effects of perception errors related to perceived misalignment between real and virtual [51], the participants were instructed to focus on aligning the virtual overlay of the needle with the virtual plan rather than the needle itself, for all evaluated approaches.

Five pre-planned insertions were prepared for each user: two trajectories targeting the ventricles (one from the front and one from the back), and three trajectories were virtual insertions with a longer depth distance. The distances from the entry points on the surface of the head phantom to the end points were 35mm, 60mm, 75mm, 70mm, and 60mm respectively.



(a) Smartphone (Smart2).

(b) 2D Window (Win2).



(c) Fully aligned (Align3).

(d) Rotationally aligned (Off3).

Fig. 5: The training session under the different conditions

3.3.2 Training

In this study, we assumed that the participants are unfamiliar with the task and/or the AR approach used. Consequently, the participants underwent a training session where the time-to-completion (tt) was recorded. The purpose of the training session was to get them familiar with the VE visualization, the insertion task, and the handling and navigation of the needle using the various AR approaches (see Fig. 5).

For the training, the users were presented with a green trajectory and were asked to virtually align the needle (augmented in pink). The alignment is accepted if the tip-to-tip distance is less than 1 mm and the rotational error is less than 1°.

3.3.3 Data collection

During the experiment, needle position and orientation (reflected in the position and orientation of its virtual overlay) from the beginning of the insertion to the end were acquired, and after each visualization condition, feedback forms were filled in by the users.

The *objective measures* used for each condition were:

- Positional error $\varepsilon_{P.e}$ and $\varepsilon_{P.x}$: represent the distance from the needle to the planned trajectory at the entry point, and the tip to tip distance at the end point respectively.
- Positional error over time $\varepsilon_{P.m}$ and $\varepsilon_{P.std}$: represent the mean distance between the needle tip and the pre-planned axis from the beginning till the end of the insertion, and the standard deviation from this mean.
- Angular error $\varepsilon_{\alpha.e}$ and $\varepsilon_{\alpha.x}$: the angular error between the needle and the planned trajectory at the beginning and at the end of the insertion respectively.
- Angular error over time $\varepsilon_{\alpha.m}$ and $\varepsilon_{\alpha.std}$: the mean angular error between the planned trajectory and the needle from the beginning to the end of the insertion, and the standard deviation from this mean.
- *tt*: the amount of time spent to finish the training session.

The *subjective measures* were collected for each condition based on System Usability Scale (SUS) [11] and Nasa Task Load Index (NASA-TLX) [22] inquiries. From NASA-TLX forms, mental demand (**MD**) and frustration (**FR**) levels are highlighted further since they relate more to the tasks performed. Additionally, the ranking of the conditions based on performance and preference, and the experienced advantages and disadvantages were collected at the end of the experiments.

3.3.4 Statistics

The normality of data was checked using QQ-plots in combination with Shapiro-Wilk tests. For data that was non-normally distributed, statistical significance was assessed with Wilcoxon-signed-rank dependent tests ($\alpha = 0.05$), which were run pairwise between the various conditions. As for NASA and SUS scores, the data was normally distributed, and therefore the statistical significance was verified using paired t-tests. The results for all metrics are reported as median [IQR]. All the analyses and statistics were performed using python and the python scipy package.

3.3.5 Hypotheses

The hypotheses for our study are:

H1. 3D approaches improve the insertion accuracy compared to 2D approaches.
Rationale: The provided 3D perception, by looking from various directions using 3D approaches, permits good performance in needle insertion compared to traditionally CT-guided insertion [50].

H2. 3D approaches are easier to learn than 2D approaches. Therefore, we expect a lower training session (tt) for 3D approaches than 2D approaches.
Rationale: the easier localization of targets and the intuitive hand-eye coordination in 3D approaches could be a major factor in achieving a fast training. Long et al. [43] have noticed that using HL2 reduced the insertion time significantly.

H3. 3D approaches are more usable than 2D approaches for needle insertions.
Rationale: 2D approaches could still show some of the challenges present in conventional navigation systems such as the hand-eye coordination which has a higher learning curve. This makes it hard to use for most inexperienced users.

H4. 2D approaches are more mentally demanding, and more frustrating to use for the insertion tasks than 3D approaches.
Rationale: The hand-eye coordination issue and the mental mapping of the 3D aspects in the 2D approaches can make the navigation and localization of instruments and plannings difficult.

3.4 Experimental procedure

Fig. 6 is a flowchart representing the steps that participants went through during the experiment. When a volunteer comes in, a consent form with a definition of the task, statement of anonymization, and consent for data use needed to be signed. Subsequently, the volunteer's general information such as gender, age, background and familiarity with OST HMDs (Sect. 4.1) was collected. Subsequently, the volunteer performed a standard eye calibration [39] as provided by the HL2.

In the experiment, the AR conditions were randomly assigned to each volunteer following a Latin Square Design to prevent bias caused by the learning effects resulting from going over the different approaches [53]. The Latin Square was balanced for both order effects and carry-over (preceding and succeeding) effects [10, 31]. Each visualization condition started with its training session (Sect. 3.3.2); and at the end of the training, a reminder of the requested task was given orally allowing for the real experiment to begin. The participants had to perform five insertions for each visualization condition. After each confirmed (by the participant) alignment, the next planned trajectory was visualized. For each insertion, the participants had to start inserting the needle after an initial alignment (with the tip on the model). The participants were asked to inform the investigators when they were ready to insert the needle (i.e., after right orientation with the tip at the model surface), and when they were done with the insertion (i.e., tip at the right position inside the model). At the end of the five insertion tasks (for each AR condition), the participants had to fill two forms: SUS and NASA-TLX, and then the next AR approach was assigned to them until they have gone through all the conditions.

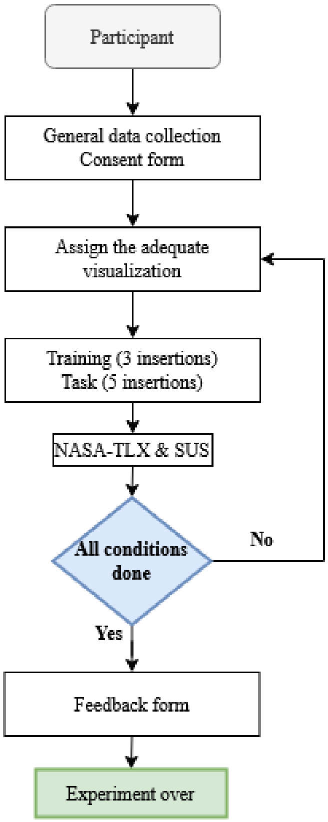


Fig. 6: Experiment steps for a given participant.

The participants were guided throughout all the experiment to ensure that they follow the protocol. They were allowed to use both hands, to rotate the skull if they see fit (for *Smart2*, *Align3* and *Off3*), and were permitted to move, rotate or scale the 2D window based on their needs (for *Win2*).

The final part consisted of filling a form requesting feedback on their experience, preference, advantages and disadvantages of using each of the AR approaches during the insertion tasks.

4 RESULTS

4.1 participants

An announcement was made among employees and students in our institution, after which 32 participant volunteered to be part of this study. The characteristics of the population are as follows:

- gender: 15 Female and 17 Male,
- age range: $[20-29]$ ($n = 23$), $[30-39]$ ($n = 5$), $[40-49]$ ($n = 3$) and over 50 years old ($n = 1$),
- background: 14 technical, 2 medical, 13 both technical and medical, 3 other,
- familiarity with OST HMDs: only 14 volunteer reported their familiarity with OST HMDs.

At the end of the experiments, 640 insertion tasks were performed (five insertions under four conditions performed by 32 participants).

Table 1: The results obtained under the various conditions presented as median [IQR]. (-): no statistical significance, (*): $p < 0.05$, (**): $p < 0.01$, (***): $p < 0.001$

Metrics	$\varepsilon_{P.e}$ (mm)	$\varepsilon_{P.x}$ (mm)	$\varepsilon_{P.m}$ (mm)	$\varepsilon_{\alpha.e}$ $(^\circ)$	$\varepsilon_{\alpha.x}$ $(^\circ)$	$\varepsilon_{\alpha.m}$ $(^\circ)$	tt (s)	SUS	NASA-LTX	MD	FR
Smart2	2.2 [4.7]	2.8 [5.3]	3.1 [5.2]	1.8 [3.4]	1.7 [2.8]	2.2 [2.5]	88 [153]	62 [20]	63 [25]	12 [6]	11 [7]
Win2	1.4 [1.7]	3.1 [4.1]	2.8 [3.1]	1.3 [1.7]	2.3 [2.9]	2.5 [2.2]	137 [258]	56 [17]	68 [18]	14 [5]	12 [7]
Off3	1.7 [1.8]	2.0 [2.6]	2.0 [1.9]	1.0 [1.6]	1.6 [1.6]	1.8 [1.2]	60 [97]	78 [18]	52 [27]	9 [8]	7 [6]
Align3	1.6 [2.4]	2.0 [2.5]	1.9 [2.5]	1.3 [2.1]	1.4 [1.7]	1.8 [1.4]	76 [125]	79 [18]	48 [22]	9 [8]	7 [7]
Smart2 vs. *Win2*	***	-	*	***	-	-	*	**	*	**	-
Smart2 vs. *Align3*	***	***	***	**	*	**	-	***	***	***	***
Smart2 vs. *Off3*	***	***	***	***	**	***	*	***	***	**	***
Win2 vs. *Align3*	*	***	**	-	***	***	**	***	***	***	***
Win2 vs. *Off3*	-	***	***	***	***	***	***	***	***	***	***
Align3 vs. *Off3*	-	-	-	***	-	-	-	-	-	-	-

4.2 Positional and angular error (H1)

Table 1 shows the median and interquartile range (IQR) of the subjective and objective metrics collected in this study, including the positional and angular error. The table shows a statistically significant smaller positional and angular error for the 3D approaches (Off3 and Align3) compared to the 2D approaches (Smart2 and Win2). However, there is no significant difference between Off3 and Align3 . This can be observed for the mean and end-point positional and angular error (shown in Fig. 7 and Fig. 8). For the entry-point, the positional and angular error for the smartphone (Smart2) were significantly higher than those for Win2, Off3 and Align3.

The standard deviation of the mean positional error and angular error shown in the bottom-right of Fig. 7 and Fig. 8 respectively, indicates a higher occurrence in changing the insertion angle and distance from planned axis for 2D approaches compared to 3D approaches.

4.3 Training time (H2)

Table 1 shows the time required for the participants to finish the training task. It shows that the participants were significantly faster in finishing the alignment using 3D approaches: 60 $[97]s$ and 76 $[125]s$ for Off3 and Align3 respectively, than when using Win2 (137 $[258]$ s). The training time for Smart2 (88 $[153]s$) is slightly higher than that of Align3 but not statistically significant. Furthermore, no statistically significant difference in training time was observed between Off3 and Align3.

4.4 Usability (H3)

Fig. 9 top-left shows the box plots of the SUS scores for the evaluated visualization approaches. The figure shows that Align3 achieved the highest SUS score followed closely by Off3 and then Smart2 and finally Win2. The figure also shows that the difference between 2D approaches and 3D approaches was statistically significant, however no significant difference was found within the 2D or 3D approaches. This indicates a higher usability for Align3 and Off3 compared to Smart2 and Win2. The latter two approaches also did not score above the SUS average score (68), where Win2 (56 $[17]$) was statistically significantly smaller than Smart2 (62 $[20]$).

4.5 Workload (H4)

The same order in the usability results (Sect. 4.4) was also reflected for NASA-TLX, where Align3 (48 $[22]$) ranked the lowest in terms of perceived workload and Win2 ranked the highest (68 $[18]$). According to Fig. 9, users experienced significantly higher frustration levels using 2D approaches compared to 3D approaches, with no significant difference between Smart2 and Win2 or Off3 and Align3. However, the use of 2D window (Win2) shows a statistically significant higher mental demand compared to the use of Smart2, Align3 and Off3.

4.6 Preference vs performance

In addition to filling the SUS and NASA-TLX forms, the participants ranked the four visualization approaches based on preference (which visualization they liked the most), and based on performance (on which visualization they thought they performed the best). Fig. 10 shows the number of times each visualization was ranked first by the participants in terms of preference and performance, and compares it to the number of times each visualization was actually ranked first based on the tip-to-tip distance error at the end of the insertion task ($\varepsilon_{P.x}$). Most participants (n=29), preferred the 3D visualization approaches Align3 (n=14) and Off3 (n=15) over the smartphone (Smart2) and the 2D window (Win2). It can also be seen that most participants felt that they performed better using the 3D approaches (n=28) and were correct about their predictions. The remaining participants (n=4) felt that they performed better using the smartphone and 2D window while they achieved better performance using either Align3 or Off3.

4.7 Participants' feedback

The participants' feedback on each of the visualization approaches was collected and summarized in the **supplementary material S1**.

Most participants agreed that the smartphone, Off3 and Align3 are easy to use (high occurrence of similar comments) which is not the case for the 2D window (Win2). The window visualization was perceived as a complicated approach that is hard to learn and use (25 remarks). The participants however, did not agree on in their preference for Off3 versus Align3 (i.e., liking and disliking the direct overlay).

5 DISCUSSION

In this within-subject study, we investigated the effect of AR visualization approach, 2D vs. 3D, on needle insertion tasks in the context of EVD and VS catheter placement. The study investigated the use of a smartphone, a 2D navigation window, a fully aligned and a rotationally aligned patient model seen through an OST HMD (HL2).

5.1 Positional and angular error

The results obtained for the positional and angular error support the hypothesis (**H1**). 3D approaches were more practical for the insertion task than the 2D approaches and allowed achieving smaller errors. The significantly lower $\varepsilon_{P.std}$, $\varepsilon_{P.m}$, $\varepsilon_{\alpha.m}$, and $\varepsilon_{\alpha.std}$ suggest that from the entry point till the end point for most participants, the needle did not get far from the insertion axis and therefore, less adaptations of the trajectory were required along the way. This is probably due to the consequences of an improved depth perception: the 3D approaches (Align3 and Off3) provide a higher visual feedback allowing the localization of the entry and exit points more accurately.

For further analysis, we decomposed the positional error at the insertion end-point to lateral and depth error (see Fig. 11). Lateral error was defined as the distance between the end-point of the planning trajectory and the intersection point of the longitudinal axis of the needle and a plane that is orthogonal to the longitudinal axis of the planned trajectory and passing by the planning end-point. The depth error was defined as the distance between the planning end-point and the intersection point of a plane orthogonal to the longitudinal axis of the planning trajectory and passing by the needle end-point. Fig. 12 shows the depth and lateral positional error for all evaluated approaches. From

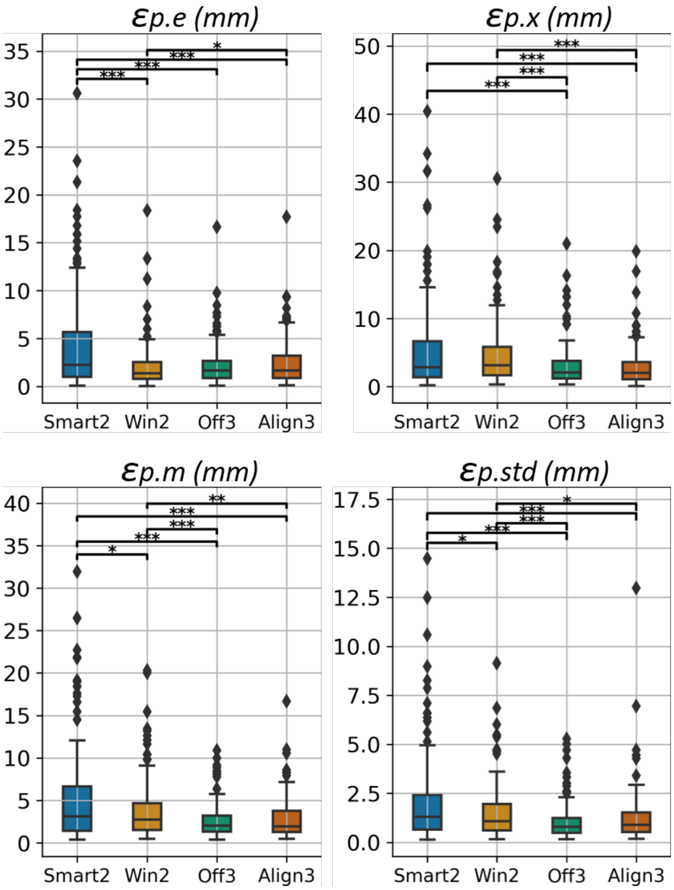


Fig. 7: Positional Error (*mm*): (top-left) at entry point, (top-right) at exit point, (bottom-left) mean over insertion, (bottom-right) std over insertion

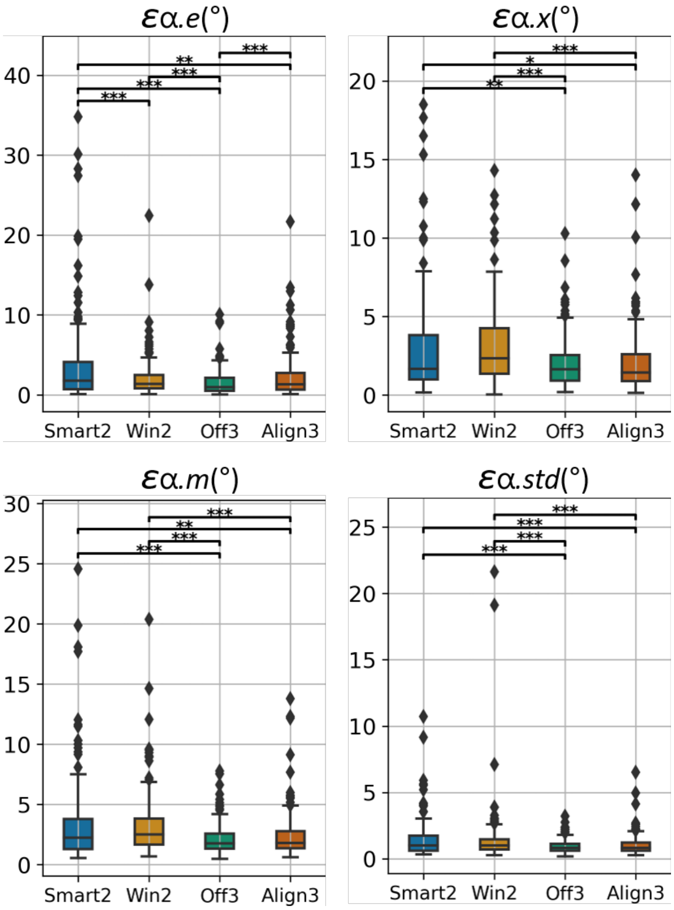


Fig. 8: Angular Error (°): (top-left) at entry point, (top-right) at exit point, (bottom-left) mean over all the insertion, (bottom-right) std over all the insertion

the figure, it can be observed that all approaches achieved very low depth error, median 0.5*mm*. This is probably due to the effectiveness of the virtual elements (VE) that were added to the augmentation of the needle and planned trajectory. Using these VE helped the users identify when the end-point of the needle arrives at the correct depth of the planned trajectory. However, They provided no help in ensuring the tip-to-tip lateral alignment, hence the high lateral error. Perhaps other forms of VE can be investigated in future studies to give more clues and guidance for tip-to-tip lateral alignment. Moreover, participants had limited needle adjustments inside the phantom due to the gelatin material inside, making it hard to make a lateral movement to correct for the lateral error.

5.2 Training time

The results from the training task show a shorter time-to-completion for the Off3 and Align3 visualizations. In contrast, Smart2 and Win2 show a higher time-to-completion with a higher IQR. The reason behind this increase from 3D to 2D, likely is the depth perception lacking in the 2D images. It is hard, as stated by most participants, to gain enough visual feedback to locate the needle and the planning accurately in 3D space using any of the 2D approaches (Win2 and Smart2), especially when the requirements to pass each alignment in the training session are 1*mm* and 1°. Moreover, it has been reported by the participants that the 2D window has multiple views which require an adaptation time to understand the hand movements and their implications on the visualization. This difficult interpretation, in combination with the switch of focus from one view to the other has significantly hardened the task.

5.3 Usability

The 3D approaches had a significantly higher usability compared to the 2D approaches which confirmed (**H3**). However, no statistically significant difference was observed between the rotationally aligned and fully aligned approaches. This could be the result of the very similar perceived visual feedback from the 3D aspects of the model and the insertion trajectory. Thus, the only difference is the user preference which is explained below (Sect. 5.5). On the other hand, 2D approaches (Smart2 and Win2) show significantly poor usability scores where the usage of the window (Win2) was the worst. Obviously, the lack of depth perception in the smartphone and the Window made it hard to locate the 3D objects in 3D space which required more adjustments to the needle position. Moreover, the smartphone can be intrusive between the operator and the operation site. It can limit the operator's hand movements, which was reported by multiple participants, it can provoke the continuous adaptation of the device's position if a flexible arm is used, and/or incite the need for an assistant operator when more actions are needed on the AR interface. Considerations about the usability are highlighted in Sect. 5.6.

5.4 Workload

The results from the NASA-TLX forms follow the same trend as the SUS scores. The 3D approaches seem to be less demanding for needle insertion confirming (**H4**): in particular, lower frustration levels and lower mental demand were observed. These approaches facilitated the viewing of the patient model and the insertion trajectories from different angles, allowing a faster and intuitive interaction and perception of the projected virtual elements. The participants on the other hand,

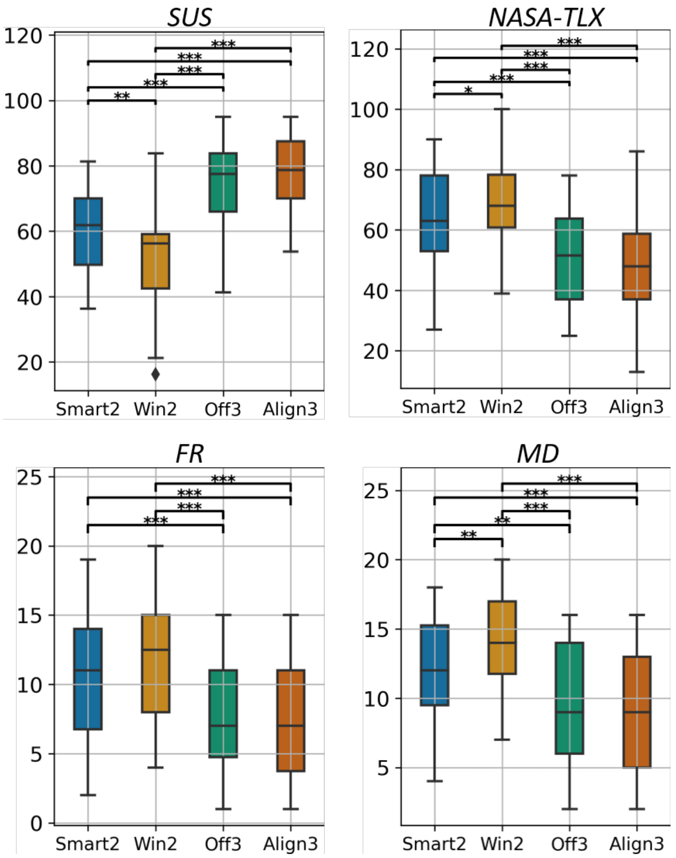


Fig. 9: Subjective measures: (FR) frustration level, (MD) mental demand

experienced higher frustration levels and mental demands when using the 2D approaches. This was reflected by the higher mean and standard deviation of both angular and positional errors. This imply that users were not confident in their orientation, that they had to adapt the needle frequently. This is due to the difficulty to perceive depth using only 2D views for both the smartphone (Smart2) and the window (Win2), and the simultaneous alignment on one view and misalignment in another view of the 2D window. These are two reasons why the task performed with 2D visualization is more mentally demanding, raising the time needed to locate insertion points (also reflected by the training time), and leading to an increased frustration levels.

5.5 Preference vs performance

The results shown in Fig. 10 indicated a higher preference for the 3D visualization approaches, with Off3 being preferred by slightly more participants (n=15) compared to Align3 (n=14). Both approaches were reported to be more intuitive and easy to use compared to the 2D approaches. Some participants favored Align3 due to the direct overlay of the virtual plan so that their attention is focused on the patient, while some others found the direct overlay sub-optimal as the physical and the virtual blend together making it harder to focus on each. The feedback of the participants regarding each of the visualization approaches is outlined in the next section. Furthermore, Fig. 10 reflects the confidence of the participants about their performance when using Align3 and Off3 to visualize the virtual planning, where the majority of them predicted that they performed the best using the 3D visualizations and were correct about their predictions. It is also noted that all participants performed the best using Align3 and Off3, while no participant performed best using Smart2 and Win2.

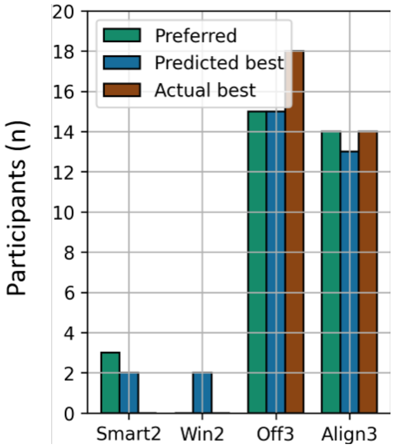


Fig. 10: Participants ranking of the four visualizations, in terms of preference (green), predicted best performance (blue), and actual best performance (brown)

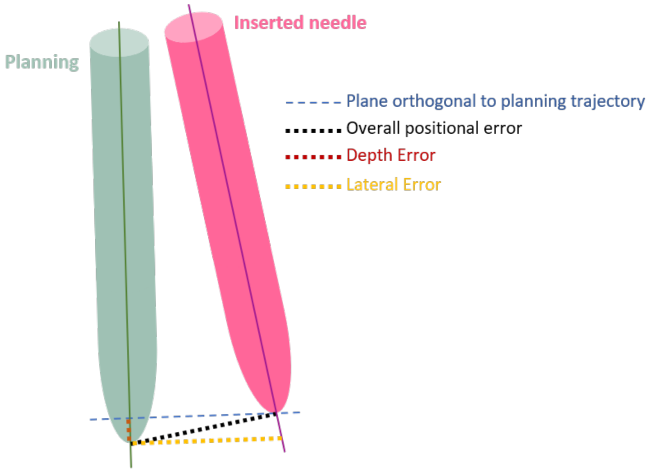


Fig. 11: The calculation of the depth and lateral positional errors for the inserted needle tip

5.6 Participants' feedback

The participants provided feedback on what they liked and disliked for every visualization approach (see table **supplementary material S1**).

Most complaints about the smartphone are related to the missing of the 3D information: only one viewpoint is available. Additionally, the small screen, the movement limitation, and the uncertainty of performance were highlighted issues. Therefore, the implementation of a smartphone AR needs to consider enough training time with the right visualization of the trajectories, which allows to see depth with minimal operator movement, the use of a joint-based arm holder to stabilize the device at the right poses. Moreover, because the phone is a physical object between the patient and the surgeon, a good balance between smartphone size and hands access to the surgical site needs to be found. This can be based on iterative testing with the main operators during a training phase.

The usage of the 2D window was very criticized by the participants. The main complain is that it is hard to interpret the 3D movements visible in the screen, and to translate them in the world space. It is expected that the mastery of this approach takes longer times. The FoV of the headset can be a limitation to such an approach: the full window would not be fully apparent to the user unless they move their head. It is therefore important to allow for the modification of the window's size and position for a constant visibility of the window, while minimizing the eyes' switch of focus between the interface and the surgical site.

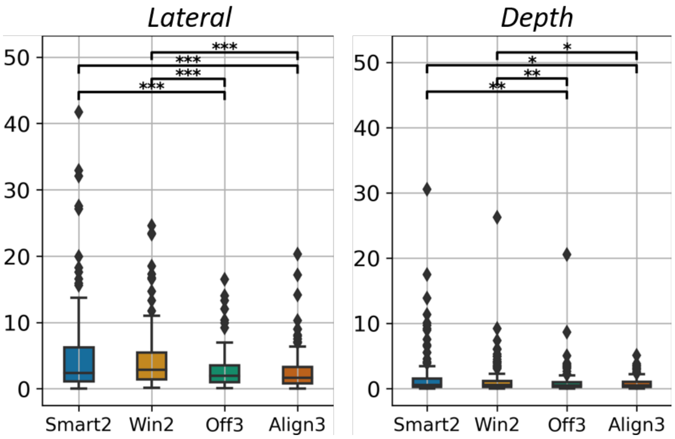


Fig. 12: Tip-to-tip positional error $\varepsilon_{P.x}$ decomposed into (left) lateral error, and (right) depth error in (*mm*)

Half of the participants preferred the rotationally aligned approach stating that they liked the separation between real and virtual objects, while some other participants complained about having to look away from the surgical site to perform the insertion. The field of view of the HMD can cut the visualization of the virtual model, which forces the participants to adjust their heads. Therefore the offset created needs to be adjustable according to user preference: depending on surgical site location with respect to the operator's position and height.

The reason for which the rotationally aligned approach was appreciated, was also the reason why the fully aligned approach was disliked. The participants found the overlap distracting and confusing. The confusion was the result of how participants were not able to differentiate between virtual and real objects, making it hard to focus on one or the other. For this, some participants experienced ”blurred images” feeling. This could mean that it might take time for some users to get familiar with the OST HMD even to perceive depth in details: some user suggested the use of another perspective to be certain of a good alignment which can be solved for example by the use of augmented mirrors [45, 46].

It is important to mention that not every participant enjoyed wearing a headset. An OST HMD can have negative effects on some users (e.g., blurred appearance). Therefore, in case of the need to choose between any of the approaches, it is important to investigate if the OST device provides enough contrast for the application, if the user can disregard the blending between the virtual and real world and consider it as one space without changing the eyes' focus, and if the 3D aspect is well perceived.

5.7 Limitations and future work

This user study compared four different AR visualization approaches in needle or catheter insertion tasks. For a fair comparison, the same visualization of the virtual model and the planning was used for all of the tested approaches. However, a few restrictions and or adaptations were needed for the different approaches. For example, the smartphone needed to be mounted on a smartphone-holder and was not allowed to be moved during the experiment. This was needed due to the requirement of maintaining the line-of-sight with the phantom and QR-code, as well as to keep the hands of the user free to perform the insertion task. This, in turn, limits the user from looking at the patient from different views and have a better depth perception. However, in this study we tried to mitigate this by allowing the user to rotate the phantom to get the best view, which may not be feasible with real patients.

For visualizing the virtual planning on a 2D window, there are different ways of displaying the planning model and navigation data. It is hard to design a navigation 2D window that works best for all applications. In this study, we used three orthogonal camera views of the virtual model, which is a common approach used in conventional navigation systems (usually with preoperative data). Other ways of presenting the planning in the 2D window can be used and therefore the performance is closely related to the design of the 2D window. However, in general, 2D window visualization shares the lack of depth perception and therefore further considerations needs to taken into account to compensate for that (see Sect. 5.6).

In this study, the participants went though a training session for each of the visualization approaches before the actual insertion task. Due to the time limitation of the experiment, the participants had to perform only three alignments of the needle with the virtual plan. Therefore, it is possible that some participants would have benefited from longer training session for some of the visualizations (especially Smart2 and Win2) and would have perhaps been able to perform better. However, the findings in this study also reflect the learning curve of the four visualization approaches, where users have shown to perform well with 3D AR visualizations, even when they are not familiar with them.

The training time *tt* in Table 1 shows that participants generally took longer to finish the training session for 2D approaches compared to 3D approaches. The IQR of the 2D approaches also reflect higher variance between the participants, which could be attributed to many factors, one of which is the visuospatial ability of the user. It can be interesting to see if this high variance in 2D approaches is affected by the visuospatial ability of the user, and whether the difference in visuospatial abilities among users has lesser effect in 3D approaches. A mental rotation test (MRT) [71] can possibly be used to investigate this.

Since the main objective of this study was to assess 2D and 3D AR visualization approaches in needle/catheter insertion tasks, such as EVD and VS, the targeting accuracy of the physical needle was not measured with respect to the patient anatomy. Only the alignment accuracy of the virtual overlay of the needle with respect to the planning was measured, and that is to focus on assessing the perception and depth estimation errors for the different visualization approaches (given that all other sources of errors such as calibration, registration and tracking are being equal for all approaches). The system's calibration, registration and tracking accuracies have been assessed in previous studies and with other applications [7, 8, 67]. For clinical validation of the system for catheter placement, all the sources of errors need to be taken into account and the targeting accuracy should be measured. The system needs to be validated with experienced neurosurgeons to measure the added value of using AR guidance over free hand catheter insertion in EVD and VS procedures, which is an objective for future work.

6 CONCLUSION

Studies comparing different AR approaches for needle insertion tasks are lacking. Subsequently, in this study, we conducted a user study with the aim to compare four different AR approaches, namely, smartphone-based AR, 2D window viewed in an OST device, and fully aligned and rotationally aligned 3D patient models. To that end, 32 users had to perform multiple guided needle insertions using the four AR approaches, which were then assessed based on alignment accuracy and users' preference and feedback.

In this study, the 3D approaches (Align3 and Off3) achieved better alignment accuracy compared to the 2D approaches (Smart2 and Win2). Hence, given an insertion trajectory visualized using these AR approaches, it is easier to achieve better performance under 3D approaches. Our study shows no statistically significant difference between Align3 and Off3 in terms of alignment accuracy, therefore it comes down to the operator's preference and the specific application for which to choose. In contrast, 2D approaches were shown to be more difficult to learn and less preferable, where the confidence on performance drops. Furthermore, this study provides considerations and recommendations on the evaluated approaches based on the users feedback, which we hope would help in improving the design of future AR systems for insertion tasks such as external ventricular drain or ventricular shunt placement.

REFERENCES

[1] https://developer.vuforia.com.
[2] https://www.blender.org.

[3] https://www.mevislab.de.

[4] A. AlAzri, K. Mok, J. Chankowsky, M. Mullah, and J. Marcoux. Placement accuracy of external ventricular drain when comparing freehand insertion to neuronavigation guidance in severe traumatic brain injury. *Acta neurochirurgica*, 159(8):1399–1411, 2017.

[5] E. Azimi, Z. Niu, M. Stiber, N. Greene, R. Liu, C. Molina, J. Huang, C. M. Huang, and P. Kazanzides. An Interactive Mixed Reality Platform for Bedside Surgical Procedures. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12263 LNCS, pp. 65–75. Springer Science and Business Media Deutschland GmbH, 2020. doi: 10.1007/978-3-030-59716-0_7

[6] G. Badiali, V. Ferrari, F. Cutolo, C. Freschi, D. Caramella, A. Bianchi, and C. Marchetti. Augmented reality as an aid in maxillofacial surgery: Validation of a wearable system allowing maxillary repositioning. *Journal of Cranio-Maxillofacial Surgery*, 42(8):1970–1976, dec 2014. doi: 10.1016/j.jcms.2014.09.001

[7] M. Benmahdjoub, W. J. Niessen, E. B. Wolvius, and T. Van Walsum. Virtual extensions improve perception-based instrument alignment using optical see-through devices. *IEEE Transactions on Visualization and Computer Graphics*, 27(11):4332–4341, nov 2021. doi: 10.1109/TVCG.2021.3106506

[8] M. Benmahdjoub, W. J. Niessen, E. B. Wolvius, and T. van Walsum. Multimodal markers for technology-independent integration of augmented reality devices and surgical navigation systems. *Virtual Reality*, 1:1–14, may 2022. doi: 10.1007/S10055-022-00653-3/FIGURES/16

[9] M. Benmahdjoub, T. van Walsum, P. van Twisk, and E. Wolvius. Augmented reality in craniomaxillofacial surgery : added value and proposed recommendations through a systematic review of the literature. *International Journal of Oral & Maxillofacial Surgery*, (November), 2020. doi: 10.1016/j.ijom.2020.11.015

[10] J. V. Bradley. Complete Counterbalancing of Immediate Sequential Effects in a Latin Square Design. *Journal of the American Statistical Association*, 53(282):525–528, 1958. doi: 10.1080/01621459.1958.10501456

[11] J. Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.

[12] I. M. Butaslac, Y. Fujimoto, T. Sawabe, M. Kanbara, and H. Kato. Systematic Review of Augmented Reality Training Systems. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–20, aug 2022. doi: 10.1109/TVCG.2022.3201120

[13] F. Y. Chai, F. Farizal, and T. Jegan. Coma due to malplaced external ventricular drain. *Turkish neurosurgery*, 23(4):561–563, 2013. doi: 10.5137/1019-5149.JTN.5724-12.1

[14] K. Cleary and T. M. Peters. Image-Guided Interventions: Technology Review and Clinical Applications. *http://dx.doi.org/10.1146/annurev-bioeng-070909-105249*, 12:119–142, jul 2010. doi: 10.1146/ANNUREV-BIOENG-070909-105249

[15] W. O. Contreras López, P. A. Navarro, and S. Crispin. Intraoperative clinical application of augmented reality in neurosurgery: A systematic review. *Clinical Neurology and Neurosurgery*, nov 2018. doi: 10.1016/J.CLINEURO.2018.11.018

[16] P. Debenham, G. Thomas, and J. Trout. Evolutionary augmented reality at the Natural History Museum. *2011 10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2011*, pp. 249–250, 2011. doi: 10.1109/ISMAR.2011.6092400

[17] W. Deng, F. Li, M. Wang, and Z. Song. Easy-to-Use Augmented Reality Neuronavigation Using a Wireless Tablet PC. *Stereotactic and Functional Neurosurgery*, 92(1):17–24, jan 2014. doi: 10.1159/000354816

[18] M. Doughty, N. R. Ghugre, and G. A. Wright. Augmenting performance: A systematic review of optical see-through head-mounted displays in surgery. *Journal of Imaging*, 8(7):203, 2022.

[19] D. García-Mato, R. Moreta-Martínez, M. García-Sevilla, S. Ochandiano, R. García-Leal, R. Pérez-Mañanes, J. A. Calvo-Haro, J. I. Salmerón, and J. Pascau. Augmented reality visualization for craniosynostosis surgery. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, 9(4):392–399, 2021. doi: 10.1080/21681163.2020.1834876

[20] K. Gavaghan, T. Oliveira-Santos, M. Peterhans, M. Reyes, H. Kim, S. Anderegg, and S. Weber. Evaluation of a portable image overlay projector for the visualisation of surgical navigation data: Phantom studies. *International Journal of Computer Assisted Radiology and Surgery*, 7(4):547–556, jul 2012. doi: 10.1007/s11548-011-0660-7

[21] A. Halbig, S. K. Babu, S. Gatter, M. E. Latoschik, K. Brukamp, and S. von Mammen. Opportunities and Challenges of Virtual Reality in Healthcare – A Domain Experts Inquiry. *Frontiers in Virtual Reality*, 0:14, mar 2022. doi: 10.3389/FRVIR.2022.837616

[22] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati, eds., *Human Mental Workload*, vol. 52 of *Advances in Psychology*, pp. 139–183. North-Holland, 1988. doi: 10.1016/S0166-4115(08)62386-9

[23] F. Heinrich, L. Schwenderling, F. Joeres, and C. Hansen. 2D versus 3D: A Comparison of Needle Navigation Concepts between Augmented Reality Display Devices. *Proceedings - 2022 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2022*, pp. 260–269, mar 2022. doi: 10.1109/VR51125.2022.00045

[24] F. Heinrich, L. Schwenderling, F. Joeres, K. Lawonn, and C. Hansen. Comparison of Augmented Reality Display Techniques to Support Medical Needle Insertion. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3568–3575, 2020. doi: 10.1109/TVCG.2020.3023637

[25] A. Hersh, S. Mahapatra, C. Weber-Levine, T. Awosika, J. N. Theodore, H. M. Zakaria, A. Liu, T. F. Witham, and N. Theodore. Augmented Reality in Spine Surgery: A Narrative Review. *HSS Journal*, 17(3):351, oct 2021. doi: 10.1177/15563316211028595

[26] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Am. A*, 4(4):629–642, Apr 1987. doi: 10.1364/JOSAA.4.000629

[27] F. Incekara, M. Smits, C. Dirven, and A. Vincent. Clinical Feasibility of a Wearable Mixed-Reality Device in Neurosurgery. *World Neurosurgery*, 118:e422–e427, oct 2018. doi: 10.1016/j.wneu.2018.06.208

[28] P. L. Ingrassia, G. Mormando, E. Giudici, F. Strada, F. Carfagna, F. Lamberti, and A. Bottino. Augmented Reality Learning Environment for Basic Life Support and Defibrillation Training: Usability Study. *Journal of Medical Internet Research*, 22(5), may 2020. doi: 10.2196/14910

[29] H. Iqbal, F. Tatti, and F. R. y Baena. Augmented reality in robotic assisted orthopaedic surgery: A pilot study. *Journal of Biomedical Informatics*, 120:103841, 2021.

[30] T. Jiang, M. Zhu, G. Chai, and Q. Li. Precision of a Novel Craniofacial Surgical Navigation System Based on Augmented Reality Using an Occlusal Splint as a Registration Strategy. *Scientific Reports*, 9(1):501, dec 2019. doi: 10.1038/s41598-018-36457-2

[31] A. Jones, J. E. Swan, G. Singh, and E. Kolstad. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. *Proceedings - IEEE Virtual Reality*, pp. 267–268, 2008. doi: 10.1109/VR.2008.4480794

[32] K. T. Kahle, A. V. Kulkarni, D. D. Limbrick Jr, and B. C. Warf. Hydrocephalus in children. *The lancet*, 387(10020):788–799, 2016.

[33] J. Kandasamy, C. Hayhurst, S. Clark, M. D. Jenkinson, P. Byrne, K. Karabatsou, and C. L. Mallucci. Electromagnetic stereotactic ventriculoperitoneal csf shunting for idiopathic intracranial hypertension: a successful step forward? *World neurosurgery*, 75(1):155–160, 2011.

[34] H. G. Kenngott, M. Wagner, F. Nickel, A. L. Wekerle, A. Preukschas, M. Apitz, T. Schulte, R. Rempel, P. Mietkowski, F. Wagner, A. Termer, and B. P. Müller-Stich. Computer-assisted abdominal surgery: new technologies. *Langenbeck's Archives of Surgery*, 400(3):273–281, apr 2015. doi: 10.1007/S00423-015-1289-8

[35] Y. Kolstee and W. Van Eck. The augmented Van Gogh's: Augmented reality experiences for museum visitors. *2011 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities, ISMAR-AMH 2011*, pp. 49–52, 2011. doi: 10.1109/ISMAR-AMH.2011.6093656

[36] D. Konik. https://www.mirage-app.com.

[37] H. B. Kwon, Y. S. Park, and J. S. Han. Augmented reality in dentistry: a current perspective, oct 2018. doi: 10.1080/00016357.2018.1441437

[38] C. R. Larsen, J. L. Soerensen, T. P. Grantcharov, T. Dalsgaard, L. Schouenborg, C. Ottosen, T. V. Schroeder, and B. S. Ottesen. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. 338, 2009. doi: 10.1136/bmj.b1802

[39] J. R. Lewis, Y. Wei, R. L. Crocco, B. I. Vaught, K. S. Perez, and A. A.-A. Kipman. Aligning inter-pupillary distance in a near-eye display system, 2011.

[40] Y. Li, X. Chen, N. Wang, W. Zhang, D. Li, L. Zhang, X. Qu, W. Cheng, Y. Xu, W. Chen, and Q. Yang. A wearable mixed-reality holographic computer for guiding external ventricular drain insertion at the bedside. *J Neurosurg*, p. 1, 2018. doi: 10.3171/2018.4.JNS18124

[41] C. Lin, D. Andersen, V. Popescu, E. Rojas-Munoz, M. E. Cabrera, B. Mullis, B. Zarzaur, K. Anderson, S. Marley, and J. Wachs. A First-

Person Mentee Second-Person Mentor AR Interface for Surgical Telementoring. *Adjunct Proceedings - 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2018*, pp. 3–8, jul 2018. doi: 10.1109/ISMAR-ADJUNCT.2018.00021

[42] K. Liu, Y. Gao, A. Abdelrehem, L. Zhang, X. Chen, L. Xie, and X. Wang. Augmented reality navigation method for recontouring surgery of craniofacial fibrous dysplasia. *Scientific Reports 2021 11:1*, 11(1):1–7, may 2021. doi: 10.1038/s41598-021-88860-x

[43] D. J. Long, M. Li, Q. M. De Ruiter, R. Hecht, X. Li, N. Varble, M. Blain, M. T. Kassin, K. V. Sharma, S. Sarin, V. P. Krishnasamy, W. F. Pritchard, J. W. Karanian, B. J. Wood, and S. Xu. Comparison of Smartphone Augmented Reality, Smartglasses Augmented Reality, and 3D CBCT-guided Fluoroscopy Navigation for Percutaneous Needle Insertion: A Phantom Study. *CardioVascular and Interventional Radiology*, 44(5):774–781, may 2021. doi: 10.1007/s00270-020-02760-7

[44] M. Mahan, R. F. Spetzler, and P. Nakaji. Electromagnetic stereotactic navigation for external ventricular drain placement in the intensive care unit. *Journal of Clinical Neuroscience*, 20(12):1718–1722, 2013.

[45] A. Martin-Gomez, J. Fotouhi, U. Eck, and N. Navab. Gain A New Perspective: Towards Exploring Multi-View Alignment in Mixed Reality. In *Proceedings - 2020 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2020*, pp. 207–216. Institute of Electrical and Electronics Engineers Inc., nov 2020. doi: 10.1109/ISMAR50242.2020.00044

[46] A. Martin-Gomez, A. Winkler, K. Yu, D. Roth, U. Eck, and N. Navab. Augmented Mirrors. In *Proceedings - 2020 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2020*, pp. 217–226. Institute of Electrical and Electronics Engineers Inc., nov 2020. doi: 10.1109/ISMAR50242.2020.00045

[47] A. Meola, F. Cutolo, M. Carbone, F. Cagnazzo, M. Ferrari, and V. Ferrari. Augmented reality in neurosurgery: a systematic review. *Neurosurgical review*, 40(4):537, oct 2017. doi: 10.1007/S10143-016-0732-9

[48] K. Meyerbröker and N. Morina. The use of virtual reality in assessment and treatment of anxiety and related disorders. *Clinical Psychology & Psychotherapy*, 28(3):466, may 2021. doi: 10.1002/CPP.2623

[49] M. C. Palumbo, S. Saitta, M. Schiariti, M. C. Sbarra, E. Turconi, G. Raccuia, J. Fu, V. Dallolio, P. Ferroli, E. Votta, E. De Momi, and A. Redaelli. Mixed Reality and Deep Learning for External Ventricular Drainage Placement: A Fast and Automatic Workflow for Emergency Treatments. pp. 147–156, 2022. doi: 10.1007/978-3-031-16449-1_15

[50] B. J. Park, S. J. Hunt, G. J. Nadolski, and T. P. Gade. Augmented reality improves procedural efficiency and reduces radiation dose for CT-guided lesion targeting: a phantom study using HoloLens 2. *Scientific Reports 2020 10:1*, 10(1):1–8, oct 2020. doi: 10.1038/s41598-020-75676-4

[51] E. Peillard, F. Argelaguet, J. M. Normand, A. Lecuyer, and G. Moreau. Studying exocentric distance perception in optical see-through augmented reality. In *Proceedings - 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2019*, pp. 115–122. Institute of Electrical and Electronics Engineers Inc., oct 2019. doi: 10.1109/ISMAR.2019.00-13

[52] G. Pellegrino, C. Mangano, R. Mangano, A. Ferri, V. Taraschi, and C. Marchetti. Augmented reality for dental implantology: A pilot clinical report of two cases. *BMC Oral Health*, 19(1), jul 2019. doi: 10.1186/s12903-019-0853-y

[53] D. A. Preece. *Latin Squares, Latin Cubes, Latin Rectangles*. American Cancer Society, 2014. doi: 10.1002/9781118445112.stat00867

[54] L. Qian, A. Barthel, A. Johnson, G. Osgood, P. Kazanzides, N. Navab, and B. Fuerst. Comparison of optical see-through head-mounted displays for surgical interventions with object-anchored 2D-display. *International journal of computer assisted radiology and surgery*, 12(6):901, jun 2017. doi: 10.1007/S11548-017-1564-Y

[55] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147:103778, 2020. doi: 10.1016/j.compedu.2019.103778

[56] R. Radkowski, J. Herrema, and J. Oliver. Augmented Reality-Based Manual Assembly Support With Visual Features for Different Degrees of Difficulty. *International Journal of Human-Computer Interaction*, 31(5):337–349, may 2015. doi: 10.1080/10447318.2014.994194

[57] E. Rae, A. Lasso, M. S. Holden, E. Morin, R. Levy, and G. Fichtinger. Neurosurgical burr hole placement using the Microsoft HoloLens. In *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*, p. 20. SPIE, mar 2018. doi: 10.1117/12.2293680

[58] A. V. Reinschluessel, T. Muender, D. Salzmann, T. Döring, R. Malaka, and D. Weyhe. Virtual Reality for Surgical Planning – Evaluation Based on Two Liver Tumor Resections. *Frontiers in Surgery*, 9:110, feb 2022. doi: 10.3389/FSURG.2022.821060/XML/NLM

[59] G. Scavo, F. Wild, and P. Scott. The GhostHands UX: telementoring with hands-on augmented reality instruction. *Ambient Intelligence and Smart Environments*, 19:236–243, 2015.

[60] T. Sielhorst, M. Feuerstein, and N. Navab. Advanced medical displays: A literature review of augmented reality. *IEEE/OSA Journal of Display Technology*, 4(4):451–467, dec 2008. doi: 10.1109/JDT.2008.2001575

[61] O. Suess, T. Kombos, R. Kurth, S. Suess, S. Mularski, S. Hammersen, and M. Brock. Intracranial Image-Guided Neurosurgery: Experience with a new Electromagnetic Navigation System. *Acta Neurochirurgica 2001 143:9*, 143(9):927–934, 2001. doi: 10.1007/S007010170023

[62] X. Sun, S. B. Murthi, G. Schwartzbauer, and A. Varshney. High-Precision 5DoF Tracking and Visualization of Catheter Placement in EVD of the Brain Using AR. *ACM Transactions on Computing for Healthcare*, 1(2), mar 2020. doi: 10.1145/3365678

[63] H. Sveistrup. Motor rehabilitation using virtual reality. *Journal of NeuroEngineering and Rehabilitation 2004 1:1*, 1(1):1–8, dec 2004. doi: 10.1186/1743-0003-1-10

[64] J. E. Swan, G. Singh, and S. R. Ellis. Matching and Reaching Depth Judgments with Real and Augmented Reality Targets. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1289–1298, nov 2015. doi: 10.1109/TVCG.2015.2459895

[65] L. B. Tabrizi and M. Mahvash. Augmented reality–guided neurosurgery: accuracy and intraoperative application of an image projection technique. *Journal of Neurosurgery*, 123(1):206–211, jul 2015. doi: 10.3171/2014.9.JNS141001

[66] Z. N. Tang, L. H. Hu, H. Y. Soh, Y. Yu, W. B. Zhang, and X. Peng. Accuracy of Mixed Reality Combined With Surgical Navigation Assisted Oral and Maxillofacial Tumor Resection. *Frontiers in Oncology*, 11:5755, jan 2022. doi: 10.3389/fonc.2021.715484

[67] A. Thabit, M. Benmahdjoub, M.-L. C. Van Veelen, ·. Wiro, J. Niessen, E. B. Wolvius, and T. Van Walsum. Augmented reality navigation for minimally invasive craniosynostosis surgery: a phantom study. *International Journal of Computer Assisted Radiology and Surgery 2022*, pp. 1–8, may 2022. doi: 10.1007/S11548-022-02634-Y

[68] H. M. Tully and W. B. Dobyns. Infantile hydrocephalus: a review of epidemiology, classification and causes. *European journal of medical genetics*, 57(8):359–368, 2014.

[69] P. C. van de Woestijne, W. Bakhuis, A. H. Sadeghi, J. J. Peek, Y. JHJ Taverne, and A. JJC Bogers. 3D Virtual Reality Imaging of Major Aortopulmonary Collateral Arteries: A Novel Diagnostic Modality. doi: 10.1177/21501351211045064

[70] F. Van Gestel, T. Frantz, C. Vannerom, A. Verhellen, A. G. Gallagher, S. A. Elprama, A. Jacobs, R. Buyl, M. Bruneau, B. Jansen, J. Vandemeulebroucke, T. Scheerlinck, and J. Duerinck. The effect of augmented reality on the accuracy and learning curve of external ventricular drain placement. *Neurosurgical Focus*, 51(2):E8, aug 2021. doi: 10.3171/2021.5.FOCUS21215

[71] S. G. Vandenberg and A. R. Kuse. Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*, 47(2):599–604, 1978.

[72] S. Webel, U. Bockholt, T. Engelke, M. Peveri, M. Olbrich, and C. Preusche. Augmented Reality Training for Assembly and Maintenance Skills. *BIO Web of Conferences*, 1(4):00097, dec 2011. doi: 10.1051/bioconf/20110100097

[73] R. Woll, T. Damerau, K. Wrasse, and R. Stark. Augmented reality in a serious game for manual assembly processes. *2011 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities, ISMAR-AMH 2011*, pp. 37–39, 2011. doi: 10.1109/ISMAR-AMH.2011.6093654

[74] B. Xie, H. Liu, R. Alghofaili, Y. Zhang, Y. Jiang, F. D. Lobo, C. Li, W. Li, H. Huang, M. Akdere, C. Mousas, and L.-F. Yu. A Review on Virtual Reality Skill Training Applications. *Frontiers in Virtual Reality*, 0:49, apr 2021. doi: 10.3389/FRVIR.2021.645153

[75] J. Yasuda, T. Okamoto, S. Onda, Y. Futagawa, K. Yanaga, N. Suzuki, and A. Hattori. Novel navigation system by augmented reality technology using a tablet PC for hepatobiliary and pancreatic surgery. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 14(5):e1921, oct 2018. doi: 10.1002/RCS.1921

[76] M. J. Zinser, R. A. Mischkowski, T. Dreiseidler, O. C. Thamm, D. Rothamel, and J. E. Zöller. Computer-assisted orthognathic surgery: waferless maxillary positioning, versatility, and accuracy of an image-guided visualisation display. *British Journal of Oral and Maxillofacial*

Surgery, 51(8):827–833, dec 2013. doi: 10.1016/J.BJOMS.2013.06.014