virtual worlds

MDPI

Article

An Initial Qualitative Investigation of Augmented Reality Interactions with AR-Novice Radiologists

Jacob Hobbs * and Christopher Bull

Open Lab, School of Computing, Newcastle University, Newcastle upon Tyne NE4 5TG, UK; christopher.bull@newcastle.ac.uk
* Correspondence: j.hobbs11@ncl.ac.uk

Abstract

The need for efficiency in radiology has been made clear and augmented reality (AR) has the potential to address this by aiding with issues around irregular lighting conditions impacting the accuracy of reports, while also speeding up the process with improved interactions. However, some interaction questions are still very much open, and interactions for radiological applications deserve more attention. This short-term exposure study involved a series of seven interviews with practising radiologists and radiology registrars, most of whom were AR-novices. Participants engaged in AR tasks using the Microsoft HoloLens 2 and Meta Quest 3, running software provided by GigXR, as a basis for rich dialogue around the interaction value of AR for radiological tasks. Reflexive thematic analysis was used to analyse the resulting data, generating five themes, with the NASA-TLX providing additional context. The themes were mapped to three design considerations that stand as the contribution of this work and highlight how radiological AR requires interaction design distinguishing intended from unintended movement, standardised hardware requirements for clinical-grade HMDs, and versatile, workflow-integrated deployment to overcome resistance to clinical adoption. These are embodied in the speculation of an AR-First world that illustrates key requirements.

Keywords: augmented reality; mixed reality; radiology; interaction

Academic Editor: Daniel R. Mestre

Received: 9 July 2026
Revised: 25 August 2026
Accepted: 31 August 2026
Published: 9 September 2026

1. Introduction

Augmented reality (AR) is a maturing technology with a long history and a range of proposed applications in medical environments, including radiology. Within radiological practice, the literature suggests that AR has the potential to improve efficiency and diagnostic accuracy while reducing cognitive load, therefore enhancing overall performance and throughput. This may contribute to addressing the workflow capacity crisis identified by the Royal College of Radiologists [1]. These potential benefits are likely rooted in the interactions that AR enables. However, AR– interaction design remains an evolving area of research [2]. Fundamental tasks such as text input continue to have multiple competing solutions, and the relative advantages of hand gestures compared to alternative input modalities remain under discussion. This lack of consensus is reflected in commercially available head-mounted displays (HMDs), where devices such as the Meta Quest series incorporate controller-based input, while devices including the Microsoft HoloLens and Apple Vision Pro rely on hand gesture-based interaction. Despite the advantages described in the literature, AR has seen little adoption in day-to-day radiological practice, and there is currently no clear agreement on which interaction techniques best support clinical users.

In this study, a series of interviews were conducted with radiology consultants and registrars. Participants engaged directly with AR technology and took part in structured discussions exploring the potential impact of AR interactions on clinical radiology work. The aim was to investigate how AR interactions might support improved efficiency and comprehension in radiological tasks. This study is a short-term, single-exposure investigation involving participants with largely no prior AR experience. Findings should therefore be read as preliminary relative to what sustained, post-training use in real clinical practice might show. To facilitate this, both the Microsoft HoloLens 2 and Meta Quest 3 were used in conjunction with software provided by GigXR [3], with participants completing a series of guided activities. A think-aloud protocol was employed to capture participants' experiences and perspectives, enabling the identification of both opportunities and challenges associated with AR in this context [4]. Sessions were audio- and video-recorded, and the resulting data were analysed using reflexive thematic analysis to build themes across the dataset [5,6]. In addition to this qualitative analysis, the NASA Task Load Index (NASA-TLX) was used to provide an indication of perceived task difficulty when using AR systems [7,8].

The reflexive thematic analysis resulted in five core themes, each representing a distinct aspect of participant experience during the interviews. These themes were subsequently mapped to a set of three design considerations, which constitute the primary contribution of this study. These design considerations cover application-specific interaction design, validation and implementation of visualisation hardware requirements, and multi-application integration. This work aims to align future AR interaction design with the needs and priorities of practising radiologists, offering guidance on how such systems may support clinical workflows. In doing so, it contributes to understanding how AR might form part of a broader solution to the workflow capacity challenges identified by the Royal College of Radiologists [9]. To address this aim, the following research questions are posed.

1. How could the interaction advantages of AR afford the opportunity for increased efficiency or comprehension for radiologists?
    1.1 Could AR interactions contribute to greater efficiency in a radiological workflow?
    1.2 Could AR interactions contribute to greater comprehension a radiological workflow? What would be needed to support this?
2. What challenges or opportunities do AR interactions provide radiologists over current systems or capability?

2. Background

AR is not a new concept, but this emerging technology is being explored in many different contexts and is striving to find its place in the world [10–12]. A resurgence in interest around XR technologies has been ongoing since around 2016 with the release of the Microsoft HoloLens, along with the continued development of hardware, and now with a variety of HMDs commercially available, applications are spreading into a breadth of operational areas [13–15]. The potential of AR spans the breadth of healthcare specialties with applications suggested across fields such as surgical guidance, post-operative care, medical training, and radiology [16–18]. However, few AR applications have made it into regular use within tertiary care (highly specialised care), as a number of significant issues remain unsolved [19,20].

Whilst definitions of AR and mixed reality (MR) vary, it is generally agreed that AR superimposes digital objects into the users' view in real-time using a headset or other device [21–23]. The aim is to add virtual components to the user's field of view in order to provide them with additional information while carrying out a task. A widely accepted approach to defining AR and where it sits relative to virtual reality (VR) and MR was laid out by Milgram and Kishino [24] in 1994. They put forward the Reality–Virtuality Continuum, which has since been iterated on by [25]. On this continuum, augmented reality sits towards the 'real' end, with mixed reality describing everything between fully 'real' and fully 'virtual'. The work described in this paper is focused on AR but at times broadens references to XR technologies where tying to AR specifically is inappropriate.

There is extensive dialogue across the literature regarding AR, and more broadly XR, in diagnostic radiology [26], with Douglas et al. [26] amongst those suggesting that using AR could improve diagnostic accuracy and speed up the diagnostic process when viewing cross-sectional images such as CT and MRI. Pelargos et al. [27] state that "Surgical planning is inherently a 3D task" and that VR and AR technologies could help by improving the understanding of the complex anatomical relationships. These tools have the potential to offer better visualisation of areas of interest and therefore improve the understanding and the speed at which decisions can be made [28,29]. Trestioreanu et al. [30] argue that AR and VR have the potential to improve radiology healthcare by improving the cognitive experience, by reducing the cognitive load that a clinician undergoes when viewing 2D slices of 3D anatomy. They go on to suggest that while a few 3D visualisation methods currently exist, they do not offer the increased practicality or ergonomics that AR approaches could offer.

The need for efficiency in radiological workflows is clear and outlined by the Royal College of Radiologists [1]. Modern imaging technology has made cross-sectional scanning cheaper and quicker to acquire which in turn has resulted in a huge uplift in the number of scans being taken. A defensive stance in medicine meaning scans are requested to avoid doubt has also contributed to the uplift. There are now more scans being taken than can be reported, and, as such, wait times for results are extending [1,9]. Technology, and XR technology in particular, could offer some relief [17].

Across the literature the attitudes towards AR, and XR generally, in diagnostics and surgical planning is that these are inherently 3D tasks but they are performed in 2D. Pelargos et al. [27] state that *"Surgical planning is inherently a 3D task"*, and Trevia and Pinho [31] note It is remarkable that medical images which are inherently 3D in nature are mostly visualized in clinical practice by physicians and radiology technicians in reading rooms, using multiple 2D displays among 1D or 2D input media" [32–34]. This is the expected value that XR could bring, and it extends to the potential cited for improved interactions with XR over current methods by allowing users to "fly through" 3D models using an input device such as a joystick. It has been suggested that XR could improve accuracy due firstly to this improved visualisation, but also by addressing the irregular ambient lighting conditions in the rooms where images are reported having a negative effect [17,35]. Poor ergonomics and room conditions, such as ambient lighting, in radiologists' offices is a key issue stated in the literature that XR has the potential to solve with positive early work [36]. This technology could reduce costs due to the redundancy of the specialist monitors currently required, and is a quicker and more flexible alternative to 3D printing which has been suggested in areas like surgical planning [37]. The proposed impact of this improved interaction, negated ambient lighting issues, and cheaper equipment for diagnostics is fewer misdiagnoses, increased accuracy, and lower cost [17,31].

There is a model of interaction that "sees interaction as a cycle of communication acts channelled through input/output from the machine perspective, or perception/action from the human perspective", this dialogue metaphor is helpful to consider in this context. Modern AR technology has reached a point where a selection of commercially available HMDs are on offer and despite these HMDs offering advanced capability they are still restricted, with some users still suffering motion sickness, headaches, fatigue and other physical symptoms [38,39]. The vergence–accommodation conflict (VAC) (caused by the eye's two mechanisms of focusing competing against one another) and focal rivalry are common perceptual issues that arise from inaccurately presented content and have been documented to cause physical symptoms such as motion sickness, nausea, and visual fatigue [40,41].

Research has investigated individual perceptual cues (such as occlusion and binocular disparity (stereopsis), motion parallax, and relative size) with some positive results, for example, with VR, reduced depth errors have been shown when true occlusion is achieved [42]. However, satisfying multiple cues remains difficult and it is generally accepted that modern HMDs cannot faithfully represent all the cues that enable a user to interpret their environment [40,43].

Considering the other side of the interaction dialogue, AR interactions can be supported by modalities such as speech, gaze, and hand gestures as well as hardware such as controllers or gloves [44,45]. The literature describing interactions for manipulating objects in virtual environments is contradictory, suggesting both direct manipulation of virtual objects with a hand and indirect manipulation with, for example, a controller have value [46,47]. There has been little synthesis, and research continues to explore interactions with virtual objects aiming for intuitive and immersive movements [48]. Some interaction questions are still very much open, with efficient text input in virtual environments a key example [49]. With many interaction modalities and hardware options available and some interaction questions unanswered, it can be argued there has not been any suggestion to match the versatility and accuracy of a keyboard and mouse.

This work engages with radiologists to understand potential positive and negative impacts of AR interactions on their workflows with a view for efficiency and cognitive load, and it makes suggestions for any ability that could be afforded by AR for these users.

3. Methodology

The interviews described here were run to investigate two lines of research as two studies, one of which is detailed here, while the companion study is currently under review and is available as a preprint [50]. One series of interviews was used for both studies because of the overlap in potential participants and the similarity of the activities that could be used in the interview. Aspects specific to each study as well as broader questions that applied to both were asked throughout the interviews. This yielded raw data relevant to the research questions of each study that could then be interpreted and analysed separately. Participants gave informed consent for the interviews, which covered both this study and the companion study (Hobbs and Bull [50]), and were informed that they could withdraw at any time. This work was granted ethical approval by the Newcastle University ethics committee, ref: 61692/2023.

For this study, one-on-one interactive interviews were conducted with the identified stakeholders, and the participants were involved using two different AR HMDs and software provided by GigXR to have detailed conversation about the interactions and their experience with the technology. The think-aloud approach was used to put the participants first, and data were collected through video and audio sources. These data were then analysed using reflexive thematic analysis as described below [3–5].

3.1. Recruitment Process and Participants

For this work, seven radiologists, four registrars, and three consultants from North Eastern NHS trusts were recruited. Six of these participants were male, which illustrates the issues with recruiting women in a male-dominated field such as this, with women making up less than 30% of the radiological workforce [51,52]. This is discussed in more depth in the Limitations Section (Section 5.4). The consultants were associated with a variety of sub-specialties, and while registrars are general, their interests were varied. Five participants sat within the 25–34 age bracket while two consultants were in the next two brackets. Participant demographic details and summarised in Table 1. Participants were asked to rate their prior experience with XR from one to five with the labels shown in Table 2. Six participants listed their experience as one or two out of five with the final participant rating their experience as five, owning a Meta Quest 3. A convenience sampling approach was employed as is appropriate in a qualitative HCI study with a demographically and geographically local sample universe [53]. This restricts the generalisability of this study to the local level and we make no decontextualised abstract claims as is discussed further in the Limitations Section (Section 5.4). Two of the participants participated in a previous study by Hobbs et al. [54] as part of the same thread of research. The remaining five participants recruited were through snowballing from participants' contacts and by reaching out directly via the Newcastle University Medical School. Upwards of 100 total potential participants were reached either through direct contact or advertisement of the study via mailing lists. To be eligible for this study, participants had to be currently practising as a clinical radiologist or radiology registrar within the NHS. Interviews were conducted between 4 July and 18 September 2025, with interviews lasting from 90 min to two hours. While Braun and Clarke [55] argue saturation as a concept is at odds with reflexive thematic analysis, our participants possessed high information power characterised by their relevant experience, specialised knowledge, and direct engagement with the research topic, and as such a smaller sample size yielded substantial depth [56].

Table 1. Workshop study's participant demographic information.

ID	Age Range	Role	Experience in Field	AR Experience
A	25–34	Consultant Radiologist	5–15 years	1
B	25–34	Radiology Registrar	<5 years	5
C	25–34	Radiology Registrar	<5 years	2
D	25–34	Radiology Registrar	<5 years	2
E	25–34	Radiology Registrar	<5 years	2
F	35–44	Consultant Radiologist	>15 years	1
G	45–54	Consultant Radiologist	>15 years	2

Table 2. AR experience ratings.

Experience Rating	Description
1	Heard of it but never used it
2	Experienced it once or twice; e.g., a demo
3	Occasional user
4	Frequent user
5	Own an AR or VR device

3.2. The Interviews

These interviews were designed to enable us to present the participants with a variety of different situations in AR in a short space of time, closely enough aligned with their current workflows and goals that meaningful discussions could be made around the opportunities and challenges these AR tools provided. Most participants had very little experience of AR and the first objective of the workshop was to demonstrate AR to the participants in a way representative of a longer more extensive AR trial. The aim here was to allow the discussions had with the participants to go beyond the initial learning curve associated with using AR, and indeed any new technology, and on to a more realistic, holistic view of the role it could play. These seven interviews were run one-on-one with the lead researcher and a participant.

The aim of these interviews was to qualitatively investigate the impact of AR interactions for the integration into radiological reporting, with the research question of how could the interaction advantages of AR afford the opportunity for increased efficiency or comprehension for radiologists? Various forms of data were collected through the interviews, principally the point of view (POV) video recording of what the participant was seeing and the audio recording of the conversation. The two HMDs were set up to cast this POV to a Windows PC. A Microsoft Teams call was set up and the screen sharing and meeting recording functions were used to achieve this. The participant was seated at one end of a desk in front of a PC but space was deliberately provided along the desk to the left and around the room further to the left and behind them, should they have wanted to use the space, as shown in Figure 1. A GoPro was positioned with the participant in frame to record how the participant used the wider space and any bigger actions that would not be caught on the POV footage. The researcher was seated at a second desk to the right of the participant with a second monitor, keyboard, and mouse to the PC, with the cast POV from the HMD displayed throughout the workshop. Figure 1 shows the experimental set up from the view of the GoPro camera, and Figure 2 shows a floor plan diagram of the set up. The researcher made some additional observational notes throughout the interviews where necessary, generally to guide the analysis and remember key moments in interviews rather than to add another form of data to collect. Finally the NASA-TLX was used to collect initial indicative quantitative data about how difficult the participants found the tasks. The NASA-TLX is a subjective assessment designed to estimate the mental workload of a participant during a task [8].

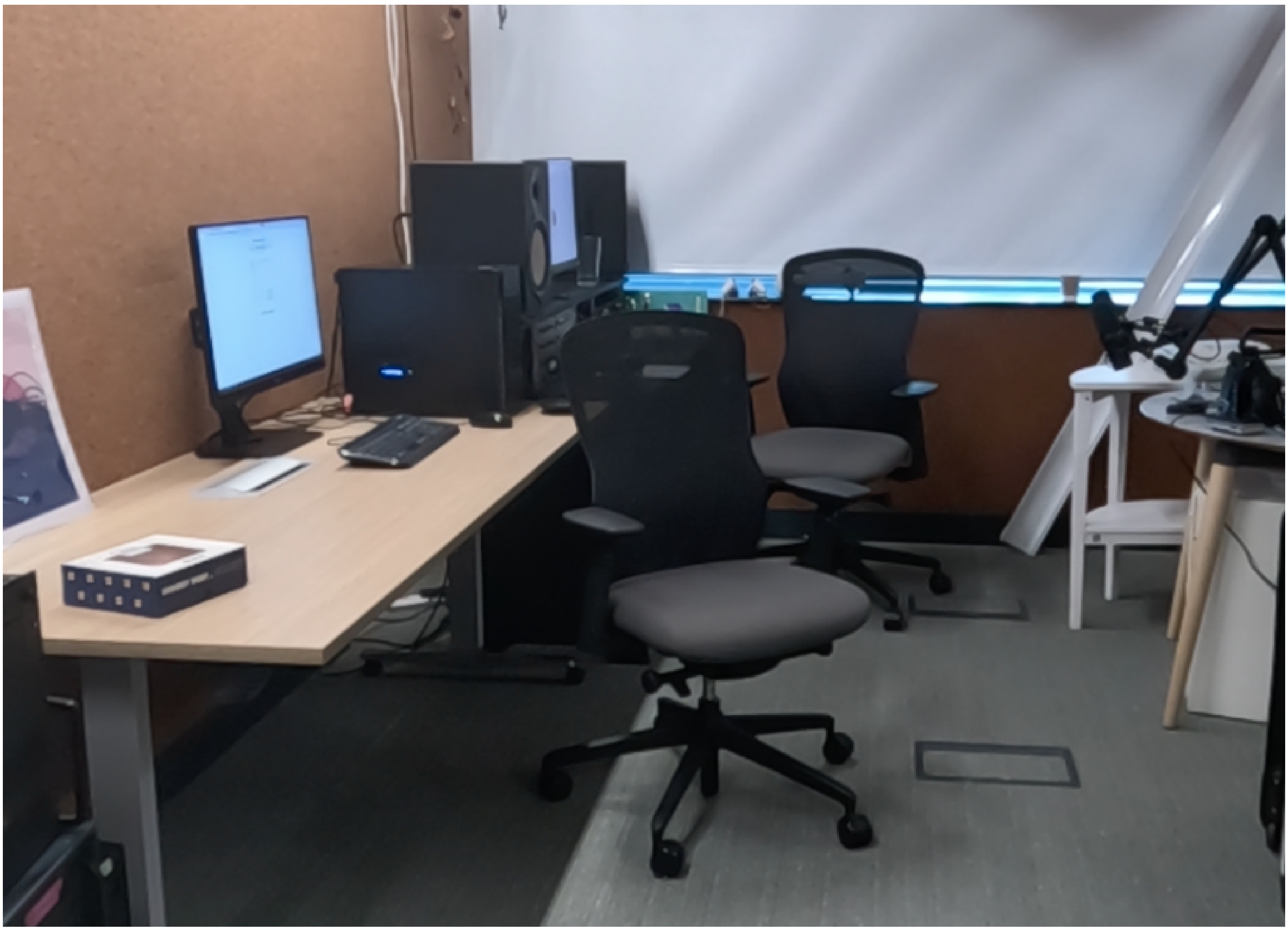

Figure 1. Workshop setup. Participant closest to camera; researcher to their right.

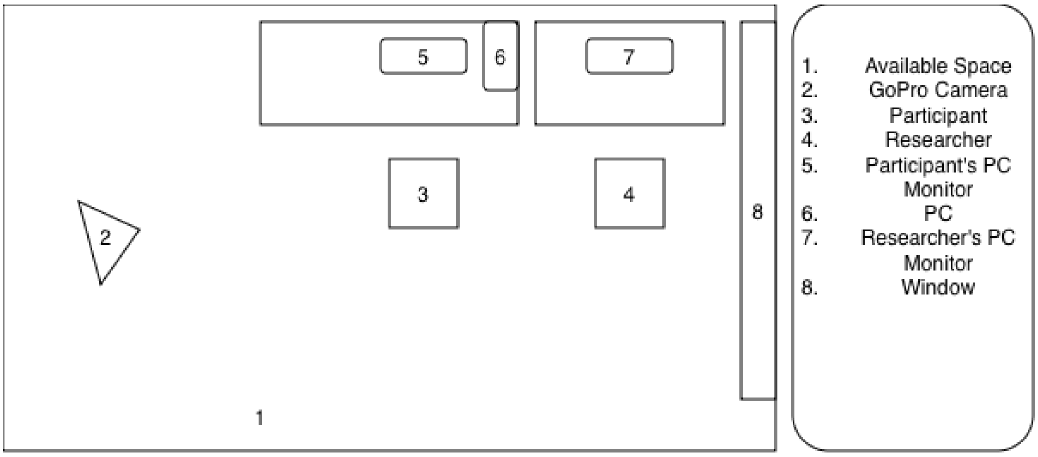

Figure 2. Workshop floor plan.

Think aloud was the principle technique used to enable organic conversation but keep to the core focus of the interviews. Think aloud is a commonly used evaluation technique for assessing a systems usability. This technique encourages users to "verbalise their experiences, thoughts, actions, and feelings whilst interacting with the interface" [4]. While we were not looking to evaluate the applications used in the interviews, think aloud was an appropriate technique as it allows the expertise of the participant to lead the conversation with the researcher filling the gaps and driving the conversation forwards.

Two HMDs were used for the interviews with two different applications, both supplied by GigXR [3]. GigXR are a software company that make and sell XR software for medical education. Their core applications aim to facilitate clinical skills practice and anatomy learning for students. Their products are designed to enhance and extend existing pedagogical approaches to support active and collaborative learning. The Microsoft HoloLens 2 was used for the majority of the workshop with the GigXR DICOM XR Library application [57]. While the HoloLens is now an older device, it was once an industry leader and has had a significant impact on multiple industries, applications, and XR research and development [58]. The GigXR DICOM XR Library application was chosen for these interviews as it is a very flexible tool with a variety of built-in features that position it well for radiological discussion. This application has a considerable number of medical cases that have been built up into 3D models and the anatomy segmented. Tools enabling the inclusion or removal of individually segmented anatomy, placing pin markers, and a cutting plane to slice through the model and see the internal structures are provided. These tools make this application valuable in demonstrating to radiologists what a 3D AR representation of the scans they look at day-to-day could look like and how they may be able to interact with them. This enables a solid platform to have in-depth conversations about issues they have or can foresee if bespoke applications were to be designed. As an educational tool, GigXR's DICOM XR Library is not designed to match the diagnostic feature set of clinical-grade DICOM PACS software, and it does not aim to replicate a diagnostic workflow. This was appropriate for the aims of this study. Participants were not asked to perform or evaluate diagnostic tasks, but to interact with a 3D AR representation of familiar scan data as a basis for discussing the interaction value AR could offer if purpose-built radiological tools were developed. The tool's role in this study was as a discussion aid and elicitation device, not as a candidate diagnostic instrument. Figures 3–5 show the DICOM XR Library functionality and illustrate the interaction capability the participants faced in the interviews (e.g., the cut-view feature, model manipulation, and available menu options), rather than documenting specific experimental tasks.

The Meta Quest 3 was used additionally for a small portion at the end of the workshop to contrast to the HoloLens and enable the researchers to begin to isolate which issues or advantages were associated with the HoloLens and which were resolved or provided with the newer and very different viewing mechanism of the Quest. The DICOM XR Library application is not available for the Meta Quest 3 so the GigXR HoloHuman+ application was used instead [59]. While this application is designed more explicitly for anatomy education with a more extensive but simplified and customisable model, it gave the participants the opportunity to experience the different viewing mechanism of the Quest and the different ways of interacting (i.e., with the controllers), and it allowed for reflection regarding what they saw as advantages and disadvantages of each HMD. Figures 6 and 7 show the HoloHuman+ application to illustrate some differences between the the two device-application pairs.

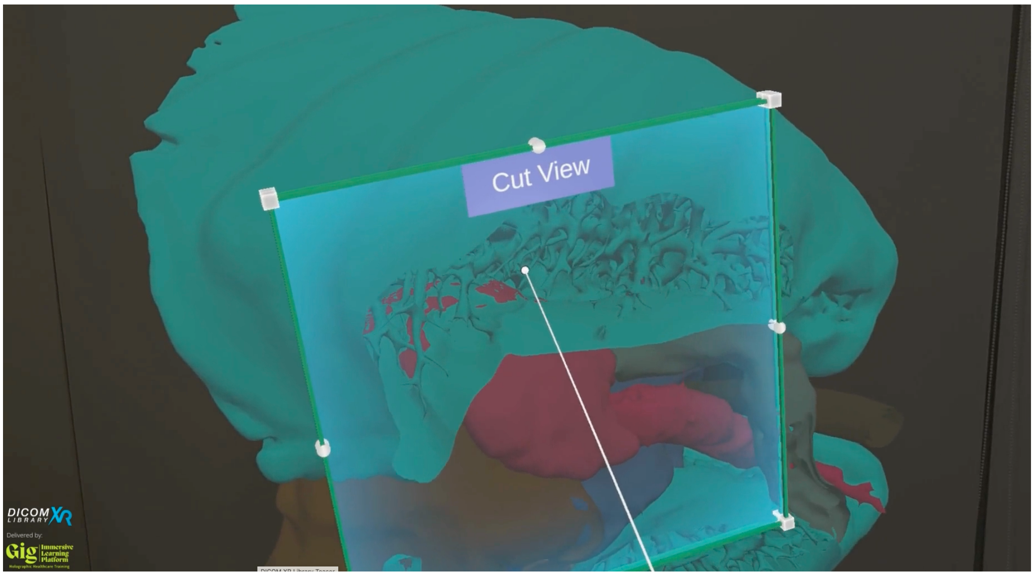

Figure 3. GigXR DICOM XR Library demonstrating cut view [57].

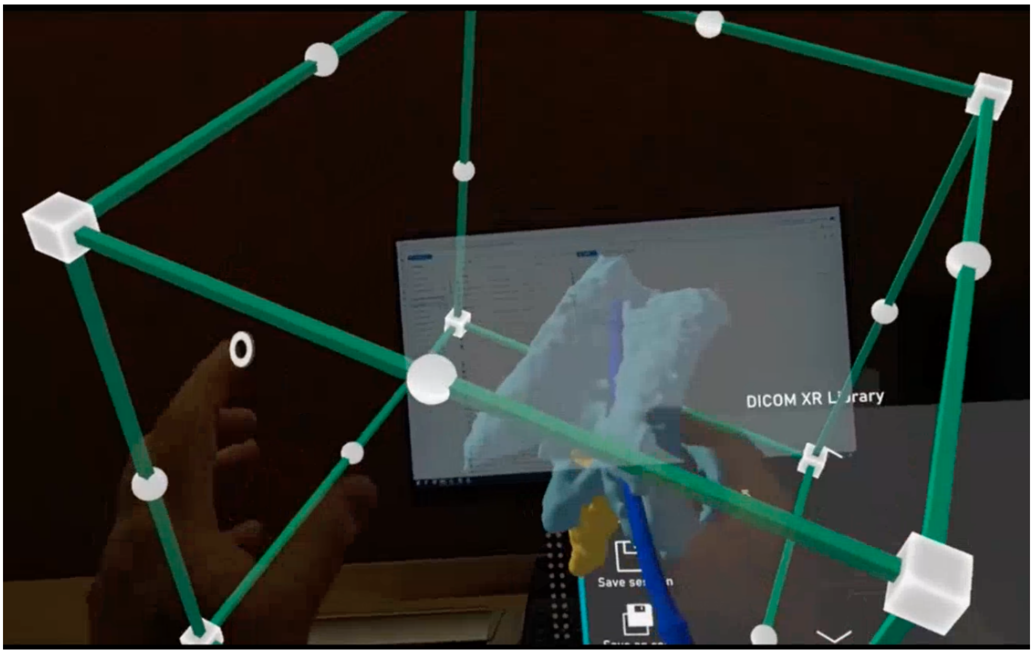

Figure 4. Participant manipulating brain model.

Figure 5. GigXR DICOM XR Library menu [57].

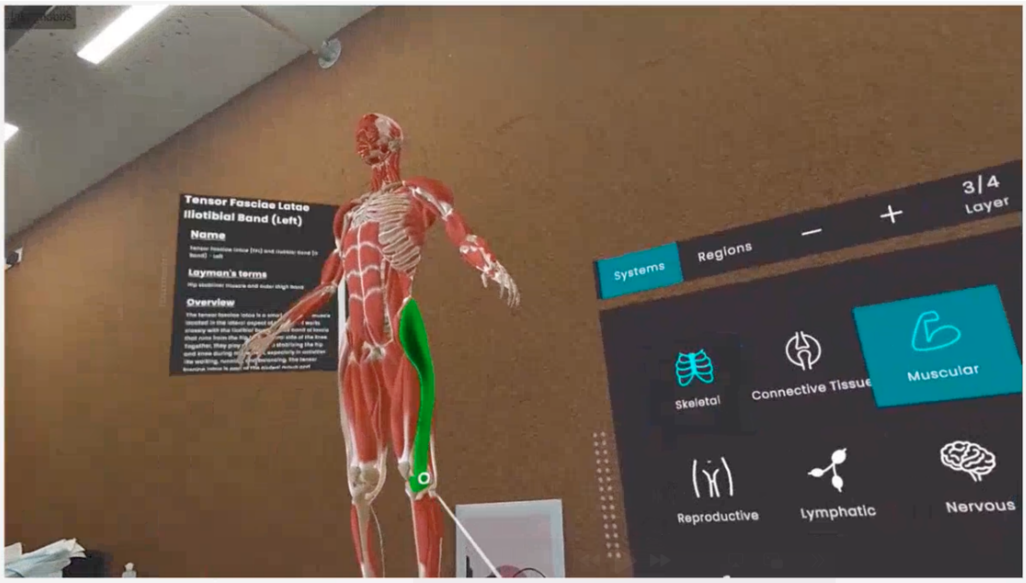

Figure 6. Participant exploring HoloHuman+ [59].

The interactions required by the participants revolved around the "point-and-grab" hand gestures commonly seen across modern XR HMDs. These interactions enabled the participants to point at and select or drag virtual content within both of the GigXR applications. There is crossover in how interactions are performed between the HoloLens and Quest albeit mostly performed with controllers when using the Quest. Interaction research shows hand gestures as an intuitive form of interaction and GigXR have relied on this capability from both devices [46].

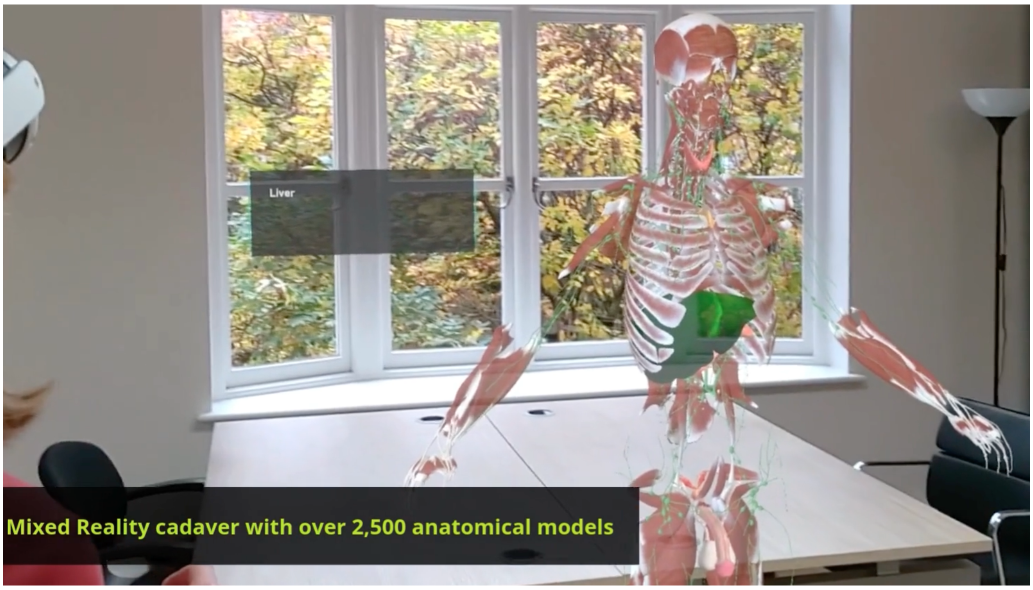

Figure 7. GigXR HoloHuman+ demonstration video [60].

The Interview Activities

The interviews consisted of five activities with the HoloLens and two with the Quest, with these activities completed in the same order with each participant. The first activity was preceded only by a very brief explanation for the researcher of how to put the HoloLens on, what the buttons on the HMD do, the fact that the visor tips up, and how to click or select apps. Once the participant was comfortable with the HMD, they were instructed to select the Tips app and go through the hand gestures training path of the Tips app provided with the HoloLens by Microsoft; the steps are shown in Table 3. This gave the participants, who had likely not used XR much before, an opportunity to understand and get used to the interactions without the medical context and before diving into the body of the workshop. This had the advantage of also triggering the HoloLens to realise a different person was using the device and completing the calibration. After this was complete the participant was told to go back the Start menu open the DICOM XR Library app and, once logged in, select a case from the library that appealed to their interests or sub-specialty. Participants were left free to choose whichever case from the library was relevant to their field or interests as the tasks were compatible with all models. With this first model participants were invited to have a play with the model and describe what their first impressions were, and any initial thoughts about the model manipulation. Additional tools were then introduced by the researcher to continue the conversation. Following this it became quite natural with some participants when they wanted to move to a different model. Generally a new model was chosen for each task but this was left flexible to respond to the attitudes in the workshop.

With a new model the participant was then asked a question from the facilitator guide about the case in front of them, which was provided along with the GigXR DICOM XR Library software. This question was relevant to the model chosen but was often along the lines of "describe the abnormal finding", which the researcher then asked the participant to talk them through what it was like using the AR to answer this question. This was to try and explore how the participant responded to having to obtain clinical information about the case from the AR software. For the next activity participants were asked to role play measuring a structure that was presented on the model, often a tumour, with two or more pin markers. This activity was inspired by the common task for radiologists of measuring structures. The DICOM XR Library software does not provide a measuring tool but does allow the user to create and place pin markers to point at and mark out structures. This exercise forced the participant to manipulate objects accurately and tested the nature of precision interactions. It also often meant leaning on multiple tools which encouraged a more holistic use of the software.

The final task with the HoloLens involved role-playing writing an email to a colleague on the PC about the case the participant was viewing in AR. This did not involve writing a full report of the case but the participant was asked to type out an email with the physical keyboard on the PC as if they were, for example, asking a colleague for a second opinion. This activity was used to explore how the participants found using the PC and HMD at the same time and discuss any interaction issues or opportunities that arose in that regard.

For the Quest part of the workshop there were two activities, accompanied by questions following up conversations that were had when using the HoloLens, aiming to understand continuity or discontinuity. The first activity with the Quest was much like the HoloLens, i.e., having a play with the software, exploring its capabilities, and having the participant narrate what they were doing, aiming to do, liked, and did not like. Secondly, while this software does not have pin markers like the DICOM XR Library, the participant was asked to use one of the small finger bones to point at other structures on the model like a marker. This, like on the HoloLens, was used to explore precise interactions, interactions with small structures, and interactions with multiple virtual components colliding. Two devices were used, in part to relieve focus from one particular device–application pairing, but this study was not designed or intended as a controlled comparison between the HoloLens 2 and Meta Quest 3. The two devices ran different applications with differing numbers of tasks, and the Quest activities were conducted later in each session, introducing a potential order effect. As such, any observations regarding device experience should be read as an exploratory comparison of user experiences across device–application pairings, rather than a controlled test of hardware capability, with these confounds discussed further in the Limitations (Section 5.4).

The interview procedure used by the researchers to guide the structure of the interviews is shown in Table A1. This appendix also shows and the focus of the questions was split between this study and the companion study with the radiological education focus.

To conclude the interviews, the participants were asked if there was any further reflections or considerations they would like to put forward before completing the NASA-TLX questionnaire.

Table 3. Microsoft HoloLens 2 Tips app hand gestures' training path.

Task	Hand Gesture
Tap three holograms of gems	Select virtual objects within reach
'Air tap' three holograms of gems	Select virtual objects out of reach
Move 3 holograms of flowers into a circle	Moving virtual objects from 'A' to 'B'
Make 3 holograms of flowers bigger or smaller	Change virtual object size
Rotate 3 holograms of flowers until they open	Rotating virtual objects
Open and close the start menu with two or just one hand	N/A
Exit an immersive app	N/A

3.3. Analysis Process

The POV video and audio from the interviews were recorded via the Microsoft Teams call which automatically produced a transcript of the conversation. The POV video along with the transcripts were then subjected to reflexive thematic analysis which allowed the realisation of trends, contradictions, and insights across the whole data corpus. Reflexive thematic analysis is an interpretive qualitative analysis approach, developed by Braun and Clarke [61], that engages critical reflection from the researchers to analyse and identify patterns across a qualitative dataset. Reflexivity is integral to this analysis method and the researchers' perspectives and biases are used as tools for analysis. As such, it is important to understand these biases in order to understand the context used to support and influence the conclusions made.

This study adopts Braun and Clarke's updated version of reflexive thematic analysis [61] which builds on their original work [6]. For this study, the analysis was approached from a critical realism ontological position and utilised a contextualism epistemology. This means that the researchers could explore the meaning from the participants in context and be directed by this, constructing meaning and evidence through the analysis. This is opposed to a more traditional realist post-positivist approach where it is considered that a single objective truth exists within the data and it is the researcher's job to find it [61,62].

An inductive coding process was used in this study which enabled the focus to be put on the participants' experiences and opinions, allowing themes across the data to be built from participant experiences. This inductive process was coloured by inherent epistemological and ontological assumptions as "you cannot enter a theoretical vacuum when doing thematic analysis" [63]. A combination of both semantic and latent codes were used throughout the process to capture the explicit, surface-level as well as the deeper, more implicit points being made. This combination allowed for all aspects of the interviews to be captured in the analysis.

The GoPro footage from the side of the workshop activity space was coded separately to the main data of the POV video and audio recordings as it allowed for a secondary perspective of the interviews. The codes from the GoPro analysis were then used to inform the theming of the core data. This GoPro footage also allowed observations to be gained around how the participants used the space. These observations are integrated into the results below to reinforce the themes built from the core data.

3.3.1. Six Phases of Reflexive Thematic Analysis

Braun and Clarke detail a six-phase approach for reflexive thematic analysis [61]: familiarisation, coding, initial theme generation, developing and reviewing themes, refining defining and naming themes, and writing up. For this study the familiarisation phase was achieved in two ways, firstly by facilitating the interviews there is an initial exposure to all of the data in the context they were given. Secondly, through the transcription process. The automatically generated transcript from Microsoft Teams was used as a base but the transcripts were checked against the workshop recordings. This ensured that the transcripts were accurate while also contributing to the familiarisation phase of the analysis. The coding and theme generation were primarily carried out by the lead author, with the second author offering opinions and challenging decisions after each round. Two full coding rounds were completed and theme generation was completed over two iterations. The second author reviewed codes and themes and provided analytical feedback after the first round of coding, after all interviews had been coded, and between iterations of theme development. Where the first and second author's interpretations diverged, this was resolved through discussion until a shared interpretation was reached, informed by returning to the original data. This discussion itself formed part of the reflexive process, surfacing assumptions that fed into the refinement of codes and themes, rather than being treated as an error requiring correction. This was part of the reflexive work of the coding and theme generation phases of the analysis as it was an opportunity for the challenging of assumptions and biases. This provoked further reflection on the codes and themes throughout the analysis process and meant that assumptions could be challenged, resulting in well-founded themes being developed.

While Nvivo is an obvious choice for thematic coding of textual data, it was not appropriate for this study as there were textual and video data to code [64]. Instead, a virtual whiteboard on FigJam was used, with sticky notes used to represent codes [65]. Each sticky note had the code name, participant ID, timestamp from the POV or GoPro video, and often the quote from the transcript. This made the theming stages intuitive as it was easy to see the codes individually and holistically. It also meant links between themes and distinctions were easier to see as the theming process iterated. Initial groupings could then be discussed coherently between authors to challenge the assumptions and biases that helped build them.

Following the two coding rounds, codes were grouped into preliminary collections based on shared underlying meaning, rather than surface-level topic overlap, before being reviewed collectively to identify connections and patterns across the interviews. These collections formed the basis of initial candidate themes, which were then individually interrogated by the first author and subsequently reviewed jointly by both authors, challenging assumptions and refining boundaries between themes, before being consolidated into the final themes presented in the Results.

3.3.2. Positionality Statement

Here, the authors consider our positionality and discuss how it will have impacted this work. We are computer scientists based in Open Lab, a Human–Computer Interaction lab in the School of Computing at Newcastle University, UK. We have some knowledge of digital health and HCI health-tech but no formal medical training. Our expertise lies in qualitative methods and designing technologies for specialist user groups. This places us as outsiders to the clinical contexts examined in this work, and our engagement with participants was shaped by this status.

During the interviews we adopted an investigative stance, using our outsider position to ask foundational or clarifying questions, including asking the "stupid question", in order to reduce assumptions and fully understand clinical processes. A medically trained facilitator might have elicited different kinds of narratives or prioritised different aspects of the discussion; our disciplinary lens inevitably shaped the direction and content of data generation.

The theoretical positioning for this work has been based around critical realism, being attentive to how meaning was shaped through interaction. Our technological and systems focus also influenced the generation of codes and themes, often drawing attention toward workflow, processes, and technological touchpoints within clinical practice.

The authors engaged in reflexive practice throughout the analysis, principally through ongoing discussions in which we examined the assumptions guiding decisions and reconsidered how our positionality influenced coding and theme development. We also reflected on how our epistemological positioning shaped what we expected to find in the data and how these expectations changed during analysis. We acknowledge that the data and the final themes are shaped by our backgrounds, disciplinary training, theoretical commitments, and presence in the research process.

4. Results

The results from this study are presented here as the five themes developed through the reflexive thematic analysis process. These themes are Frustrating Object Manipulation, which talks about the extensive issues with virtual object manipulation. Improved Understanding and Communication encapsulates the two key advantages seen by participants. Hardware Constraints on Precision covers the hardware limitations and their impact. 2D Images Are Not a Restriction challenges the potential misconception that the three standard planes used in current PACS are a limitation. Finally, Misalignment with 2D Practice captures the lack of advantage many participants saw. Figure 8 shows a theme map with themes shown in red and their associated sub-themes in blue, with arrows pointing outward from each theme to indicate that the theme functions as the central organising concept, with each sub-theme representing a distinct facet of it. Observations made by the lead researcher during the interviews and in the analysis of the footage from the GoPro camera have been integrated throughout the themes. Most of these observations were clarified in conversation and therefore have quotes to support them directly. The NASA-TLX results are covered first, in Section 4.1, along with what these quantitative results illustrate for this study.

Figure 8. Theme map showing themes and sub-themes.

4.1. NASA-TLX

The NASA-TLX is a method of quantifying how strenuous a user found a task overall and across the six subscales. It was initially developed by NASA to help develop the cockpit controls for easier use by pilots but has been shown to be much more widely applicable. While there has been some argument against its validity, it is still widely regarded as a good benchmark. At its core, the NASA-TLX consists of six Likert scales which can be considered in isolation as well as the making of the whole TLX score. We used the NASA-TLX in this study to gain an understanding of how difficult participants found the tasks and to reinforce the opinions expressed by the participants, framed below in the five themes. We asked the participants to focus on the section of the interview using the HoloLens when completing the questionnaire as this was the more significant part of the workshop and most representative of the work radiologists do. While there was a limited number of participants and statistical significance is difficult to identify from this small set of results, we consider the NASA-TLX scores here anyway, as an indication of the workload of the task and as an argument close at hand with the themes from the reflexive thematic analysis.

Due to the small sample size ($n = 7$) and the questionnaire's primary focus on the HoloLens activities, the results of this NASA-TLX should be interpreted cautiously and are considered exploratory. The NASA-TLX was used to give an indication of how challenging participants found the tasks in the interviews and which aspects of the tasks were most and least taxing. For this purpose, these results can be considered as descriptive statistics providing value in showing how participants considered the tasks.

NASA-TLX was administered using the standard weighted procedure. Participants completed the 15 pairwise comparisons between the six subscales, and each subscale's weight (0–5, based on the number of pairwise comparisons won) was applied to its raw rating before averaging, giving an overall workload score out of 100. The Performance subscale followed the standard NASA-TLX orientation, with "Excellent" (0) and "Poor" (100) as endpoints. As such, a low Performance score therefore indicates a positive self-assessment, consistent with the direction of the other five subscales, and no reverse-scoring adjustment was required.

Results are represented in Figures 9 and 10 and are reported using medians and interquartile ranges (IQRs) as appropriate for a small sample size, alongside the means and standard deviations for descriptive purposes. In these graphs the six subscales are shown independently as well as the overall weighted TLX score.

The overall rounded NASA-TLX score had a median of 53.5 (IQR = 31.5–69.0, range = 6–78), indicating moderate perceived workload. Among the subscales, Effort showed the highest median (63.5, IQR = 43.5–84.8), followed by Performance (56.0, IQR = 22.3–79.8) and Frustration (53.5, IQR = 38.5–83.5). Mental (28.5), Physical (21.0), and Temporal Demand (23.5) were comparatively lower.

Temporal demand was expected to be relatively lower as there was no explicit or restricting time limit set on tasks, just an expected time frame for the workshop overall. It is encouraging to see a low physical demand but this is contradictory of the few participants that raised issues with having their arms raised for longer periods of time, and one participant was observed to be supporting one arm with their other at one point. Similarly, a low mental demand can be considered a positive indication of little difficulty understanding how to approach the tasks and how to interact with the virtual content. It is commensurate with the qualitative results to see the scores for Effort and Frustration, in particular, are higher. The frustration with the virtual object manipulation as discussed in the next sub-section is seen in the quantitative data here.

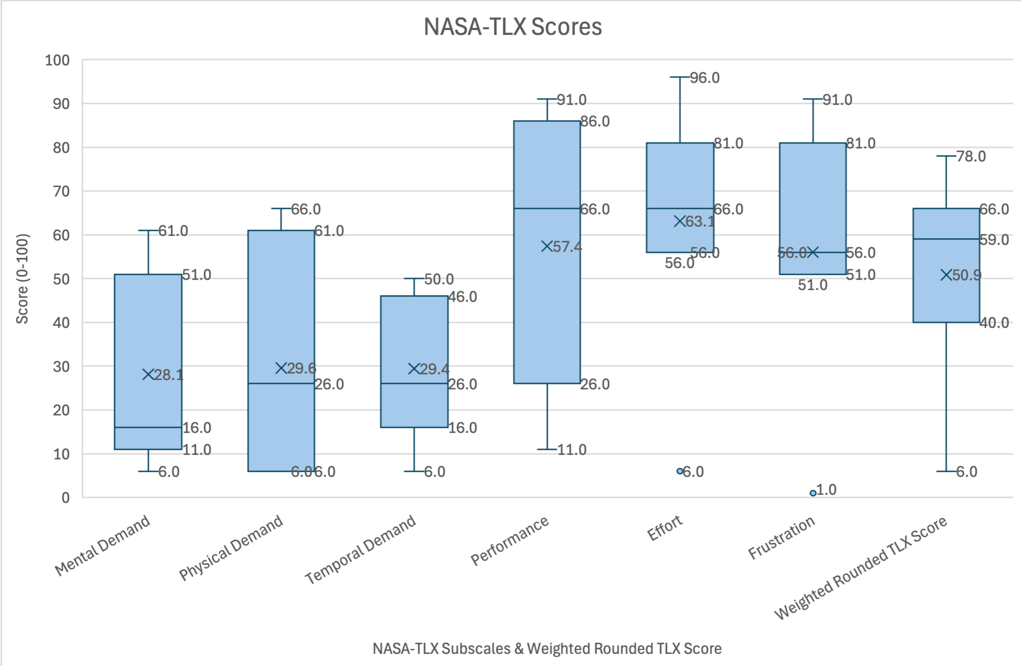

Figure 9. Box and whisker plot of NASA-TLX results.

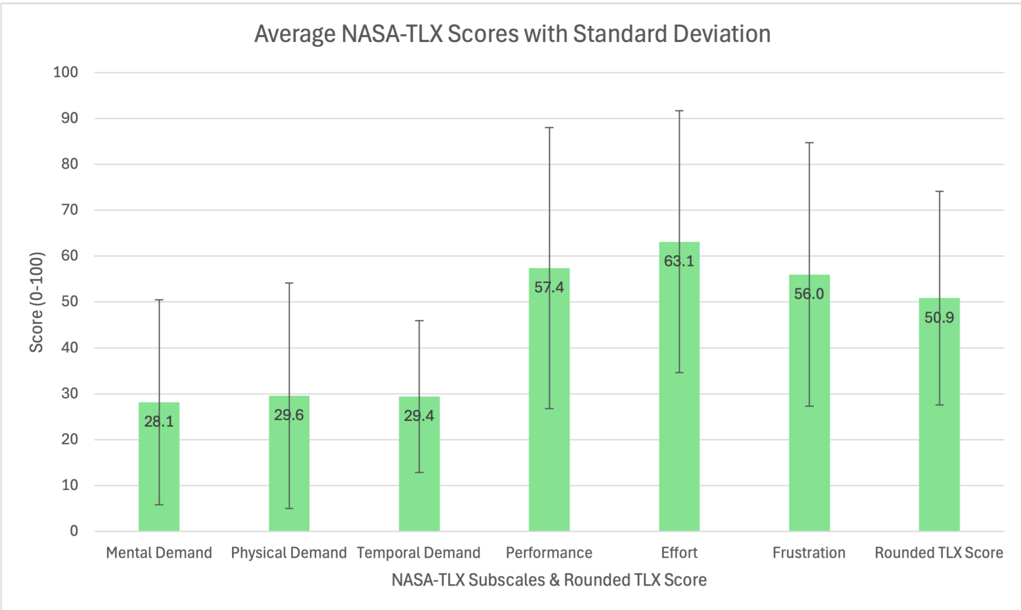

Figure 10. Bar chart of mean NASA-TLX results with standard deviation.

4.2. Frustrating Object Manipulation

This first theme captures the plethora of issues around virtual object manipulation experienced by the participants throughout the interviews. Broadly, virtual object manipulation is slow, frustrating, and therefore limits utility. While the learning curve was perceived to be shallow and participants got used the interactions fairly quickly, there was persistent difficulty with the object manipulation and consistent complaints of flawed interactions for this application Interactions with the Quest were better than the HoloLens in some ways but equally poor in others. This effect can be seen in the Frustration and Effort components of the NASA-TLX.

4.2.1. Frequent Misclicks

Misclicks in varying different ways were a constant issue for every participant throughout the duration of the workshop with the controllers of the Quest reducing these to some extent but not removing the problem. The most frustrating and most limiting type of misclick was where participants knew what they want to achieve but could not, for example, because they thought they had pressed a button but had not, or they could not find the grabbable part of a virtual object, or just because the hand tracking did not register the action they were intending to perform. This resulted in participants having to attempt the same action multiple times and say things like "the click misregistration stuff is a bit frustrating" (participant D), "it seems to be the clicking on things that is the most difficult thing" (participant F) and "Why am I not managing to even change pages? It's like I can't get close enough" (participant G). Further misclicks occurred where participants moved a virtual object accidentally, for example, participant A said "I want to have this where it can just stay put for a minute, and I keep accidentally moving it". The controllers of the Quest cut down on these misclicks and a few participants speculated that this would be the case even before they had used the Quest, which suggests there is intuition in using a physical device to manipulate virtual content, much like a mouse; "again, the misclicking. Yeah, I think that's the thing. Yeah, it's a bit annoying. Maybe actually having a controller would actually be useful in that regard" (Participant D). This was then reinforced by other participants using the controllers, with comments around the benefit of the haptic feedback of having a physical button to press. "The controllers feel actually quite intuitive and [. . .] Like things like scrolling up and down that feels easier with the controllers [. . .] if you've actually got a button to press, it gets around that problem [misclick]" (participant A). "the clicking and holding these is much more useful than what we have before from the point of view of, just using your hands" (participant F).

4.2.2. Within-Interview Shallow Learning Curve

Multiple participants said things like "it does get easier as you use it, though" (participant F) and "I'm definitely getting more used to manipulating it, the more I do it" (participant E), which was also observed and made clear with how the speed and confidence of participants hand gestures grew just through the 90 min usage in the interviews. Participant B, who owned a Meta Quest, was immediately more confident with hand gestures as well and looked to have fewer misclicks, but he still struggled with the same types of virtual object manipulations as all other participants. Within the interviews, a perceived shallow initial learning curve for the AR interactions was consistently observed with participants becoming comfortable with the core mechanics quickly within the session. However, participants also noted a persistent difficulty achieving precision and avoiding misclicks that did not appear to resolve as the session progressed. Some participants struggled at the very beginning or the first time a new type of virtual object manipulation was required, likely because it was not obvious how to interact with the virtual content and they had never seen anything similar so had no experience to call on. Participant A said "Once you know what you're doing it's intuitive. It's just the first time you do it where it's a bit- Not quite sure what's going on". However, this was also seen with participant B, despite owning an XR HMD, as he said "not quite sure how to how to interface with this one" when a completely new type of object manipulation was required and he did not know what to expect or understand how to interact. This suggests that where interactions in AR are very different from previous experiences guidance or explicit suggestion of how to manipulate virtual content could be valuable. It was also observed that participants used more horizontal and vertical space as the interviews progressed, as if learning how wide the scope of where they could place virtual content was.

4.2.3. Clumsy Object Selection and Deselection

Picking up and putting down virtual objects was one of the biggest causes of imprecision and therefore frustration, particularly with the HoloLens, with participants saying things like "Once I've picked it up, very easy. The difficulty is in picking it up" (participant A) and "Once you've grabbed it it's much easier" (participant D). This was due to the natural movement of the hand and wrist when performing the pinching or opening in order to grab or yield a virtual object. No matter how still participants tried to stay or however they attempted to adjust how they were performing the pinch, there was still some movement that dragged the virtual object off the desired position. "When you're rotating it and let go of it, it does an extra little bit of movement that that you would just it would infuriate you if you're using it for a report" (participant F); "The problem is that whenever you pinch your fingers, it moves away from that [position]" (participant G). However, while the controllers were better in terms of being able to select the correct part of the virtual object, there was little improvement in the precision of picking up and putting down virtual objects exactly where the participants wanted, maintaining that frustration. This improvement when using the Quest came as the controllers and physical buttons required less movement to actuate a response but the unintended movement issue still remained. Participant C said "the problem is by the act of squeezing and releasing there's sort of- by necessity there's movement by the wrist in order to change my grip shape and it's picking up on that". This links into the next sub-theme around requests for mouse-like stability when clicking and the advantage of a mouse being against a solid surface for a reference point.

4.2.4. Lack of Precision & Feedback

This sub-theme captures details of further aspects about the object manipulation that led to imprecision and a desire for the advantages that a mouse provides. Firstly, there were repeated complaints about the movements being too sensitive and therefore jittery and any shaking or tremor in the participant's hand being picked up on unwillingly and translating into a shaky virtual object. This was only an issue with the small, more precise actions but made it difficult to have confidence in the placement of virtual objects relative to one another, and position things accurately, easily. "It's quite nice to manipulate, but very hard to manipulate accurately" (participant E); "I'm under no illusion I think I'm holding my hand steady. I know I'm not. But it's then also picking up on those little tremors that I have. And maybe a little bit of smoothing would help. And sort of it being able to discern what are the purposeful movements. And what is me just being a little bit caffeinated?" (participant C).

This was made easier when using the Quest and the controllers, and participants felt more comfortable being more accurate with statements like "I think if you asked me to measure with this, it would be more accurate" (participant D) and "It's definitely a lot easier" (participant B). However, there were still issues. The small movements with small virtual objects were still difficult to make accurately. "That was pot luck, whatever that one was. Yeah, but for most of them, large, medium and little larger small size structures, it's fine, yeah. Not perfect but fine" (participant F) and "Getting a bit fiddly isn't it" (participant G). It was clear to observe that, when using the Quest, the participants arm movements seemed much smaller as the controller allows larger movements by just twisting the wrist. However, these smaller movements were still difficult with both modalities.

Part of the struggle with manipulating virtual objects was the lack of "physicality" and haptic feedback. Participants commented on this as an issue in a range of scenarios such as when trying to move two virtual objects relative to one another and knowing when they collided and no haptic or even audio feedback when pressing a button. Participants said

"I was having to line it up. Not like- you know in 3D space without any tactile feedback. Without knowing when I was actually making contact with the structure" (participant C); "if I was to do this in reality, I would probably want it to be a bit more kind of almost feel as if it was a physical object" (participant G); and "I guess it's more accuracy if when you click your fingers and when you don't click your fingers isn't as there's no if there was some sort of sound to tell you that you've done it, that would be useful because you get a click, don't you with a mouse and then you know you've done it" (participant F). This desire for mouse-like characteristics when manipulating the virtual content is continued with participant D saying "It's got to be a bit more mouse-like in the sense that like, you know, like you can just click away" and participant F saying "If you have like a tool like a mouse, you can see those things there, something like that that you could essentially move in small proportions that did something on your virtual reality screen [. . .] I think that'd be much better than having to deal with your hands".

It is clear, then, why keyboard and mouse are so ubiquitous and that the methods of interaction here are not matching this capability. This theme has covered the range of issues participants raised with virtual object manipulation, in sum common misclicks, difficulty with precision movements, and lack of feedback.

4.3. Improved Understanding and Communication

This second theme discusses the key advantages of AR for radiology as acknowledged by the participants and the reasons for them. Namely, the enhanced ability to understand the complex 3D anatomy, and the relationships between structures, and the communication ability AR could afford them. These advantages are discussed in isolation in the literature, but we highlight them here to build the empirical understanding of the advantage AR brings to radiology and link this empirical understanding of the advantages with the interactions AR affords.

4.3.1. Enhanced Spatial Understanding

Throughout the interviews it was unanimous that the advantage of using the 3D holograms to view the models was the ability to understand the complex 3D structures and the relationships between anatomy. This is reflected in the relatively low Mental Demand component of the NASA-TLX. Participants commended the enhanced ability to visualise structures with comments such as "I really quite like that. Sometimes it's quite hard to visualise. Because everything we do is 2D images of the 3D structure, and sometimes it's quite hard in your head to build that 3D picture. So it's kind of nice to be able to do this and just cut through it to get an idea of exactly what's going on" (participant E), and "[. . .] Then you can actually see like the hole for the ventricle. So where that's where like the fluid comes out the CSF and surrounds. Yeah, I think the model's great" (participant D). The enhanced ability to understand the positioning and relationships of different structures was exemplified by participant A. When he was using the model to find the tip of a shunt tube in a CSF ventricle in the brain, he said "Yeah, you can see so nice, this really is so nice- 3-dimensional location of the exact tip of the shunt tubing. [. . .] I quite like seeing just that little tip of the tube there. Amazing. I've never really been able to imagine, you know, what this looks like in this sort of way before".

It was brought up by multiple participants that 3D techniques are currently used in practice for niche scenarios or clinical questions, and participants mentioned that using AR in those scenarios could be of value. Participants B and E both mentioned this and the advantage here, common between the heart and the brain, is where the structures of interest are of complex shapes that do not follow any one of the three planes, principally vessels. Participant B said "one of the most common uses of it [3D tools] is when they're looking at the vessels, intracranial vessels, so looking for small aneurysms. It's a lot easier to let the software render a 3D model and then you effectively can do the same thing where you create your own segmentations to remove the bits of anatomy that you don't need and then focus on the bits that you're interested in. Something like that this would be great for. This is a much better interface" and participant E similarly said "we do lots of VRTs so they are 3D reconstructions of, for example, the heart and the coronary arteries. We partly do that because it's easier, because the coronary arteries curve around the heart and go all the way around, and aren't really ever in an axial sagittal or coronal plane, it's just easier to pick them out and figure out where they are when they are in 3D".

4.3.2. Enhanced Communication

The second key benefit acknowledged through the interviews was the ability to communicate more effectively when using the AR as a tool to show the area of interest and demonstrate abnormality. Participant G highlighted this as the value that they saw "it's more about that sort of the added value of the 3D information [communication]". There was some contradiction in the value of a tool to aid communication in multidisciplinary team meetings (MDTs), with some participants commending it for a role like this to improve how cases are presented and others opposing this as current methods are good enough for the purpose that these meetings serve, which is more around making decisions based on group expertise and information than an in-depth look into the images. Participant B said "I think it would be- a good application of something like this would be for when we sort of present our cases in meetings", but participant G contradicted this with "Less so in an MDT because an MDT is really about decision making involving the wider team, but on a one to one level where it's a case review more so".

When talking about the communication value that AR could afford, many participants mentioned other clinicians such as surgeons as the likely group they would use a tool like this to communicate with. Participant A said "improving the accuracy and the quality and the communication of my description of the findings to other people like surgeons. You know, that aspect would be improved", and participant G said "So this would have a more niche value. And as you say, possibly be into that working with and communicating with surgeons to help them understand or if there's a more unusual case that you feel you need to go through together so that you get onto the same page". This commonality ties in with the sub-theme "Radiologists think in 2D" in Theme 5, as there is an implicit suggestion that radiologists would not gain from using 3D tools to communicate between each other but those less used to looking at 2D slices would gain from a 3D representation. This can be seen as well in the personal preference for radiologists using currently available 3D renders, as participant B said "Similarly for the CT colons with a 3D volume rendering, not everyone uses it, even though technically, as per the guidelines they're then meant to at least review it. But because most people but especially the older radiologists who've been practising for a decade or two before you had readily accessible 3D volume rendering, they're just way much more accustomed to dealing with the 2D imaging". This more in-depth one-on-one communication and discussion of the scans between a radiologist and another clinician is reflected in what participant G said about the utility of AR tools in MDTs above "Less so in an MDT because an MDT is really about decision making involving the wider team, but on a one to one level where it's a case review more so". Here, then, there was no consensus across participants about who AR could help radiologists communicate with but this value in communication was undisputed and directly linked to the method of interacting that AR provides.

4.4. Hardware Constraints on Precision

This theme captures the effects that the limitations of the hardware have on the use of an AR tool for radiology. Field of view, resolution, and brightness were amongst the key issues raised, each affecting the utility of such an AR radiology tool differently.

4.4.1. Limited FOV Discourages Wider Use of the Space

Here, we observed across the participants the effect that the narrow FOV of the HoloLens has on using an AR tool like this. Repeatedly throughout the interviews, there were participants that lost buttons or virtual objects in their periphery because of the narrow FOV and some participants voiced this as an issue. Participant A said "I was struggling to see the whole screen [menu] all at once" and "At this point it's it feels a little bit cluttered having several different things going on at once, like we've got the cutting plane, you know, we've got the brain box and now I've got this big pin thing. There's menus and stuff like that", and participant G said "So I wasn't kind of seeing the four- [QR code scanner edges for log in]". However, moving to the Quest, participant E noted how much more free he felt to use the space versus being stuck looking forward with the HoloLens. "I don't know why, but I feel like from this one I'm far more willing to turn around and look everywhere, whereas I felt like I was stuck looking forward on that one, and although I do know and maybe I could have moved the stuff around like- I don't know. It feels easier to do this. [Researcher asked if that was due to the FOV] Maybe that's partly it, I don't know because- I guess in that one [the HoloLens] I felt maybe it did just have a bit of tunnel vision.". Similarly Participant A had a moment of realisation when using the HoloLens that and began to move virtual objects in to much more of the space "the thing I was saying was it feels very cluttered. I've made it cluttered for myself by having everything just here. So why not make it uncluttered by moving things". This was also observed in the participants' behaviour in the interviews with some participants seeming to use more of the vertical space once they had stood up and then sat back down as if realising that they could use more of the space. This is in line with previous research that shows the restrictive effect of a narrow field of view and the additional mental load it takes to use more of the space if the FOV is narrow [66,67]. We acknowledge this here as an observation of the study and as an aspect of application design to be considered forward when designing AR systems in this clinical radiological context.

4.4.2. Perceptual Issues Limit Utility

Brightness, or the lack of, was an issue raised by multiple participants when using the HoloLens, with participant F asking for the PC monitor behind the virtual objects to be turned off and participant G, amongst others, asking to close the blind to make the room darker. Participant D said "I think you end up getting strained because they're like two different like light sources of two different intensity in front of you". This was resolved when using the Quest with the added benefit of the true occlusion that the VST technology provides. However, other issues were then raised as a result of the hardware specs of the Quest. Participant A said "It's double-edged sword [. . .] I think it is better looking at the thing you know looking at the virtual thing. That is better. But obviously it's not as good looking at the non-virtual thing", and participant E said "I do like that it's solid rather than the lens although, the slight haziness of the real world is a little bit distracting". This continued with participant B placing the virtual content behind him against a white wall as much as possible. Participant F also commented on the Quest making him feel unwell almost as soon as he put the HMD on, concluding that "It's not something I'd do again. It makes me feel unwell". We are not trying to measure if the Quest makes more people feel unwell than the HoloLens, but it cannot be ignored that one of the participants experienced this and its therefore a consideration for AR tools in this context. However, it is clear that the resolution and brightness of the virtual content is better with the Quest, as is the provision of true occlusion, but the real world resolution of the Quest is not good enough.

4.5. 2D Images Are Not a Restriction

This theme challenges the assumption outlined in the Background that looking at 3D objects in 3D is an advantage, and it illustrates why the use of axial (top down), sagittal (side to side), and coronal (front to back) planes is a tool for radiological reporting rather than a restriction.

4.5.1. Slow Setup

A common objection to the application used in the workshop was the six degrees of freedom that AR enables. The unlimited ability to move the model to whatever angle, rotation, and position desired was seen largely as a negative, which opposes the intuition discussed in the Background section and what we hypothesised at the start of this paper. Participant C encapsulated this nicely by saying "If you want to confuse a radiologist show them something off axis". It could be argued here then that this could be just because using the three planes is how radiologists have been trained, and if they were trained with 6 degrees of freedom this would not be an issue; indeed he followed this up with "because we're so used to working on those three reels, and it's only when we want to, as I said, when we want to demonstrate pathology that we then go off axis", highlighting this and that the ability to go off-plane exists but has only niche value for communication. However, when other participants requested the same three default planes to view the models it was shown that these hold clinical utility beyond entrenched taught technique. Participant D said "if it's not bang on [axis] then, like you don't know how large something is" and both participant C and E said that when measuring a structure it is imperative to be exactly aligned on the plane you need to ensure accuracy of measurements "when I'm doing a measurement on the system, I know that I'm only measuring on the plane that that I'm interested in. The only way for me to get a measurement out of plane is for me to put the whole screen out of plane" (participant C), "for measuring a tubular structure you have to be exactly opposite [. . .] it's fiddly to measure stuff using this" (participant E). Participant E went on to describe how symmetry is valued for spotting abnormality and being slightly off plane removes any symmetry therefore making reporting more difficult "with the brain that I'm looking at, if you've got the left half and the right half perfectly lined up rather than slightly obliqued, it's much easier to spot abnormality because it's asymmetrical".

Another impact of the application being set up to enable the free movement but participants wanting three fixed planes is that whenever a new model was loaded up there was a significant amount of time spent moving the model around to fit whichever of the three planes the participant wanted. Participants said "You kind of just want it to be in the correct plane [. . .] It took me a bit longer than I would have liked to put it in a plane that I'm happy with" (participant D); "it should just sort of default into standard projection I guess" (participant B); and "I want it to load the right way round" (participant C).

Part of this extended time during the set up is the lack of standardised or consistent orientation for the model to appear in. This means that not only does the user have to spend time reorientating themselves to the model and then the model to the position that they desire but the extra cognitive load required to identify the orientation of the model can lead to doubt. Participant B went on to say "like it should just sort of default into standard projection I guess. So just like a front, at the very least, like, you know, defaults it so that in your model viewer- It's you know, standard sort of projection that we'll be accustomed to. So like for this I don't know which is left or right [. . .] that would be immediately apparent [with current systems]", and participant F asked "Do you know, anatomically if the right is the right side? Like we get in- So normally it's reversed on CT like this would be the right side and that would be the left side. Is it the same way?". This all adds time and cognitive load to the reporting process.

4.5.2. Potential for Fusion of 2D and 3D

The desire to be able to fuse the 2D and 3D representations of the images was expressed my many participants, because they suspected that this would provide them with the detail they need to answer clinical questions and the enhanced understanding of the 3D anatomy and the relationships between structures. Participant E suggested "you could combine this with a proper axial plane or have an option to keep it in exactly a flat symmetrical plane sort of thing" and participant A's statement concurred "The problem with not having the fusion with the conventional MRI sequences is, like I said, there are some brain structures that just aren't here, in this brain, like, I would want to know, you know, where's this? Where are the vessels and stuff in relation to this metastasis? I can't really answer that question with this". Participant A went further than this, saying that a radiologist must understand anatomy in 2D cross-section as well as the anatomical 3D, "in radiology, we have to understand the anatomy, not only how it looks on a model, you know, or on a person in real life, but how it looks in two dimensions. You know, as cross-section. And so that's the only thing that this is sort of missing".

Participant G suggested that having traditional reporting techniques as the core, using this for the majority of image interpretation, but then having the AR tool ready to launch alongside. The AR tool could then launch exactly the view that was being viewed in 2D but then with the advantages of six degrees of freedom and enhanced anatomical understanding. Here, the AR technology is another tool to work alongside and integrate with current methods that can be called upon when specific cases demand it's value. "I will quite often put markers and things on a DICOM image, so again if that can be imported across to save all of this faffing around, if there's a structure of interest and you're then loading a specific case, then then that might be helpful. You know the kind of lookout for whatever." Researcher: "So that's like you've got- you've seen something on the on the DICOM images that you want to see in 3D?" Participant G: "Yeah, exactly. And it just instead of having to sort of do all of that [setup], if the precision is better on the 2D image, then just import it in as you do this".

4.5.3. Niche Clinical Value

While the main advantages discussed in the interviews are the communication and understanding of complex anatomy, covered in Theme 2, there were also niche areas of clinical practice that participants saw potential use in AR technology for. These were often areas where 3D techniques are already used and this was seen as a better interface, where the clinical question requires the radiologist to move off the three standard planes, or in particularly complex cases with challenging anatomy. "It can be quite hard to figure out exactly which plane and something like a cancer is biggest in. Whereas with this because you can just twist your wrist around that's quite easy [. . .] We just manipulate the planes manually in the PACS software, but it just feels a little bit easier to very quickly do it with this by twisting your wrist" (participant E); "I think this would definitely, for example, things like the GI tract we do- There's something called a CT colonoscopy, [. . .] you can do a 3D reconstruction of it. And it sort of simulates you tracking through the bowel where you're looking for any signs of polyps or malignancy. I feel like this would be a good application, especially for those because those can take quite some time and they're very tricky. So effectively when you do the 3D reconstruction of it, it looks kind of like this as if you're within the bowel and travelling through. So it would be similar to this, but this would just be a much more easier sort of interface to look through it, at least for me." (participant B); and "for some of the complex, you know like paediatric brain tumours and that sort of thing where we might be spending half an hour, you know reporting the imaging then you know with the practise maybe in those situations might make it a bit quicker. Yeah, So for the complex cases, perhaps" (participant A). However, these are niche clinical situations and AR was not seen to be some magical answer, but "another tool in the box" ("I don't know whether this would definitely be easier, but, it's another tool in the box I guess, if you're trying to answer a difficult question" participant E). This empirically links the interaction advantages of AR with specific, niche use cases in clinical radiology.

4.6. Misalignment with 2D Practice

This final theme encapsulates the conceptual blockers of implementing AR into radiological workflows, the practical and personal reasons against the use of 3D renders and the impact of this on any AR radiological application. The sub-themes here capture the core negative reasons participants were against using AR in practice, aside from the object manipulation issues.

4.6.1. 3D Renders Do Not Have the Detail for Reporting

This sub-theme covers this issues encountered with the level of detail in the virtual models used in the interviews and the wider issue of using 3D renders for diagnostics. The first issue to come up regarding 3D renders was the known fact of surface renders of structures not being diagnostically useful, because the job of radiologists is to look inside structures. Introducing the cutting plane in the interviews helped with this but the low level of detail in the segmented structures presented meant that participants struggled to navigate specifics around the model and could not answer certain clinical questions. "There are some brain structures that just aren't here, in this brain, like, I would want to know, you know, where's this? Where are the vessels and stuff in relation to this metastasis? I can't really answer that question with this" (participant A); "so if it's all the one colour, because it's been categorised as liver tissue, you lose one of the key things that you're looking at which is sort of the texture of liver" (participant C); and "this is at the moment quite crude an analysis. It's not really got all the structures that myself or the surgeon would want to see" (participant G). To some extent this may have been a limitation of the software, and it was discussed with participants if their opinions on the level of detail would change if the models were in a more raw from, i.e., if the raw DICOM images had been stacked into a 3D render without any segmentation, abstraction, or smoothing which participants were more receptive to. However, this was only speculation. A lack of detail in the 3D model was a big complaint from every participant and an issue that rivalled that of the difficulties manipulating virtual objects; this went hand in hand with the low resolution and resulted in a projected total lack of confidence in results if participants were asked to properly evaluate and report on a case diagnostically.

However, the difficulty in getting 3D renders to reach diagnostic precision surfaced. Due to the way images are taken, in axial slices, the computation of stacking them into a 3D render and the AI models used to do this mean that you cannot guarantee their accuracy. Because the image has been computed rather than being presented in its raw form, the models cannot be relied on in the same way as a 'raw' 2D DICOM cross-section. There is such a large amount of anatomical variance meaning that segmentation defects are too common. "The difficulty of the 3D models is that they won't be like they won't be as accurate because we can't get diagnostic precision on them because it's still computed [. . .]

you see, like all of these segmentation defects. So like it's not going to be perfect, it's nice, but like even then this part, this wouldn't happen in a liver like in real life" (participant D).

4.6.2. Unclear Added Value

This sub-theme describes the general opinion that most participants shared to some degree that, while they can acknowledge some minor benefits in niche areas or for specific tasks, they cannot see enough value in AR technologies for radiological workflows to overcome the barriers they face and issues they introduce. Participants talked about the time that using this technology could add, and participants felt they would be quicker with traditional methods: "If you were going through these lungs, you would have done it on the screen within you know seconds. Whereas with this, I get there would be a little bit of practise, but it's also. I've got to click three times to try and get it to move in the right direction. You know, you would have looked through it by that point. So I think it- It's not fast by any means" (participant F); "There's nothing here that would speed up my reporting [. . .] If I'm honest, it seems like this, which, you know, just add time for these applications" (participant A); and "that obviously took a few minutes, partly because of the technical difficulties with actually getting the pin, but it does just feel like it would be quicker to just click on two spots" (participant E).

More broadly than speed, there was a generally appreciation of a lot of potential in the technology but also that it did not add anything compared to traditional methods. "why would that [a 3D model to rotate] be better and actually physically moving around the back rather than just having both views already available?" (participant C); "Again, what's the difference with a screen? [. . .] unless it was a hologram in the middle and it made everyone be able to see it easier. But then one person's looking at one side and one person's looking at the other. So I don't know what- you probably don't gain. Unless it made viewing easier for those people. Or everyone has an individual hologram but then everyone could have an individual screen as opposed to the individual hologram" (participant F); and "having to decide what exactly it is that adds the value. Using 3D printing as the analogy, we haven't really found a niche whereby that improves efficiencies, safety, or whatever else on a case by case basis or in a collective kind of case series. As far as I'm aware, to justify it being embedded into practise" (participant G). This lack of seeing any direct path for AR into clinical practice was unanimous across participants. This links to the following sub-theme as a possible cause of the opinions covered in this sub-theme.

4.6.3. Radiologists Think in 2D

Across the interviews there was an implicit indication, at times becoming more explicit, that radiologists think in 2D and they build up their own 3D models in their heads from an understanding of the 2D slices. "we're constantly moving the images because basically what we're doing as radiologists is that we're creating this model in our brain" (participant D); "But if you learnt this sort of stuff early on and started to think in 3D in medical school [. . .]" (participant E); and a further suggestion that the virtual 3D models used in the workshop were 'too real' for a radiologist, "if you look at the American journals, they love 3D reconstructions of pathology. But that's not the realm that the radiologist works in, they don't work with the raw image, but they don't work with this sort of like 3D, as if it's real life either they work with the sort of the grey area between and this is trying too much to be what the surgeon sees" (participant C).

To strengthen this point it was made clear that older, more experienced radiologists are less inclined to use currently available 3D tools, "for the CT colons with a 3D volume rendering. Not everyone uses it, even though technically, as per the guidelines they're then meant to at least review it. But because most people but especially the older radiologists who've been practising for a decade or two before you had readily accessible 3D volume rendering, they're just way much more accustomed to dealing with the 2D imaging" (participant B). And this in turn suggests what participant A put well, "this feels more intuitive for somebody who doesn't have much experience of radiology", suggesting that the 3D nature would be of much more value to, for example, surgeons or patients. "I also think the surgeons would quite like it because again, they're used to actually dealing with real lungs, not a slice of lung. And so if you're able to show them create something using this that they can then interact with, that might be quite good for a surgeon" (participant E); "I have to remember that I've had 15 years of looking at 2D slices, so it's very difficult. Whereas someone coming and having to look at 2D slices or looking at this, I suspect they'd like this because the anatomy is easier to distinguish. So for someone- for a surgeon who isn't maybe used to looking through 2D slices, they always say they prefer 3D rendered things" (participant F); and "you don't need to be a radiologist to see there's something going on here" (participant A).

4.6.4. Calling for Familiar PC-like Features

As is likely to be expected, there was a lot of call for carry-over of tools and capability that currently exist within the systems used for radiological reporting. These things ranged from specific tools such as the ability to cut away or segment off specific irrelevant anatomy to provide the desired view to windows snapping into place like they do on Windows when you pull a window against the side of the monitor. While some of these things were complaints about the AR not doing something the participant wanted it to do, others were commending the ability to use the space more effectively. For example, the ability to move the model into the each of the three planes. Despite not being limited to the three standard planes, generally, participants chose to view the models roughly in line with one of these planes, but there was a comment about both the ease and the hassle of moving between planes. Participant E said "it's quite easy to put into coronal and axial planes and again, just because that's the way I'm used to looking at stuff I'm trying, I am just defaulting to try and put it in kind of sagittal and kind of axial, it's quite nice to just to be able to grab and flick through and look at what I want to look at", but participant C said "maybe having some sort of snapping into place or the menu that was to my left having you know your axial, coronal, sagittal views and then it just sort of snapping into that".

The specific points participants mentioned could likely be used as part of wider data gathering when defining requirements for an AR system for radiological reporting, but for this study, the interesting part of this theme is more subtle. This calling for PC-like features could be viewed as a lack of imagination in the participants, or the participants not embracing the scope of the capability of an AR tool. However, it could also point to how refined an experience users receive from PCs because of their pervasiveness and how long researchers and designers have spent developing interactions with keyboard, mouse, and monitor. Keyboard and mouse are ubiquitous, and this sub-theme acknowledges that if AR is going to have a positive impact in this space, it needs to complement this capability rather than expect to immediately replace it.

5. Discussion

The five themes above illustrate the attitudes and opinions of the participants throughout the interviews regarding AR interactions being harnessed and integrated into their workflows. There are several ideas running across the themes as well as contradicting opinions and views. This study has explored the opportunities and challenges of AR in this area and has laid out the findings as the five themes above. In this section, these themes are mapped to three design considerations that stand as this study's contribution and give support to future work in this area. In offering these design considerations, we answer our research questions and contribute actionable steps forward in future work.

5.1. Interaction Development

Our analysis suggests that the advantage of AR tools in radiology hinges around the interactions AR affords, and this manifests as enhanced comprehension of complex 3D structures and improved communication. In addition, there was acknowledgement of potential value as a tool to answer a few niche clinical questions. It is clear across all three of these that while there is potential advantage; it is a small advantage but one rooted in the interactions. The difficult object manipulation was a subject extensively discussed with each participant and was a key source of discontent, but despite this, value rooted in the interactions came through. This was clear from the qualitative reflexive thematic analysis findings and is given supporting context by the low Mental Demand but high Effort and Frustration components of the NASA-TLX. From this, we contend that the value from AR interactions lies in waiting and needs application-specific interaction research to design systems that will allow radiologists to exploit the advantages of the technology in their work.

Based on the conversations with participants in the interviews and the reflexive thematic analysis, we suggest here three requirements that future interaction designers should incorporate into AR tools for clinical radiology. Firstly, the distinction between intended and unintended movement. This is mostly a reference to any tremor or unconscious arm movement that, with current interactions, is passed on to the virtual object being manipulated. This factor was present with the HoloLens and the Quest with controllers, and it was one inhibiter to the required precision and thus performance. Secondly, the selection and deselection of objects must be guaranteed. Participants struggled just to click buttons when using the HoloLens and even more when trying to pick up or put down virtual objects they wanted to manipulate. This was improved with the Quest controllers due to the presence of a physical button, but this then makes moving between the AR world and a PC keyboard and mouse more difficult, and there was still no haptic feedback on successful selection. Finally, smaller movements must be easier to achieve to enable the required precision. In part due to the two previous issues, small object selection and manipulation were by far the most difficult for all participants and this was noted as a restriction on the precision that was needed.

Some of these have precedence in other medical contexts, principally the Da Vinci surgical robot. This was raised in the interviews as a potential source of inspiration as one of the things the Da Vinci robot does is translate large unstable movements of the operators hands into small, controlled, deliberate movements of the instruments. If these three requirements can be achieved for AR object manipulation, with inspiration taken from cases like the Da Vinci robot, radiological use of AR technology will likely be an important step closer to realisation. This is reflected more generically in the AR literature with the open question around efficient text input in AR.

Developing these interactions to exploit the value acknowledged in this study, namely, the comprehension of complex 3D anatomy and the ease of communication, could take AR a step closer to being ready to deploy in clinical radiology. Value was recognised by radiologists in this study while using an older HMD and an application not designed specifically for radiology. Carrying forward the design recommendations stated here and incorporating them into a radiology-specific AR tool could amplify the value for the clinical radiologists. This first design implication then calls for application-specific interaction design incorporating the three aspects stated here, distinction between intended and unintended movement, selection and deselection of objects must be guaranteed, and smaller movements must be easy to achieve.

5.2. A View for Standardisation and Integration

This study was conducted primarily with the HoloLens which, while once the market leader, is now an outdated device. The Meta Quest 3 was also used and which has stepped towards resolving many of these issues such as brightness and virtual object resolution. A key reason for using two HMDs in this study was to attempt to isolate the results away from any one particular device, and while the results show that the Quest was an improvement in many ways, it is not perfect, and the more fundamental question of 'what is the value AR brings to radiology' is still very much unanswered. This study cannot answer this question here but our analysis suggests that the root of any benefit is in the comprehension of complex 3D structures as a result of the interactions afforded. Communication is an immediate secondary benefit of this enhanced comprehension. If this is the core value provided by AR in clinical radiology, the next steps should be in working out how to exploit this value effectively and remove the obstacles to this. To look at this more practically, the cost–benefit ratio of introducing AR technology into some part of a clinical radiological workflow is too low. The value gained from the technology is small and only in a niche number of scenarios, and then there is not only the cost of buying the devices and the training but also the time taken to learn how to use the technology. Even once integrated, the time and effort taken to put on a HMD, boot up, load the correct case, and set everything up is far too high of a cost for the value gained from a slightly better understanding of the anatomy.

The hardware restrictions detailed in the results are applicable to any precision task in AR and the reasons they were deemed unacceptable by participants for clinical radiology were explored in the results as well. Aspects such as these will inevitably be improved with following generations of HMD but to ensure this happens these requirements of higher resolution, wider FOV, and using VST to achieve true occlusion must be acknowledged by manufacturers and designers. Two of the newest HMD offerings, the Apple Vision Pro and the Samsung Galaxy XR, may have already improved on some of these issues but here we acknowledge the high standard of HMD that is required and the ability that must be afforded. From this, we call for a set of standards to be developed that recognises the hardware requirements such as resolution and brightness for AR HMDs for radiological use. Much as the reporting PC monitors currently used must adhere to standards of resolution and luminance, amongst others, similar standards should be developed to definitively ascertain when HMDs have reached an acceptable level of capability [68]. It may be that the Apple Vision Pro and/or the Samsung Galaxy XR have reached some of these standards already, but having a definitive quantitative yardstick to aim for greases the wheels for adoption.

Here, then, is the second design consideration. In tandem with developing the object manipulation, the other side of the interaction dialogue should be considered and brought up to par. As discussed, this satisfactory level may have been reached by state of the art HMDs, but standardisation of appropriate hardware requirements for HMDs used in clinical radiology is required in order to provide a definitive yardstick when developing in this space. While the question of the value AR provides in radiology will not be answered by this it removes a distracting barrier in this investigation and therefore allows freer research of this more difficult question. This will aid targeted research aiming to achieve AR adoption in radiology.

5.3. An AR-First World

Here, we build on an AR-first world discussed in Hobbs et al. [54], where an AR headset could be an extension to a desktop environment, the discussed limitations (of cost, learning curve, and clinical practice adjustment) are negated, and the cost–benefit ratio of AR in these situations becomes more amenable. If HMDs like the Apple Vision Pro and the Samsung Galaxy XR help mitigate some of the interaction issues and hardware constraints, there is still the question of integration. We acknowledge here the clinical proof that would be required to use HMDs like these in practice but take this as a given for the purposes of this speculation exercise.

The Apple Vision Pro gives us an example of what an AR-first world could look like and how AR could integrate with an office environment with the Mac virtual display capability. This feature lets the user use the HMD as a wireless, spatial external monitor for the Mac. Once connected, the Mac's screen appears as a large, floating display in the virtual environment that they can resize, reposition, and view in different aspect ratios. The key point here for interactions is the user can still control the Mac using the keyboard, trackpad, or mouse, while having the immersive, adjustable workspace provided by Vision Pro. This then offers an expansive workspace ideal for multitasking. Initial minor benefits for radiologists include the private nature and the portability which could lend to a more flexible working environment. This would also likely eradicate inconsistencies from ambient conditions in the environment around the user, as discussed in the literature [36]. However, the experience still depends on wireless stability and user comfort for extended use.

We extend the AR-first world described in Hobbs et al. [54] as we can speculate a world where the AR HMD is the sole display, untethered but harnessing the computing power of a PC. This removes any time lost due to switching device, putting on and setting up the HMD to view a case is carried out only when one would log into a PC, and there is no added friction to using the AR device. The other major advantage that could address some of the object manipulation interaction issues discussed at length in our analysis is the use of the normal keyboard and mouse. The familiarity, flexibility, accuracy, and precision of the keyboard and mouse can be harnessed in this new environment instead of introducing an extended learning curve of getting used to a lot of new hand gestures. Further than this, as the virtual display can be used as a normal monitor for many tasks (for example, sending emails), the only difference is the viewing medium; the act of sending an email is unchanged. This can then be continued for a reporting workflow for a radiologist. Any tasks that are better with 2D slices can continue to be carried out so with a keyboard and mouse on a '2D screen' and with as many screens as the particular radiologist wants with no further cost. However, a 3D representation can be ready to run alongside these 2D images and can be referred to when the clinical question necessitates it or when the radiologist would benefit from viewing some complex part of the anatomy in its 3D form. Further tools could then be built on top of this as the utility of a 3D model in this environment is explored, with tools such as the ability to measure structures added.

The Apple Vision Pro sets the precedence for online communication in this way as well with its integration of Facetime. As a lot of the communication a radiologist performs, such as participating in MDT meetings, is carried out online, this integration of this part of the workflow into AR is clear. Integrating AR in this way then could remove many of the blockers that currently mean the cost–benefit of using AR technology is too low.

However, there are questions that remain. As previously mentioned, in the interviews, using the Quest made one of the participants feel sick. This leads to the bigger question of are these devices usable all day every day? This raises issues with everything from the weight and therefore comfort of the device, if it is worn for eight hours a day, to the impact on a human's vision if using an AR device for a whole career, to the question of being made to feel sick. If any of these potential issues hold true in an AR-first world, questions of accuracy in the radiologists' work will be raised. These issues of comfort and extended use are discussed in the literature, but it is important to consider the implications of these issues for this medical setting.

Additionally, and opposed to the privacy advantage mentioned above, an AR-first world has the potential to force communication online. In a world where no standard monitors are used, if another clinician were to come into a radiologists office in-person ad hoc and ask a question about a case, there would then be no way for the radiologist to share the images with the clinician to have a discussion. There is then the potential impact of forcing communication online to consider.

As suggested in [54], there are the major practicalities of moving to a new technology to consider, with a large upfront monetary and time cost. There is the initial strain of rewriting procedures around the new technology and the learning curve of engaging with the new medium, as well as all of the software that is currently used having to be ported to be AR compatible. As a lot of software is provided by private companies and many NHS trusts use software from multiple different providers; this is no small task. However, if the value is proven, this could be an achievable activity.

Here then we put forward the final design consideration, that of versatility and integration. It is integration that will be the real test of AR technology in this space. Doctors are resistant to change, medical regulations are strict, and money is tight. The value of AR could be maximised by its wide application, which in turn also lowers the obstacles, such as the effort of putting on and booting up an HMD. Many of the technical and application-based issues raised in the interviews are approachable with intelligent, collaborative, iterative design. But tackling the integration of the technology into any radiological workflow will be the make or break.

5.4. Limitations

This qualitative analysis aims to provide a representative insight into the views and opinions of radiologists in the United Kingdom on the role AR could play in radiological imaging. However, we must acknowledge the limitations of both the methodology and the dataset.

Our participants were radiologists and radiology registrars from NHS trusts in the North East of England. We successfully recruited a range of participants with a range of specialisms to provide a variety of views and differing contexts, which adds strength and breadth to this work. However, a potential shortcoming of this participant pool was our ability to only recruit one woman, which could introduce bias. Where possible, we took appropriate steps to try and recruit women, but in part due to this being a very male-dominated field, we were unable to, and this disparity is represented in the radiological workforce [51,52]. A convenience sampling approach was used, appropriate for a qualitative study with a demographically and geographically local sample universe [53]. One consequence of this approach is that our sample skews toward relatively junior participants with limited prior XR experience, and it does not include more senior radiologists with extensive clinical experience or participants with moderate-to-high familiarity with XR technology. Rather than treating this as a defect to be corrected through purposive or persona-based sampling, we treat it as a boundary on the sample universe itself. Our findings speak to the perspectives of this particular group of radiologists and registrars, not to radiologists in general, and any transferability beyond this group should be assessed by the reader with this composition in mind. We did not set out to construct a sample representative of the full range of experience and XR familiarity across the profession, and we do not present our findings as such. Future work with a larger, more purposively or persona-based sampled cohort deliberately including more experienced clinicians, greater gender balance, and participants with moderate-to-high prior XR exposure, drawn from a wider geographic pool, would be a valuable complement to this study and could test whether the patterns identified here hold, change, or diminish across a broader range of radiological experience.

A further limitation concerns the comparison between the HoloLens 2 and Meta Quest 3. Each device was paired with a different application and set of activities, meaning hardware and software were confounded. Consequently, observed differences cannot confidently be attributed to the headsets themselves and may instead reflect differences in the applications, activities, or their interaction with the hardware. The findings should therefore be interpreted as comparisons of device-application pairings rather than headset capability alone. A future study with controlled comparison using the same application and equivalent activities would be required to isolate hardware effects.

This study used GigXR's DICOM XR Library, an educational tool not designed to replicate a diagnostic workflow or match the feature set of clinical-grade DICOM PACS software. As discussed in the Methodology, this was appropriate to the study's aim of eliciting discussion around AR's interaction value, rather than evaluating a candidate diagnostic tool. However, this means our findings speak to participants' experience of this particular application, and any broader claims about AR's suitability for radiological practice should be read with this software choice in mind. Findings such as those relating to model segmentation accuracy should therefore be understood as observations about this specific tool, not as general claims about the segmentation capabilities achievable in a purpose-built radiological AR application.

Overall, these limitations primarily affect the transferability and interpretation of the findings rather than their value as exploratory qualitative insights. The range of specialisms provides breadth, but the limited gender representation, regional sample, and hardware-software confounding mean that conclusions should be interpreted within the context of this study.

6. Conclusions

In this paper, we have presented the results of a thematic analysis of interviews with radiology consultants and registrars in order to qualitatively investigate the impact of AR interactions for the integration into clinical radiology workflows. We contribute three design considerations for AR systems within radiological imaging workflows based on the results of this qualitative analysis, situating them within the literature and giving direction for future research. The interviews described here were an opportunity to join the technology with the target users, and the qualitative approach allowed rich discussion on the important interaction and integration challenges.

Responding to our research questions, the clearest opportunity identified was in supporting comprehension of complex 3D anatomy, with more limited potential for efficiency gains or advantage over current systems. The most significant barriers identified were object manipulation and workflow integration. Building on this, we contribute an empirical understanding of attitudes towards AR interactions in clinical radiology in its current form. We highlight how radiological AR requires application-specific interaction design, particularly in being able to distinguish intended from unintended movement. We call for the standardisation of hardware requirements for clinical-grade HMDs, and in doing this we call for specific work addressing both sides of the interaction dialogue. Finally, we underscore the value of versatile, workflow-integrated deployment to overcome resistance to clinical adoption. Our speculation of an AR-First world illustrates key requirements and demonstrates these design considerations in a way that future work could take forward.

Author Contributions: Conceptualization, J.H. and C.B.; methodology, J.H. and C.B.; formal analysis, J.H.; investigation, J.H.; data curation, J.H.; writing—original draft preparation, J.H.; writing—review and editing, J.H. and C.B.; visualization, J.H.; supervision, C.B.; project administration, J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and the ethical approval granted from Newcastle University, reference: 61692/2023, approval date 17 June 2025.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw dataset from this study will not be made publicly available due to anonymity concerns.

Acknowledgments: The authors thank GigXR for providing free access to the DICOM XR Library and HoloHuman+ software used in this work.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Interview Procedure

Table A1. Interview procedure.

Content	Notes
Preliminaries • PRESS RECORD • Welcome etc. • Provisional questionnaire done?	Common

1. Intro
Workshop format
• Four parts (intro, HoloLens, Quest, outro).
• This is a tool (or two) developed by GigXR.
  – It's not designed for this specifically but it's quite flexible so should serve the purpose.
• I don't have any medical training—forgive my ignorance.
• When wearing the headsets, I can see what you are seeing, albeit with a bit of a delay. If you want to point at anything while you're talking to me, go ahead.
• You're not limited to staying in this chair or facing the desk at any point.
Study outline
• Two studies
• Aim: To qualitatively investigate the impact of AR interactions for the integration into radiological reporting, and radiological training
• Focussing on the interactions, and the integration of the technology onto these tasks
  – Thinking about the current workflows, benefits, concerns, limitations
• Will be using some classic HCI techniques throughout
	Common

Table A1. *Cont.*

Content	Notes

The HoloLens
• How to put it on
• Volume & brightness buttons
• Visor tips up
• Readjust IPD?
• Overview of interactions
  – This is how you select things—be overly positive with actions
• Tips App
  – Hand gestures
• Questions?
	Common

2. GigXR DICOM Library • Navigate to it • Log in • Open model 1	

Model 1—Have a play
• Think aloud
  – What are you doing?
  – What are you thinking?
  – Why did you do that?
  – What were you expecting to happen?
  – What are you trying to do?
  – Was that what you expected to happen?
  – What are you confused about?
  – Can you describe your thought process?
  – Tell me more about that.
• Introduce cutting plane, markers, segmentation etc.
	Common

Model 2—Questions
• Questions from the facilitator guide
• e.g., What is the abnormal finding?
  – There is a mass in the left occipital lobe, right parietal lobe, and right frontal lobe, with significant deformation of the right lateral ventricle.
  – etc.
• Can you explain this to me as if you were looking at this with one of your registrars?
• Can you explain this to me as a non-medical professional?
• Does having the 3D model change how you would explain these points to either group? How so?
• How was it using the app to get the answers to these questions?
• Were there any issues or struggles?
• Given time to practice with the app do you see your views changing?
• What would be needed to improve it?
• How do you think it would affect the speed or accuracy you could report a scan?
	These questions are targeted to either the clinical professional or a radiological trainer/trainee

Table A1. *Cont.*

Content	Notes
Model 3—Measuring with markers • Measure this with markers • Think aloud • Thinking about the training side, what do you think you would have made of this if it had been used in your training? • Is there a particular level of trainee you think it would be more applicable to, or more or less useful for? • Do you think using this would have helped your comprehension of anatomy (or other) during your training?	Questions targeted towards training. Task relevant to clinical professionals and discussion open around this.
Model 4—Email a colleague • Email a colleague (say a neurologist?) a short report of this scan. • (Halfway through) copy in a trainee and add some extra description for them.	Clinical professional focus.
3. The Meta Quest 3 • Quick intro, interactions, using the controllers • HoloHuman+ is wrong for the application but tell me your experiences using the Quest after the HoloLens – Have a play with the model in front of you but tell me what you're thinking. – Tell me what it's like to use one of the small finger bones to point at other parts of the model.	Common
4. Outro • Return headsets • NASA-TLX • Closing comments? • Do you think there are any questions I didn't ask or angles I didn't explore? • Can you refer me to colleagues?	Common (NASA-TLX clinical focus only)

References

1. The Royal College of Radiologists. *Clinical Oncology Census Reports*; The Royal College of Radiologists: London, UK, 2024. Available online: https://www.rcr.ac.uk/news-policy/workforce-censuses/2025-clinical-oncology-workforce-census-report/2024-clinical-oncology-workforce-census-report/ (accessed on 30 July 2026).
2. Hughes, C.L.; Karwowski, W. Evaluating interaction design and user experience in augmented reality: A systematic review. *Front. Virtual Real.* **2025**, *6*, 1710161. [CrossRef]
3. GigXR. *GigXR: Holographic Healthcare Training*; GigXR: Santa Monica, CA, USA, 2024.
4. Alhadreti, O.; Mayhew, P. Rethinking Thinking Aloud: A Comparison of Three Think-Aloud Protocols. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, New York, NY, USA*; ACM: New York, NY, USA, 2018; CHI '18, pp. 1–12. [CrossRef]
5. Braun, V.; Clarke, V. Reflecting on reflexive thematic analysis. *Qual. Res. Sport. Exerc. Health* **2019**, *11*, 589–597. [CrossRef]
6. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
7. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*; Advances in Psychology; Hancock, P.A., Meshkati, N., Eds.; North-Holland: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183. [CrossRef]
8. Cao, A.; Chintamani, K.K.; Pandya, A.K.; Ellis, R.D. NASA TLX: Software for assessing subjective mental workload. *Behav. Res. Methods* **2009**, *41*, 113–117. [CrossRef] [PubMed]
9. The Royal College of Radiologists. *Clinical Oncology Census Reports*; The Royal College of Radiologists: London, UK, 2023. Available online: https://www.rcr.ac.uk/news-policy/latest-updates/2023-clinical-radiology-and-clinical-oncology-workforce-census-reports/ (accessed on 30 July 2026).

10. Dhar, P.; Rocks, T.; Samarasinghe, R.M.; Stephenson, G.; Smith, C. Augmented reality in medical education: Students' experiences and learning outcomes. *Med. Educ. Online* **2021**, *26*, 1953953. [CrossRef] [PubMed]
11. Machała, S.; Chamier-Gliszczyński, N.; Królikowski, T. Application of AR/VR Technology in Industry 4.0. *Procedia Comput. Sci.* **2022**, *207*, 2990–2998. [CrossRef]
12. Paulson, S.; Figueiredo, D.; Matre, S. Understanding perception of the radiology community concerning virtual reality (VR) and augmented reality (AR) technology in radiology education. *Egypt. J. Radiol. Nucl. Med.* **2025**, *56*, 12. [CrossRef]
13. Hobbs, J.; Bull, C. Designs and Interactions for Near-Field Augmented Reality: A Scoping Review. *Informatics* **2025**, *12*, 77. [CrossRef]
14. Vertucci, R.; D'Onofrio, S.; Ricciardi, S.; De Nino, M. History of Augmented Reality. In *Springer Handbook of Augmented Reality*; Nee, A.Y.C., Ong, S.K., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2023; pp. 35–50. [CrossRef]
15. Mendoza-Ramírez, C.E.; Tudon-Martinez, J.C.; Félix-Herrán, L.C.; Lozoya-Santos, J.d.J.; Vargas-Martínez, A. Augmented Reality: Survey. *Appl. Sci.* **2023**, *13*, 10491. [CrossRef]
16. Iqbal, A.I.; Aamir, A.; Hammad, A.; Hafsa, H.; Basit, A.; Oduoye, M.O.; Anis, M.W.; Ahmed, S.; Younus, M.I.; Jabeen, S. Immersive Technologies in Healthcare: An In-Depth Exploration of Virtual Reality and Augmented Reality in Enhancing Patient Care, Medical Education, and Training Paradigms. *J. Prim. Care Community Health* **2024**, *15*, 21501319241293311. [CrossRef] [PubMed]
17. Kukla, P.; Maciejewska, K.; Strojna, I.; Zapał, M.; Zwierzchowski, G.; Bąk, B. Extended Reality in Diagnostic Imaging—A Literature Review. *Tomography* **2023**, *9*, 1071–1082. [CrossRef] [PubMed]
18. Sun, P.; Zhao, Y.; Men, J.; Ma, Z.R.; Jiang, H.Z.; Liu, C.Y.; Feng, W. Application of Virtual and Augmented Reality Technology in Hip Surgery: Systematic Review. *J. Med. Internet Res.* **2023**, *25*, e37599. [CrossRef] [PubMed]
19. Eckert, M.; Volmerg, J.S.; Friedrich, C.M. Augmented Reality in Medicine: Systematic and Bibliographic Review. *JMIR mHealth uHealth* **2019**, *7*, e10967. [CrossRef] [PubMed]
20. Edwards, P.J.; Chand, M.; Birlo, M.; Stoyanov, D. The Challenge of Augmented Reality in Surgery. In *Digital Surgery*; Atallah, S., Ed.; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; pp. 121–135. [CrossRef]
21. Morimoto, T.; Kobayashi, T.; Hirata, H.; Otani, K.; Sugimoto, M.; Tsukamoto, M.; Yoshihara, T.; Ueno, M.; Mawatari, M. XR (Extended Reality: Virtual Reality, Augmented Reality, Mixed Reality) Technology in Spine Medicine: Status Quo and Quo Vadis. *J. Clin. Med.* **2022**, *11*, 470. [CrossRef] [PubMed]
22. Speicher, M.; Hall, B.D.; Nebeling, M. What is Mixed Reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, New York, NY, USA*; ACM: New York, NY, USA, 2019; CHI '19; pp. 1–15. [CrossRef]
23. Vinci, C.; Brandon, K.O.; Kleinjan, M.; Brandon, T.H. The clinical potential of augmented reality. *Clin. Psychol. Sci. Pract.* **2020**, *27*, e12357. [CrossRef] [PubMed]
24. Milgram, P.; Kishino, F. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Inf. Syst.* **1994**, *77*, 1321–1329.
25. Skarbez, R.; Smith, M.; Whitton, M.C. Revisiting Milgram and Kishino's Reality-Virtuality Continuum. *Front. Virtual Real.* **2021**, *2*, 647997. [CrossRef]
26. Douglas, D.B.; Wilke, C.A.; Gibson, J.D.; Boone, J.M.; Wintermark, M. Augmented Reality: Advances in Diagnostic Imaging. *Multimodal Technol. Interact.* **2017**, *1*, 29. [CrossRef]
27. Pelargos, P.E.; Nagasawa, D.T.; Lagman, C.; Tenn, S.; Demos, J.V.; Lee, S.J.; Bui, T.T.; Barnette, N.E.; Bhatt, N.S.; Ung, N.; et al. Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. *J. Clin. Neurosci.* **2017**, *35*, 1–4. [CrossRef] [PubMed]
28. Elsayed, M.; Kadom, N.; Ghobadi, C.; Strauss, B.; Al Dandan, O.; Aggarwal, A.; Anzai, Y.; Griffith, B.; Lazarow, F.; Straus, C.M.; et al. Virtual and augmented reality: Potential applications in radiology. *Acta Radiol.* **2020**, *61*, 1258–1265. [CrossRef] [PubMed]
29. Dallas-Orr, D.; Penev, Y.; Schultz, R.; Courtier, J. Comparing Computed Tomography–Derived Augmented Reality Holograms to a Standard Picture Archiving and Communication Systems Viewer for Presurgical Planning: Feasibility Study. *JMIR Perioper. Med.* **2020**, *3*, e18367. [CrossRef] [PubMed]
30. Trestioreanu, L.; Glauner, P.; Meira, J.A.; Gindt, M.; State, R. Using Augmented Reality and Machine Learning in Radiology. In *Innovative Technologies for Market Leadership: Investing in the Future*; Glauner, P., Plugmann, P., Eds.; Future of Business and Finance; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 89–106. [CrossRef]
31. Trevia, R.; Pinho, M.S. Virtual Reality in Radiology: A Systematic Mapping Study. In *Proceedings of the 23rd Symposium on Virtual and Augmented Reality*; Association for Computing Machinery: New York, NY, USA, 2022; SVR '21; pp. 177–181. [CrossRef]
32. Locuson, M.D.; Lecakes, G.D.; Aita, A.; Goldman, H.W.; Mandayam, S.; Lalovic-Hand, M. A virtual scalpel for visualizing patients in a three-dimensional, immersive, navigable and interactive virtual reality environment. In *Proceedings of the 2015 IEEE Sensors Applications Symposium (SAS)*; IEEE: New York, NY, USA, 2015; pp. 1–4. [CrossRef]
33. Wirth, M.; Gradl, S.; Sembdner, J.; Kuhrt, S.; Eskofier, B.M. Evaluation of Interaction Techniques for a Virtual Reality Reading Room in Diagnostic Radiology. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*; ACM: New York, NY, USA, 2018; pp. 867–876. [CrossRef]

34. Card, S.K.; Mackinlay, J.D.; Robertson, G.G. The design space of input devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Empowering People—CHI '90'*; ACM Press: New York, NY, USA, 1990; pp. 117–124. [CrossRef]
35. Samei, E.; Badano, A.; Chakraborty, D.; Compton, K.; Cornelius, C.; Corrigan, K.; Flynn, M.J.; Hemminger, B.; Hangiandreou, N.; Johnson, J.; et al. Assessment of display performance for medical imaging systems: Executive summary of AAPM TG18 report. *Med. Phys.* **2005**, *32*, 1205–1225. [CrossRef] [PubMed]
36. Sousa, M.; Mendes, D.; Paulo, S.; Matela, N.; Jorge, J.; Lopes, D.S. VRRRRoom: Virtual Reality for Radiologists in the Reading Room. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery*; ACM: New York, NY, USA, 2017; CHI '17; pp. 4057–4062. [CrossRef]
37. Łoginoff, J.; Augustynowicz, K.; Świąder, K.; Ostaszewska, S.; Morawski, P.; Pactwa, F.; Popińska, Z. Advancements in Radiology and Diagnostic Imaging. *J. Educ. Health Sport* **2023**, *33*, 45–51. [CrossRef]
38. Curran, V.R.; Xu, X.; Aydin, M.Y.; Meruvia-Pastor, O. Use of Extended Reality in Medical Education: An Integrative Review. *Med. Sci. Educ.* **2023**, *33*, 275–286. [CrossRef] [PubMed]
39. Jung, C.; Wolff, G.; Wernly, B.; Bruno, R.R.; Franz, M.; Schulze, P.C.; Silva, J.N.A.; Silva, J.R.; Bhatt, D.L.; Kelm, M. Virtual and Augmented Reality in Cardiovascular Care. *JACC Cardiovasc. Imaging* **2022**, *15*, 519–532. [CrossRef] [PubMed]
40. Erkelens, I.M.; MacKenzie, K.J. 19-2: Vergence-Accommodation Conflicts in Augmented Reality: Impacts on Perceived Image Quality. *SID Symp. Dig. Tech. Pap.* **2020**, *51*, 265–268. [CrossRef]
41. Gabbard, J.L.; Mehra, D.G.; Swan, J.E. Effects of AR Display Context Switching and Focal Distance Switching on Human Performance. *IEEE Trans. Vis. Comput. Graph.* **2019**, *25*, 2228–2241. [CrossRef] [PubMed]
42. Sielhorst, T.; Bichlmeier, C.; Heining, S.M.; Navab, N. Depth Perception—A Major Issue in Medical AR: Evaluation Study by Twenty Surgeons. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006*; Larsen, R., Nielsen, M., Sporring, J., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2006; pp. 364–372. [CrossRef] [PubMed]
43. Cutolo, F.; Fida, B.; Cattari, N.; Ferrari, V. Software Framework for Customized Augmented Reality Headsets in Medicine. *IEEE Access* **2020**, *8*, 706–720. [CrossRef]
44. Pham, D.M.; Stuerzlinger, W. Is the Pen Mightier than the Controller? A Comparison of Input Devices for Selection in Virtual and Augmented Reality. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology, New York, NY, USA*; ACM: New York, NY, USA, 2019; VRST '19, pp. 1–11. [CrossRef]
45. Kim, H.; Kwon, Y.T.; Lim, H.R.; Kim, J.H.; Kim, Y.S.; Yeo, W.H. Recent Advances in Wearable Sensors and Integrated Functional Devices for Virtual and Augmented Reality Applications. *Adv. Funct. Mater.* **2021**, *31*, 2005692. [CrossRef]
46. Kang, H.J.; Shin, J.h.; Ponto, K. A Comparative Analysis of 3D User Interaction: How to Move Virtual Objects in Mixed Reality. In *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*; IEEE: New York, NY, USA, 2020; pp. 275–284. [CrossRef]
47. Prilla, M.; Janssen, M.; Kunzendorff, T. How to Interact with Augmented Reality Head Mounted Devices in Care Work? A Study Comparing Handheld Touch (Hands-on) and Gesture (Hands-free) Interaction. *AIS Trans. Hum. Comput. Interact.* **2019**, *11*, 157–178. [CrossRef]
48. Argelaguet, F.; Andujar, C. A survey of 3D object selection techniques for virtual environments. *Comput. Graph.* **2013**, *37*, 121–136. [CrossRef]
49. Wang, Y.; Wang, Y.; Chen, J.; Wang, Y.; Yang, J.; Jiang, T.; He, J. Investigating the Performance of Gesture-Based Input for Mid-Air Text Entry in a Virtual Environment: A Comparison of Hand-Up versus Hand-Down Postures. *Sensors* **2021**, *21*, 1582. [CrossRef] [PubMed]
50. Hobbs, J.; Bull, C. Qualitatively Investigating the Impact of Augmented Reality Interactions for Radiological Training. *Preprints* **2026**. [CrossRef]
51. Barnes, K.L.; McGuire, L.; Dunivan, G.; Sussman, A.L.; McKee, R. Gender Bias Experiences of Female Surgical Trainees. *J. Surg. Educ.* **2019**, *76*, e1–e14. [CrossRef] [PubMed]
52. Jessica, C.; Mary, W.; Pauline, G.; Robyn G., R. Lessons learned the hard way: Sharing experiences from female radiologists regarding gender inequality. *Curr. Probl. Diagn. Radiol.* **2025**, *54*, 40–44. [CrossRef]
53. Robinson, O.C. Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. *Qual. Res. Psychol.* **2014**, *11*, 25–41. [CrossRef]
54. Hobbs, J.; Bull, C.; Claisse, C.; Elameer, M.; Davison, R. Role of Augmented Reality in Tertiary Care: Qualitative Investigation Using Thematic Analysis. *JMIR XR Spat. Comput. (JMXR)* **2025**, *2*, e68810. [CrossRef] [PubMed]
55. Braun, V.; Clarke, V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qual. Res. Sport. Exerc. Health* **2021**, *13*, 201–216. [CrossRef]
56. Malterud, K.; Siersma, V.; Guassora, A.D. Sample Size in Qualitative Interview Studies. *Qual. Health Res.* **2016**, *26*, 1753–1760. [CrossRef] [PubMed]
57. GigXR. *GigXR DICOM XR Library*; GigXR: Santa Monica, CA, USA, 2024.

58. Karunaratne, D.; Whittock, J.; Moore, A.; Dasigan, K.; Chevolleau, J.; Bartholomew, B.; Kelly, N.; Cohen, C.E. Exploring the Use of an Augmented Reality Device Learning Tool for Multidisciplinary Staff Training on Domestic Abuse and Sexual Violence: Postintervention Qualitative Evaluation. *JMIR Form. Res.* **2025**, *9*, e60075. [CrossRef] [PubMed]
59. GigXR. *GigXR HoloHuman+*; GigXR: Santa Monica, CA, USA, 2024.
60. GigXR. *GigXR HoloHuman+ Vimeo*; GigXR: Santa Monica, CA, USA, 2024.
61. Braun, V.; Clarke, V. *Thematic Analysis: A Practical Guide*; SAGE Publications: Thousand Oaks, CA, USA, 2021.
62. Creswell, J.W.; Poth, C.N. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*; SAGE Publications: Thousand Oaks, CA, USA, 2016; Google-Books-ID: DLbBDQAAQBAJ.
63. Braun, V.; Clarke, V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qual. Res. Psychol.* **2021**, *18*, 328–352. [CrossRef]
64. Lumivero. *Nvivo by Lumivero*; Lumivero: Denver, CO, USA, 2026.
65. Figma. *FigJam: The Online Collaborative Whiteboard for Teams*; Figma: San Francisco, CA, USA, 2026.
66. Trepkowski, C.; Eibich, D.; Maiero, J.; Marquardt, A.; Kruijff, E.; Feiner, S. The Effect of Narrow Field of View and Information Density on Visual Search Performance in Augmented Reality. In *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*; IEEE: New York, NY, USA, 2019; pp. 575–584. [CrossRef]
67. Ren, D.; Goldschwendt, T.; Chang, Y.; Hollerer, T. Evaluating wide-field-of-view augmented reality with mixed reality simulation. In *Proceedings of the 2016 IEEE Virtual Reality (VR)*; IEEE: New York, NY, USA, 2016; pp. 93–102. [CrossRef]
68. The Royal College of Radiologists. *Picture Archiving and Communication Systems (PACS) and Guidelines on Diagnostic Display Devices*; The Royal College of Radiologists: London, UK, 2019.

https://doi.org/10.3390/virtualworlds5030045