informatics

MDPI

Article

Designs and Interactions for Near-Field Augmented Reality: A Scoping Review

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@newcastle.ac.uk

Abstract

Augmented reality (AR), which overlays digital content within the user's view, is gaining traction across domains such as education, healthcare, manufacturing, and entertainment. The hardware constraints of commercially available HMDs are well acknowledged, but little work addresses what design or interactions techniques developers can employ or build into experiences to work around these limitations. We conducted a scoping literature review, with the aim of mapping the current landscape of design principles and interaction techniques employed in near-field AR environments. We searched for literature published between 2016 and 2025 across major databases, including the ACM Digital Library and IEEE Xplore. Studies were included if they explicitly employed design or interaction techniques with a commercially available HMD for near-field AR experiences. A total of 780 articles were returned by the search, but just 7 articles met the inclusion criteria. Our review identifies key themes around how existing techniques are employed and the two competing goals of AR experiences, and we highlight the importance of embodiment in interaction efficacy. We present directions for future research based on and justified by our review. The findings offer a comprehensive overview for researchers, designers, and developers aiming to create more intuitive, effective, and context-aware near-field AR experiences. This review also provides a foundation for future research by outlining underexplored areas and recommending research directions for near-field AR interaction design.

Keywords: augmented reality; design; interaction; head-mounted display; near field; scoping review

check for updates

Academic Editor: Long Jin

Received: 13 June 2025
Revised: 11 July 2025
Accepted: 24 July 2025
Published: 1 August 2025

Citation: Hobbs, J.; Bull, C. Designs and Interactions for Near-Field Augmented Reality: A Scoping Review. *Informatics* **2025**, *12*, 77. https://doi.org/10.3390/informatics12030077

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Augmented reality (AR) has rapidly emerged as a technology with wide-reaching potential, enhancing the physical world by overlaying digital content in real time [1,2]. As AR applications proliferate across domains such as education, healthcare, manufacturing, and entertainment, the quality of user interactions—particularly in near-field environments—has become critical to the effectiveness of these systems [3,4]. Designing effective AR interactions presents unique challenges due to the integration of digital and physical spaces, which directly impacts user engagement, task performance, and overall experience [5].

Although modern commercially available head-mounted displays (HMDs) offer robust capabilities, their perceptual limitations are well documented (as discussed in the following section) [6–8]. While ongoing research aims to address these limitations, developers must, in the meantime, find practical strategies to work within the constraints of

current hardware [9–11]. However, there is a noticeable gap in the literature regarding the techniques accessible to developers that mitigate these perceptual issues in near-field AR interactions using commercial HMDs.

In this paper, we investigated existing strategies to overcome, adapt to, or alleviate perceptual challenges in near-field AR environments. To map the current state of knowledge in this niche area, we conducted a scoping literature review of full conference papers and journal articles published since 2016. Following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Reviews) [12,13] guidelines (the process suggested by Cooper et al. [14] and the Joanna Briggs Institue [15]), we searched four academic databases—the ACM Digital Library, IEEE Xplore, PubMed, and ScienceDirect—yielding an initial set of 780 unique items. After applying inclusion and exclusion criteria, seven papers were retained for the final analysis. We employed both deductive coding and inductive latent pattern content analysis to extract meaningful insights.

Our findings provide a categorisation and visualisation of the current literature landscape, identifying key themes and recurring challenges. We emphasise the roles of embodiment and immersion in near-field AR interactions and highlight the absence of a consistent definition for the near-field range. We propose a defined range of 25 to 100 cm for future studies investigating near-field interactions based on our analysis and offer recommendations for future research directions.

2. Background

In this section, we begin by discussing the relevant concepts and definitions before going on to address the need and rationale for this work.

2.1. Augmented Reality

Augmented reality was conceptualised as early as 1965 but has seen a boom in interest in recent years [16,17]. While the definition of AR is somewhat fluid, and there is some argument among researchers around the differences between AR and mixed reality (MR), it can generally be agreed that AR superimposes digital objects into the users' view in real time using a headset or other device [18–20]. 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. Thomas Caudell at Boeing coined the term augmented reality in 1992, but it is since around 2016 that a new and continuing wave of development has changed the face of AR [21–23].

The year 2016 saw several significant events in the field, such as the release of the first generation of Microsoft HoloLens [24], which was the first fully untethered, standalone HMD. As one of the first of the modern breeds of HMDs, HoloLens had much more freedom than previous devices and was much more flexible in a larger range of environments. Google Scholar searches including the term "augmented reality" exhibited peaks in 2016 and 2017, as shown in Figure 1. This chart shows the search interest in the given term relative to the highest point on the graph [25]. The first HTC Vive [26] virtual reality (VR) HMD and Pokemon Go! [27] were also released in 2016, with other developments such as IKEA place [28] and the Enterprise Edition of Google Glass [29] released in 2017, which shows a wider and consistent heightened interest in AR and XR (extended reality) more generally.

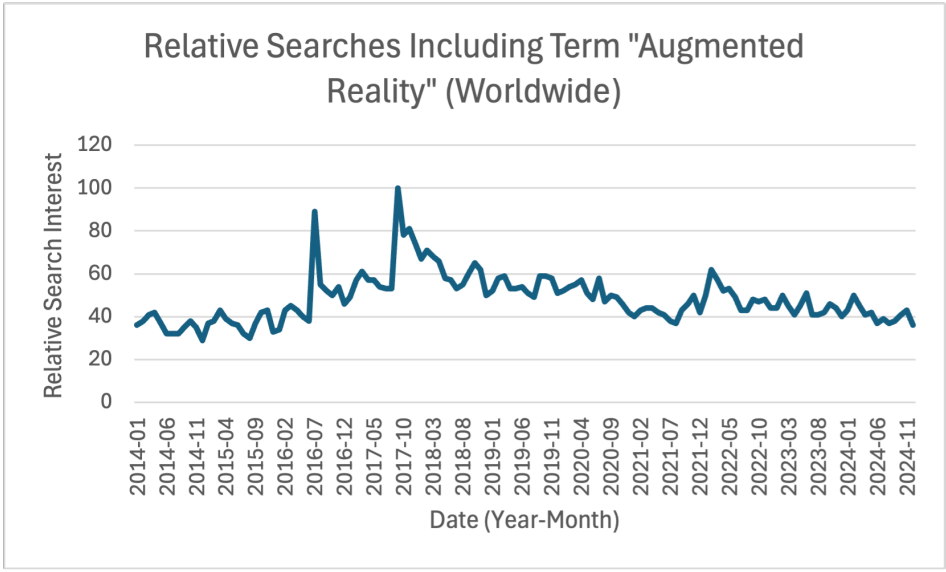


Figure 1. Google Scholar relative search interest—search interest relative to the highest point—of the term "augmented reality" in 2014–2025.

2.2. The Near Field and Perception

The interactions that AR affords can be considered one of the core advantages of the technology, and there has been considerable research in this area [30]. A subset of this area is specifically interactions in near-field ranges, that is, the immediate area around the user. However, definitions of the near field are varied and no more specific than "within arm's reach" [31,32]. Speaking more generally than XR, Cutting and Vishton [33] defined personal space in this way in 1995, going on to "arbitrarily delimit this space to be within two metres".

2.2.1. Perceptual Issues in the Near Field

Modern commercially available HMDs are very capable devices, but there are still limitations to their use [34–36]. One of these limitations that has various knock-on effects is the fixed focal depth, which, for most HMDs, is around two to three metres, meaning that, regardless of the depth that the virtual content is intended to be, it is always projected at a distance of around two metres [9,10,37]. This can induce a number of perceptual issues, principally the vergence–accommodation conflict (VAC), which is a result of the users' eyes receiving mismatched cues [10]. The eyes' two mechanisms of focusing (converging onto the target and the lenses accommodating to the same target) compete against one another and can result in discomfort, simulation sickness, and an inaccurate perception of the virtual content. This affect is more common or more pronounced when the distance between the viewing target and the two-metre focal depth is greater, for instance, in near-field interactions [38]. Focal rivalry can also cause similar discomfort in situations like this, where a user attempts to keep two objects, virtual and physical, at different depths in focus at the same time [39]. In 2013, Argelaguet and Andujar [40] stated that a major block for effective user interaction in virtual environments (VEs) is that the technology fails to faithfully reproduce the physical constraints of the real world. While improvements have been made since 2013, modern HMDs still cannot provide users with all the cues that they need to interpret the environment accurately [37].

Edwards et al. [41] provided an overview of the challenges of AR for surgery and suggested that "visualisation and issues surrounding perception and interaction are in fact the biggest issues facing AR surgical guidance". It is likely then that this can be generalised to other manual tasks with similar requirements.

2.2.2. The Challenge of Researching Perception

Perceptual issues and how design or interactions affect them are difficult to investigate for two key reasons: Firstly, there are a large number of contributors to perception that all have complex effects and interactions with each other, which makes it difficult to isolate one contributor to perception, evaluate it, and measure what affects it [7,42]. Secondly, everyone perceives things slightly differently, and the results must be empirical; only the participant of a study can see what they are seeing, and the researcher can only measure the effect of an action that the participant makes as a result of perceiving something [43,44]. This makes perception a challenging area to research.

2.2.3. Design

As stated by Cutting and Vishton [33], the strongest perceptual cue at arm's length is occlusion, followed by binocular disparity (stereopsis), motion parallax, relative size, illumination and shadows, convergence, and accommodation. However, it is generally accepted that modern HMDs cannot faithfully represent all these cues that enable a user to interpret their environment [36,38]. Various design techniques that can be used to work around these issues have been investigated, such as using virtual apertures in an occluding surface when focusing on an object beneath—X-ray vision is an example here [45]. True occlusion is not possible with optical see-through (OST) devices but can be achieved with video see-through (VST) HMDs, and reduced depth errors have been shown in VR when true occlusion is achieved [46].

Perceptual issues have some commonality between different HMDs, but some variation can be observed when moving between different display technologies, that is, OST and VST [47,48]. These two technologies have fundamentally different ways of presenting content to the user, and, as such, the perception of that content varies between the two. These perceptual differences have been investigated somewhat, but there has been much less evaluation of their effects when interactions move into the near field [47].

2.2.4. Interaction

AR interactions can be supported by modalities such as speech, gaze, hand gestures, and additional hardware such as controllers or gloves [49,50]. The body of work investigating interactions for manipulating objects in virtual environments (VEs) in the near field generally concludes that "natural", direct manipulation is preferred [51]. However, there is research supporting the contrary, and many of the studies conducted are very specific to one task [52]. There has been comparatively less generalisable evaluations, and research continues to extensively explore intuitive and immersive interactions with virtual objects [40].

In VR, studies show that increasing embodiment with high-fidelity avatars can reduce mental load, increase performance on a task, and improve depth estimation accuracy, a subject that is gaining traction in AR [53]. The extent to which avatarisation affects embodiment in AR and interaction is not fully understood, but recent research shows its potential [54].

2.3. The Rationale, Research Gap, and Need for This Review

As discussed above, interaction is a key area of XR research, with many issues clearly identified and work carried out to combat them [5,30]. However, to the best of our knowledge, there is relatively little work and no synthesis of the work around what AR application developers can do to address these issues while working within the constraints of commercial HMDs. The idea of using design and interaction techniques to optimise the use of commercial HMDs is necessary in lieu of adjacent (but likely slower) work going into

developing HMDs to overcome these issues fundamentally [11,55]. The aim of this review is to map the current research in this area to aid developers and researchers in this matter and then to suggest directions for future work to take this further.

3. Methodology

We aim to identify and map the literature that discusses design or interaction techniques to address perceptual challenges in near-field AR using commercially available HMDs. As such, we choose to conduct a scoping literature review. Sutton et al. [56] define a scoping review as "a review that seeks to explore and define conceptual and logistic boundaries around a particular topic with a view to informing a future predetermined systematic review or primary research". Systematic reviews are defined by their "comprehensive search approach", asking one specific question and carrying out a formal synthesis, while scoping reviews instead explore a range of evidence and present the included literature to provide a structured overview of the landscape [56]. Our structured analysis process and presentation of results are inspired by existing scoping reviews such as those by Grave et al. [57] and Hirzle et al. [58]. We follow the Preferred Reporting Items for Systematic Reviews and Meta Analyses for Scoping Reviews (PRISMA-ScR) guidelines and the PRISMA-ScR checklist [12].

3.1. Research Questions and Aim

This scoping review is focused on answering the following research question:

RQ1: What virtual design or interaction techniques are used to alleviate perceptual issues in near-field AR?

The following two sub-questions are also addressed:

RQ1.1: Is near-field perception better in VST than OST?

RQ1.2: How are these techniques being used to work within commercial HMDs?

This work is motivated in part by experientially seeing the lack of research looking at how perceptual issues can be addressed while still using commercially available "off-the-shelf" HMDs. As discussed in the previous section, most commercial HMDs have similar hardware restrictions magnified in the near field, but there is little guidance on how to work around these. There is comparatively more work looking at how the hardware within HMDs can be improved, but we are interested in the things that designers and software developers can do when building AR applications to help alleviate some of the effects that the known perceptual issues can cause [59].

Therefore, the aim of this review is to understand and map the existing literature that uses design and interaction to address near-field perceptual issues while using commercially available HMDs. We focus our search on how researchers are working within the bounds of commercial hardware to alleviate and overcome these perceptual issues, using design and interaction techniques.

3.2. Search Strategy

3.2.1. Definition of Keywords

To define the keywords that would be used to construct the search term, we used the PCC (Participants, Concept, Context) framework recommended by the Joanna Briggs Institute [15]. This process structures the definition of keywords into three groups, enabling the definition of keywords and then the listing of alternatives. Using this relieves some subjectivity from the keyword generation process. Four key areas of interest were identified to group the terms at this stage, derived from the research question: perception, design/interaction, AR, and the near field. The lead author was primarily responsible for this, with the second author consulted to ensure that they were in agreement with the

terms defined. Please see Table 1 for the PCC table of terms. An initial search term was then constructed using the keywords from the PCC table, combining the four key areas (perception, design/interaction, AR, and the near field) with "AND" statements and the keywords within these areas with "OR" statements. The initial search terms are listed as follows:

```
2016 to present
(Percept* OR visual OR vision) AND
(design OR interaction) AND
(‘‘mixed reality’’ OR MR OR ‘‘augmented reality’’ OR AR) AND
(‘‘near field’’ OR ‘‘near-field’’ OR ‘‘close up’’ OR ‘‘close range’’ OR
    ‘‘peripersonal’’ OR ‘‘near range’’ OR ‘‘near by’’ OR ‘‘near-by’’)
```

To ensure that no important keywords had been missed, an initial search of the ACM Digital Library and IEEE Xplore was conducted with the keywords from the PCC table. The titles, abstracts, and author keywords from the papers returned were screened for any keywords that had been missed by the search term. From this process, only "close range" was added to our list. There were many other candidate keywords but all were either deemed too broad, not specific enough to our question, or too specific to be included in our list. For example, "depth cues" would narrow the search too much, but "3D vision" would broaden into the computer vision literature. The final search terms are listed as follows:

```
2016 to present
(‘‘Percept*’’ OR ‘‘visual’’ OR ‘‘vision’’) AND
(‘‘design’’ OR ‘‘interaction’’) AND
(‘‘mixed reality’’ OR ‘‘MR’’ OR ‘‘augmented reality’’ OR ‘‘AR’’) AND
(‘‘near field’’ OR ‘‘near-field’’ OR ‘‘close up’’ OR ‘‘close range’’ OR
    ‘‘close-range’’ OR ‘‘peripersonal’’ OR ‘‘near range’’ OR ‘‘near by’’ OR
    ‘‘near-by’’)
```

Table 1. Initial PCC concept mapping.

	Main Concept	**Alternative Keywords**
Participants	Anyone	x
Concept	Perceptual Challenges Design Techniques Interaction Techniques	Visual challenges/issues Perceptual accuracy Perceptual efficiency Design methods Interaction methods
Context	Near-Field Augmented Reality	Mixed reality Augmented reality AR MR Close up Close range Peripersonal Near range Nearby

3.2.2. Search Databases and Date Range

Based on preliminary searches with a subset of the keywords defined above, four databases were identified for this review: the ACM Digital Library, IEEE Xplore, PubMed,

and Science Direct. These preliminary searches were also used to establish a date range for the final search. The year 2016 was decided to be an appropriate start date. As discussed in the Background (Section 2), this is because it was an important year for the development of XR technology, with the release of Microsoft HoloLens 1, which was the first fully untethered, stand-alone HMD. It was also a key turning point for the hype around AR, which led to increased research focus and development, as discussed in the Background Section [22–24].

3.3. Evidence Screening and Filtering

We used a two-stage process to filter and identify appropriate papers to be included. The first stage involved filtering based on the papers' title and abstract, and then the second stage involved filtering based on the whole paper. This is represented in a PRISMA-ScR diagram (Figure 2).

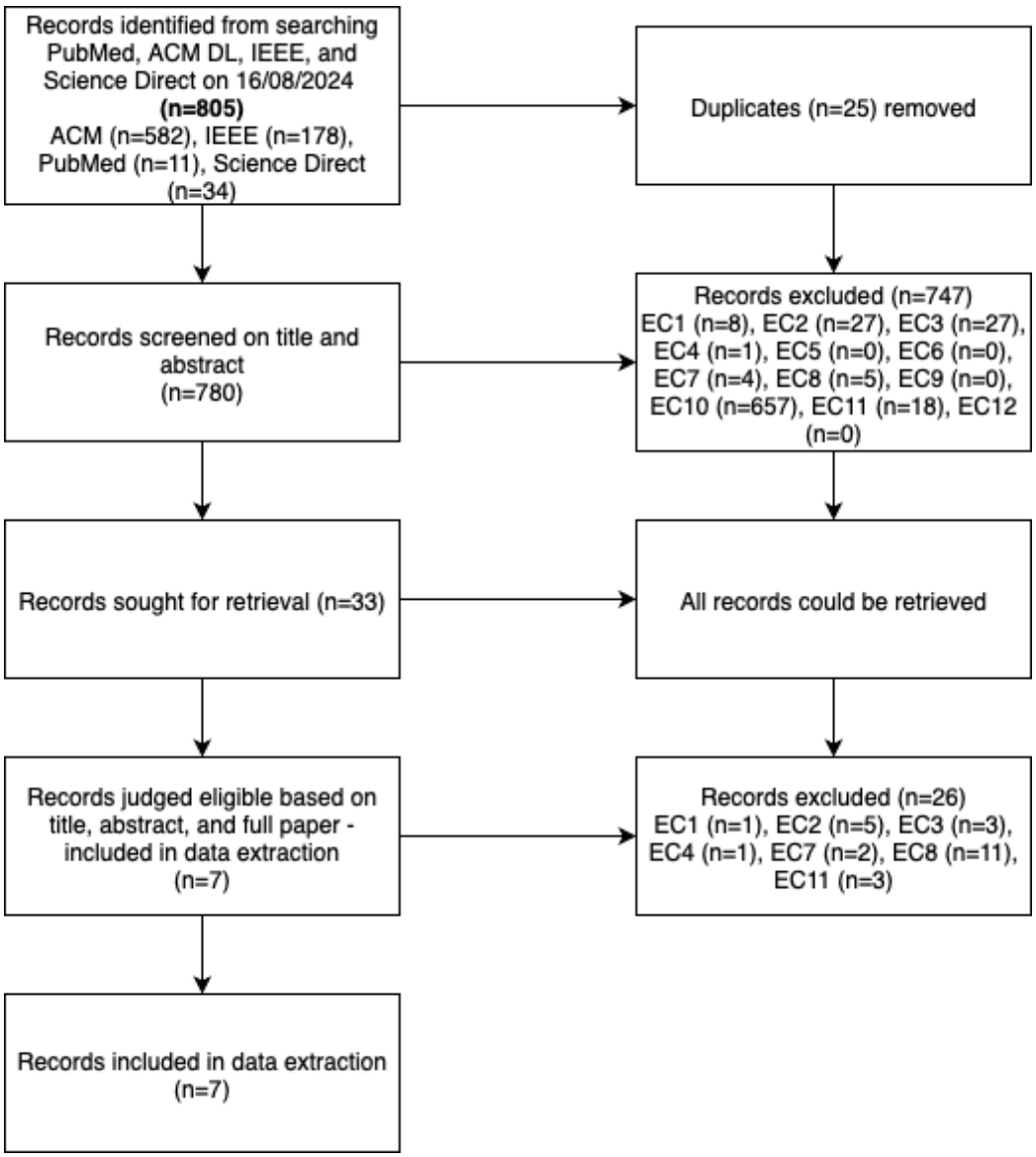


Figure 2. PRISMA-ScR diagram illustrating the scoping review process from the identification of sources to the final sample of articles included for data extraction.

3.3.1. Inclusion and Exclusion Criteria

The inclusion and exclusion criteria used to screen papers are summarised in Table 2. These criteria were reached by taking the research question(s) and breaking it down into

its components. The exclusion criteria represent a lack of each of these components; to be excluded, a paper must meet one or more of these exclusion criteria. An included paper must fit at least one of the inclusion criteria.

Table 2. Inclusion and exclusion criteria.

ID	Description
IC1	Requires further reading
IC2	Near-field perceptual problem investigated in a design or interaction context
IC3	Interaction technique(s) applied to a perceptual problem in the near field
IC4	Design technique(s) applied to a perceptual problem in the near field
EC1	Mobile AR (including headset with phone in)
EC2	Virtual reality
EC3	Hardware suggestion, e.g., new lens tech
EC4	Not in main proceedings
EC5	Not in English
EC6	Missing relevance to AR
EC7	Not human perception
EC8	Missing relevance to design/interaction technique(s)
EC9	Not published in 2016–2025
EC10	False positive, i.e., keyword used in a different context
EC11	Missing relevance to the near field
EC12	Review paper

These criteria deliberately exclude any studies that contribute new hardware. This is principally any suggested lens technology, where researchers suggest and evaluate a new form or component of AR lens technology. This choice is made to align with the research questions and aim of understanding what techniques have been suggested to work with commercially available HMDs. This does restrict the criteria and reduces the number of papers that were included, but it aligns with the research goals of this study.

3.3.2. Running the Query

The final search was run on the 28th of February 2025 (28/02/2025), with **805 total records** returned (ACM: 582, IEEE: 178, PubMed: 11, Science Direct: 34). Once collated, there were **25 duplicates**, leaving **780 records** for title and abstract filtering. A total of **747 records were excluded** at the title and abstract filtering stage, leaving **33 records** to be retrieved and read in full. A total of **26 records were excluded** after reading them in full, leaving **7 records for the final data extraction**. The PRISMA-ScR diagram (Figure 2) details how many records were included and excluded at each stage in line with the inclusion and exclusion criteria.

3.4. Data Extraction

Data extraction was, again, a two-stage process, with the first involving deductive coding and categorisation and the second involving inductive content analysis. This approach was chosen in line with other scoping reviews and recommendations from PRISMA-ScR and the JBI [58,60]. It allowed us to develop a rich understanding of the work included in our corpus and generate a thorough construction of the conceptual and logical boundaries within the content.

3.4.1. Stage One—Deductive

This first stage involved developing a codebook (a set of deductive codes) and coding the seven included papers against this codebook. Each paper was read twice during this stage to ensure that all elements were coded appropriately. The codes were developed first

by the lead author by breaking down the research questions and objectives of the review. The four sections of the search query (perception, design/interaction, AR, and the near field) also came into the code development to ensure that all bases had been covered and to maximise the amount of appropriate data that could be collected. After the first round in this stage, the codes were reviewed and discussed with all of the authors to fill any gaps and address the appropriateness of the codes taken forwards.

Each code either had a list of values to choose from to categorise (e.g., C1, C4, and C6) or the text was copied and summarised directly from the paper (e.g., C2, C3, C5, and C7). Please see Table 3 for a full list of the 18 codes and predefined values. By the nature of the research question and therefore the inclusion criteria, either C14 or C15 had to be complete; i.e., the paper must have investigated either design or interaction. This deductive coding stage also acted as a familiarisation phase for the next stage of the analysis.

Table 3. Deductive codes—[. . .] denotes copying from the text.

Code	Description	Values
C1	Category	Interaction technique investigated in the context of perception, Design technique investigated in the context of perception, Design technique applied to perceptual issue, Interaction technique applied to perceptual issue
C2	Research Question/Objective	[. . .]
C3	Contribution	[. . .]
C4	Contribution Type	Empirical, Applications, Methodological, System/Artefact, Theoretical
C5	Limitations	[. . .]
C6	Type of User Study	Yes Empirical, Yes Expert Evaluation, Yes Field study, Yes Workshop, No, Other
C7	Purpose of User Study	[. . .]
C8	Metric for Evaluation	[. . .]
C9	Field Study Task Type	[. . .]
C10	Study Details (e.g., Participant Demographics, Target Users, Sample Size)	[. . .]
C11	User Task	[. . .]
C12	Type of AR	OST AR, VST AR
C13	Device	[. . .]
C14	Interaction Technique	[. . . , n/a]
C15	Design Technique	[. . . , n/a]
C16	Perceptual Issue Addressed/Investigated	[. . .]
C17	Performance Improvement	[. . . , Area Investigated]
C18	Definition of Near Field	[. . .]

3.4.2. Stage Two—Inductive Content Analysis

Stage two of the data extraction involved inductive coding, where codes were developed throughout the process. Latent pattern content analysis [61] was chosen as an appropriate method for this, as the core advantage of content analysis is that information within and across the data corpus is condensed. The latent pattern part of this method means that the researcher plays an active role in interpreting and finding meaning within the data while seeking to establish a pattern of characteristics within the data. It is more than just collecting overt, surface-level meaning, and it is inherently linked to the researcher and their interpretation. Latent pattern analysis goes beyond the surface level to identify patterns within the data, but, while the researcher is integral, it is not reliant on the position or lived experiences of the analyst to establish the patterns [61]. The patterns are recognisable within the context.

Again, each paper was read twice in this stage of the analysis to ensure scrutiny and to close any gaps. The previous deductive analysis acted as a familiarisation stage, which aided the beginning of this process. The lead author conducted the content analysis, but the other authors were met with after each round of reading to discuss how the papers were being coded, any issues, and the codes themselves.

3.4.3. Critical Appraisal, Limitations, and Potential Bias

While scoping reviews typically include all types of evidence, we chose to restrict our search to full paper proceedings and journal articles in order to set a high level of research rigour. The result of this was a small data corpus of just seven papers. But this in itself was the motivation for this review—the lack of research in this area. By conducting this review, we hope to help direct future work in this space.

A potential limitation of this review is the chance of missed keywords and therefore missed papers. We endeavoured to avoid this by using the PCC method to collate keywords for the search term; however, this is not unassailable. We do however believe that we were thorough enough to cover enough of the data in this space to be representative. This enables us to maintain confidence in the results of our review.

One challenge during the evidence screening was the large number of papers returned from the query that were deemed irrelevant, specifically that fell into exclusion criterion 10 (EC10), false positive. We believe that this was due to the imprecise nature of the keywords that define this space. Nearly all of the keywords that should have narrowed down the search (presented in the PCC table in Table 1) are used in different contexts. The result of these keywords having multiple meanings or a wide meaning is a large number of papers that the query returns correctly (because the keywords are present) but are not relevant to our research question. As discussed previously, this was part of the motivation for not including any more keywords returned after the initial search to validate search terms. Maintaining these wide-reaching keywords, however, increased the likelihood that we found all relevant papers.

We recognise that there is the potential for bias throughout the process of choosing keywords, screening evidence, etc., but, through the process described here, we endeavoured to limit the effects of this bias. By using methods such as PCC, we removed a level of subjectivity when establishing keywords. Similarly, by defining thorough inclusion and exclusion criteria ahead of screening and by discussing papers where there was any doubt with all authors, we ensured rigour in our work.

3.4.4. Synthesis

While a formal synthesis is not usually included in a scoping review, the Results Section below presents the findings of the data extraction from both the deductive and inductive coding in order to harmonise the meaning. The results from the deductive round are presented first, largely with tables and graphs illustrating the categorisation of the included papers into different groups. The results of the inductive round are presented as themes that represent a collection of related codes recognising a pattern in the data.

4. Results

This section is split in line with the two rounds of analysis, the deductive coding and the inductive content analysis. The deductive results give an outline of the focuses of the included papers and the metrics that describe the data corpus. The results of the content analysis are presented as the themes that were generated through this process.

Two key findings that will be discussed further in line with the themes are highlighted here. The first observation is that of the two competing goals when developing near-field

AR experiences: One goal is to replicate the real world faithfully to provide the user with all of the cues that they would receive and use to make decisions for real-world interactions. Conversely, the other goal is to build on the real world using AR to achieve what cannot be achieved in the physical world, interactions and experiences that are only possible due to the technology. The second observation concerns the way that design and interaction techniques are applied. Design techniques are employed to help replicate the real world or work around scenarios where true replications cannot be achieved; interaction techniques are applied to work around obstacles to realism or go beyond realistic capabilities and provide experiences impossible without the technology [62,63].

We present answers to our main research question and second sub-question (RQ1.2) but found insufficient evidence to provide an answer for the first sub-question (RQ1.1). The Discussion Section describes links between points in the results, their impact, and what this means going forwards.

4.1. Deductive Codes

The deductive codes allow us to break down the data corpus into groups, which enables us to illustrate how different categories and factors are represented within the included papers. Table 4 gives an overview of the included papers' characteristics, and Table 5 highlights the key concepts addressed in each study.

Firstly, codes C1, C15, and C16 show us that there was a near 50:50 split between papers of the two categories, as well as interaction techniques vs. design techniques. Four papers fell into the category of *"Design technique applied to perceptual issue"*, and three fell into the category of *"Interaction technique applied to perceptual issue"*; this was similarly reflected between C15 and C16.

In opposition to this 50:50 split, of the included papers, only one used a VST HMD; the rest used optical see-through HMDs. Of these six, five used a Microsoft HoloLens (with four studies using the second generation), and the other two used other HMDs (Meta DK1 and Epson Moverio BT200).

All seven studies in the data corpus conducted an empirical field study with participants. This is likely due to the nature of research in the area of perception. As everyone perceives things slightly differently and it is impossible to know exactly what someone else in seeing, one can only measure the result of an action upon what they are seeing—an empirical study is necessitated [44]. This builds on the challenges of researching this area, as discussed in the Background Section.

There was a lot of variability in how the near field was presented and the boundaries placed upon it, with many studies loosely defining it as "within arm's reach". All studies either defined the near field or conducted their study between 25 cm and 100 cm in front of the participant.

Figure 3 shows the distribution of publication years across the data corpus, with most (three) papers published in 2023 and two papers published in 2024. Figure 4 illustrates C8 describing the physicality of the objects used in the study task. The tasks in four of the studies fell into "virtual-to-physical" category; i.e., the participant was asked to move a virtual object relative to a physical target. The tasks in three studies included both virtual-to-physical and virtual-to-virtual elements. No studies investigated purely virtual-to-virtual stimuli.

Figure 5 shows the number of participants enrolled in each study, with a mean average of 28. This includes one study enrolling just 4 participants, while all others had 13 or more participants. This difference in the number of participants was due to the specialised and intensive nature of the user task, involving participants specific to the application and a much longer engagement.

Table 4. Summary of included papers.

Title	Paper Number	Metrics of Evaluation	Task Type	Device	Interaction Technique	Design Technique
"Evaluation of Different Visualization Techniques for Perception-Based Alignment in Medical AR" [62]	1	Orientation error; distance error; alignment time; user feedback	Perceptual mapping	Microsoft HoloLens 1	–	Outline; semi-transparent; wireframe; replicate
"The Impact of Occlusion on Depth Perception at Arm's Length" [64]	2	Alignment accuracy; confidence in choice	Depth estimation	Custom/HoloLens 2	–	Occluding surface and illumination variables: none, gray and bright, gray with bright and hole, golf ball pattern and bright, golf ball pattern and bright with hole
"SharpView: Improved clarity of defocused content on optical see-through head-mounted displays" [65]	3	Pupil diameter, focal blur size	"Adjustment"	Epson Moverio BT-200 OST HMD	–	Sharpness/blur
"Visuotactile integration for depth perception in augmented reality" [66]	4	Final depth estimation bias (mean and standard deviation), task duration (mean)	Depth estimation	Meta DK1	Visuotactile—haptics integrated for depth perception	–
"Investigating the Effects of Avatarisation and Interaction Techniques on Near-field Mixed Reality Interactions with Physical Components" [67]	5	Accuracy, efficiency, perceived usability	Hoop on peg target task	Microsoft HoloLens 2	Pinch-to-grab; stick-on-touch	No Augmented Avatar (No-Av), Augmented Avatar (Av), Augmented Avatar with Translational Gain (AvG)
"Give Me a Hand: Improving the Effectiveness of Near-field Augmented Reality Interactions By Avatarizing Users' End Effectors" [68]	6	Performance, perceived real hand visibility, perceived usability	Collision avoidance-based object retrieval task	Microsoft HoloLens 2	Avatarisation of end effectors (hands): no avatarisation, Iconic Augmented Avatar, Realistic Augmented Avatar	Virtual light intensity, high and low
"A Fitts' Law Study of Gaze-Hand Alignment for Selection in 3D User Interfaces" [63]	7	Task completion time, throughput, effective width, error rate, hand movement, usability questionnaire	Point and select/depth estimation	Microsoft HoloLens 2	3D selection techniques: Gaze & Handray, Gaze & Finger vs. baseline Gaze & Pinch, Handray, HeadCrusher	–

Table 5. Summary of concepts of included papers.

Paper Number	Tactile Feedback	Depth	Multi-Modal Interactions	Occlusion	Light Conditions	Avatarisation	Focus Blur	Parallax	VAC	Perception Action Coordination
1		■		■	■				■	
2		■		■						
3		■					■		■	
4	■	■	■							
5						■				■
6					■	■				■
7		■	■					■		

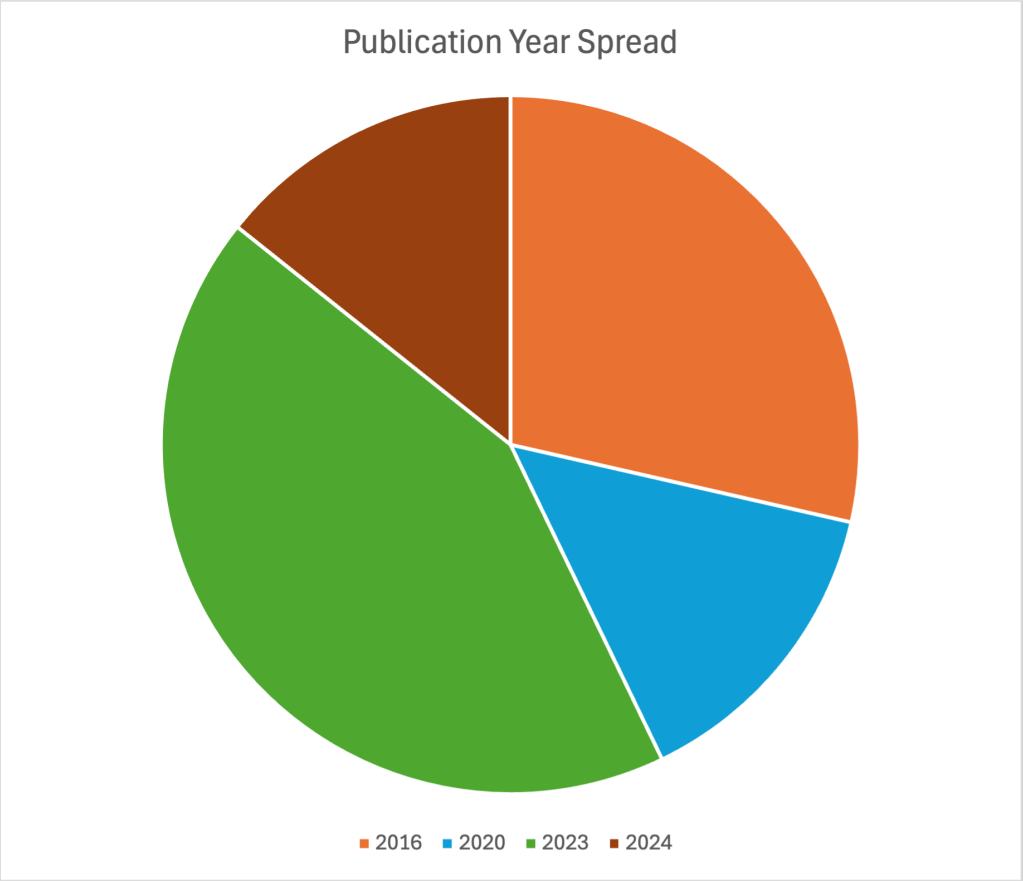


Figure 3. Pie chart showing spread of publication dates.

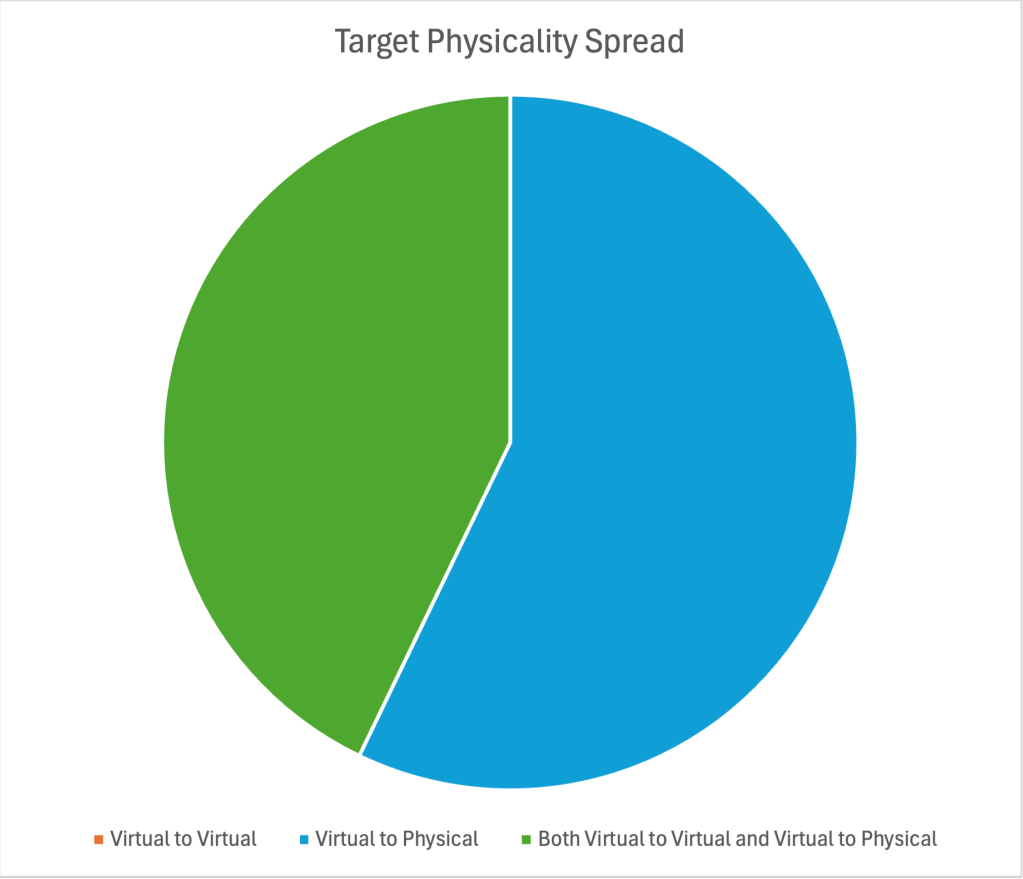


Figure 4. Pie chart showing target physicality of the user task.

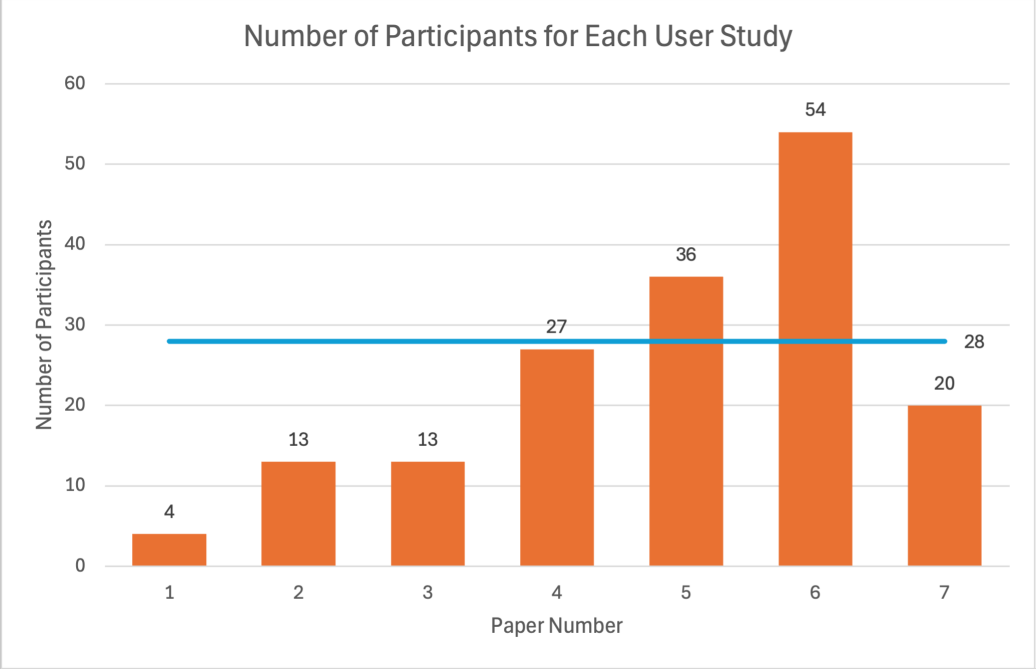


Figure 5. Bar chart showing the number of participants enrolled in each study.

There was little cohesion between the types of task that each study used in their investigations. Some asked users to move a virtual object relative to another, following or taking inspiration from blind reaching (a common technique used to study hand–eye coordination [69]), while others asked users to interact more directly with a single virtual object. Furthermore, there was not much consistency in the metrics used for evaluation. Task completion time was the most common metric, and, often, a combination of quantitative and qualitative assessments was used. Quantitative metrics were very specific to the task and the research questions, whereas qualitative metrics tended to be usability questionnaires or the NASA-TLX [70] or similar.

A wide range of perceptual issues were investigated in the studies included, with depth perception being the most common. Others included focal rivalry and VAC, perception–action coordination, and parallax.

4.2. Inductive Analysis

4.2.1. Hardware Effects or Limitations on Near Field Perception

It is well known that there are limitations to experiences laid on by the hardware of commercially available HMDs; what we found here supports this and gives further support and context to this point for near-field applications. As discussed in the Background Section, the source of a significant issue is the fixed focal depth that all virtual content is displayed at, and, when virtual content moves away from this point, perceptual challenges arise. This fixed depth is often around two metres, which therefore exaggerates issues for near-field tasks, as the interaction point is at least a metre away from this fixed focal depth (as indicated by the definition of the near field above).

Also highlighted by this theme is the difference in hardware limitations between OST and VST technologies. Firstly, we obtained further support of the statement that OST and VST pose different design challenges. This is a known entity and is intuitive, but our review reaffirms this in the context of near-field interactions. Additionally, we highlight two key issues with OST technologies that hinder immersion and exacerbates perceptual issues. One is the inability of OST HMDs to provide true occlusion. While it is stated within the contribution of one of our included papers that occlusion is not necessary for accurate

depth perception or near-field interactions if all other depth cues are preserved, occlusion is a strong depth cue that users are often reliant on, especially as current HMDs do not preserve other depth cues 100% accurately. Cutting and Vishton [33] ranks occlusion as the most important depth cue for near-field perception. This means that depth perception overall is decreased as the user often cannot decipher which object is in front of another. In addition, and as discussed in the background, HMDs still cannot faithfully represent all the cues needed to perceive an environment [40].

The second issue is the effect of the restricted field of view (FOV) of commercial HMDs. This is particularly an issue when the amplitude of interactions is high or higher than the FOV [63]. We theorise that this is in part due to immersion being reduced when the harsh boundaries of a narrow FOV interrupt natural interactions. This notion of immersion continues into the next themes.

4.2.2. Real Interactions Are Multi-Modal

This theme discusses the observations and opportunities of multi-modal near-field interactions, where two modes of interaction (for example, gaze and pinch) are used together with the aim of being quicker or more accurate than a uni-modal alternative. It is known that multi-modal interactions can perform better than uni-modal interactions, but we confirm this for near-field interactions, provide specificity of what multi-modal interactions were investigated in our data corpus, and theorise a reason for this.

Gaze and tactile feedback were the key additional input modalities explored by studies in our data corpus, where each was combined with hand gestures and evaluated relative to a uni-modal counterpart. Overall, an increase in performance was observed but with various nuances in the effect. Gaze is naturally linked to hand gestures as a precursor interaction, and using gaze to initiate an interaction with a hand gesture as a trigger has been shown to be effective [63]. It is, however, necessary to choose an appropriate combination of gaze and hand gesture to suit the task presented. Tactile feedback was shown to improve depth perception [66]; however, one observation with tactile feedback was that a trade-off developed between speed and accuracy. Different forms of tactile feedback could be chosen for tasks that require higher accuracy or faster completion [66].

There are two main points to draw from this theme. The first centres around how interactions are task-dependent, and, while the studies that we included in our data corpus present some suggestion, there has been no work that evaluates AR interactions' situational pros and cons or that present any kind of taxonomy for how to choose them appropriately for a given task.

Secondly, we come back to immersion. Real-world interactions are multi-modal [71], and, as summarised at the start of the discussion, there are two competing goals for developing AR experiences: replicating the real world and building on the real world. Interactions in AR often try to do both of these at the same time: real-world interactions naturally have tactile feedback, which is simulated somehow in AR, but also gaze is used in combination with hand gestures to provide the user with interactions that are not possible in the real world. Sometimes gaze may be used to simulate a real-world interaction or as a workaround for when a direct simulation is not possible, and these two ideas are not always distinct, but, in every situation, the aim (indirectly perhaps) is to increase immersion, which, in turn, increases performance.

4.2.3. Despite More Accurate Options, Hand Gestures Are Preferred for Near-Field Interactions

As discussed in the previous theme, there are a multitude of different interactions available when designing for AR, each with their own pros and cons; two interaction methods could outperform one another on different criteria. There is, however, consensus

that natural hand gestures (a wide group) are the preferred mode of interaction for near-field applications, despite becoming less effective at further ranges. This is still the case even when more accurate options, such as longer precise interactions or haptic hand-tracking gloves, are available. The ease, naturalness, and speed of natural hand gestures is preferred over any accuracy gained by more precise but longer interactions or the use of wearables.

This raises two considerations: Firstly, it reaffirms the need to choose interactions appropriate to the task, and there will never be one interaction for every task in every situation. Secondly, some level of inaccuracy is tolerable or even preferred if the alternative is better accuracy but longer completion times or additional hardware.

This somewhat goes against the previous theme and multi-modal interactions, as without any additional hardware, haptic feedback (shown to be effective at improving depth perception) cannot be provided. However, if a "good enough" interaction can be achieved using a different second interaction mode or simply using hand gestures alone, then the experience is not hindered by the subtle inaccuracy. For surgical applications, however, or any other scenario with a manual task and where accuracy is of highest priority, the extra time or additional hardware may still be worth the cost to the experience.

Avatarisation, Perception, and Embodiment

It was made clear from the studies within our data corpus that a significant source of error for interactions in OST AR is a difference in the interacting layer and the visible layer. The visible layer is the effector that the user can see when performing an action, and the interacting layer is what the HMD "sees" and what is used to determine whether or not a collision has occurred between the effector and the object. This is not an issue that occurs in VST AR, as the visible layer is replicated from the interacting layer and presented to the user. It is also an issue likely exaggerated in near-field interactions due to the interaction point being much closer than the focal depth of the HMD, as discussed previously.

Studies within our data corpus investigated avatarising users' end effectors (hands) in OST AR to approach this issue and came away concluding that avatarising end effectors can improve interaction performance regardless of the avatar fidelity. Even a crude avatar overlay over the user's hand could improve performance due to the visible and interaction layers aligning [67,68].

This links back to the points made on immersion in the first two themes. Avatarising effectors can be argued to improve immersion and embodiment, which has positive effects on interaction efficacy, especially as the effectiveness of interaction techniques has been shown to change with the degree of physicality [67,68]. This is evidence that the more a user is immersed in an AR experience, the more they feel that they embody the "character" that they play in the AR experience, the better the interactions, and the better the experience that they have. It can be argued that all the techniques covered in this review aim to increase a user's level of immersion, both the design techniques employed to replicate realistic cues and the interaction techniques working around obstacles to realism or going beyond realistic bounds, as mentioned at the start of this section.

Further work needs to be carried out to investigate the full effects of avatarisation on interaction efficacy in the near field, but the extent to which it improves levels of immersion and, in turn, interaction efficacy is growing.

4.2.4. Depth Is the Main Contributor to Perceptual Inaccuracy, Particularly in the Near Field

Depth perception was one of the main issues investigated by the studies in our data corpus, and we suggest it here as the main contributor to perceptual inaccuracy for near-field interactions. This is an effect of near-field interactions being a long way from the fixed focal depth of commercial HMDs, as discussed in the Background Section and throughout

this review. Challengingly, a depth bias has been noted in both positive and negative directions. This means that an object can appear either nearer or further away than it actually is.

Design Techniques Used to Alleviate Depth Estimation Errors in the Near Field

Design techniques are largely focused on replicating the real world to improve realism and immersion in AR experiences, and it is the perception of depth where their efforts have most frequently been applied. Depth cues are difficult to replicate for virtual objects in AR; therefore, design techniques are employed to either better replicate real-world depth cues or to provide the user with additional information with which they can make a judgement on the depth of an object [62,64]. However, design techniques are also used to more directly address the clarity of the virtual components [65]. There is little cohesion in the techniques employed within the studies included in our data corpus, and more research is required to be able to suggest any kind of guidelines for what design techniques to use in a given situation.

Some key findings are listed as follows:

- Correct estimation of depth is possible without occlusion (although with lower confidence) if all other depth cues are preserved.
- Virtual objects that are less opaque and have a higher contrast are easier to align with physical counterparts.
- Virtual object size and brightness can affect depth estimation.
- Sharpening algorithms can aid the perception of virtual objects that are much closer than the fixed focal depth.
- Complex occluding surfaces negatively impact the perception of objects beneath, but virtual holes in the surface can help.

4.2.5. Perception Is Personal and Made Up of a Multitude of Factors

This final theme reiterates the known entity of how personal perception is and brings it into the context of the near field. The first point here observed through our analysis is that perception is a personal experience with many factors effecting it, but, crucially, it cannot be measured directly, only via a participant's response. The researcher is never going to know exactly what the participant sees; instead, they must measure the effects of empirical observations from a participant. Further to this, there are a large number of factors that affect a user's perception, not only physical variations such as age but also personal choices and experiences. The preferred interaction technique for a situation may vary between users; some interaction techniques can cause physical exertion, a factor that will affect users, and therefore interaction efficacy, differently. Interaction techniques can also have different affects on perception, which adds further variability.

Beyond the personal nature of perception, it is a difficult area to research: it has many components, and it is difficult to isolate and investigate them individually. Although this goes for perceptual cues individually, we also found that design and interaction techniques are linked, making it difficult to investigate either independently. This links to the previous theme and the use of design techniques to aid depth estimation; as there are many factors involved when investigating this area, it is challenging to make strong conclusions and suggestions.

5. Discussion

Here, we summarise our results and how they address our research question, highlight our contributions, and suggest directions for future work.

5.1. Summary of Results and Contributions

Our results show that, of the studies included in this review, there was a near 50:50 split between the two categories *design technique applied to perceptual issue* and *interaction technique applied to perceptual issue*, with all papers involving empirical user studies. A variety of HMDs were used, but the Microsoft HoloLens 2 was mostly used, with all but one study focusing on OST technology [66]. The definition of the near field varied, but all studies conducted their tasks within 25 and 100 cm in front of the user.

Through inductive content analysis of the data, we presented five themes and two sub-themes, as listed below:

1. Hardware effects or limitations on the near field perception.
2. Real interactions are multi-modal.
3. Despite more accurate options, hand gestures are preferred for near-field interactions.
    - Avatarisation, perception, and embodiment.
4. Depth is the main contributor to perceptual inaccuracy, particularly in the near field.
    - Design techniques used to alleviate depth estimation errors in the near field.
5. Perception is personal and made up of a multitude of factors.

For this review, we asked the following research question: "What virtual design or interaction techniques are used to alleviate perceptual issues in near-field AR?". The above themes characterise various concepts that contextualise insights in line with this question and aim to answer it. Some elements discussed confirm some known concepts in the context of near-field interactions, as well as addressing the research question more directly.

Embodiment and immersion are raised throughout, and it is clear that a more immersive experience contributes to improved interactions. While this may not directly address the research question, it can be used as a proxy or adjacent goal for AR experiences. It is design and interaction techniques that address perceptual issues or enhance an experience and a more immersive experience with a stronger feeling of embodiment that affords users more natural, precise interactions.

Our results show how design and interaction techniques are applied differently to effectively address different goals. It was observed that building AR experiences comes with two goals, replicating the real world to enable believable immersive experiences and enhancing real-world capabilities via AR technology; the use of design and interaction techniques can be seen to coincide with this. Design techniques tend to be employed to better represent the real world or work around hardware restrictions to provide the user with more information to understand their experience. Conversely, interaction techniques generally enable experiences that are not possible in the real world, affording the user the opportunity to interact in a different way. There is crossover here, for example, where interactions that are not possible in the same way in AR (such as the natural feedback from touching an object) are worked around with different interaction techniques enabled by AR technology [66].

Here, we highlight two findings from this review: the adoption of immersion and embodiment as a key goal when evaluating interactions in near-field augmented reality and the acknowledgement of two competing goals for AR experiences—replication and advancement. In addition, we found no further definition for near-field interactions across our data corpus other than the "arm's length" or up to two metres suggested by Cutting and Vishton [33], with studies conducted at a range of distances. In response to this, we advocate for future research to converge on using a range of 25–100 cm in front of the user for studies involving near-field interactions. Closer than this, and even without the use of AR, reading becomes difficult; further than this is further than the average person's arm length and is closer to the two-metre fixed focal depth used by most commercial HMDs

to reduce perceptual issues [38,72]. This range is indicative of the ranges that the studies included in this review were conducted within. Similarly, the upper boundary of the "comfortable viewing zone" [73] recommended by Microsoft for the HoloLens 2 is 125 cm and is thus commensurate.

For this review, we also asked two sub-questions: "Is near-field perception better in VST than OST?" and "How are these techniques being used to work within commercial HMDs?". This first question we are not in a position to answer, as only one paper included in our data corpus used a VST HMD. This is an area that needs further work, as discussed in the next section. We asked the second question with the aim of offering some initial advice to developers and designers of AR applications in this space. We attempted to answer this question throughout the Results Section, but it is clear that there is a lot more research to do in this area before clear guidance can be given for these developers and designers.

5.2. Research Opportunities

As illustrated by the small number of studies included in our data corpus, this is an under-researched area that would benefit from increased attention. Going forwards, using commercially available HMDs is going to be the most accessible option, and, therefore, until hardware issues and limitations such as the fixed focal depth (amongst other things) can be overcome, we must work within their bounds. Determining how design and interaction techniques can be employed to work around these limitations and create better experiences will be invaluable.

Secondly, and in tandem with this first point, a wide-scale, generalisable evaluation of different interaction techniques, their pros and cons, and their applicability to different scenarios would be valuable to designers creating AR experiences in making informed decisions about what techniques should be used in different situations. This could result in a taxonomy being produced that designers of AR experiences could call upon when developing for different scenarios.

Finally, it is clear that the commercial market is moving away from OST HMDs and towards VST devices [74]; however, the studies available at the time of this review are very OST-focused. This illustrates a wider shift in the AR scene around how Microsoft (amongst others) grabbed the AR scene with the HoloLens in 2016, but, since then, the world has shifted away from this technology towards VST instead. Microsoft has dropped the HoloLens 2 and any plans for a third-generation device, and the market leaders now mostly produce VST devices, Meta with the Quest 3 and Apple with the Vision Pro being the biggest examples. How current research in this area and the taxonomy called for above translates to VST is unclear. Can VR guidelines be drawn on? How do they need to be adapted for VST AR? How does existing research in OST design and interaction transfer to VST? Future research must investigate this.

6. Conclusions

We present a scoping literature review of seven studies that investigate design or interaction techniques to address the perceptual challenges of near-field augmented reality. We reviewed the papers with a two-phase approach, with both a deductive code book and inductive content analysis. We contribute a series of key findings that address our research question, and we go on to suggest three areas for future research and give reasons for them.

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.; supervision, C.B.; project administration, J.H. and C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable—study did not involve humans or animals.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Syed, T.A.; Siddiqui, M.S.; Abdullah, H.B.; Jan, S.; Namoun, A.; Alzahrani, A.; Nadeem, A.; Alkhodre, A.B. In-Depth Review of Augmented Reality: Tracking Technologies, Development Tools, AR Displays, Collaborative AR, and Security Concerns. *Sensors* **2023**, *23*, 146. [CrossRef]
2. Arena, F.; Collotta, M.; Pau, G.; Termine, F. An Overview of Augmented Reality. *Computers* **2022**, *11*, 28. [CrossRef]
3. Reljić, V.; Milenković, I.; Dudić, S.; Šulc, J.; Bajči, B. Augmented Reality Applications in Industry 4.0 Environment. *Appl. Sci.* **2021**, *11*, 5592. [CrossRef]
4. 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]
5. Papadopoulos, T.; Evangelidis, K.; Kaskalis, T.H.; Evangelidis, G.; Sylaiou, S. Interactions in Augmented and Mixed Reality: An Overview. *Appl. Sci.* **2021**, *11*, 8752. [CrossRef]
6. Cooper, E.A. The Perceptual Science of Augmented Reality. *Annu. Rev. Vis. Sci.* **2023**, *9*, 455–478. [CrossRef] [PubMed]
7. Bremers, A.W.D.; Yöntem, A.Ö.; Li, K.; Chu, D.; Meijering, V.; Janssen, C.P. Perception of perspective in augmented reality head-up displays. *Int. J. Hum.-Comput. Stud.* **2021**, *155*, 102693. [CrossRef]
8. Bhowmik, A.K. Virtual and augmented reality: Human sensory-perceptual requirements and trends for immersive spatial computing experiences. *J. Soc. Inf. Disp.* **2024**, *32*, 605–646. [CrossRef]
9. Wang, Y.J.; Lin, Y.H. Liquid crystal technology for vergence-accommodation conflicts in augmented reality and virtual reality systems: A review. *Liq. Cryst. Rev.* **2021**, *9*, 35–64. [CrossRef]
10. Itoh, Y.; Langlotz, T.; Sutton, J.; Plopski, A. Towards Indistinguishable Augmented Reality: A Survey on Optical See-through Head-mounted Displays. *ACM Comput. Surv.* **2021**, *54*, 120:1–120:36. [CrossRef]
11. Diaz, C.; Walker, M.; Szafir, D.A.; Szafir, D. Designing for Depth Perceptions in Augmented Reality. In Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Nantes, France, 9–13 October 2017; pp. 111–122. [CrossRef]
12. Tricco, A.C.; Lillie, E.; Zarin, W.; O'Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.; Horsley, T.; Weeks, L.; et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [CrossRef]
13. PRISMA. PRISMA for Scoping Reviews (PRISMA-ScR). Available online: https://www.prisma-statement.org/scoping (accessed on 9 June 2025).
14. Cooper, S.; Cant, R.; Kelly, M.; Levett-Jones, T.; McKenna, L.; Seaton, P.; Bogossian, F. An Evidence-Based Checklist for Improving Scoping Review Quality. *Clin. Nurs. Res.* **2021**, *30*, 230–240. [CrossRef]
15. Joanna Briggs Institue. JBI Manual for Evidence Synthesis—JBI Global Wiki. Available online: https://jbi-global-wiki.refined.site/space/MANUAL (accessed on 9 June 2025).
16. Sutherland, I. The Ultimate Display. In Proceedings of the Congress of the Internation Federation of Information Processing (IFIP), New York, NY, USA, 24–29 May 1965.
17. McCarthy, C.J.; Uppot, R.N. Advances in Virtual and Augmented Reality—Exploring the Role in Health-care Education. *J. Radiol. Nurs.* **2019**, *38*, 104–105. [CrossRef]
18. 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]
19. Speicher, M.; Hall, B.D.; Nebeling, M. What is Mixed Reality? In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; CHI '19; pp. 1–15. [CrossRef]
20. Vinci, C.; Brandon, K.O.; Kleinjan, M.; Brandon, T.H. The clinical potential of augmented reality. *Clin. Psychol. Sci. Pract.* **2020**, *27*, e12357. [CrossRef]
21. Caudell, T.; Mizell, D. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences, Kauai, HI, USA, 7–10 January 1992; Volume 2, pp. 659–669. [CrossRef]

22. Javornik, A. The Mainstreaming of Augmented Reality: A Brief History. Available online: https://hbr.org/2016/10/the-mainstreaming-of-augmented-reality-a-brief-history (accessed on 9 June 2025).
23. 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: Cham, Switzerland, 2023; pp. 35–50. [CrossRef]
24. Microsoft. HoloLens (1st gen) Hardware. Available online: https://learn.microsoft.com/en-us/previous-versions/mixed-reality/hololens-1/hololens1-hardware (accessed on 9 June 2025).
25. Google. Google Trends: Understanding the Data—Google News Initiative. Available online: https://newsinitiative.withgoogle.com/en-gb/resources/trainings/google-trends-understanding-the-data/ (accessed on 13 June 2025).
26. VR Compare. HTC Vive: Full Specification. Available online: https://vr-compare.com/headset/htcvive (accessed on 9 June 2025).
27. Pokémon. Pokémon GO. Available online: https://pokemongo.com/ (accessed on 9 June 2025).
28. IKEA Global. Launch of New IKEA Place App. Available online: https://www.ikea.com/global/en/newsroom/innovation/ikea-launches-ikea-place-a-new-app-that-allows-people-to-virtually-place-furniture-in-their-home-170912/ (accessed on 9 June 2025).
29. BBC. Google Glass Smart Eyewear Returns. Available online: https://www.bbc.com/news/technology-40644195 (accessed on 9 June 2025).
30. Dargan, S.; Bansal, S.; Kumar, M.; Mittal, A.; Kumar, K. Augmented Reality: A Comprehensive Review. *Arch. Comput. Methods Eng.* **2023**, *30*, 1057–1080. [CrossRef]
31. Babu, S.V.; Huang, H.C.; Teather, R.J.; Chuang, J.H. Comparing the Fidelity of Contemporary Pointing with Controller Interactions on Performance of Personal Space Target Selection. In Proceedings of the 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Singapore, 17–21 October 2022; pp. 404–413. [CrossRef]
32. Babu, S.; Tsai, M.H.; Hsu, T.W.; Chuang, J.H. An Evaluation of the Efficiency of Popular Personal Space Pointing versus Controller based Spatial Selection in VR. In Proceedings of the ACM Symposium on Applied Perception 2020, Virtual, 12–13 September 2020; SAP '20. [CrossRef]
33. Cutting, J.E.; Vishton, P.M. Chapter 3—Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth. In *Perception of Space and Motion*; Epstein, W., Rogers, S., Eds.; Handbook of Perception and Cognition; Academic Press: San Diego, CA, USA, 1995; pp. 69–117. [CrossRef]
34. Xu, J.; Doyle, D.; Moreu, F. State of the art of augmented reality capabilities for civil infrastructure applications. *Eng. Rep.* **2023**, *5*, e12602. [CrossRef]
35. Mutis, I.; Ambekar, A. Challenges and enablers of augmented reality technology for in situ walkthrough applications. *J. Inf. Technol. Constr. (ITcon)* **2020**, *25*, 55–71. [CrossRef]
36. Cutolo, F.; Fida, B.; Cattari, N.; Ferrari, V. Software Framework for Customized Augmented Reality Headsets in Medicine. *IEEE Access* **2020**, *8*, 706–720. [CrossRef]
37. Yin, K.; He, Z.; Xiong, J.; Zou, J.; Li, K.; Wu, S.T. Virtual reality and augmented reality displays: Advances and future perspectives. *J. Phys. Photonics* **2021**, *3*, 022010. [CrossRef]
38. 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]
39. Blignaut, J.; Venter, M.; van den Heever, D.; Solms, M.; Crockart, I. Inducing Perceptual Dominance with Binocular Rivalry in a Virtual Reality Head-Mounted Display. *Math. Comput. Appl.* **2023**, *28*, 77. [CrossRef]
40. Argelaguet, F.; Andujar, C. A survey of 3D object selection techniques for virtual environments. *Comput. Graph.* **2013**, *37*, 121–136. [CrossRef]
41. Edwards, P.J.E.; Chand, M.; Birlo, M.; Stoyanov, D. The Challenge of Augmented Reality in Surgery. In *Digital Surgery*; Atallah, S., Ed.; Springer International Publishing: Cham, Switzerland, 2021; pp. 121–135. [CrossRef]
42. Reichelt, S.; Häussler, R.; Fütterer, G.; Leister, N. Depth cues in human visual perception and their realization in 3D displays. In Proceedings of the Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV, Orlando, FL, USA, 5–9 April 2010; SPIE: Bellingham, WA, USA, 2010; Volume 7690, pp. 92–103. [CrossRef]
43. Dror, I.E.; Schreiner, C.S. Chapter 4—Neural Networks and Perception. In *Advances in Psychology*; Jordan, J.S., Ed.; North-Holland: Amsterdam, The Netherlands, 1998; Volume 126, pp. 77–85. [CrossRef]
44. Marcos, S.; Moreno, E.; Navarro, R. The depth-of-field of the human eye from objective and subjective measurements. *Vis. Res.* **1999**, *39*, 2039–2049. [CrossRef] [PubMed]
45. Ellis, S.R.; Menges, B.M. Localization of Virtual Objects in the Near Visual Field. *Hum. Factors* **1998**, *40*, 415–431. [CrossRef] [PubMed]
46. Sielhorst, T.; Bichlmeier, C.; Heining, S.M.; Navab, N. Depth Perception—A Major Issue in Medical AR: Evaluation Study by Twenty Surgeons. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006, Proceedings of the 9th International Conference, Copenhagen, Denmark, 1–6 October 2006*; Larsen, R., Nielsen, M., Sporring, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 364–372. [CrossRef]

47. Ballestin, G.; Solari, F.; Chessa, M. Perception and Action in Peripersonal Space: A Comparison Between Video and Optical See-Through Augmented Reality Devices. In Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, 16–20 October 2018; pp. 184–189. [CrossRef]
48. Clarke, T.J.; Mayer, W.; Zucco, J.E.; Matthews, B.J.; Smith, R.T. Adapting VST AR X-Ray Vision Techniques to OST AR. In Proceedings of the 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Singapore, 17–21 October 2022; pp. 495–500. [CrossRef]
49. 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, Parramatta, Australia, 12–15 November 2019; VRST '19; pp. 1–11. [CrossRef]
50. 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]
51. 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), Atlanta, GA, USA, 22–26 March 2020; pp. 275–284. [CrossRef]
52. 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]
53. Seinfeld, S.; Feuchtner, T.; Pinzek, J.; Müller, J. Impact of Information Placement and User Representations in VR on Performance and Embodiment. *IEEE Trans. Vis. Comput. Graph.* **2022**, *28*, 1545–1556. [CrossRef] [PubMed]
54. Genay, A.; Lécuyer, A.; Hachet, M. Being an Avatar "for Real": A Survey on Virtual Embodiment in Augmented Reality. *IEEE Trans. Vis. Comput. Graph.* **2022**, *28*, 5071–5090. [CrossRef]
55. Hammady, R.; Ma, M.; Strathearn, C. User experience design for mixed reality: A case study of HoloLens in museum. *Int. J. Technol. Mark.* **2019**, *13*, 354–375. [CrossRef]
56. Sutton, A.; Clowes, M.; Preston, L.; Booth, A. Meeting the review family: Exploring review types and associated information retrieval requirements. *Health Inf. Libr. J.* **2019**, *36*, 202–222. [CrossRef]
57. Grave, R.B.d.; Bull, C.N.; Monteiro, D.M.d.S.; Margariti, E.; McMurchy, G.; Hutchinson, J.W.; Smeddinck, J.D. Smartphone Apps for Food Purchase Choices: Scoping Review of Designs, Opportunities, and Challenges. *J. Med Internet Res.* **2024**, *26*, e45904. [CrossRef]
58. Hirzle, T.; Müller, F.; Draxler, F.; Schmitz, M.; Knierim, P.; Hornbæk, K. When XR and AI Meet—A Scoping Review on Extended Reality and Artificial Intelligence. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023; CHI '23. [CrossRef]
59. Oun, A.; Hagerdorn, N.; Scheideger, C.; Cheng, X. Mobile Devices or Head-Mounted Displays: A Comparative Review and Analysis of Augmented Reality in Healthcare. *IEEE Access* **2024**, *12*, 21825–21839. [CrossRef]
60. Pollock, D.; Peters, M.D.J.; Khalil, H.; McInerney, P.; Alexander, L.; Tricco, A.C.; Evans, C.; de Moraes, É.B.; Godfrey, C.M.; Pieper, D.; et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. *JBI Evid. Synth.* **2023**, *21*, 520–532. [CrossRef]
61. Kleinheksel, A.J.; Rockich-Winston, N.; Tawfik, H.; Wyatt, T.R. Demystifying Content Analysis. *Am. J. Pharm. Educ.* **2020**, *84*, 7113. [CrossRef]
62. Fischer, M.; Leuze, C.; Perkins, S.; Rosenberg, J.; Daniel, B.; Martin-Gomez, A. Evaluation of Different Visualization Techniques for Perception-Based Alignment in Medical AR. In Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Recife, Brazil, 9–13 November 2020; pp. 45–50. [CrossRef]
63. Wagner, U.; Lystbæk, M.N.; Manakhov, P.; Grønbæk, J.E.S.; Pfeuffer, K.; Gellersen, H. A Fitts' Law Study of Gaze-Hand Alignment for Selection in 3D User Interfaces. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023; CHI '23. [CrossRef]
64. Fischer, M.; Rosenberg, J.; Leuze, C.; Hargreaves, B.; Daniel, B. The Impact of Occlusion on Depth Perception at Arm's Length. *IEEE Trans. Vis. Comput. Graph.* **2023**, *29*, 4494–4502. [CrossRef]
65. Oshima, K.; Moser, K.R.; Rompapas, D.C.; Swan, J.E.; Ikeda, S.; Yamamoto, G.; Taketomi, T.; Sandor, C.; Kato, H. SharpView: Improved clarity of defocused content on optical see-through head-mounted displays. In Proceedings of the 2016 IEEE Symposium on 3D User Interfaces (3DUI), Greenville, SC, USA, 19–20 March 2016; pp. 173–181. [CrossRef]
66. Rosa, N.; Hürst, W.; Werkhoven, P.; Veltkamp, R. Visuotactile integration for depth perception in augmented reality. In Proceedings of the 18th ACM International Conference on Multimodal Interaction, Tokyo, Japan, 12–16 November 2016; ICMI '16; pp. 45–52. [CrossRef]
67. Venkatakrishnan, R.; Venkatakrishnan, R.; Canales, R.; Raveendranath, B.; Pagano, C.C.; Robb, A.C.; Lin, W.C.; Babu, S.V. Investigating the Effects of Avatarization and Interaction Techniques on Near-field Mixed Reality Interactions with Physical Components. *IEEE Trans. Vis. Comput. Graph.* **2024**, *30*, 2756–2766. [CrossRef]

68. Venkatakrishnan, R.; Venkatakrishnan, R.; Raveendranath, B.; Pagano, C.C.; Robb, A.C.; Lin, W.C.; Babu, S.V. Give Me a Hand: Improving the Effectiveness of Near-field Augmented Reality Interactions By Avatarizing Users' End Effectors. *IEEE Trans. Vis. Comput. Graph.* **2023**, *29*, 2412–2422. [CrossRef] [PubMed]
69. Weast, R.A.T.; Proffitt, D.R. Can I reach that? Blind reaching as an accurate measure of estimated reachable distance. *Conscious. Cogn.* **2018**, *64*, 121–134. [CrossRef]
70. Hart, S.G. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2006**, *50*, 904–908. [CrossRef]
71. Quek, F.; McNeill, D.; Bryll, R.; Duncan, S.; Ma, X.F.; Kirbas, C.; McCullough, K.E.; Ansari, R. Multimodal human discourse: Gesture and speech. *ACM Trans. Comput.-Hum. Interact.* **2002**, *9*, 171–193. [CrossRef]
72. Katz, M. Convergence Demands by Spectacle Magnifiers. *Optom. Vis. Sci.* **1996**, *73*, 540. [CrossRef]
73. Microsoft. Comfort—Mixed Reality. Available online: https://learn.microsoft.com/en-us/windows/mixed-reality/design/comfort (accessed on 13 June 2025).
74. Feld, N.; Pointecker, F.; Anthes, C.; Zielasko, D. Perceptual Issues in Mixed Reality: A Developer-oriented Perspective on Video See-Through Head-Mounted Displays. In Proceedings of the 2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Bellevue, WA, USA, 21–25 October 2024; pp. 170–175. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.