Journal on Interactive Systems, 2025, 16:1, doi: 10.5753/jis.2025.4694

An Updated Systematic Mapping Study on Usability and User Experience Evaluation of Touchable Holographic Solutions

Thiago Campos [**Federal University of Paraná** | ***contact@thiagotpc.com***]
Maria Castello [**Federal University of Paraná** | ***mariaemiliacastello@gmail.com***]
Eduardo Damasceno [**Federal University of Technology of Paraná** | ***damasceno@utfpr.edu.br***]
Natasha Valentim [**Federal University of Paraná** | ***natasha@inf.ufpr.br***]

Department of Informatics - Federal University of Paraná (UFPR), Rua Cel. Francisco H. dos Santos, 100, Centro Politécnico, Jardim das Américas, Curitiba, PR, 81531-980, Brazil, PO Box 19081.

Received: 11 July 2024 • **Accepted:** 05 December 2024 • **Published:** 01 January 2025

Abstract: This article presents an extended Systematic Mapping Study (SMS) focused on usability and user experience (UX) evaluation technologies for Touchable Holographic Solutions (THS). Given the growing integration of holograms in Augmented Reality (AR) and Mixed Reality (MR) settings, evaluating usability and UX becomes highly important. Our study expands on previous research by analyzing an additional two years of publications, covering 5429 studies, and selecting 65 that discuss 200 evaluation technologies. The main problem addressed is the gap in comprehensive evaluation frameworks that integrate usability and UX criteria. We followed systematic guidelines to identify and analyze evaluation technologies, highlighting an increased focus on UX alongside traditional usability. Key findings include the persistent emphasis on time efficiency in usability evaluations and the dominance of generic UX, usability, and pleasure/fun in UX assessments. However, unique aspects of MR, such as presence, are often overlooked. The study also reveals a preference for empirical validation through controlled experiments and case studies, although few technologies have undergone such validation. Head-mounted displays (HMDs) and smart glasses, especially Microsoft Hololens™, remain prevalent due to their advanced capabilities. Our findings underscore the need for integrated evaluation technologies and empirical validation to ensure reliability. This work contributes to the Human-Computer Interaction (HCI) area by mapping current evaluation technologies, identifying research gaps, and providing a foundation for developing innovative and effective evaluation methods for THS, thus advancing the understanding and improvement of user interaction in immersive environments.

Keywords: Usability, User Experience, Evaluation, Touchable Hologram, Mixed Reality, Augmented Reality, Systematic Mapping Study

1 Introduction

The future of hand interaction might be the direct manipulation of holograms — 3D virtual objects integrated into the real world within Augmented Reality (AR) or Mixed Reality (MR) settings, supported by wearable techs like smart glasses or head-mounted displays (HMDs). This kind of interaction, long envisioned and popularized in science-fiction narratives such as Iron Man 2 [Favreau, 2010], is a key aspect of Spatial User Interaction, using physical space properties to offer immersive and intuitive experiences [Chaoui *et al.*, 2023]. Spatial User Interfaces represent a user-centric interface approach, focusing on freedom of movement and personalized interaction design [Koçer Özgün and Alaçam, 2023].

Because it is an interactive system, touchable holography requires quality evaluation, especially with regard to usability and User Experience (UX) [Barbosa *et al.*, 2021], [ISO, 2011]. Usability evaluates ease of use [Nielsen, 2012] or effectiveness, efficiency, and satisfaction with using an artifact with a specific purpose and context [ISO, 2018]. UX is a person's quality of experience when interacting with a product [Hassenzahl, 2011]. UX focuses on user preferences, perceptions, emotions, and physical and psychological responses that occur before, during, and after use [Bevan *et al.*, 2015]. Evaluating the quality of interactive systems allows identifying problems to be corrected, reducing the development process's time, effort, and cost [Barbosa *et al.*, 2021].

Although there are known technologies for evaluating usability and UX in interactive artifacts, like *System Usability Scale* (SUS) [Brooke, 1996] and *User Experience Questionnaire* (UEQ) [Laugwitz *et al.*, 2008], interaction with touchable holography emerges as a new context of use. In this context, it is new for the user, the device (smart glasses/HMDs), the interaction objects (virtual 3D objects), and the interaction scenario (AR/MR environment without the limitation of a screen). The main novelty in this context is the interaction by mid-air gestures that simulate touch. The user naturally touches virtual objects introduced in their real environment, perceiving that objects are really there. In addition to the already-know usability and UX aspects, the interaction with an MR hologram/environment has peculiar aspects and characteristics, such as immersion and presence.

Immersion includes the key features of a virtual world, providing users with things to see, hear, and do [Mestre and Vercher, 2011]. Aspects such as screen quality and what you can see affect how immersive it feels [Slater, 2018]. Better resources, such as clearer images, wider views, and responsive controls, make the experience more immersive. These resources, such as image clarity, field of view, and the speed with which they respond, affect how real it feels [Bowman and McMahan, 2007]. Presence comes from immersion and is based on "telepresence," that is feeling like you are really there when you are doing things with good sensory feedback [Skarbez *et al.*, 2017]. Presence is the human reaction to immersion [Slater, 2003], it is the personal feeling [Berkman and Akan, 2019], as a psychological state [International Society for Presence Research, 2000]. Given the same immersive system, people can experience different levels of presence [Slater, 2003].

Therefore, our research aims to identify the technologies used to evaluate Usability and UX in touchable holograms. For this identification, we conducted a Systematic Mapping Study (SMS) aiming to provide an overview (state-of-the-art) of a research area and to identify the amount, type of research, and results available [Petersen *et al.*, 2008, 2015; Kitchenham *et al.*, 2016]. Our SMS [Campos *et al.*, 2023b] followed the guidelines of Kitchenham *et al.* [2016] and Petersen *et al.* [2008] and extracted data from publicized studies until April/2021 about the evaluation technologies (such as types of methods, which aspects they cover, and whether they were known) and holographic solutions (such as their purpose, devices involved, type of gesture, and feedback).

However, the substantial volume of publications collected in the SMS implied a data selection and extraction process that lasted almost two years. In a scenario of constant technological advancement, the evolution of publications during this period can potentially influence or even reveal new trends on the topic in question. This motivated the decision to extend the SMS, covering a more recent time window. Although the guidelines of Kitchenham *et al.* [2016] do not include an explicit workflow for SMS updates, the authors frequently emphasize the importance of well-documented and transparent systematic reviews, features that are critical to allowing future updates. The practice of updating systematic reviews is recognized as important for maintaining the relevance and accuracy of conclusions, especially in fields with rapid technological evolution.

Following the approach of Mendes *et al.* [2020], updating an SMS involves three stages. Initially, the focus is on ensuring the continued relevance of the topic, the impact of the original study, and the high-quality methodology. Despite the recent SMS lacking substantial citations in other works, it has methodological soundness and talks about a relevant topic, justifying an extension to a more recent time window. The second stage aims to identify new studies or methods after the original SMS's publication, and it was done by watching research group activities and relevant conferences. Finally, the decision to update the SMS passed by the analysis of the potential impact on conclusions, considering that including recent studies can modify or enhance the credibility of prior SMS results.

The Extended SMS identified an increase in the proportion of articles in the engineering area compared to the initial SMS, reinforcing the multidisciplinary factor and cooperation in research of this type. An increase in evaluations based on UX criteria was also noted. Neither SMS found any specific evaluation technology to evaluate usability and UX together in the context of touchable holographic solutions. The SMS identified some well-known technologies applied to the evaluation of these types of solutions (such as Think-Aloud, interviews, and the application of the SUS and UEQ questionnaires), but with limitations of not covering all possible usability or UX aspects in an evaluation, including distinctive aspects of this new context of use, such as immersion and presence. In the Extended SMS, it was also noted a tendency to use the term AR to refer to solutions of this type, as well as the predominance of the use of HMDs, in particular, the MS Hololens still dominant, but also with the first appearances of Quest equipment. This SMS allows a better understanding of the state-of-the-art of usability and UX evaluation in touchable holographic and contributes to classifying the technologies applied to these solutions.

This study contributes by: (1) updating the SMS with recent publications, ensuring relevance in the evolving field of HCI; (2) analyzing trends and gaps in usability and UX evaluation of THS, providing a foundation for identifying research opportunities and guiding the development of more comprehensive evaluation frameworks and methodologies; and (3) providing insights into the evolution of evaluation technologies in AR/MR environments, fostering innovation and interdisciplinary progress.

In the following section, we present other works that sought to classify holographic solution evaluation research. In Section 3, we present information about the planning, execution, and reporting of the SMS conducted. In Section 4, we present the results in relation to the research sub-questions, obtained in both SMS. Section 5 provides a discussion of the results. Section 6 presents limitations and threats to the study. Section 7 presents conclusions and future work.

2 Related Work

Swan and Gabbard [2005] categorized user-based studies conducted using AR, covering publications from 1992 to 2004 from three conferences and one journal. Of the 1104 publications analyzed, 266 contained some aspect of AR research, and 21 publications described user-based experiments. They classified the experiments into three categories: studies on understanding human perception and cognition in AR contexts, investigating user task performance in specific AR applications, and examinations of generic user interaction and communication involving multiple participants.

In parallel, Dünser *et al.* [2008] reviewed publications from 1992 to 2007 using the Web of Science library. They used two classification schemes to categorize the publications. The first scheme classified the publications based on their area and type of evaluation, adding the Usability category to the classification previously established by Swan and Gabbard. Of 161 publications, 41 were classified as Usability type. The second scheme focused on the methods and approaches used, categorizing the evaluations into objective measurements, subjective measurements, qualitative analysis, usability evaluation techniques, and informal evaluations. Among the publications, 7 of 161 were classified using usability evaluation techniques, including heuristic evaluation, task analysis, Think Aloud, and Wizard of Oz methods. Other usability evaluation methods, such as observation, questionnaires, and interviews, were placed in different categories. The review indicated that, at the time, formal qualitative analysis and more general usability assessment techniques were uncommon.

Bai and Blackwell [2012] reviewed papers from the International Symposium of Mixed and Augmented Reality (ISMAR) proceedings between 2001-2010. They focused on evaluating goals, measurements, and methods. The usability research was classified into performance, perception and cognition, collaboration, and UX. They analyzed the balance of evaluation approaches and identified performance goals, metrics, UX factors, and evaluation methods for each AR application. They found 17 papers that evaluated UX through questionnaires, physiological measurements, observation, feedback, and user attitude. The study highlighted technological limitations, the lack of a unified strategy for function-related UX evaluation, and the challenge of learning due to the novelty and limited experience as the main challenges in UX evaluation for AR.

Merino *et al*. [2020] reviewed publications from four conferences between 2009 and 2019. The review categorized publications based on paper type, research topic, evaluation scenario, cognitive aspects, and context. Two main groups emerged: technology-centric evaluations focusing on algorithm performance and human-centric studies analyzing application implications, design, and human factors. The review noted a growing number of evaluation papers, with user performance and UX becoming more common evaluation scenarios. Usability, presence, and cognitive load were evaluated using methods such as Post Experience Questionnaire (PEQ) [**?**], Igroup Presence Questionnaire (IPQ) [Schubert *et al*., 2001], and NASA Task Load Index (NASA-TLX) [Hart and Staveland, 1988]. Additionally, a portion of the studies examined participants' emotions and presence.

Recently, Marques *et al*. [2024] investigated the evaluation of UX in the context of immersive experiences. The authors conducted an SMS to explore current UX evaluation techniques and the key UX dimensions for immersive experiences, such as engagement, presence, and immersion. The study's primary outcome was a theoretical model based on literature definitions and relationships designed to assist in understanding UX dimensions and developing new evaluation techniques. The authors recognize that traditional UX assessment methods may not be adequate for immersive contexts. The findings emphasize the predominance of questionnaire and interview-based evaluation methods and highlight the main UX dimensions relevant to immersive experiences. Examples of the immersive experiences examined include Immersive VR, 360-degree videos, and the use of HMDs.

Our review differs from previous ones, focusing on touchable holographic solutions rather than generic augmented or mixed-reality environments. Unlike previous reviews, which were limited to specific conferences, our SMS covers a broader range of sources. Furthermore, our approach includes different classifications compared to the previous works. Apart from documenting the usability and UX evaluation methods, we also investigate the specific technologies employed and the usability or UX aspects covered by these evaluation methods. Additionally, we classify the experiments based on the methodology, number of participants, and type of analysis. We categorize solutions based on holographic display and interaction detection technologies, gesture types, holographic quality, and user feedback.

During the preparation of our studies, searches were carried out in the Journal of Interaction Systems (JIS) and SBC Open Lib (SOL) using keywords such as ["systematic mapping" OR "literature review"] AND ["augmented reality" OR "mixed reality" OR "holog*" and [evaluation OR assessment] AND ["mixed reality" OR "augmented reality" OR "holog*"]. These searches did not identify systematic reviews or mappings directly related to the evaluation of usability and UX in THSs. Some secondary studies found, such as "Use of VR and AR in exergames for the elderly" [Pereira *et al*., 2021], are relevant in the general context of AR and MR, but do not address direct interaction with the hand in virtual elements created in AR or MR. This gap somewhat reinforces the originality of the present work, which focuses on the analysis of SHTs in AR and MR contexts, highlighting the need for integrated frameworks for usability and UX assessment, an approach investigated in the literature.

3 Systematic Mapping Study

3.1 Initial and Extended

To facilitate the presentation of results and discussions, the first study will be called Initial or SMS1 and the second, Extended or SMS2. Both studies shared the same objectives, research questions, data sources, search string, and selection criteria, and were performed using the same support tool. The only difference between them was the time window of publications, given by the submission date of the search string.

3.2 Planning

3.2.1 Goal

The SMS goal was based on Goal-Question-Metrics (GQM) paradigm [Basili *et al*., 1994]: **Analyze** scientific publications **for the purpose of** identifying and characterizing **with respect to** usability and UX evaluation technologies applied to touchable holographic solutions **from the point of view of** Human-Computer Interaction (HCI), Software Engineering (SE), AR and MR researchers **in the context of** available publications at the ACM Digital Library, IEEE Xplore, and Elsevier Scopus digital libraries.

3.2.2 Research Questions

The main question of this SMS was "What technologies are used to evaluate usability and UX in touchable holographic solutions and interfaces?". The aspects covered in the evaluation and the main characteristics and context of the evaluated holographic solution are expected to be understood by identifying the technologies used. A series of sub-questions were created to answer this main research question (Tables 4, 5 and 6).

The word "technology" is delimited in its intellectualism, utilitarian, and instrumentalist sense [Veraszto *et al*., 2009]. Therefore, technology was understood as practical knowledge derived directly and exclusively from the development of scientific theoretical knowledge, as a synonym for technique, and; as a simple tool or artifact built for a task. "Technology" generalizes procedures, tools, techniques, methodologies, methods, and other proposals [Santos *et al.*, 2012].

3.2.3 Data Sources

The defined search string was applied to the ACM Digital Library[1], IEEE Xplore[2] and Elsevier Scopus[3] digital libraries. The first is specialized in publications in Computing; the second, in Engineering, Computing, Information Technology, among others and; the third covers several interdisciplinary bases, totaling approximately 240 disciplines. In this way, we sought relevant publications in Computing without discarding the possibility of finding relevant publications in multidisciplinary studies or studies carried out outside Computing. No other filters were applied.

3.2.4 Search String

The chosen terms were based on the previous knowledge of the researchers involved in this SMS. Based on the *Population, Intervention, Comparison, Outcome, and Context* (PICOC) [Kitchenham *et al.*, 2016] strategy, the final string contained the Population, Intervention, and Outcome components 1. The Comparison and Context components were not used because the SMS goal was to characterize and not compare evaluation technologies. The boolean operator "OR" was used to indicate synonyms or alternatives, and the boolean operator "AND" was used to join the terms in each component and all the components. The search string was ((“holographic” OR “holography” OR “hologram” OR “augmented reality” OR “extended reality” OR “mixed reality”) AND (“touchable” OR “tangible” OR “touchless” OR “mid-air”)) AND (“method” OR “methodology” OR “approach” OR “technique” OR “solution” OR “framework” OR “proposal” OR “tool”) AND ((“usability” OR “user experience” OR “UX” OR “user satisfaction”) AND (“evaluation” OR “assessment”)).

3.2.5 Criteria

Some criteria were established for excluding and including publications to select publications that contain helpful information for extracting data and answering research subquestions. A publication was excluded if it met any exclusion criteria and included if it met any inclusion criteria. The inclusion criteria were I1: The publication evaluates touchable holographic solutions in terms of UX or usability; I2: The publication presents a new UX and/or Usability evaluation proposal for a touchable holographic solution; and E3: The publication evaluates a new UX and/or Usability evaluation proposal for a touchable holographic solution.

The exclusion criteria were E1: Publication not available in English or Portuguese; E2: Publication is not peer-reviewed; E3: The publication is not a full paper and may not contain research details; E4: Publication is not a single article; E5: Publication requires payment for access and is not accessible through institutional proxies available to SMS researchers; E6: The publication has already been included by another digital library used in this SMS. It is a duplicate publication; E7: Publication does not report a primary study but is a secondary one (systematic review, mapping, or other surveys); and E8: The publication does not meet any inclusion criteria. It does not meet the scope and main research question.

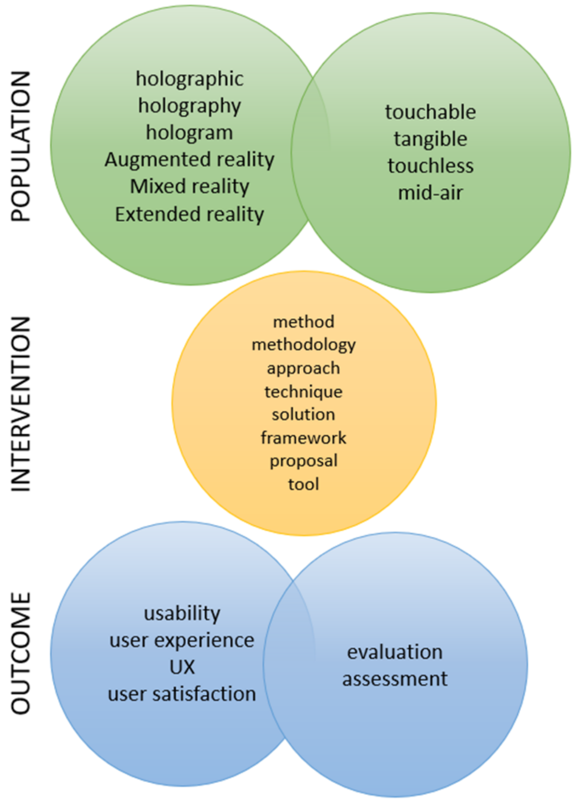

Figure 1. Search String Composition

3.3 Execution

3.3.1 Publication selection

The studies used the collaborative tool Porifera [Campos *et al.*, 2022]. The publication selection process had two stages (two filters), run collaboratively and with predefined criteria (Figure 2). Three researchers participated in the publication selection process, aiming mainly at reducing the bias that would occur if only one carried out the entire process. A doctoral student, his co-advisor, and her advisor carried out the Initial SMS. The Extended SMS was conducted by an undergraduate in her Completion of Coursework, with the participation of the same doctoral student and advisor from the previous study.

In the **first stage**, the publications were analyzed based on their metadata (title, abstract, type of publication, among others). Each of the three researchers individually evaluated the first 15% of the publications and assigned an inclusion or exclusion criterion without knowing the criteria applied by their colleagues. If there was agreement, the publication moved on to the next phase. If there was disagreement, the researchers discussed it in a meeting.

This process was limited to 15% of the publications in

[1] https://dl.acm.org/
[2] https://ieeexplore.ieee.org/
[3] https://www.scopus.com/

the first filter due to the high number of publications obtained. Review by all researchers on all items would take more time because it would require, in addition to the individual assessment of each researcher, meetings to discuss and establish decision-making consensus. These first evaluations served to equalize knowledge and adjust the application of the criterion among researchers, obtaining a reliable level of agreement based on the Kappa Fleiss coefficient [Fleiss, 1971], which allowed the continuation of the evaluation of the remaining 85% of the publications by only one of the researchers to accelerate the process. In SMS 1, the postdoctoral student completed the reviews on the publications of the first filter and, in SMS 2, the undergraduate finished the reviews on publications of first filter..The simple agreement and the Kappa index are calculated based on the nature of the criteria assigned by each researcher.

The Initial SMS collected 3531 publications from data sources, of which 266 (about 7.5%) were accepted in the first stage. The first stage raised 0.7767 Kappa Fleiss index and 92.08% of simple agreement. The Kappa Cohen coefficients, which consider evaluators in pairs, had values of 0.723 for Researcher 1 (doctoral student) and Researcher 2 (advisor), 0.798 for Researcher 1 and Researcher 3 (co-advisor) and; 0.805 between Researcher 2 and 3. These values are classified as substantial byLandis and Koch [1977] and good, by Altman [1990].

The Extended SMS collected 1878 publications from data sources, of which 100 (about 5,3%) were accepted in the first stage. The first stage raised 0.6213 Kappa Fleiss index and 91.49% of simple agreement. The Kappa Cohen coefficients had values of 0.658 for Researcher 1 and Researcher 2, 0.606 for Researcher 1 and Researcher 4 (undergraduate) and; 0.607 between Researcher 2 and 4. These values are classified as substantial byLandis and Koch [1977] and good, by Altman [1990] for the first pair of researchers, and moderate in both interpretations [Landis and Koch, 1977; Altman, 1990] for the second and third pair.

In the **second stage**, all the publications accepted in the first filter were evaluated by all researchers. In this step, the full text of the publications was accessed, and the publications were analyzed in light of their entire content. In the same way, each of the three researchers made their assessment individually, and if there was agreement, the publication moved on to the data extraction phase. When there was disagreement, the researchers discussed and decided by consensus.

In this step, the Initial SMS selected 40 publications for data extraction, obtaining a simple agreement of 81.58%, and Kappa index of 0.5779, considered "moderated" in both interpretations [Altman, 1990; Landis and Koch, 1977]. The Kappa Cohen coefficients were 0.593 for Researcher 1 and Researcher 2, 0.669 for Researcher 1 and Researcher 3 and; 0.475 between Researcher 2 and 3. These values are classified as substantial moderate in both classifications for the first and third pairs and, substantial and good, respectively by Landis and Koch [1977] and Altman [1990] for the second researcher's pair.

The Extended SMS selected 25 publications for data extraction, obtaining a simple agreement of 88.0%, and Kappa index of 0.7847, considered "good" in interpreting by Altman [1990] and "substantial" by Landis and Koch [1977]. The Kappa Cohen coefficients had values of 0.811 for Researcher 1 and Researcher 2, 0.811 for Researcher 1 and Researcher 4 and; 0.733 between Researcher 2 and 4. These values are classified as "almost perfect" by Landis and Koch [1977] and "very good" by Altman [1990] for the first and second researcher's pairs, and substantial and good by the same authors respectively for the third researcher's pair.

Combined, both studies collected 5409 publications, of which 366 (approximately 6.8%) passed through the first filter and 65 passed through the second filter and were moved to data extraction (Tables 1 and 2).

Table 1. Selected Publications in Initial SMS

Publication	Study
[Williams *et al.*, 2020]	Initial
[Dehghani *et al.*, 2020]	Initial
[Jasche and Ludwig, 2020]	Initial
[Kim *et al.*, 2020]	Initial
[Lee *et al.*, 2020]	Initial
[Whitlock *et al.*, 2018]	Initial
[Vaquero-Melchor and Bernardos, 2019]	Initial
[Seiger *et al.*, 2021]	Initial
[Xu *et al.*, 2020]	Initial
[Kang *et al.*, 2020]	Initial
[Mahajan *et al.*, 2020]	Initial
[Plasson *et al.*, 2019]	Initial
[Aslan *et al.*, 2019]	Initial
[Qian *et al.*, 2019]	Initial
[Matsumaru *et al.*, 2019]	Initial
[Xu *et al.*, 2019]	Initial
[Munsinger *et al.*, 2019]	Initial
[Kim *et al.*, 2019]	Initial
[Becker *et al.*, 2019]	Initial
[Frutos-Pascual *et al.*, 2019]	Initial
[Dudley *et al.*, 2018]	Initial
[Whitlock *et al.*, 2018]	Initial
[Cao *et al.*, 2018]	Initial
[Lin *et al.*, 2017]	Initial
[Kim and Lee, 2016]	Initial
[Shim *et al.*, 2016]	Initial
[Brancati *et al.*, 2015]	Initial
[Caruso *et al.*, 2015]	Initial
[Adhikarla *et al.*, 2015]	Initial
[Ha *et al.*, 2014]	Initial
[Sand *et al.*, 2015]	Initial
[Lee *et al.*, 2019b]	Initial
[Wright *et al.*, 2019]	Initial
[Mau-Tsuen Yang and Wan-Che Liao, 2014]	Initial
[Zhu and Grossman, 2020]	Initial
[Weichel *et al.*, 2014]	Initial
[Lee *et al.*, 2019a]	Initial
[Higuchi and Komuro, 2013]	Initial
[Lee *et al.*, 2013]	Initial
[Huang *et al.*, 2015]	Initial

Table 2. Selected Publications in Extended SMS

Publication	Study
[Bozgeyikli and Bozgeyikli, 2021]	Extended
[Chien and Lin, 2021]	Extended
[Daskalogrigorakis *et al*., 2022]	Extended
[Flick *et al*., 2021]	Extended
[He *et al*., 2022]	Extended
[Hu *et al*., 2022]	Extended
[Jang *et al*., 2022]	Extended
[Li *et al*., 2022]	Extended
[Liu *et al*., 2023]	Extended
[Lystbæk *et al*., 2022a]	Extended
[Lystbæk *et al*., 2022b]	Extended
[McCord *et al*., 2022]	Extended
[Pei *et al*., 2022]	Extended
[Plasson *et al*., 2022]	Extended
[Qian *et al*., 2022]	Extended
[Serrano *et al*., 2022]	Extended
[Shen *et al*., 2022]	Extended
[Uzor and Kristensson, 2021]	Extended
[Venkatakrishnan *et al*., 2023]	Extended
[Wang *et al*., 2022]	Extended
[Wang *et al*., 2021a]	Extended
[Wang *et al*., 2021b]	Extended
[Weerasinghe *et al*., 2022]	Extended
[Yu *et al*., 2022]	Extended
[Zhao *et al*., 2023]	Extended

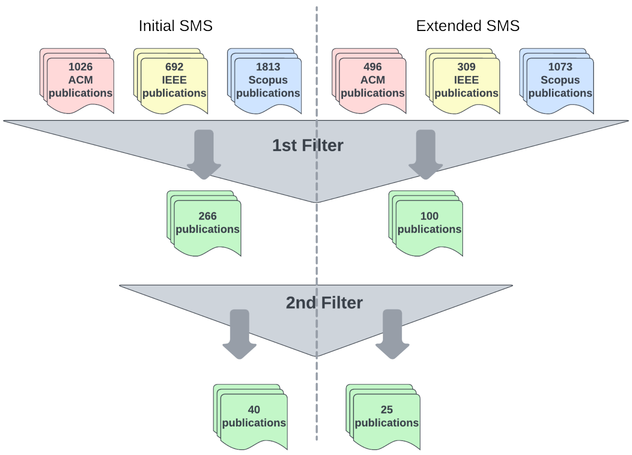

Figure 2. Publication selection

3.3.2 Data extraction

Data extraction is a phase of SMS that aims to respond to research sub-questions through information in selected publications. The extraction was carried out with the support of a template form instantiated for each publication. One of the SMS researchers performed the extraction, which was checked by the other researchers involved.

3.3.3 Data analysis

All the extracted data was tabulated in a spreadsheet to perform counting, statistical calculations, and graph plotting to better understand the numbers obtained. The following sections present the numbers and main results that were achieved.

3.4 Reporting

The string submission occurred on April 25, 2021, in the Initial SMS. The query result brought 3,531 publications, 1026 from the ACM Digital Library, 692 from IEEE Xplore, and 1813 from Scopus. In the Extended SMS, the string submission occurred on April 25, 2023, and obtained 1,878 publications, 496 from the ACM Digital Library, 309 from IEEE Xplore, and 1,073 from Scopus. Complete reports are available online for the Initial SMS [Campos *et al*., 2023a] and Extended SMS [Campos *et al*., 2024].

3.4.1 Publication years

In the Initial SMS, the selected publications were between 2013 and 2021, with the last three years (2019, 2020, and 2021) concentrating more than half of the period publications. In the Extended SMS, the search string was submitted on 25 April 2023 and only selected the publications after 25 April 2021.

The number of publications by year (Figure 3) shows significant waves. The first wave occurred between 2013 and 2015, after the launch of Google Glass™ in mid-April 2013. The production and availability of Google Glass™ were stopped in early 2015 for the public and resumed only after July /2017 for the business environment [Woolf, 2015]. This fact may explain the decline of the first wave. Before Google Glass™, producing touchable holography required creativity to combine equipment not designed for this purpose, such as the use of a Liquid Crystal Display (LCD) [Higuchi and Komuro, 2013], the use of reflective material [Weichel *et al*., 2014] or projection on a fog screen [Caruso *et al*., 2015; Sand *et al*., 2015]. Until then, most AR/MR applications were limited to the screen space of smartphones and tablets, where the main way of interaction was through the touchscreen.

The second wave occurred between 2017 and 2020. After the launch of Google Glass™, other companies sought to offer smart glasses equipment, such as MAD Gaze Ares™, announced in 2016. This was followed by the launch of HMD equipment, such as MS Hololens™ 1, Magic Leap One™ and MS Hololens™ 2, released in Mar/2016, Aug/2018 and May/2019, respectively. The availability of these new devices coincides with the beginning of the second wave of publications. In addition, in 2019, the Applied Sciences journal published a special issue entitled *Augmented Reality: Current Trends, Challenges and Prospects* [MDPI, 2024].

The COVID-19 pandemic may justify the decline in publications in 2020 and 2021 since experiments with users should take place in person, and the shared use of wearable equipment is commonly necessary. Soon after this period, a third wave started in the year 2022, which presented the highest number of publications, showing a resumption in research and confirming the topic's growth potential. The low number of publications in 2023 is justified by submitting the search string in the year's first four months of that year.

3.4.2 Publication locales

Regarding the geographical distribution of the publications (Figure 4 and Table 3), considering both studies, most are

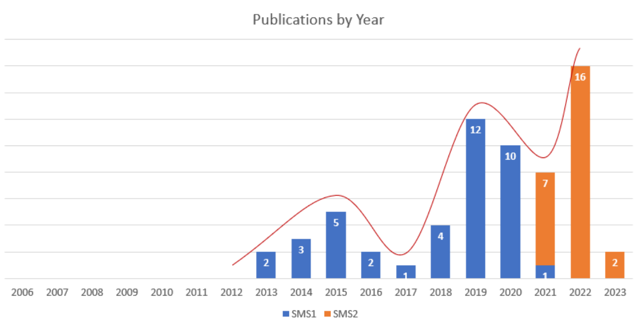

Figure 3. Publications by Year

from Europe (24 publications: United Kingdom, six; Germany, three; Denmark, Finland, France, Spain, two each and; Greece, Hungary and Slovenia, one publication each), North America (21 publications: United States, 18 and; Canada, three), and Asia (20 publications: China, seven; South Korea, five; Hong Kong, four; Japan and Taiwan, two). If the publication said the locale of the study, like a laboratory or a university, this information was considered. If not, the country of most of the article's authors or the first author was considered the origin.

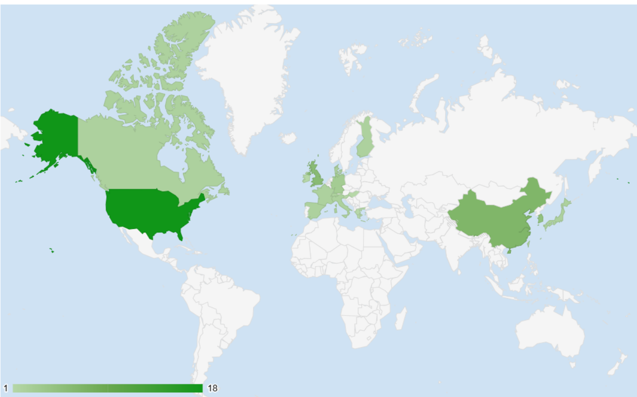

Figure 4. Publications by Country

In both studies, the continental distribution of publications did not undergo significant changes. Europe was the birthplace of most publications in both studies (37.5% and 40%, respectively). Asia ranged from 32.5% to 28%, and North America from 30% to 32%. The United States predominated leadership in both studies (25% and 32%, respectively). China's growth in publications (from 5% to 20%), as did the United Kingdom (7.5% to 12%). Of the nations that occupied a relevant position in the first study and decreased in the second, we highlight South Korea (from 10% to 4%) and Hong Kong (from 10% to 0%).

According to common sense, the countries listed are more developed and have a higher per capita income than most countries in regions such as South or Central America and Africa. In addition, some vendors do not make devices available globally, concentrating their sales in the so-called first-world countries. Also, there are many well-structured laboratories and research groups dedicated to AR/MR/VR research in the countries from which the publications originated, for example, the Ubiquitous Virtual Reality (UVR) Lab in South Korea [4]; the MIT Media Lab [5] in the United States; and the Digital Media Technology (DMT) Lab, in the UK [6].

[4]https://uvrlab.org

[5]https://www.media.mit.edu/

Table 3. Publications by Locales

Locale	**SMS1**	**SMS2**	**Sum**
Europe	14	10	25
North America	13	8	21
Asia	13	7	20
Countries			
United States	10	8	18
China	2	5	7
United Kingdom	3	3	6
South Korea	4	1	5
Hong Kong	4	0	4
Germany	2	1	3
Canada	3	0	3
Denmark	0	2	2
Finland	2	0	2
France	1	1	2
Italy	2	0	2
Japan	2	0	2
Spain	1	1	2
Switzerland	2	0	2
Taiwan	1	1	2
Greece	0	1	1
Hungary	1	0	1
Slovenia	0	1	1
Total	**40**	**25**	**65**

3.4.3 Publication venue

Both studies showed similar distributions in terms of the publication vehicle (event or journal). In the Initial SMS, almost two-thirds of the publications (65%, 26 publications) took place in events (conferences, symposiums, among others), and about a third (35%, 14 publications) were in journals. In the second SMS, there were 15 (60%) publications in events and 10 (40%) in journals.

About events (Figure 5), considering both studies, 53.6% (22 of 41) of the publications focused on four events: at the IEEE Conference on Virtual Reality and 3D User Interfaces (IEEEVR); ACM Conference on Human Factors in Computing Systems (ACM CHI); International Symposium of Mixed and Augmented Reality (ISMAR); and ACM Symposium on User Interface Software and Technology (ACM UIST). These four events are references in the VR/AR/MR (IEEEVR and ISMAR) and HCI (ACM CHI and ACM UIST). ACM CHI is the premier conference in the field of HCI, organized by the Special Interest Group on HCI (Special Interest Group on Human-Computer Interaction, SIGCHI). IEEEVR emerged from the union of two events, one dedicated to VR and the other to 3D User Interface. ISMAR emerged from the merger of two events, one dedicated to AR and the other to MR, and is also supported by ACM SIGCHI. Finally, UIST brings together researchers and practitioners from diverse fields that include graphical and web user interfaces, tangible and ubiquitous computing, VR and AR, and multimedia, among other topics. The Extended SMS registered a growth of publications at the ISMAR conference (2 to 6 publications).

[6]https://www.bcu.ac.uk/computing/research/digital-media-technology

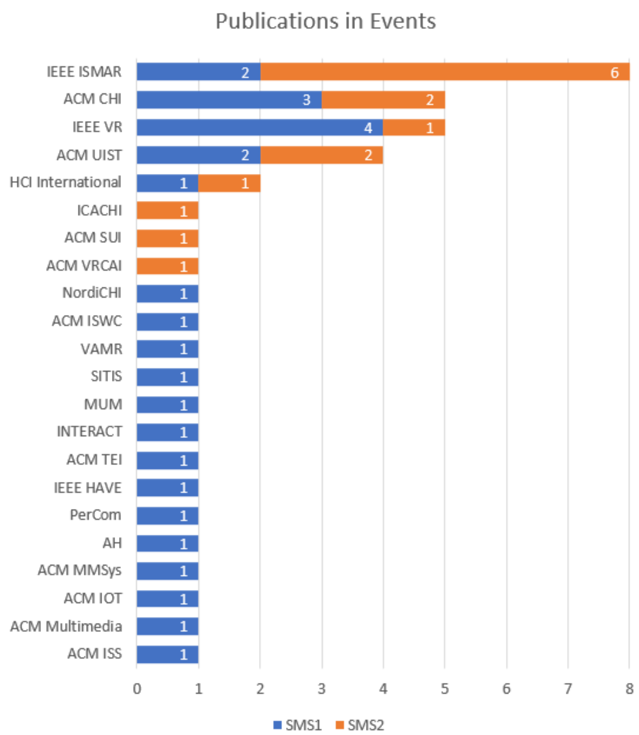

Figure 5. Publications in Events

Regarding the publications in journals (Figure 6), the second study brought a substantial contribution from the IEEE Transactions on Visualization and Computer Graphics, which became the most frequent publication vehicle in the sum of the studies. In second place comes the Applied Sciences, which stood out in the first study due to a special edition called "*Augmented Reality: Current Trends, Challenges, and Prospects* (2009)" and a new publication selected in the second study. Other vehicles with more than one publication were Proceedings of the ACM on Human-Computer Interaction (PACMHCI), ACM Transactions on Computer-Human Interaction, and Multimedia Tools and Applications. The other publications were distributed in journals with different focus, including VR, Computer Graphics and Visualization, Applied Ergonomics, HCI, and Technology in Society.

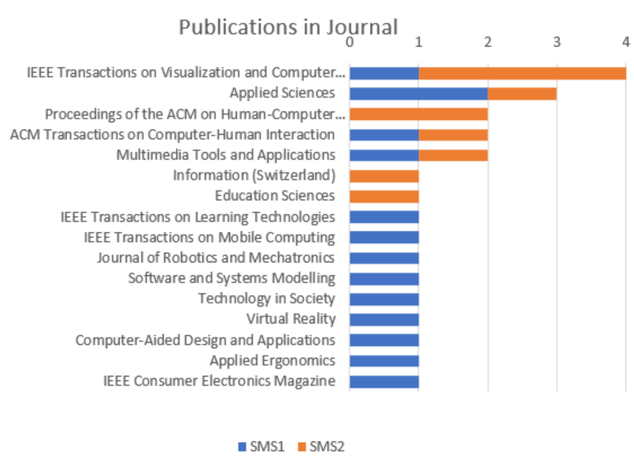

Figure 6. Publications in Journals

4 Findings

4.1 SQ1 - Publication Contribution Type

As proposed by Wieringa *et al*. [2006], the type of contribution can be evaluation research, proposal of solution, validation research, philosophical, opinion, or personal experience. Knowing the publication's purpose helps to understand the research stage within the engineering cycle. All publications in both studies were of the validation research contribution type. In this contribution type, the investigation focuses on the properties of a proposal not yet implemented in practice. Possible research methods are experiments, simulation, prototyping, mathematical analysis, and mathematical proof of properties, among others [Wieringa *et al*., 2006].

In these studies, the selected publications focused on investigating the properties of a touchable holographic solution in a laboratory or controlled environment by user experimentation and applying usability or UX evaluation technologies. For example, Williams *et al*. [2020] investigated how users manipulate virtual objects in AR using multimodal (gesture+voice) and unimodal (gestures only) interaction. Using an HMD, each participant performed some manipulation tasks and then answered a NASA Task Load Index (NASA-TLX) questionnaire [Hart and Staveland, 1988]. The paper presents the most used gestures and the least effort way to interact.

Only one publication (1.5%, 1 of 65) of Initial SMS presented a proposal of solution. A proposal of solution offers a technical solution (new or an improvement of an existing technique) and defends its relevance through a proof of concept, example, or solid argument [Wieringa *et al*., 2006]. For example, Dehghani *et al*. [2020] presented a theoretical framework to evaluate solutions on the Windows Mixed Reality (WMR) platform regarding user satisfaction and behavioral intention based on components such as perceived functional benefits, perceived trust, perceived visual appeal, perceived immersion, perceived autonomy and interaction. This publication also presented a study in which participants performed tasks using HMD (MS Hololens™) and answered a questionnaire. The result verified the reliability and validity of the hypotheses of the theoretical model.

No publications' contributions were classified as evaluation, philosophical, opinion, or personal experience research papers.

4.2 SQ2 - Research Area

We tried to identify the origin of the publication's authors to determine what area (Administration, Computing, Engineering, among others) carried out the research. Solutions created outside the areas that usually study human-computer or human-machine interaction could be focused on practical applications of touchable holography.

In the results, the area of Computer Science dominated, with 53.8% (35 of 65). For example, Munsinger *et al*. [2019] is a publication from the Dept. of Computer Science at the University of Texas at San Antonio and Jang *et al*. [2022] is a publication from the School of Computing at KAIST in Daejeon, South Korea. However, the number of publications

Table 4. Summary Results - sub-questions related to publications

sub-question	**Possible Answers**	**SMS1**		**SMS2**		**Sum**	
SQ1. What is the type of contribution of the article? According to Wieringa *et al*. [2006]	Validation research	40	100%	25	100%	65	100%
	Proposal of solution	1	2.5%	0	0%	1	1.5%
SQ2. In which area was the research carried out?	Computing	26	65.0%	9	36%	35	53.9%
	Engineering	11	27.5%	11	44%	22	33.8%
	Other	3	7.5%	5	20%	8	12.3%

Table 5. Summary Results - sub-questions related to evaluations technologies

sub-question	**Possible Answers**	**SMS1**		**SMS2**		**Sum**	
SQ3. Which quality criteria is the focus of the evaluation technology?	Usability	85	80,2%	53	56.4%	138	69.0%
	UX	20	18,9%	39	41.5%	59	29.5%
	Usability and UX	1	0,9%	2	2.1%	3	1.5%
SQ4. What is the type of evaluation technology? For usability, according Ivory and Hearst [2001] For UX, according Roto *et al*. [2009]	Usability - test	33	38,4%	20	36.4%	53	37.6%
	Usability - inspection	1	1,2%	0	0.0%	1	0.7%
	Usability - inquiry	52	60,4%	35	63.6%	87	61.7%
	UX - survey	21	100%	41	100%	62	100%
SQ5. What aspects or metrics were considered in the evaluation technology?	Open answers						
SQ6. Was the identified assessment technology based on any existing technology, or was it created for the study?	Yes, fully	42	39,6%	28	29.7%	70	35.0%
	Yes, partially	7	6,6%	4	4.3%	11	5.5%
	No	57	53,8%	62	66.0%	119	59.5%
SQ7. Was the technology empirically evaluated?	Yes	1	0,9%	0	0%	1	0.5%
	No	105	99,1%	94	100%	199	99.5%

from Computer Science decreased from 65% (26 of 40) to 36% (9 of 25) between the Initial and the Extended SMS.

Another area that stood out was Engineering (Industrial, Electrical, Mechanical, and Biomedical), with 33.8% (22 of 65) publications. For example, Kim and Lee [2016] is a publication from the Dept. of Industrial Engineering at Chonnam National Univ. in Gwangju, South Korea, and McCord *et al*. [2022] is a publication from different departments of Engineering, like School of Sustainable Engineering and the Built Environment of the Arizona State University, Department of Civil and Environmental Engineering of University of Nebraska-Lincoln and, Department of Engineering Education of Virginia Polytechnique Institute, among others involved centers.

Other distinct areas were identified (12.3%, 8 of 65), like the Administration/Business [Dehghani *et al*., 2020], the Design [Kang *et al*., 2020], the "Culture Technology" in Humanities and Social Convergence Sciences [Ha *et al*., 2014] and the Medical Science [Yu *et al*., 2022] areas. In the Initial SMS were three publications (7.5% from 40) from other areas. This number increased in the Extended SMS to five publications (20% from 25).

4.3 SQ3 - Quality Criteria

Knowing the focus of Evaluation Technology (ET) can help to understand the goal of the ET and the aspects that the technology was looking to cover. The denomination given by the publication's authors was respected to determine the ET criterion. Some publications did not explicitly use usability or UX terms. In this case, it was classified as usability if the aspects or metrics described were related to pragmatic attributes or the classic components of usability, such as effectiveness and efficiency. It was classified as UX if the aspects or metrics described were hedonic attributes. For example, in Vaquero-Melchor and Bernardos [2019], a test-based ET, the metrics were "*time to perform the process*" and "*size error*". These metrics are common for effectiveness and efficiency aspects related to usability. Therefore, although not named by the authors, this ET was classified as a usability ET.

Both SMS identified 200 applications of evaluation technologies in the collected studies, 106 in the Initial SMS and 94 in the Extended SMS. Usability is the most present criterion, found exclusively in 69% (138 of 200) of the ETs. For example, in Mahajan *et al*. [2020], the solution was evaluated twice through a usability test with registered data and the SUS questionnaire. The Extended SMS result indicated a decrease in technologies exclusively focused on usability, from 80.2% (85 of 106) to 56.4% (53 of 94).

Almost one-third (29.5%, 59 of 200) of technologies focused exclusively on UX. This focus had an increment from Initial to Extended SMS, from 18.9% (20 of 106) to 41.5% (39 of 94). For example, in Xu *et al*. [2020], the UEQ questionnaire was applied to evaluate the UX of the touchable holographic solution, while, in Qian *et al*. [2019], a customized questionnaire, with Likert scale, was created to evaluate the UX of the solution presented by the authors.

Only three technologies (1.5%) addressed Usability and UX criteria together. In the Initial SMS, one ET was found, and in the Extended SMS, two ETs were found. This happened in Dehghani *et al*. [2020] through a prepared questionnaire by the authors, part of a framework to evaluate WMR solutions, in Plasson *et al*. [2022] when the participants filled out a questionnaire with open questions like "Are they hard?" and "Are they frustrating?" and in Li *et al*. [2022] where a 7-point Likert scale questionnaire encompassed issues about ease of learning, ease of performance, naturally intuitive, comfort, enjoyment, satisfaction, tiring, frustration, suitability for the task, and efficiency.

Table 6. Summary Results - sub-questions related to touchable holographic solutions

sub-question	Possible Answers	SMS1		SMS2		Sum	
SQ8. Has the holographic solution been subjected to an empirical study?	Yes	40	100%	25	100%	65	100%
	No	0	0%	0	0%	0	0%
SQ8.1 What is the empirical method used?	Controlled experiments	26	65.0%	20	80.0%	46	70.6%
	Case study	14	35.0%	6	24.0%	20	30.7%
SQ8.2 How many people participated in the study?	Average	17.6	—	18.4	—	—	—
	Median	13	—	15	—	—	—
	Min	4	—	4	—	—	—
	Max	98	—	54	—	—	—
SQ8.3 What was the type of analysis of the study results?	Quantitative	23	57.5%	8	32.9 %	31	47.7%
	Qualitative	1	2.5%	0	0.0 %	1	1.5%
	Both (quanti+quali)	16	40.0%	17	68.0%	33	50.8%
SQ9. How was holography named in the study?	Augmented Reality	29	72.5%	23	92.0%	52	80.0%
	Mixed Reality	9	22.5%	2	8.0%	11	16.9%
	Other	2	5.0%	0	0.0 %	2	3.1%
SQ10. How is holography presented to the user?	By screen	36	90.0%	25	100.0%	61	93.9%
	By optical-physical phen.	4	10.0%	0	0.0%	4	6.1%
SQ11. What type of technology is used to (re)produce holography?	Head-mounted display	23	57.5%	25	100.0%	48	73.8%
	Smart glass	7	17.5%	0	0.0%	7	10.8%
	Smartphone or tablet	3	7.5%	0	0.0%	3	4.6 %
	Others	3	7.5%	0	0.0%	3	4.6 %
	Reflexive material	2	5.0%	0	0.0%	2	3.1 %
	Fog screen projector	2	5.0%	0	0.0%	2	3.1 %
SQ11.1. Which manufacturer and model of technology(s) are used for projection?	Open answers.						
SQ12. What type of technology is used to detect interaction (touch)?	Near Infrared (NIR) external camera	9	22.5%	2	8.0%	11	16.9%
	RGB-D built-in camera	19	47.5%	23	92.0%	42	64.6%
	RGB-D external camera	7	17.5%	1	4.0 %	8	12.3%
	RGB built-in camera	4	10.0%	0	0.0%	4	6.1%
	RGB external camera	3	7.5%	0	0.0%	3	4.6%
SQ12.1 What manufacturer and model of technology(s) were used to detect the interaction?	Open answers.						
SQ13. For what purpose was the THS created?	Open answers.						
SQ14. Which touch gesture styles are allowed by the solution? According to Aigner *et al*. [2012]'s classification	Pointing	40	100,0%	25	100.0%	65	100.0%
	Pantomimic	26	65.0%	8	32.0%	34	52.3%
	Manipulation	15	37.5%	1	4%	16	24.6%
SQ15. What is the quality of the holographic projection image touched by the user?	WIMP	23	57.5%	15	60.0%	38	58.4%
	3D RAW object	17	42.5%	9	36.0%	26	40.0 %
	3D with light, shadow or texture	6	15.0%	5	20.0%	11	16.9%
	Not available	1	2.5%	0	0.0%	1	1.5%
SQ16. Is there feedback from the touch action? By which means?	Yes, visual	39	97.5%	23	92.0%	62	95.4%
	Yes, ultrassonic (haptic)	2	5.0%	2	8.0%	4	6.1%
	Yes, auditive	1	2.5%	8	32.0%	9	13.8%
	Not available	1	2.5%	1	4.0%	2	3.1%

4.4 SQ4 - Type of Evaluation Technology

According to Ivory and Hearst [2001], usability methods can be testing, inspection, inquiry, analytical modeling, or simulation. In UX, according to Roto *et al*. [2009], the methods can be a laboratory study, field study, survey, and expert evaluation. This sub-question is expected to obtain the most common methods and those that are less used in the context of touchable holographic solutions. The results for this question had little variation between the two studies, maintaining almost the same proportion between the results.

Regarding usability ETs, the most used method was the inquiry (61.7%, 87 of 141), when users give feedback through surveys, interviews, questionnaires, or related ways [Ivory and Hearst, 2001]. For example, in Weichel *et al*. [2014], the authors used the interview to evaluate the usability. In Shim *et al*. [2016], the authors applied the USE questionnaire to evaluate the solution. In Wang *et al*. [2021b], the authors applied the SUS and a custom 5-point Likert scale questionnaire to evaluate the solution's usability.

The second method was testing (37.6%, 53 of 141), when a researcher observes users interacting with the interface to determine usability issues [Ivory and Hearst, 2001]. For example, in Brancati *et al*. [2015], user testing with the Think-Aloud protocol was used to evaluate the usability. User testing was performed in Wright *et al*. [2019] and in Uzor and Kristensson [2021], with data recorded (like completion time and error rate) for further analysis.

Only one technology (0.7%), found in the initial SMS, focusing on usability, used the inspection method when an evaluator uses a set of criteria or heuristics to identify potential usability problems in the interface [Ivory and Hearst, 2001]. This happened at Jasche and Ludwig [2020], where experts inspected a holographic interface for 3D printing control. The Extended SMS found no other inspection method-based technology.

There were no analytical modeling or simulation methods. Analytical modeling is when an evaluator employs user and interface models to generate predictions about usability, and; simulation is when an evaluator uses user and interface models to simulate interaction and report the results [Ivory and Hearst, 2001].

Regarding UX ETs, the only method used was the survey (100%, 62 of 62). In surveys, user feedback is collected through interviews and questionnaires answered [Roto *et al*., 2009]. For example, in Kang *et al*. [2020], three assessments were performed using the survey. The first used a questionnaire with a Likert scale created for the solution, the second used an adaptation of the Technology Acceptance Method (TAM) to evaluate the perceived level of enjoyment, and the last was through user interviews [van der Heijden, 2004]. In Liu *et al*. [2023], the authors used the UEQ and a structured interview to assess the user experience.

Looking specifically at the Usability inquiry and UX survey types (Figure 7), which can be carried out through interviews or questionnaires, for example, we noticed that the questionnaire-based method was the most abundant among the evaluation technologies found (60%, 120 of 200 ETs). The interviews appeared more among the technologies that evaluated the UX criterion (33%, 21 of the 62 ETs that evaluated only UX or Usability+UX). In Extended SMS, there was a proportional increase in interview use in UX assessments, jumping from 23.8% (5 of 21) to 39% (16 of 41) of the technologies (Figure 8).

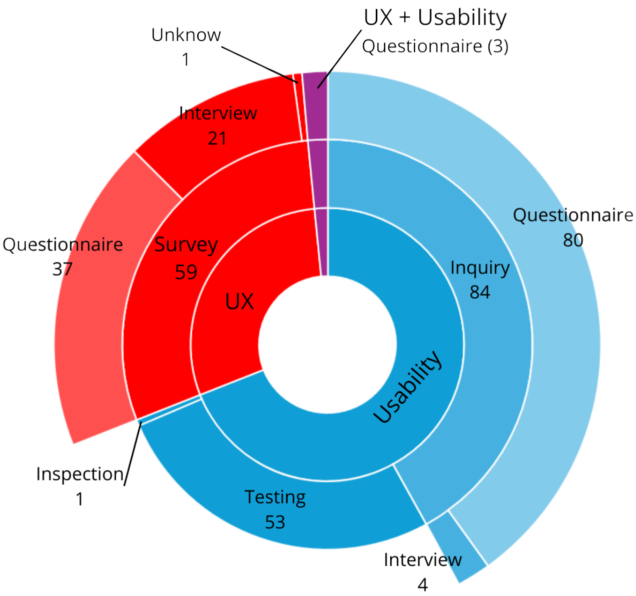

Figure 7. Evaluation Methods (joined SMSs)

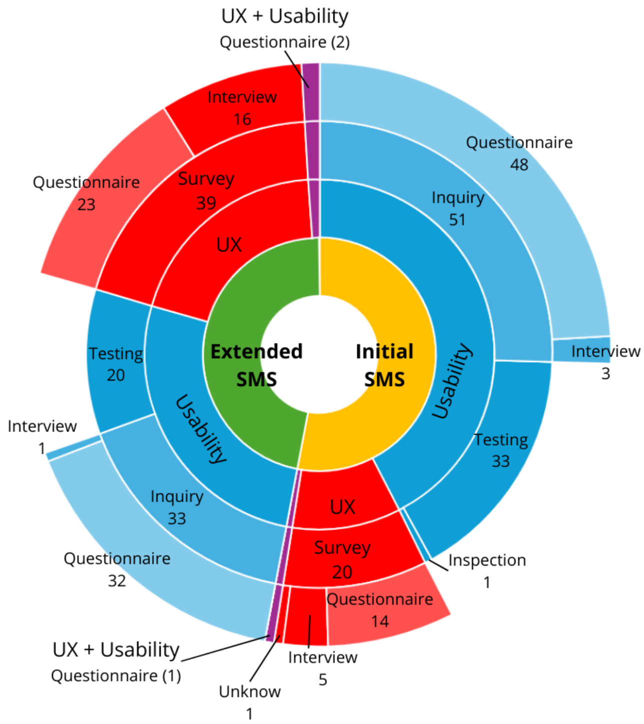

Figure 8. Evaluation Methods (separated SMSs)

4.5 SQ5 - Covered Aspects by the ET

The term "aspect" refers to a component or dimension within the usability or UX criteria. We prioritized identifying aspects such as effectiveness, efficiency, satisfaction, and enjoyment. When this was impossible, the evaluation's metrics, like time to complete a task, were used to determine the aspects. These identified aspects helped classify the ET's coverage concerning the UX and usability criteria. The following process was used to determine which usability or UX aspects an evaluation technology covers. An exploratory search first identified the main aspects of usability and UX in the literature. Then, each metric or aspect found in the publications was mapped or reclassified to match the identified aspects of the exploratory search. For example, various technologies had UX aspects labeled as "amusement," "enjoyment," "perceived enjoyment," and "perceived level of enjoyment," which were reclassified under "Pleasure/Fun".

Regarding the aspects present in the Usability ET (Table 7), the most frequent aspect (39.7%, 36 of the 141 technologies) was efficiency about time (*Efficiency/Time*), with metrics such as "task completion time", "entry rate", "speed" and "words per minute". This aspect was covered mainly through test sessions with data recording. The second most frequent aspect (31.9%, 45 of 141) was *Overall Usability*, evaluated mainly through the SUS questionnaire and the Think-Aloud method. Also appeared with good frequency the *Effectiveness/Accuracy* (29.8%, 42 of 141), *Satisfaction/Comfort* (27%, 38 of 141), *Satisfaction* (25.5%, 36 of 141), *Efficiency/Physical Effort* (23.4%, 33 of 141) and *Efficiency/Mental Effort* (19.9%, 28 of 141) aspects.

Both studies presented the same aspect with the highest frequency (*Efficiency/Time*) and the same group with six aspects with the highest frequency, varying the order internally

(*Effectivenes/Accuracy*, *Efficiency*, *Efficiency/Time*, *Overall Usability*, *Satisfaction* and *Satisfaction/Comfort*). In relation to the Initial SMS, the Extended SMS presented a significant increase in technologies covering the aspects of *Efficiency*, *Efficiency/Physical Effort*, *Satisfaction* and *Satisfaction/Comfort*. However, the *Overall Usability* aspect had less appearance in the Extended SMS.

Each technology evaluated between two and three (2.65) aspects on average. The evaluation technology considered to have the greatest coverage of aspects was the Usefulness, Satisfaction, and Ease of Use Questionnaire (USE) [Lund, 2001], which covered 11 aspects. A group of aspects that appeared frequently was *Effectiveness/Accuracy* together with *Efficiency/Time*, both appearing in 37 ETs, with the two aspects appearing alone in 25, all in usability tests with recorded data.

Regarding the aspects present in UX ET, the most frequent in both SMS were *Generic UX* (present in 70.9% of UX ETs, 44 of 62), *Usability* (32.2%, 20 of 62) and *Pleasure/Fun* (24.2%, 15 of 62). These three aspects appeared more frequently in both SMSs, only reversing the second and third placed order. These three aspects also appear in 9 evaluations that used the same ET, the UEQ [Laugwitz *et al.*, 2008]. This ET also addresses the *Stimulation*, *Trustworthness* and *Creative/Novelty* aspects, which in the studies appeared with the frequency, respectively, of 16.1% (10 of 62), 16.1% (10 of 62) and 14.5% (9 of 62).

Each technology evaluated two (2.08) aspects on average. The evaluation technology considered to have the greatest coverage of aspects was the UEQ, which covered 6 aspects. Almost half (40 out of 62) of the technologies evaluated only one aspect (median: 1), and the *Generic UX* aspect was the one that most appeared as the only aspect of evaluation. This happened in 32 technologies, with greater frequency in those found in Extended SMS (25 vs. 7 technologies in Initial SMS) and, 20 of which occurred through interviews. The aspects that also appeared in isolated evaluations were *Presence* (3x) and *Pleasure/Fun* (2x).

4.6 SQ6 - Existing/Known Technology

The answers to this question can indicate what common technologies are used to evaluate the usability or UX of touchable holographic solutions.

Regarding the nature of evaluation technologies applied (Table 9), 60% (120 of 200) were not based on any known technology. This group includes technologies with test and investigation methods exclusively created for the presented solution, called ad-hoc. For example, Weichel *et al.* [2014] presented a solution based on screen and semi-mirror material for 3D modeling using hands. This solution was submitted to two usability evaluations. In one, a questionnaire with eight sentences on a Likert scale was applied, containing terms/phrases such as “Usefulness of Existing Objects”, “Useful of Gestural Icons”, “No arm fatigue,” and “Easy to use”. Considering only ad-hoc technologies, 39.1% (47 of 120) were based on data recording and applied in usability tests. Another 27.5% (33 de 120) were applications of questionnaires on a Likert scale, 20.8% (25 de 120) were applications of interviews, 6.6% (8 de 120) were applications of questionnaires with open questions, 2.5% (3 de 120) were applications of questionnaires using other scales, such as numerical ranking, and 3.3% (4 de 120) were applications of questionnaires with an unknown scale (not reported in the publication).

Of the 200 technologies, 72 (36%) were assessments performed with known technologies, such as Think-Aloud [Nielsen, 1993], SUS [Brooke, 1996], NASA-TLX [Hart and Staveland, 1988], and UEQ [Laugwitz *et al.*, 2008]. The remaining 4% (8 of 200) were assessments performed with adaptations or part of known technologies. For example, Xu *et al.* [2020] investigated users’ awareness of the interaction limitation given by the limit of the field of view in HMD-based solutions. This publication related seven assessments, including six based on existing evaluation technologies. Think-Aloud was used to raise users’ challenges during the interaction and possible improvements in the system. The SUS, NASA-TLX, Borg CR-10 [Borg, 1998], Computer Vision Syndrome Questionnaire (CVS-Q) [Seguí *et al.*, 2015], and UEQ questionnaires were also applied for this solution. The other evaluation of the solution was through an ad-hoc usability test based on data records. Ordered by most frequently, the main known technologies (Table 11) found in both SMS were: NASA-TLX (24 times, including RAW-RLX [Hart, 2006]); SUS (21 times); UEQ (9 times); Think-Aloud protocol (6 times); UEQ (5 times); USE (3 times) and; Borg CR-10 (3 times).

4.7 SQ7 - Empirical Assessment of ET

This sub-question aimed to know if the applied ET was empirically evaluated and if this evaluation was reported in the publication. In the Initial SMS, only one (2.5%) of the 106 performed ET was an original proposal submitted to an empirical evaluation. This ET Dehghani *et al.* [2020] was created to evaluate the experiences using WMR solutions in retail services to verify the factors that affect the user’s attitude towards using WMR in shopping experiences. The model investigated whether Perceived Functional Benefits, Perceived Trust, Perceived Visual Appeal, Perceived Immersion, Perceived Autonomy, and Interactivity could influence user satisfaction and behavioral intention in adopting WMR. This ET is a questionnaire with sentences grouped on these factors. Each group of sentences was created based on several other studies and existing technologies [Shareef *et al.*, 2018; Huang *et al.*, 2017; Marasco *et al.*, 2018; tom Dieck *et al.*, 2018; Toet *et al.*, 2021; Yim *et al.*, 2017; Venkatesh *et al.*, 2012]. In the Extended SMS, none of the evaluation technologies found were subjected to an empirical evaluation; that is, the studies’ objective was limited to evaluating the SHT and not the technology that evaluated that holographic solution.

4.8 SQ8 - Empirical Study

In this group of sub-questions, the goal was to classify the study undergone in terms of study type (SQ8.1, according to Easterbrook *et al.* [2008]), number of participants (SQ8.2), and type of analysis (SQ8.3, qualitative/quantitative). These

Table 7. Usability Aspects

Aspect	SMS1			SMS2			Sum		
	Count	%	Order	Count	%	Order	Count	%	Order
Effectiveness	8	9.3%		5	9.1 %		13	9.2%	10
Effectiveness \Accuracy	24	27.9%	3	18	32.7%	2	42	29.8%	3
Effectiveness \Completeness	4	4.47%		0	0%		4	2.8%	
Efficiency	13	15.1%		14	25.5%	5	27	19.1%	8
Efficiency \Time	36	41.9%	1	20	36.4%	1	56	39.7%	1
Efficiency \Physical Effort	18	20.9%	5	15	27.3%	4	33	23.4%	6
Efficiency \Mental Effort	16	18.6%		12	21.8%		28	19.9%	7
Learnability	10	11.6%		5	9.1%		15	10.6%	9
Memorability	3	3.5%		1	1.8%		4	2.8%	
Prevent and Recovery Error	3	3.5%		0	0%		3	2.1%	
Controllability	5	5.8%		2	3.6%		7	5.0%	
Satisfaction	18	20.9%	5	18	32.7%	2	36	25.5%	5
Satisfaction \Usefulness	9	10.5%		1	1.8%		10	7.1%	
Satisfaction \Comfort	21	24.4%	4	17	30.9%	3	38	27.0%	4
Satisfaction \Trust	3	3.5%		1	1.8%		4	2.8%	
Satisfaction \Pleasurable	7	8.1%		2	3.6%		9	6.4%	
Overall Usability	31	36.0%	2	14	25.5%	5	45	31.9%	2

Table 8. UX Aspects

Aspect	SMS1			SMS2			Sum		
	Count	%	Order	Count	%	Order	Count	%	Order
Creative and Novelty	5	23.8%		4	9.8%	4	9	14.5%	
Desirability	2	9.5%		0	0%		2	3.2%	
Emotional	0	0%		1	2.4%		1	1.6%	
Immersion	0	0%		4	9.8%	4	4	6.4%	
Stimulation	5	23.8%		5	12.2%		10	16.1%	4
Pleasure \Fun	9	42.9%	2	6	14.6%	3	15	24.2%	3
Presence	4	19.0%		2	4.9%		6	9.7%	
Trustworthiness	6	26.8%	4	4	9.8%	4	10	16.1%	4
Usability	8	38.1%	3	12	29.3%	2	20	32.2%	2
Usefulness	3	14.3%		2	4.9%		5	8.0%	
Generic UX	13	61.9%	1	31	75.6%	1	44	70.9%	1

Table 9. Nature of ETs

Criteria	Know Full	Know, modified	Ad-Hoc
Usability	58	6	74
UX	14	1	44
Usability and UX	0	1	2

Table 10. Type of Ad-Hoc ETs

Type	Usability	UX	Usability + UX	Total	Porc
interview	4	21	0	25	20.8%
Likert scale	20	12	1	33	27.5%
open questions	0	7	1	8	6.6%
other scales	1	2	0	3	2.5%
record data	47	0	0	47	39.1%
unknow	2	2	2	4	3.3%

sub-questions are relevant for comparing studies and identifying the main types of analysis used. This information can serve as a reference for future studies.

In both studies, all selected solutions were subjected to empirical studies (SQ8). Concerning the empirical method (SQ8.1), the controlled experiment was used by almost two-thirds of the studies (70.6%, 46 of 65 in the sum of both studies). A controlled experiment investigates a testable hypothesis about the relationship between cause-effect variables in a controlled environment [Easterbrook *et al*., 2008]. This method was used, for example, in Kim *et al*. [2020], which evaluated the biomechanical stress of the neck and shoulders during AR interactions using HMD, and in [Zhao *et al*., 2023], which evaluated the movement time to interact with

Table 11. Know ETs

Know technologies	Full	Adapted	Criteria
NASA-TLX, including Raw-TLX	22	2	Usability
SUS	20	1	Usability
UEQ	9		UX
Think Aloud	6		Usability
Borg CR-10	3		Usability
USE	2	1	Usability
MSAQ	2		Usability
TAM	1	1	UX
CSV-Q	1		Usability
DAQ, ISO 9241-9:2000 part 9 annex C	1		Usability
Post-Study System Usability Questionnaire (PSSUQ)	1		Usability
ITC-Sense of Presence Inventory	1		UX
Questionnaire on Current Motivation (QCM)	1		UX
Slater-Usoh-Steed	1		UX
UAHIS	1		UX
IKC-Q (ISO 9241-410)		1	Usability
ISO/TS 9241-411, Annex C		1	Usability
CAMBS, REE, RNGVT, DVE, RSPPA, IVP, CAUIT		1	Usability + UX

Note, the criteria column represents the classification given to the ET according to SQ3. Some ETs, for example, evaluate only one or some aspects of a criterion, such as Slater-Usoh-Steed, which evaluates Presence, and QCM, which evaluates Motivation.

a target in an AR environment. The remaining 30.7% (20 of 65) of the studies used the case study method. Plasson *et al*. [2022] had both kinds of empirical methods. In the case study, a phenomenon is investigated within its real context. This method usually uses intentional rather than random sampling, selecting the most relevant cases for the study [Easterbrook *et al*., 2008]. It was also classified as a case study when it was carried out without presenting a hypothesis and cause-effect variables. For example, in Whitlock *et al*. [2020], the authors presented a tool for prototyping the organization of museum display objects. Part of the design and evaluation of the solution took place at the Museum of Natural History, curated by the University of Colorado. No studies were classified as survey, ethnography, or action research. In Initial SMS, the ratio between controlled experiments and case studies was 65%/35% (26/24 publications). In the Extended SMS, this ratio was 80%/24% (20/6).

Regarding the count of participants (SQ8.2), both studies had similar results. In the Initial SMS, each experiment had 17 to 18 participants on average (median 13). The lowest participation was 4 participants [Dudley *et al*., 2018], and the highest, 98 [Dehghani *et al*., 2020]. One publication did not indicate the number of participants [Mau-Tsuen Yang and Wan-Che Liao, 2014]. In Extended SMS, each experiment had 18 to 19 participants on average (median 15). The lowest participation was 4 participants [Plasson *et al*., 2022], and the highest was 54 [Venkatakrishnan *et al*., 2023]. Combined (Figure 9), the average was 17.97 participants (median 14).

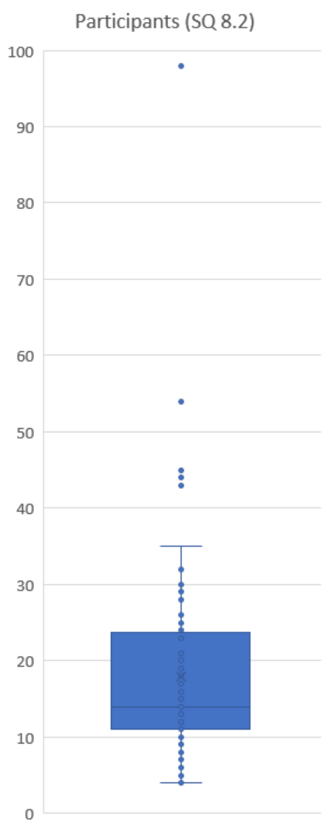

Figure 9. Participants Count in Empirical Studies

Regarding the nature of the analysis in the experiments (SQ8.3), 57.5% (23 of 40) of the experiments performed only a quantitative analysis. For example, Munsinger *et al*. [2019] conducted user tests and presented data on the number of errors and execution time. One study performed only qualitative analysis. This was the case of Jasche and Ludwig [2020], who conducted user tests using the Think-Aloud protocol, interviews, and inspection with experts. 40.0% (16 of 40) of the experiments performed quantitative and qualitative analyses.

4.9 SQ9 - Denominations for Holography

The answers to this sub-question could be AR, MR, or another denomination. The answers intend to verify how holography was classified within the extended reality spectrum [Milgram and Kishino, 1994]. In the Initial SMS, most of the authors (72.5%, 29 of 40) denominated holography as a solution in an AR environment. Almost a quarter (22.5%, 9 out of 40) of the authors classified it as MR, and in two (5%) publications, the solution was not classified as "reality". In one of them, the solution was named as *Three-Dimensional Aerial Image Interface (3DAII)* [Matsumaru *et al*., 2019] and in another as *Light Field Display (LFD) screen projection* [Adhikarla *et al*., 2015]. The study by Matsumaru *et al*. [2019] presented a prototype with a parabolic mirror on a flat panel display (FPD), which uses LCD technology or light-emitting diode (LED) to project a hologram in the air.

In the Extended SMS, 23 publications (92%) classified holography as Augmented Reality and only two as Mixed Reality. For example, Shen *et al*. [2022] introduces AdaptiKeyboard, a mid-air gesture keyboard utilizing multi-objective Bayesian optimization to dynamically adjust layout size for optimizing speed and accuracy in AR. On the other hand, [Li *et al*., 2022] explores the design and experience of object-centered user interfaces (UI) in MR environments with head-worn, summarizing four design recommendations for future everyday use. Both these studies were conducted using MS Hololens™.

4.10 SQ10 - Holography Display Mode

The answers to this sub-question classify the presentation mode of holography as a physical optical phenomenon or projection on screen/display. This information is intended to differentiate solutions created through reproducing images on screens from those that tried to obtain holography from its physical concept. Regarding this sub-question, in the Initial SMS, 90% (36 of 40) of the solutions used holography projection directly onto a screen (LCD, HMD, smart glasses, smartphone, tablet, LFD, and transparent OLED). For example, Frutos-Pascual *et al*. [2019] used projection with MS Hololens™ and Meta 2 AR™ HMDs. The remaining 10% (4 of 40) of the solutions resorted to optical phenomena supported by artifacts other than the screen, such as the use of a reflective mirror in a parabolic shape [Matsumaru *et al*., 2019], glass in an angular position [Weichel *et al*., 2014] and projection on a fog screen [Caruso *et al*., 2015; Sand *et al*., 2015]. The predominance of screen projection solutions indicates a trend or pattern of choice for touchable holographic solutions. In the Extended SMS, all solutions (25) used projection on a screen/display. For example, in McCord *et al*. [2022], the solution used the MS Hololens™, and in Bozgeyikli and Bozgeyikli [2021], the Magic Leap One was used. These devices project holography onto a transparent display. In Pei *et al*. [2022], the Oculus Quest was used, which projects a video onto the screen with holography mixed with the real environment.

4.11 SQ11 - Projection Technology

In this sub-question, the type of holographic display technology was classified as reflective material, fog screen + projector, holographic display, smartphone or tablet, HMD, smart glass, or another. The aim was to list the most common devices and materials that enable holographic projection.

HMDs were the device most used for holographic solutions, totaling 48 solutions (73.8%) in two combined studies. This group has the study presented by Whitlock *et al*. [2018], which examines the efficacy and usability of different interaction modalities, including multimodal voice, embodied free-hand gesture and handheld devices, for AR interactions at a distance. In the Initial SMS, 57.5% of the solutions (23 of 40) used HMDs. Meanwhile, all solutions (25) used HMDs in the Extended SMS. For example, McCord *et al*. [2022] used HMD to investigate the effectiveness of 2D versus AR formats for documenting construction sequences in civil engineering education.

Smart glasses were used in second place (10.8%, 7 of 65), all in Initial SMS. In this group are the solutions presented by Lin *et al*. [2017] and Lee *et al*. [2020], for example. Combining HMD and smart glasses means that almost five of six (84.6%) devices used for holographic projection are wearable, reinforcing the expectation that such equipment will be increasingly present and be the primary option for interaction with holographic solutions.

On the other hand, the less common devices, with one appearance each, were the holographic LFD screen [Adhikarla *et al*., 2015], the use of a simple LCD screen [Higuchi and Komuro, 2013] and the use of a transparent screen [Lee *et al*., 2013]. These last three studies are among the ten oldest selected (two dated 2013 and one from 2015).

When observing the projection device model and manufacturer (SQ11.1), MS Hololens™ stands out in the HMD category (73.9%, 17 of the 23) and Google Glass™ in the smart glass category (four of the seven solutions using the device). Other HMDs used in solutions were the Magic Leap One™ (5 solutions) and Meta 2 AR™ (2 solutions). Beyond Google Glass, the Mad Gaze ARES was the device most used (2 solutions). Table 12 shows all devices found in both SMS. Some solutions were used with more than one device.

It is important to clarify that Meta 1 and Meta 2 AR HMDs were launched by Meta, founded in 2012. This company focused specifically on AR products, distinguishing it from the current tech giant Meta Platforms (formerly Facebook), which is currently involved with Meta Quest 2, Meta Quest 3, and Meta Quest Pro HMDs. In addition, the Oculus Rift VR is a tethered VR HMD that requires a connection to a powerful PC for operation. Oculus Quest 1 and Oculus Quest 2 were standalone VR HMDs. These three devices (Oculus) were launched by the Oculus VR LLC company, in time a subsidiary of Facebook, Inc. After Facebook rebranding as Meta, Oculus Quest 2 was rebranded as Meta Quest 2.

The total number of devices exceeds 65 because, in some studies, more than one device was used. The identification "set" corresponds to studies that used projection mode by the

Table 12. Devices for Holographic Projection

Device	Type	SMS 1	SMS 2	Sum
MS Hololens	HMD	17	20	37
Magic Leap One	HMD	2	3	5
Google Glass	smart glasses	4		4
different set of devices	reflexive material or fog screen	4		4
Mad Gaze ARES	smart glasses	2		2
Meta 2 AR	HMD	2		2
Project North Star	HMD		1	1
Oculus Quest	HMD		1	1
Oculus Quest 2	HMD		1	1
Samsung S9+	smartphone/tablet	1		1
Oculus Rift with ZED Mini	HMD	1		1
Epson Moverio BT-200	smart glasses	1		1
Samsung Galaxy Tab S 10.5	smartphone/tablet	1		1
Oculus Rift VR	HMD	1		1
Nexus 5	smartphone/tablet	1		1
Meta 1 AR	smart glasses	1		1
ACCUPIX mybud	HMD	1		1
Century Corporation 4.3” display	LCD	1		1
Samsung Transparent Display prototype	OLED transparent	1		1
not available	smart glasses	1		1

optical phenomenon. In these, a set of equipment was used, such as projector + fog screen [Caruso *et al.*, 2015; Sand *et al.*, 2015] or screen + reflective material [Matsumaru *et al.*, 2019; Weichel *et al.*, 2014]. One of the studies did not identify which equipment was used for holographic projection, only indicating that it was an HMD [Mau-Tsuen Yang and Wan-Che Liao, 2014]

Note that some equipment was initially designed to serve VR, as in the case of the Oculus Rift. However, according to their authors, such equipment was used for AR [Shim *et al.*, 2016] and MR [Kang *et al.*, 2020] solutions. This fact confirms that it is not the device that determines the experienced environment but its use made of it.

4.12 SQ12 - Touch Detection Technology

This sub-question was intended to identify which type of camera could be used to detect the user’s gestures. The types of technology were classified as external NIR camera, built-in RGB-D camera, external RGB-D camera, built-in RGB camera, external RGB camera, or other.

In the selected solutions, the RGB-D depth camera was the most used. 76.9% (50 of 65 solutions) used this technology, either through a camera built into the HMDs or external cameras added to the solution’s system. For example, Seiger *et al.* [2021] presented a solution that used MS Hololens™ and the device’s RGB-D camera for touch detection, while Caruso *et al.* [2015] presented a solution whose holographic presentation was by projection in fog screen and detection via MS Windows Kinect™.

The external NIR camera was another frequently (16.9%, 11 of 65 solutions) technology used for hand detection. For example, Kim *et al.* [2019] presented a solution whose holographic vision was provided by a smart glass Epson Moverio BT-200™ and touch detection through the Leap Motion Controller™.

Not all HMD-based solutions use the device’s camera to detect touch. In Initial SMS, six of 23 HMD-based solutions used external cameras (RGB-D, 3; NIR 2, and; RGB, 1) to perform gesture detection. For example, Shim *et al.* [2016] brought a solution that used Oculus Rift VR™ combined with an RGB-D Softkinect DephSense DS325 camera™, and Kang *et al.* [2020] used the same Oculus Rift combined with the Leap Motion Controller™ for gesture detection. In the Extended SMS, only two of 25 HMD-based solutions used external cameras (RGB-D, 1 and; NIR 2). The solution proposed by Qian *et al.* [2022], for example, introduces ARnnotate, an AR interface that allows end users to create custom datasets to estimate 3D pose based on vision by manipulating virtual bounding boxes and physical objects. This solution uses the Oculus Quest 2 HMD, and hand detection is done by a ZED Dual AMP camera and a Leap Motion controller™.

The solutions that use a built-in RGB camera (4 of 65) use Google Glass™ or Mad Gaze ARE Glass™, which does not have a depth camera. RGB cameras are combined with computer vision algorithms for hand/finger tracking. These were the cases of Lee *et al.* [2020], Lin *et al.* [2017], Lee *et al.* [2019a] and Huang *et al.* [2015]. Some solutions (2 of 65) used more than one technology to perform the detection. One combines internal and external depth cameras [Kim *et al.*, 2020] and another combines internal and external RGB cameras [Lee *et al.*, 2019a].

Upon examining the model and vendor of each detection device (SQ12.1), it is noticed that among the built-in depth cameras, most come with the MS Hololens™ device (83.3%, 35 of 42). Among external depth cameras, the MS Kinect for Windows™ was the most used (37.5%, 3 of 8). The 11 solutions that used the infrared camera were through the Leap Motion Controller™, which is the second most common device, behind the MS Hololens™. Among internal RGB cameras, Google Glass was used in all solutions of this type of technology, and Mad Gaze ARE Glass was used by three of the four solutions that used this type of technology (there were four studies, three of which used both devices). One of the studies did not identify which external RGB camera was used. The study authors only indicated that it was combined with a computer vision algorithm [Mau-Tsuen Yang and Wan-Che Liao, 2014].

The MS Kinect for Windows actually falls under the category of an RGB-D camera because it combines a regular RGB camera (capturing color information) with an NIR camera to generate depth data, which, when combined with the RGB image, allows for features like 3D scanning and motion tracking. Other RGB-D cameras in this SMS used depth sensing based on stereo vision (like Optitrack and ZED dual cameras). This approach uses two synchronized cameras capturing the scene from slightly different angles. Similar to how our eyes perceive depth, the software analyzes the disparity between the two images to calculate the distance to various points in the scene. Some RGB-D cameras, like the Softkinect DepthSense DS325 camera, offer a hybrid approach.

4.13 SQ13 - Purpose of the solution

In this sub-question, we tried to identify whether the holographic solution had a specific purpose. The objective was to discover practical applications of touchable holographic solutions. From the answers provided, 33,8% (22 of 65) of the solutions were created to investigate the interaction of the user and its characteristics, such as gesture elicitation [Williams *et al.*, 2020], awareness of the limit of the field of view [Xu *et al.*, 2020], depth perception [Ha *et al.*, 2014], effectiveness of haptic feedback [Vaquero-Melchor and Bernardos, 2019] , and selection techniques [Plasson *et al.*, 2022] . Nine solutions (13.8%, 9 of 65) were created to input data as a virtual keyboard in an AR/MR environment [Lee *et al.*, 2020, 2019b,a; Dudley *et al.*, 2018; Sand *et al.*, 2015; Lystbæk *et al.*, 2022a].

Other purposes included determining the preference between the mode of interaction (voice, gesture, or other) [Aslan *et al.*, 2019; Qian *et al.*, 2019; Matsumaru *et al.*, 2019; Xu *et al.*, 2019], simulation of control of IoT or home utility equipment [Seiger *et al.*, 2021; Becker *et al.*, 2019; Whitlock *et al.*, 2018], equipment assembly [Kim and Lee, 2016], map manipulation [Mau-Tsuen Yang and Wan-Che Liao, 2014], programming learning [Mahajan *et al.*, 2020], museum display prototyping [Whitlock *et al.*, 2020], learning vocabulary Weerasinghe *et al.* [2022], light painting photography [Wang *et al.*, 2021b], among others.

The answers to this sub-question showed that many studies are still seeking to understand the characteristics of interaction with hands in holography and few end-user solutions.

4.14 SQ14 - Gesture Types/Styles

According to Aigner *et al.* [2012]'s classification, only pointing, pantomimic or manipulation gestures were considered. In this sub-question, no information about iconic and semaphore gestures was extracted because these two types do not represent gestures where the user touches or holds an object. The studies could present more than one gesture type in the solution. The answers to this question could indicate the accuracy and sophistication of hand interactions in a solution. All (65) solutions implemented the pointing gesture. Another 34 solutions (52.3% in 65) implemented pantomime-type gesture detection, 26 in Initial SMS (65%, 26 of 40 studies in that SMS) and 8 in Extended SMS (32%, 8 of 25 studies in that SMS). 16 solutions (24.6% in 65) identified manipulation-type gestures, 15 in the Initial SMS and only one in the Extended SMS. Twelve solutions (27.5%) detected all three gesture styles. For example, Frutos-Pascual *et al.* [2019] presented a solution submitted to experimentation where the tasks resided in performing movements to select (pointing), move (pantomimic), rotate (pantomimic), and resize (manipulation) a virtual object.

4.15 SQ15 - Holographic Image Quality

This sub-question was answered by observing the images linked in the publication, and its answers indicate the degree of quality with which the holographic objects are presented and mixed with the real environment. The majority (58.4%, 38 of 65) of the solutions had a hologram as the WIMP interface. For example, Becker *et al.* [2019] presented an interface for controlling a domestic music player appliance, in which the user had user interface controls to increase/decrease volume, move forward/backward between songs, and modify the brightness and colors of the interface itself. The other 26 solutions (40%) presented objects in raw 3D format, as in Frutos-Pascual *et al.* [2019], where 3D objects were only used to evaluate movement, rotation and resizing. In just 11 solutions (16.9%), more sophisticated virtual objects were presented, using texture or lighting to get close to the appearance of what a real object would be, as in Whitlock *et al.* [2020], where images of museum objects could be placed in the environment to plan an exhibition. One of the [Dehghani *et al.*, 2020] publications did not present an image of the projected hologram, not allowing it to be classified.

4.16 SQ16 - Presence of Feedback

This sub-question aimed to verify if the feedback given to the user focused only on offering visual stimuli that are usually found in user interfaces or if there were other ways to make him perceive the results of his interaction with holographic objects. This sub-question allowed multiple answers: no, visual, auditory, ultrasonic, and other. Two studies [Deghani *et al.*, 2020; He *et al.*, 2022] did not bring images or details of the holography. 62 (95.4%) solutions use visual feedback during the interaction. 9 solutions (13.8%) combine auditory feedback with visual feedback. For example, in Weerasinghe *et al.* [2022], the evaluated solution, which aims to translate and teach languages, offers visual feedback when manipulating 3D objects and buttons on a WIMP interface and also offers auditory feedback to hear the word the user wants to translate. Four solutions (6.1%) included haptic feedback via ultrasound. In one of them, ultrasound was the only feedback to the user [Pei *et al.*, 2022]. In Vaquero-Melchor and Bernardos [2019], researchers investigated how many ultrasound points were needed for users to recognize

Table 13. Devices for Hand Detection

Device	Type	SMS 1	SMS 2	Sum
MS Hololens built-in camera	RGB-D	15	20	35
Leap Motion controller	NIR	9	2	11
Magic Leap built-in camera	RGB-D	2	3	5
Google Glass camera + Computer Vision	RGB	4		4
Mad Gaze ARES Glass Camera + Computer Vision	RGB	3		3
MS Kinect for Windows	RGB-D	3		3
Optitrack camera	RGB-D	2		2
Meta 2 AR built-in camera	RGB-D	2		2
Oculus Quest built-in camera	RGB-D		1	1
ZED dual AMP camera	RGB-D		1	1
Softkinect DepthSense DS325	RGB-D	1		1
Firefly MV from Point Grey Research Inc. + Computer Vision	RGB	1		1
Creative Interaction Gesture Camera	RGB-D	1		1
ASUS Xtion Pro Live	RGB-D	1		1
not available	RGB	1		1

touch feedback in geometric shapes properly. One (2.5%) of the solutions indicated using auditory feedback in addition to visual [Matsumaru *et al.*, 2019]. It may be that the solutions classified by only providing visual feedback also have another form of feedback, such as auditory feedback. However, the classification given by this sub-question considered what was explained in the publication's text.

5 Discussion

Based on both SMS results, there might be a window of opportunity to propose new evaluation technologies for THS, since there were few publications in this line of contribution (SQ1). The extended study noted that validation research continues to be prevalent. The theme seems highly relevant to Computing, given the number of publications by authors in this discipline (SQ2) and the frequency of publications in events and journals in the area and related areas. On the other hand, the extension of the study observed a higher proportion of articles in the area of Engineering compared to the previous study, as well as articles from other areas, including a primary study in the context of Medical Sciences, reinforcing that the topic moves researchers from different areas, even providing cooperation in multidisciplinary research.

The Extended SMS reveals notable shifts in the focus of ETs for THS. Usability remains the predominant criterion, yet its exclusivity in evaluation technologies has decreased from 80.2% in the Initial SMS to 56.4% in the Extended SMS. This decline suggests an evolving landscape where usability is being increasingly complemented by other criteria, allowing for growth in the use of assessment technologies focused on the UX criterion. The UX criterion, which previously held a smaller share, has seen significant growth, now present in 41.5% of ETs compared to 18.9% previously. This shift indicates a growing recognition of the importance of hedonic attributes and overall user experience in evaluating THS. Only 1.5% of ETs integrated both usability and UX criteria, highlighting an ongoing challenge in creating comprehensive evaluation tools that address both pragmatic and hedonic aspects.

The type of ETs used to evaluate usability and UX has shown remarkable consistency between the two SMS periods. Among the methods used, the preference for those in which the user gives feedback through a questionnaire was predominant. This choice may be due to the practicality of application, ease of data collection and tabulation, or the impersonality of using a tool that can evaluate or diagnose the evaluated object. Inquiry methods, including interviews and questionnaires, continue to dominate usability evaluations, accounting for 61.7% of usability ETs. This preference underscores the practicality and efficiency of collecting user feedback through these methods. Testing remains the second most popular method, utilized in 37.6% of usability evaluations. Observation tests, which collect data by recording different variables related to effectiveness and efficiency, allow evaluators to capture user behavior, which can be fundamental to understanding how interaction progresses in a spatial environment. Both SMS showed that this combination of testing and inquiry/survey methods remains an empirically observed preference among evaluators, indicating a stable methodological approach.

The aspects covered by ETs show both continuity and evolution. However, the identification of usability and UX aspects reveals that the coverage achieved by the evaluation technologies could be more comprehensive.

In usability evaluations, the most common aspect remains efficiency related to time, present in 39.7% of usability ETs. This focus on time efficiency indicates the ongoing emphasis on optimizing task completion speed in THS. Other important aspects include overall usability, effectiveness, and various dimensions of satisfaction and comfort. Extended SMS shows a notable increase in technologies that cover efficiency, physical effort, satisfaction and comfort aspects, while overall usability saw a decline. This shift suggests a broadening of the criteria considered in usability evaluations, reflecting a more holistic approach to assessing user interaction.

The most common aspects in UX evaluations were Generic UX, Usability, and Pleasure/Fun. These aspects appeared consistently in both SMS, although their order varied. The prominence of the UEQ questionnaire, which addresses multiple aspects of UX, highlights its utility in providing a comprehensive assessment of user experience. The trend of single-aspect evaluations, particularly generic UX, increased in the Extended SMS, indicating a focus on capturing overarching user impressions through interviews. Despite these aspects being recurrently evaluated, important dimensions such as Presence and other attributes related to MR are often overlooked. The Extended SMS showed a slight improvement in the variety of aspects covered, but the need for a more broader and more integrated approach remains.

The nature of the evaluation technologies used highlights a mix of known and ad-hoc approaches. The majority (60%) of ETs were ad-hoc, created specifically for the evaluated solution. This ad hoc approach allows for tailored evaluations, but can lead to inconsistencies and difficulties in comparing results across different studies. In addition, this high percentage of custom-built ETs highlights the need for more standardized and widely applicable evaluation tools in the field. The use of known technologies such as SUS, NASA-TLX, and UEQ represented 36% of the evaluations, demonstrating the dependence on established methods for robust and reliable evaluations. However, the use of known technologies does not always guarantee comprehensive coverage of relevant aspects. The combination of different application technologies can cause shadow coverage of assessment aspects and may require careful interpretation of results due to differing definitions and scopes across technologies. The ongoing challenge is balancing ad-hoc methods' flexibility with the reliability and comparability of established technologies. The remaining 4% of the evaluations used adapted versions of known technologies, indicating ongoing innovation and customization of the evaluation methodologies to meet the specific study needs better.

In general, the Extended SMS reveals a dynamic and evolving field of usability and UX evaluation for touchable holographic solutions. The increased integration of UX criteria, the consistency in preferred evaluation methods, the broadening of covered aspects, and the balance between ad-hoc and standardized technologies reflect a growing sophistication and diversification in evaluation practices. These trends underscore the importance of developing and refining evaluation methodologies to keep up with the rapid advances in holographic technology.

Only one evaluation technology in both SMS was empirically evaluated by validation research (SQ7). The others did not undergo this study and cannot be extended safely and confidently to other holographic solutions. In addition, the only evaluation technology presented as a proposed solution did not present images of the holographic projection or information about the feedback given to users in its publication, hindering a better understanding of the context of using this technology. The Extended SMS did not find new technologies that had been empirically evaluated, and there was no evidence that the technologies proposed in the first study had been empirically evaluated in subsequent years.

A positive point returned by both SMS is that all touchable holographic solutions have undergone an empirical study with users through controlled experiments or case studies (SQ8, SQ8.1, SQ8.2). The number of usability evaluations may justify the predominance of quantitative analysis (SQ8.3) based on pragmatic attributes and aspects that can be measured by different metrics that lead to quantitative analysis. However, a combination of quantitative and qualitative analyses was present in 50.8% of the studies, increasing from 40% in Initial SMS to 68% in Extended SMS. Furthermore, most studies that presented qualitative analysis (26 of 34) included UX evaluation. Combining these types of analysis can help discover issues and understand the interaction scenario.

Based on the results from the Initial SMS, we concluded that holography was still an open or evolving concept. As the results demonstrated, researchers still needed to reach a consensus on the designations of holographic solutions belonging to AR or MR (SQ9). For example, a touch-based text input solution from a virtual keyboard added to the real environment was classified as AR by Dudley *et al*. [2018] and MR by Lee *et al*. [2019a]. That means such a lack of definition does not make the proposal of holographic solutions unfeasible, nor does it underestimate the solutions and studies carried out and identified in SMS. In contrast, it demonstrated that we were in the middle of discovering new solutions and applications that accompanied the growing progress of technological device manufacturers. In the Extended SMS, there appears to be a consolidation of the classification of holographic solutions as AR. The preference for this name appeared in 92% (23 of 25) of the solutions.

The SMS also showed that there are a good number of devices that can be used or combined to produce touchable holographic solutions (SQ10, SQ11, SQ11.1, SQ12, SQ12.1) from the most sophisticated (such as HMDs and RGB-D), even the most common in our daily lives (smartphones and RGB cameras). HMDs and smart glasses are the best way to do this, making it possible to integrate projection and detection in a single device, facilitating mobility. The race to bring equipment to the market (SQ11.1) with the necessary resources for a good immersion provides an exciting range for researchers. However, MS Hololens™ is predominant. This current predominance may be related to the fact that, due to the high acquisition cost, researchers prefer to acquire those offered by companies already consolidated and better able to support and guarantee the continuity of solutions for long-term research. In addition, some manufacturers may have good relationships with research centers or even specific channels to encourage researchers' interest. Other factors, such as the registration/acquisition of patents, may favor one or more manufacturers in the race to offer the best option to interact with holography. In addition, Extended SMS revealed that in recent years, HMDs were the preferred device for touchable holographic interaction, being the only device type found in the study.

Few touchable holographic solutions found in this SMS were created for end-use (SQ13). Now, the research focuses on understanding the characteristics and limits of the interaction provided by the available devices. Among the solutions seen, there is good coverage of the main styles of gestures that correspond to touch (SQ14), mainly pointing. This coverage tends to increase by covering more styles, such as manipulation and pantomime, as devices and algorithms for detection become more effective and efficient. Likewise, the quality of holographic images (SQ15) tends to improve as the equipment can provide better resolution, a field of view, and a refresh rate, among other attributes. The predominance of touchable WIMP-type holographs is just the beginning of a process, since these interfaces are simpler to design and do not lack so much rule (coherence) to adapt to an MR environment. Finally, feedback (SQ16) should be an essential factor for the user to understand the outcome of their interaction. At this point, visual feedback still predominates (those that demonstrate through visual perception, for example, a selected item or the response to a user action). Still, other means, such as auditive and, in some cases, haptic, are expected to be added to give the user a richer interaction experience.

6 Limitations

According to [Kitchenham *et al.*, 2016] in an SMS, Construct Validity concerns how well the study design can address the research question. Internal Validity is affected by the conduct of the study, mainly related to data extraction and synthesis, and whether there are factors that might have caused some degree of bias in the overall process. The External Validity should be based on assessing the range covered by the primary studies regarding their settings, materials, and participants. Conclusion Validity in SMS concerns how reliably we can draw conclusions about the link between a case and its outcomes. It concerns the synthesis part of a study and how well it supports the conclusions of the review. In SMS, internal and conclusion validity are closely related because both depend on the strength of the synthesis and its ability to support reliable conclusions.

For these SMSs, we sought to ensure the Construct Validity through the definition of a protocol that followed, as closely as possible, the main guidelines for planning and executing an SMS in the SE area. A potential threat in the protocol would be the absence of a consensus or a clear definition of the "holography" term. Some publications could report solutions of interest to us without citing that it is a holography. For this reason, the search string also predicted the use of the terms "augmented reality", "mixed reality" and "extended reality". The same happened with the "touchable" word, which was followed by "tangible", "touchless", and "mid-air", even if some of these could lead to some primary studies unrelated to holographic touch, such as voice or gaze interaction.

Regarding Internal and Conclusion Validity, the data extraction in each publication was reviewed by two experts in the Initial SMS. Any potential bias in the extraction and synthesis process may be related to the authors' familiarity with the research context. To mitigate this limitation, we combined experts in HCI, Usability, UX, VR, and AR applications, along with a third with mixed knowledge in both areas. Since the second SMS was conducted by an undergraduate, a graduate student, and only one expert in Usability and UX, although the graduate student gained knowledge over time, this new team configuration with one less expert constitutes a new threat to validity.

Furthermore, another potential threat to validity arises from having only one researcher complete the reviews for 85% of the publications in the first filter. Although initial evaluations by multiple researchers helped establish consistent criteria, the reliance on a single researcher for the majority of the process introduces a risk of subjective bias. This is particularly relevant as individual interpretations or oversights could influence the inclusion or exclusion of publications, potentially impacting the comprehensiveness and reproducibility of the SMS.

Regarding the External Validity, we can assume that, since the analyzed studies were up to April/2023 and, the research covers a theme in constant technological evolution, it is likely that the results arising from the interpretation of this publication set will not whether it applies or does not correspond to the primary studies conducted from April/2023.

7 Conclusion and Future Works

This paper presents an updated SMS that extends the previous investigation on usability and UX evaluation technologies for THS. The updated SMS covers an additional two years of publications, examining a total of 5429 publications and selecting 65 that present 200 evaluation technologies. The findings provide a comprehensive overview of the current state of the art, highlighting new trends and persistent gaps in evaluating THS.

The updated study reinforces the original finding that usability remains the primary focus of evaluations, but there is a noticeable increase in attention to UX criteria. This shift suggests a growing recognition of the importance of a holistic user experience in evaluating THS. However, integrating usability and UX criteria into a single evaluation instrument remains limited, indicating an opportunity to develop comprehensive evaluation frameworks that simultaneously address both aspects.

The updated study analysis indicates that the efficiency related to time remains the most frequent aspect in usability evaluations, present in 39.7% of usability ETs. Other important aspects include overall usability, effectiveness, and various dimensions of satisfaction and comfort. The extended study shows increased technologies covering physical effort, satisfaction, and comfort, reflecting a more holistic approach to usability evaluations.

In UX evaluations, generic UX, usability, and pleasure/fun remain the most frequently assessed aspects. The prominence of the UEQ questionnaire, which addresses multiple aspects of UX, highlights its utility in providing comprehensive assessments. However, important dimensions such as the presence and other MR-specific attributes are often overlooked, indicating the need for broader and more integrated evaluation approaches.

A significant observation is the persistence of validation research, with a preference for empirical studies involving users through controlled experiments or case studies. The combination of quantitative and qualitative analyses has increased, providing a nuanced understanding of user interaction and behavior.

One of the main identified limitations is the lack of empirically evaluated technologies. Despite identifying numerous evaluation methods, only a few have undergone empirical validation, highlighting a critical gap. Future research should focus on the empirical validation of existing or new evaluation technologies to ensure their reliability and applicability across different holographic solutions.

The study also notes an emerging trend towards a higher proportion of publications from the field of Engineering and other disciplines, including Medical Sciences. This multidisciplinary interest underscores the broad applicability and potential of THS in various domains. The consolidation of the classification of holographic solutions as AR further reflects the evolving understanding and standardization within the field.

Device preferences continue to favor HMDs and smart glasses, with Microsoft Hololens™ remaining the predominant choice due to its advanced capabilities and strong support infrastructure. The preference for these devices will likely persist, as they offer superior integration of projection and detection functionalities, which are essential for creating immersive and interactive holographic experiences.

Future research should address the integration of usability and UX criteria, considering MR-specific dimensions such as presence, and developing heuristics or checklists to facilitate the inspection process. Empirical validation of the proposed evaluation technologies is crucial to ensure their effectiveness and reliability.

The findings of this updated SMS contribute significantly to the field of HCI by providing a detailed mapping of current evaluation technologies and identifying key areas where further research is needed. By highlighting trends, gaps, and opportunities in THS evaluation, this work offers valuable information to HCI researchers aiming to improve user interaction and experience. The classification and analysis of existing technologies also provide a reference point for developing new and more effective evaluation methods that foster innovation and progress in the HCI domain.

In conclusion, this work advances our understanding of usability and UX evaluation in the context of THS. It underscores the importance of continuing research and development in this rapidly evolving field. By addressing identified gaps and leveraging the insights gained, future studies can create more intuitive, immersive, and effective holographic interactions, ultimately enriching the user experience and expanding the possibilities within HCI.

Declarations

Acknowledgements

Thanks to Coordination for the Improvement of Higher Education Personnel (CAPES) - Program of Academic Excellence (PROEX).

Authors' Contributions

TC contributed to the conceptualization, investigation, and formal analysis and is the main writer of this manuscript. MC and NV also contribute to investigation and formal analysis. NV contributes to editing this article. ED and NV contribute to supervision. All authors read and approved the final manuscript.

Availability of data and materials

Complete reports are available online for the Initial SMS [Campos *et al.*, 2023a] and Extended SMS [Campos *et al.*, 2024].

References

Adhikarla, V. K., Jakus, G., and Sodnik, J. (2015). Design and Evaluation of Freehand Gesture Interaction for Light Field Display. In Kurosu, M., editor, *Human-Computer Interaction: Interaction Technologies*, volume 9170, pages 54–65. Springer International Publishing, Cham. DOI: https://doi.org/0.1007/978-3-319-20916-6_6.

Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindbauer, D., Ion, A., Zhao, S., and Koh, J. T. K. V. (2012). Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for HCI. Technical Report MSR-TR-2012-111, Microsoft Research.

Altman, D. G. (1990). *Practical Statistics for Medical Research*. Chapman and Hall/CRC, London, 1st edition edition.

Aslan, I., Dang, C. T., Schlagowski, R., Dietz, M., Brain, F., and André, E. (2019). Put that Hologram there - Probing Mobile Interaction Experiences for a Vision of Mixed Material Public Spaces. In *Proceedings of the 9th International Conference on the Internet of Things*, pages 1–4, Bilbao Spain. ACM. DOI: https://doi.org/0.1145/3365871.3365877.

Bai, Z. and Blackwell, A. F. (2012). Analytic review of usability evaluation in ISMAR. *Interacting with Computers*, 24(6):450–460. DOI: https://doi.org/0.1016/j.intcom.2012.07.004.

Barbosa, S., Silva, B., Silveira, M., Gasparini, I., Darin, T., and Barbosa, G. (2021). *Interação Humano-Computador e Experiência do Usuário*. Autopublicação, Rio de Janeiro, 1a ed. edition.

Basili, V. R., Caldiera, G., and Rombach, H. D. (1994). The Goal Question Metric Approach.

Becker, V., Rauchenstein, F., and Sörös, G. (2019). Investigating Universal Appliance Control through Wearable Augmented Reality. In *Proceedings of the 10th Augmented Human International Conference 2019*, pages 1–9, Reims France. ACM. DOI: https://doi.org/0.1145/3311823.3311853.

Berkman, M. I. and Akan, E. (2019). Presence and Immersion in Virtual Reality. In Lee, N., editor, *Encyclopedia of Computer Graphics and Games*, pages 1–10. Springer International Publishing, Cham. DOI: https://doi.org/0.1007/978-3-319-08234-9_162-1.

Bevan, N., Carter, J., and Harker, S. (2015). ISO 9241-11 Revised: What Have We Learnt About Usability Since 1998? In Kurosu, M., editor, *Human-Computer Interaction: Design and Evaluation*, Lecture Notes in Computer Science, pages 143–151, Cham. Springer International Publishing. DOI: https://doi.org/0.1007/978-3-319-20901-2_13.

Borg, G. (1998). *Borg's perceived exertion and pain scales*. Borg's perceived exertion and pain scales. Human Kinetics, Champaign, IL, US. Pages: viii, 104.

Bowman, D. A. and McMahan, R. P. (2007). Virtual Reality:

How Much Immersion Is Enough? *Computer*, 40(7):36–43. DOI: https://doi.org/0.1109/MC.2007.257.

Bozgeyikli, E. and Bozgeyikli, L. L. (2021). Evaluating Object Manipulation Interaction Techniques in Mixed Reality: Tangible User Interfaces and Gesture. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pages 778–787. ISSN: 2642-5254. DOI: https://doi.org/0.1109/VR50410.2021.00105.

Brancati, N., Caggianese, G., Pietro, G. D., Frucci, M., Gallo, L., and Neroni, P. (2015). Usability Evaluation of a Wearable Augmented Reality System for the Enjoyment of the Cultural Heritage. In *2015 11th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*, pages 768–774, Bangkok, Thailand. IEEE. DOI: https://doi.org/0.1109/SITIS.2015.98.

Brooke, J. (1996). SUS: A 'Quick and Dirty' Usability Scale. In *Usability Evaluation In Industry*, pages 207–212. CRC Press, London, 1st edition.

Campos, T., Castello, M., Valentim, N., and Damasceno, E. (2024). An Updated Systematic Mapping Study on Usability and User Experience Evaluation of Touchable Holograms: Technical Report. page 944822 Bytes. DOI: https://doi.org/0.6084/M9.FIGSHARE.26261951.

Campos, T., Valentim, N., and Damasceno, E. (2023a). A Systematic Mapping Study on Usability and User Experience Evaluation of Touchable Holograms: Technical Report. page 15558639 Bytes. DOI: https://doi.org/10.6084/m9.figshare.22114355.

Campos, T. P. D., Damasceno, E. F., and Valentim, N. M. C. (2022). Porifera: A Collaborative Tool to Support Systematic Literature Review and Systematic Mapping Study. In *Proceedings of the XXXVI Brazilian Symposium on Software Engineering*, SBES '22, pages 452–457, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/0.1145/3555228.3555273.

Campos, T. P. d., Damasceno, E. F., and Valentim, N. M. C. (2023b). Usability and User Experience Evaluation of Touchable Holographic solutions: A Systematic Mapping Study. In *IHC '23: Proceedings of the 22st Brazilian Symposium on Human Factors in Computing Systems*, IHC '23, pages 1–13, Maceio, Brazil. ACM. DOI: https://doi.org/0.1145/3638067.3638071.

Cao, Y., Xu, Z., Glenn, T., Huo, K., and Ramani, K. (2018). Ani-Bot: A Modular Robotics System Supporting Creation, Tweaking, and Usage with Mixed-Reality Interactions. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 419–428, Stockholm Sweden. ACM. DOI: https://doi.org/0.1145/3173225.3173226.

Caruso, G., Carulli, M., and Bordegoni, M. (2015). Augmented Reality System for the Visualization and Interaction with 3D Digital Models in a Wide Environment. *Computer-Aided Design and Applications*, 12(1):86–95. DOI: https://doi.org/0.1080/16864360.2014.949579.

Chaoui, K., Bouzidi-Hassini, S., and Bellik, Y. (2023). SUIL: a modeling language for spatial user interaction. *Journal of Reliable Intelligent Environments*, 9(2):161–181. DOI: https://doi.org/0.1007/s40860-021-00164-z.

Chien, P. H. and Lin, Y. C. (2021). Gesture-based head-mounted augmented reality game development using leap motion and usability evaluation: 15th International Conference on Interfaces and Human Computer Interaction, IHCI 2021 and 14th International Conference on Game and Entertainment Technologies, GET 2021 - Held at the 15th Multi-Conference on Computer Science and Information Systems, MCCSIS 2021. *15th International Conference on Interfaces and Human Computer Interaction, IHCI 2021 and 14th International Conference on Game and Entertainment Technologies, GET 2021 - Held at the 15th Multi-Conference on Computer Science and Information Systems, MCCSIS 2021*, pages 149–156.

Daskalogrigorakis, G., McNamara, A., Marinakis, A., Antoniadis, A., and Mania, K. (2022). Glance-Box: Multi-LOD Glanceable Interfaces for Machine Shop Guidance in Augmented Reality using Blink and Hand Interaction. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 315–321. ISSN: 2771-1110. DOI: https://doi.org/0.1109/ISMAR-Adjunct57072.2022.00070.

Dehghani, M., Lee, S. H. M., and Mashatan, A. (2020). Touching holograms with windows mixed reality: Renovating the consumer retailing services. *Technology in Society*, 63:101394. DOI: https://doi.org/0.1016/j.techsoc.2020.101394.

Dudley, J. J., Vertanen, K., and Kristensson, P. O. (2018). Fast and Precise Touch-Based Text Entry for Head-Mounted Augmented Reality with Variable Occlusion. *ACM Transactions on Computer-Human Interaction*, 25(6):1–40. DOI: https://doi.org/0.1145/3232163.

Dünser, A., Grasset, R., and Billinghurst, M. (2008). A survey of evaluation techniques used in augmented reality studies. In *ACM SIGGRAPH ASIA 2008 courses on - SIGGRAPH Asia '08*, pages 1–27, Singapore. ACM Press. DOI: https://doi.org/0.1145/1508044.1508049.

Easterbrook, S., Singer, J., Storey, M.-A., and Damian, D. (2008). Selecting Empirical Methods for Software Engineering Research. In Shull, F., Singer, J., and Sjøberg, D. I. K., editors, *Guide to Advanced Empirical Software Engineering*, pages 285–311. Springer, London. DOI: https://doi.org/0.1007/978-1-84800-044-5_11.

Favreau, J. (2010). Iron Man 2. Translated title: Homem de Ferro 2 IMDb ID: tt1228705 event-location: USA publisher-place: Estados Unidos.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382. Place: US Publisher: American Psychological Association. DOI: https://doi.org/0.1037/h0031619.

Flick, C. D., Harris, C. J., Yonkers, N. T., Norouzi, N., Erickson, A., Choudhary, Z., Gottsacker, M., Bruder, G., and Welch, G. (2021). Trade-offs in Augmented Reality User Interfaces for Controlling a Smart Environment. In *Proceedings of the 2021 ACM Symposium on Spatial User Interaction*, SUI '21, pages 1–11, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/0.1145/3485279.3485288.

Frutos-Pascual, M., Creed, C., and Williams, I. (2019). Head Mounted Display Interaction Evaluation: Manipulating Virtual Objects in Augmented Reality. In Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., and Zaphiris, P., editors, *Human-Computer Interaction – INTERACT 2019*, volume 11749, pages 287–308. Springer International Publishing, Cham. DOI: https://doi.org/0.1007/978-3-030-29390-1_16.

Ha, T., Feiner, S., and Woo, W. (2014). WeARHand: Head-worn, RGB-D camera-based, bare-hand user interface with visually enhanced depth perception. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 219–228, Munich, Germany. IEEE. DOI: https://doi.org/0.1109/ISMAR.2014.6948431.

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908. DOI: https://doi.org/0.1177/154193120605000909.

Hart, S. G. and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In Hancock, P. A. and Meshkati, N., editors, *Advances in Psychology*, volume 52 of *Human Mental Workload*, pages 139–183. North-Holland, Amsterdan, Netherlands. DOI: https://doi.org/0.1016/S0166-4115(08)62386-9.

Hassenzahl, M. (2011). User Experience and Experience Design. In *The Encyclopedia of Human-Computer Interaction*. Interaction Design Fundation, online, 2nd edition.

He, Y., Hu, Y., Feng, H., Li, C., and Shen, X. (2022). Comparative Analysis of 3D Interactive Modes in Different Object Layouts in Mixed Reality. In *Proceedings of the Ninth International Symposium of Chinese CHI*, Chinese CHI '21, pages 120–126, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/0.1145/3490355.3490371.

Higuchi, M. and Komuro, T. (2013). AR typing interface for mobile devices. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia - MUM '13*, pages 1–8. ACM Press. DOI: https://doi.org/0.1145/2541831.2541847.

Hu, J., Dudley, J. J., and Kristensson, P. O. (2022). An Evaluation of Caret Navigation Methods for Text Editing in Augmented Reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 640–645. ISSN: 2771-1110. DOI: https://doi.org/0.1109/ISMAR-Adjunct57072.2022.00132.

Huang, C. D., Goo, J., Nam, K., and Yoo, C. W. (2017). Smart tourism technologies in travel planning: The role of exploration and exploitation. *Information Management*, 54(6):757–770. DOI: https://doi.org/0.1016/j.im.2016.11.010.

Huang, Z., Li, W., and Hui, P. (2015). Ubii: Towards Seamless Interaction between Digital and Physical Worlds. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 341–350, Brisbane Australia. ACM. DOI: https://doi.org/0.1145/2733373.2806266.

International Society for Presence Research (2000). Presence defined.

ISO (2011). ISO/IEC 25010:2011 - Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models. Technical report, International Organization for Standardization, Geneva, Switzerland.

ISO (2018). ISO 9241-11:2018 Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. Technical report, International Organization for Standardization, Geneva, Switzerland.

Ivory, M. Y. and Hearst, M. A. (2001). The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*, 33(4):470–516. DOI: https://doi.org/0.1145/503112.503114.

Jang, J., Frier, W., and Park, J. (2022). Multimodal Volume Data Exploration through Mid-Air Haptics. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 243–251. ISSN: 1554-7868. DOI: https://doi.org/0.1109/ISMAR55827.2022.00039.

Jasche, F. and Ludwig, T. (2020). PrintARface: Supporting the Exploration of Cyber-Physical Systems through Augmented Reality. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, pages 1–12, Tallinn Estonia. ACM. DOI: https://doi.org/0.1145/3419249.3420162.

Kang, H. J., Shin, J.-h., and Ponto, K. (2020). A Comparative Analysis of 3D User Interaction: How to Move Virtual Objects in Mixed Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 275–284, Atlanta, GA, USA. IEEE. DOI: https://doi.org/0.1109/VR46266.2020.00047.

Kim, J. H., Ari, H., Madasu, C., and Hwang, J. (2020). Evaluation of the biomechanical stress in the neck and shoulders during augmented reality interactions. *Applied Ergonomics*, 88:103175. DOI: https://doi.org/0.1016/j.apergo.2020.103175.

Kim, M., Choi, S. H., Park, K.-B., and Lee, J. Y. (2019). User Interactions for Augmented Reality Smart Glasses: A Comparative Evaluation of Visual Contexts and Interaction Gestures. *Applied Sciences*, 9(15):3171. DOI: https://doi.org/0.3390/app9153171.

Kim, M. and Lee, J. Y. (2016). Touch and hand gesture-based interactions for directly manipulating 3D virtual objects in mobile augmented reality. *Multimedia Tools and Applications*, 75(23):16529–16550. DOI: https://doi.org/0.1007/s11042-016-3355-9.

Kitchenham, B. A., Budgen, D., and Brereton, P. (2016). *Evidence-based software engineering and systematic reviews*. Chapman Hall/CRC innovations in software engineering and software development. CRC Press, Boca Raton. OCLC: ocn932588149.

Koçer Özgün, F. N. and Alaçam, S. (2023). AN EVALUATION OF AUGMENTED REALITY-BASED USER INTERFACES IN THE DESIGN PROCESS. *Architecture and Planning Journal (APJ)*, 28(3). DOI: https://doi.org/0.54729/2789-8547.1234.

Landis, J. R. and Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1):159. Publisher: JSTOR. DOI: https://doi.org/0.2307/2529310.

Laugwitz, B., Held, T., and Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. In Holzinger, A., editor, *HCI and Usability for Education and Work*, Lecture Notes in Computer Science, pages 63–76, Graz, Austria. Springer. DOI: https://doi.org/0.1007/978-3-540-89350-9_6.

Lee, J., Olwal, A., Ishii, H., and Boulanger, C. (2013). SpaceTop: integrating 2D and spatial 3D interactions in a see-through desktop environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 189–192, Paris France. ACM. DOI: https://doi.org/0.1145/2470654.2470680.

Lee, L. H., Braud, T., Bijarbooneh, F. H., and Hui, P. (2019a). TiPoint: detecting fingertip for mid-air interaction on computational resource constrained smartglasses. In *Proceedings of the 23rd International Symposium on Wearable Computers*, pages 118–122, London United Kingdom. ACM. DOI: https://doi.org/0.1145/3341163.3347723.

Lee, L. H., Braud, T., Bijarbooneh, F. H., and Hui, P. (2020). UbiPoint: towards non-intrusive mid-air interaction for hardware constrained smart glasses. In *Proceedings of the 11th ACM Multimedia Systems Conference*, pages 190–201, Istanbul Turkey. ACM. DOI: https://doi.org/0.1145/3339825.3391870.

Lee, L. H., Yung Lam, K., Yau, Y. P., Braud, T., and Hui, P. (2019b). HIBEY: Hide the Keyboard in Augmented Reality. In *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom*, pages 1–10, Kyoto, Japan. IEEE. DOI: https://doi.org/0.1109/PERCOM.2019.8767420.

Li, Y., Hu, Y., Wang, Z., and Shen, X. (2022). Evaluating the Object-Centered User Interface in Head-Worn Mixed Reality Environment. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 414–421. ISSN: 1554-7868. DOI: https://doi.org/0.1109/ISMAR55827.2022.00057.

Lin, S., Cheng, H. F., Li, W., Huang, Z., Hui, P., and Peylo, C. (2017). Ubii: Physical World Interaction Through Augmented Reality. *IEEE Transactions on Mobile Computing*, 16(3):872–885. DOI: https://doi.org/0.1109/TMC.2016.2567378.

Liu, X., Meng, X., Spittle, B., Xu, W., Gao, B., and Liang, H.-N. (2023). Exploring Text Selection in Augmented Reality Systems. In *Proceedings of the 18th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*, VRCAI '22, pages 1–8, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/0.1145/3574131.3574459.

Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability interface*, 8(2):3–6.

Lystbæk, M. N., Pfeuffer, K., Grønbæk, J. E. S., and Gellersen, H. (2022a). Exploring Gaze for Assisting Freehand Selection-based Text Entry in AR. *Proceedings of the ACM on Human-Computer Interaction*, 6(ETRA):141:1–141:16. DOI: https://doi.org/0.1145/3530882.

Lystbæk, M. N., Rosenberg, P., Pfeuffer, K., Grønbæk, J. E., and Gellersen, H. (2022b). Gaze-Hand Alignment: Combining Eye Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality. *Proceedings of the ACM on Human-Computer Interaction*, 6(ETRA):145:1–145:18. DOI: https://doi.org/0.1145/3530886.

Mahajan, K., Groechel, T., Pakkar, R., Cordero, J., Lee, H., and Matarić, M. J. (2020). Adapting Usability Metrics for a Socially Assistive, Kinesthetic, Mixed Reality Robot Tutoring Environment. In Wagner, A. R., Feil-Seifer, D., Haring, K. S., Rossi, S., Williams, T., He, H., and Sam Ge, S., editors, *Social Robotics*, volume 12483, pages 381–391. Springer International Publishing, Cham. DOI: https://doi.org/0.1007/978-3-030-62056-1_32.

Marasco, A., Buonincontri, P., van Niekerk, M., Orlowski, M., and Okumus, F. (2018). Exploring the role of next-generation virtual technologies in destination marketing. *Journal of Destination Marketing Management*, 9:138–148. DOI: https://doi.org/0.1016/j.jdmm.2017.12.002.

Marques, L., Barcellos, M. P., Gadelha, B., and Conte, T. (2024). Characterizing UX Assessment in the Context of Immersive Experiences: A Systematic Mapping Study. *International Journal of Human–Computer Interaction*, pages 1–17. DOI: https://doi.org/0.1080/10447318.2024.2351711.

Matsumaru, T., Septiana, A. I., Horiuchi, K., and Graduate School of Information, Production, and Systems, Waseda University 2-7 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, Japan (2019). Three-Dimensional Aerial Image Interface, 3DAII. *Journal of Robotics and Mechatronics*, 31(5):657–670. DOI: https://doi.org/0.20965/jrm.2019.p0657.

Mau-Tsuen Yang and Wan-Che Liao (2014). Computer-Assisted Culture Learning in an Online Augmented Reality Environment Based on Free-Hand Gesture Interaction. *IEEE Transactions on Learning Technologies*, 7(2):107–117. DOI: https://doi.org/0.1109/TLT.2014.2307297.

McCord, K. H., Ayer, S. K., Perry, L. A., Patil, K. R., London, J. S., Khoury, V., and Wu, W. (2022). Student Approaches and Performance in Element Sequencing Tasks Using 2D and Augmented Reality Formats. *Education Sciences*, 12(4):247. DOI: https://doi.org/0.3390/educsci12040247.

MDPI (2024). Augmented Reality: Current Trends, Challenges and Prospects.

Mendes, E., Wohlin, C., Felizardo, K., and Kalinowski, M. (2020). When to update systematic literature reviews in software engineering. *Journal of Systems and Software*, 167:110607. DOI: https://doi.org/0.1016/j.jss.2020.110607.

Merino, L., Schwarzl, M., Kraus, M., Sedlmair, M., Schmalstieg, D., and Weiskopf, D. (2020). Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009-2019). In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 438–451, Porto de Galinhas, Brazil. IEEE. DOI: https://doi.org/0.1109/ISMAR50242.2020.00069.

Mestre, D. and Vercher, J.-L. (2011). Immersion and presence. In *Virtual Reality: Concepts and Technologies*, pages 93–104. CRC Press, Boca Raton, FL, USA.

Milgram, P. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329. Publisher: The Institute of Electronics, Information and Communication Engineers.

Munsinger, B., White, G., and Quarles, J. (2019). The Usability of the Microsoft HoloLens for an Augmented Reality Game to Teach Elementary School Children. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pages 1–4, Vienna, Austria. IEEE. DOI: https://doi.org/0.1109/VS-Games.2019.8864548.

Nielsen, J. (1993). *Usability engineering*. Academic Press, Boston.

Nielsen, J. (2012). Usability 101: Introduction to Usability.

Pei, S., Chen, A., Lee, J., and Zhang, Y. (2022). Hand Interfaces: Using Hands to Imitate Objects in AR/VR for Expressive Interactions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–16, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/0.1145/3491102.3501898.

Pereira, G. A. F., Bacha, J. M. R., Silva, I. B. A. N., Kim, D. H. C., Pompeu, J. E., and Lopes, R. d. D. (2021). Virtual Reality and Augmented Reality Exergames for older fallers: considerations about design and applicability by physical therapists. In *Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames)*, pages 949–956. SBC. DOI: https://doi.org/0.5753/sbgames_estendido.2021.19734.

Petersen, K., Feldt, R., Mujtaba, S., and Mattsson, M. (2008). Systematic mapping studies in software engineering. *12th International Conference on Evaluation and Assessment in Software Engineering, EASE 2008*, I(June):1–10. DOI: https://doi.org/0.14236/ewic/ease2008.8.

Petersen, K., Vakkalanka, S., and Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64:1–18. DOI: https://doi.org/0.1016/j.infsof.2015.03.007.

Plasson, C., Blanch, R., and Nigay, L. (2022). Selection Techniques for 3D Extended Desktop Workstation with AR HMD. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 460–469. ISSN: 1554-7868. DOI: https://doi.org/0.1109/ISMAR55827.2022.00062.

Plasson, C., Cunin, D., Laurillau, Y., and Nigay, L. (2019). Tabletop AR with HMD and Tablet: A Comparative Study for 3D Selection. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, pages 409–414, Daejeon Republic of Korea. ACM. DOI: https://doi.org/0.1145/3343055.3360760.

Qian, J., Ma, J., Li, X., Attal, B., Lai, H., Tompkin, J., Hughes, J. F., and Huang, J. (2019). Portal-ble: Intuitive Free-hand Manipulation in Unbounded Smartphone-based Augmented Reality. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pages 133–145, New Orleans LA USA. ACM. DOI: https://doi.org/0.1145/3332165.3347904.

Qian, X., He, F., Hu, X., Wang, T., and Ramani, K. (2022). ARnnotate: An Augmented Reality Interface for Collecting Custom Dataset of 3D Hand-Object Interaction Pose Estimation. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*, UIST '22, pages 1–14, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/0.1145/3526113.3545663.

Roto, V., Obrist, M., and Väänänen-vainio mattila, K. (2009). User Experience Evaluation Methods in Academic and Industrial Contexts. In *Proceedings of the Workshop UXEM'09*, volume II, page 4 p, Uppsala, Sweden. Springer.

Sand, A., Rakkolainen, I., Isokoski, P., Raisamo, R., and Palovuori, K. (2015). Light-weight immaterial particle displays with mid-air tactile feedback. In *2015 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, pages 1–5, Ottawa, ON, Canada. IEEE. DOI: https://doi.org/0.1109/HAVE.2015.7359448.

Santos, G., Rocha, A. R., Conte, T., Barcellos, M. P., and Prikladnicki, R. (2012). Strategic Alignment between Academy and Industry: A Virtuous Cycle to Promote Innovation in Technology. In *26th Brazilian Symposium on Software Engineering SBES 2012*, pages 196–200, Natal, Brazil. IEEE Computer Society. DOI: https://doi.org/0.1109/SBES.2012.31.

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281. DOI: https://doi.org/0.1162/105474601300343603.

Seguí, M. d. M., Cabrero-García, J., Crespo, A., Verdú, J., and Ronda, E. (2015). A reliable and valid questionnaire was developed to measure computer vision syndrome at the workplace. *Journal of Clinical Epidemiology*, 68(6):662–673. DOI: https://doi.org/0.1016/j.jclinepi.2015.01.015.

Seiger, R., Kühn, R., Korzetz, M., and Aßmann, U. (2021). HoloFlows: modelling of processes for the Internet of Things in mixed reality. *Software and Systems Modeling*, 20(5):1465–1489. DOI: https://doi.org/0.1007/s10270-020-00859-6.

Serrano, R., Morillo, P., Casas, S., and Cruz-Neira, C. (2022). An empirical evaluation of two natural hand interaction systems in augmented reality. *Multimedia Tools and Applications*, 81(22):31657–31683. DOI: https://doi.org/0.1007/s11042-022-12864-6.

Shareef, M. A., Baabdullah, A., Dutta, S., Kumar, V., and Dwivedi, Y. K. (2018). Consumer adoption of mobile banking services: An empirical examination of factors according to adoption stages. *Journal of Retailing and Consumer Services*, 43(C):54–67.

Shen, J., Hu, J., Dudley, J. J., and Kristensson, P. O. (2022). Personalization of a Mid-Air Gesture Keyboard using Multi-Objective Bayesian Optimization. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 702–710. ISSN: 1554-7868. DOI: https://doi.org/0.1109/ISMAR55827.2022.00088.

Shim, J., Yang, Y., Kang, N., Seo, J., and Han, T.-D. (2016). Gesture-based interactive augmented reality content authoring system using HMD. *Virtual Reality*, 20(1):57–69. DOI: https://doi.org/0.1007/s10055-016-0282-z.

Skarbez, R., Brooks, Jr., F. P., and Whitton, M. C. (2017). A Survey of Presence and Related Concepts. *ACM Computing Surveys*, 50(6):96:1–96:39. DOI: https://doi.org/0.1145/3134301.

Slater, M. (2003). A note on presence terminology. *Presence connect*, 3(3):1–5.

Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology (London, England: 1953)*, 109(3):431–433. DOI: https://doi.org/0.1111/bjop.12305.

Swan, J. and Gabbard, J. L. (2005). Survey of User-Based Experimentation in Augmented Reality. In *Proceedings 1st International Conference on Virtual Reality*, pages 1–9, Las Vegas, Nevada, USA. Mira Digital Publishing.

Toet, A., Mioch, T., Gunkel, S. N., Niamut, O., and van Erp, J. B. (2021). Assessment of presence in augmented and mixed reality. preprint, PsyArXiv.

tom Dieck, M. C., Jung, T. H., and Rauschnabel, P. A. (2018). Determining visitor engagement through augmented reality at science festivals: An experience economy perspective. *Computers in Human Behavior*, 82:44–53. DOI: https://doi.org/0.1016/j.chb.2017.12.043.

Uzor, S. and Kristensson, P. O. (2021). An Exploration of Freehand Crossing Selection in Head-Mounted Augmented Reality. *ACM Transactions on Computer-Human Interaction*, 28(5):33:1–33:27. DOI: https://doi.org/0.1145/3462546.

van der Heijden (2004). User Acceptance of Hedonic Information Systems. *MIS Quarterly*, 28(4):695. DOI: https://doi.org/0.2307/25148660.

Vaquero-Melchor, D. and Bernardos, A. M. (2019). Enhancing Interaction with Augmented Reality through Mid-Air Haptic Feedback: Architecture Design and User Feedback. *Applied Sciences*, 9(23):5123. DOI: https://doi.org/0.3390/app9235123.

Venkatakrishnan, R., Venkatakrishnan, R., Raveendranath, B., Pagano, C. C., Robb, A. C., Lin, W.-C., and Babu, S. V. (2023). Give Me a Hand: Improving the Effectiveness of Near-field Augmented Reality Interactions By Avatarizing Users' End Effectors. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2412–2422. DOI: https://doi.org/0.1109/TVCG.2023.3247105.

Venkatesh, Thong, and Xu (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, 36(1):157. DOI: https://doi.org/0.2307/41410412.

Veraszto, E. V., Silva, D. d., Miranda, N. A., and Simon, F. O. (2009). Tecnologia: buscando uma definição para o conceito. *PRISMA.COM*, 0(8):19–46. Number: 8.

Wang, G., Ren, G., Hong, X., Peng, X., Li, W., and O'Neill, E. (2022). Freehand Gestural Selection with Haptic Feedback in Wearable Optical See-Through Augmented Reality. *Information*, 13(12):566. DOI: https://doi.org/0.3390/info13120566.

Wang, T., Qian, X., He, F., Hu, X., Cao, Y., and Ramani, K. (2021a). GesturAR: An Authoring System for Creating Freehand Interactive Augmented Reality Applications. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, pages 552–567, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/0.1145/3472749.3474769.

Wang, T., Qian, X., He, F., and Ramani, K. (2021b). LightPaintAR: Assist Light Painting Photography with Augmented Reality. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI EA '21, pages 1–6, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/0.1145/3411763.3451672.

Weerasinghe, M., Biener, V., Grubert, J., Quigley, A. J., Toniolo, A., Pucihar, K. C., and Kljun, M. (2022). VocabulARy: Learning Vocabulary in AR Supported by Keyword Visualisations. arXiv:2207.00896 [cs].

Weichel, C., Lau, M., Kim, D., Villar, N., and Gellersen, H. W. (2014). MixFab: a mixed-reality environment for personal fabrication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3855–3864, Toronto Ontario Canada. ACM. DOI: https://doi.org/0.1145/2556288.2557090.

Whitlock, M., Harnner, E., Brubaker, J. R., Kane, S., and Szafir, D. A. (2018). Interacting with Distant Objects in Augmented Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 41–48, Reutlingen. IEEE. DOI: https://doi.org/0.1109/VR.2018.8446381.

Whitlock, M., Mitchell, J., Pfeufer, N., Arnot, B., Craig, R., Wilson, B., Chung, B., and Szafir, D. A. (2020). MRCAT: In Situ Prototyping of Interactive AR Environments. In Chen, J. Y. C. and Fragomeni, G., editors, *Virtual, Augmented and Mixed Reality. Design and Interaction*, volume 12190, pages 235–255. Springer International Publishing, Cham. DOI: https://doi.org/0.1007/978-3-030-49695-1_16.

Wieringa, R., Maiden, N., Mead, N., and Rolland, C. (2006). Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. *Requirements Engineering*, 11(1):102–107. DOI: https://doi.org/0.1007/s00766-005-0021-6.

Williams, A. S., Garcia, J., and Ortega, F. (2020). Understanding Multimodal User Gesture and Speech Behavior for Object Manipulation in Augmented Reality Using Elicitation. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3479–3489. DOI: https://doi.org/0.1109/TVCG.2020.3023566.

Woolf, N. (2015). Google Glass ceases production 'in present form'.

Wright, T., de Ribaupierre, S., and Eagleson, R. (2019). Leap Motion Performance in an Augmented Reality Workspace: Integrating Devices with an Interactive Platform. *IEEE Consumer Electronics Magazine*, 8(1):36–41. DOI: https://doi.org/0.1109/MCE.2018.2816302.

Xu, W., Liang, H.-N., Chen, Y., Li, X., and Yu, K. (2020). Exploring Visual Techniques for Boundary Awareness During Interaction in Augmented Reality Head-Mounted Displays. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces*

(VR), pages 204–211, Atlanta, GA, USA. IEEE. DOI: https://doi.org/0.1109/VR46266.2020.00039.

Xu, W., Liang, H.-N., He, A., and Wang, Z. (2019). Pointing and Selection Methods for Text Entry in Augmented Reality Head Mounted Displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 279–288, Beijing, China. IEEE. DOI: https://doi.org/0.1109/ISMAR.2019.00026.

Yim, M. Y.-C., Chu, S.-C., and Sauer, P. L. (2017). Is Augmented Reality Technology an Effective Tool for E-commerce? An Interactivity and Vividness Perspective. *Journal of Interactive Marketing*, 39:89–103. DOI: https://doi.org/0.1016/j.intmar.2017.04.001.

Yu, K., Eck, U., Pankratz, F., Lazarovici, M., Wilhelm, D., and Navab, N. (2022). Duplicated Reality for Co-located Augmented Reality Collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2190–2200. DOI: https://doi.org/0.1109/TVCG.2022.3150520.

Zhao, C., Li, K. W., and Peng, L. (2023). Movement Time for Pointing Tasks in Real and Augmented Reality Environments. *Applied Sciences*, 13(2):788. DOI: https://doi.org/0.3390/app13020788.

Zhu, F. and Grossman, T. (2020). BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, Honolulu HI USA. ACM. DOI: https://doi.org/0.1145/3313831.3376233.