Computers & Graphics 102 (2022) 619–633

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical Section

A critical analysis on remote collaboration mediated by Augmented Reality: Making a case for improved characterization and evaluation of the collaborative process ☆

Bernardo Marques *, António Teixeira, Samuel Silva, João Alves, Paulo Dias, Beatriz Sousa Santos

IEETA, DETI, University of Aveiro, Aveiro, 3810-193, Portugal

ARTICLE INFO

Article history:
Received 31 December 2020
Received in revised form 2 July 2021
Accepted 16 August 2021
Available online 20 August 2021

Keywords:
Remote collaboration
Augmented Reality
Mixed Reality
Evaluation and characterization
collaborative process
Critical analysis
Roadmap proposal

ABSTRACT

Remote Collaboration mediated by Mixed and Augmented Reality (MR/AR) shows great potential in scenarios where physically distributed collaborators need to establish a common ground to achieve a shared goal. So far, most research efforts have been devoted to creating the enabling technology, overcoming engineering hurdles and proposing methods to support its design and development. To contribute to more in-depth knowledge on how remote collaboration occurs through these technologies, it is paramount to understand where the field stands and how characterization and evaluation have been conducted. In this vein, this work reports the results of a literature review which shows that evaluation is frequently performed in ad-hoc manners, i.e., disregarding adapting the evaluation methods to collaborative AR. Most studies rely on single-user methods, which are not suitable for collaborative solutions, falling short of retrieving the necessary amount of contextualized data for more comprehensive evaluations. This suggests minimal support of existing frameworks and a lack of theories and guidelines to guide the characterization of the collaborative process using AR. Then, a critical analysis is presented in which we discuss the maturity of the field and a roadmap of important research actions is proposed, that may help address how to improve the characterization and evaluation of the collaboration process moving forward and, in consequence, improve MR/AR based remote collaboration.

1. Introduction

Collaboration is essential in many situations, as is the case of industrial, medical, and educational domains, among others [1–4] and can be described as the process of joint and interdependent activities between co-located or remote collaborators performed to achieve a common goal [5–9].

Collaboration scenarios have evolved from simple co-located scenarios to more complex remote collaboration, encompassing several team members with different experiences, expertise's and multidisciplinary backgrounds distributed by different geographic locations around the world. Therefore, the methods required to address such activities have been growing in terms of scale, complexity, and interdisciplinarity, entailing not only the mastery of multiple domains of knowledge, but also a strong level of proficiency in each [3,4,10].

Remote collaboration, implies that collaborators establish a joint effort to align and integrate their activities in a seamless manner. Technological support for remote collaboration has been addressed among other fields by Computer-Supported Cooperative Work (CSCW), focusing on conceptualizing, designing, and prototyping solutions for communication, cooperation, assistance, training, learning as well as knowledge sharing between distributed collaborators.

One major issue of remote collaboration is the fact that collaborators do not share a common space/world, reason for the interest in using Augmented Reality (AR) in this context [11–15]. Collaboration using AR helps distributed collaborators establish a common ground, analogous to their understanding of the physical space, allowing to inform where to act, and what to do, e.g., making assumptions and beliefs visible by providing real-time spatial information, highlighting specific areas of interest, or sharing situated information associated with relevant objects in the on-site physical environment [16–20]. Remote AR-based solutions are well suited for overlaying responsive computer-generated information on top of the real-world environment, resulting in the creation of solutions that combine the advantages of virtual environments and the possibility for seamless interaction with the real-world objects and other collaborators [6,15,

☆ This article was recommended for publication by C Sandor.

* Corresponding author.
E-mail addresses: bernardo.marques@ua.pt (B. Marques), ajst@ua.pt (A. Teixeira), sss@ua.pt (S. Silva), jbga@ua.pt (J. Alves), paulo.dias@ua.pt (P. Dias), bss@ua.pt (B.S. Santos).

https://doi.org/10.1016/j.cag.2021.08.006

17,19,21–24]. A number of studies have focused on the use of virtual annotations to augment the shared understanding, using drawings, pointers, gaze, hand gestures and others on 2D images or live video streams [3,6,9,14,22,25]. As an alternative, recent studies started to explore the use of virtual replicas [26–28], as well as reconstructions of the physical environment [29,30], although these required the existence of 3D models and additional hardware, which may limit their adoption in some scenarios of application. Using such approaches to enhance the common ground can improve efficiency and accuracy of the performed tasks by enhancing the perception of the shared understanding [6,22,24,31,32], as well as collaboration times, knowledge retention, increased problem context and awareness [16,17,33–35].

While creating the means to support collaboration clearly motivated early research, advances in AR have been limited by new technical developments, which means most of the research efforts, so far, have been focused on creating the enabling technology and propose novel methods to support its design and development [14,36,37]. On the other hand, with the growing development of CSCW, the evaluation of these solutions during the collaborative effort become an essential, but difficult endeavor [31,38,39], given the novelty of the field and the lack of methods and theories [14,25] to guide the characterization of the collaborative process, i.e., describe the contributions of AR to the collaborative work effort. In addition, scenarios of remote collaboration are multifaceted [40], which means many aspects may affect the way teams collaborate, making it difficult to identify all variables related to the collaborative process. Therefore, the integration of proper characterization and evaluation methods and guidelines is of paramount importance.

In this paper, we analyze the subject of remote collaboration supported by AR through a systematic review and investigate how characterization and evaluation of the collaborative process has been conducted during user studies to better understand their specificities, rather than focusing on the development of technology itself. In this context, we analyzed existing surveys that addressed collaborative user studies and evaluation in their reviews. Plus, we performed a literature review from 2000 to 2020 to provide a high-level overview of the field, allowing identification of strengths and weaknesses of existing methods. Based on the analyzes, we describe the challenges involved with evaluating these solutions and critically analyze the state of the field. As a result, a possible roadmap is proposed to facilitate and elicit characterization of the collaboration process using AR-based solutions, so that research and development can move forward and focus on the nuances of supporting collaboration, i.e., focus squarely on the human concerns that underlie collaboration, rather than creating the enabling technology that makes remote collaboration mediated by AR possible.

The rest of the paper is organized as follows: Section 2 presents an overview based on survey papers to update, complement and fill gaps on the current information on the state-of-the-art. Next, Section 3 details the methodology adopted to conduct our literature review and describes a high-level overview of the reviewed papers. Then, Section 4 provides a critical analysis in which we discuss the challenges associated to the characterization and evaluation of the collaborative process. Afterwards, Section 5 propose a roadmap to address these challenges. Finally, Section 6 concludes by summarizing the main outcomes.

2. Related work on user evaluation in collaborative AR

This section identifies and analyzes existing survey papers that cover relevant evaluation details, which are summarized in Table 1. Our goal was to understand how evaluation has been conducted in collaborative scenarios, allowing to compare and contrast different methods, as well as identify opportunities and limitations associated to the characterization of the collaborative process. From the list of prior surveys, the first six entries are rather general in scope, although the review of collaborative AR papers is also mentioned, despite being only a portion of the results reported [6,12,29,41–43]. While this is the case, the two last entries of the list [14,25] focus entirely on the subject of Collaborative AR and MR, including co-located and remote examples. Although these surveys primarily focused on the development of collaborative AR technology itself, some important outcomes regarding evaluation are also reported, as described below in detail. Besides, another publication [7] was considered in this analysis, that even though not strictly a survey, includes important information regarding evaluation of collaborative work in the context being addressed.

2.1. Previous surveys including user evaluation information

Duenser et al. (2008) reported on user evaluation techniques used in AR research. Then, studies that evaluate collaboration between users using AR were quite underrepresented: from a total of 161 publications included in the survey, only 10 addressed collaborative AR. Besides reporting that 8 papers were formal and 2 informal user evaluation, the survey does not present further detail on the collaborative studies [41].

In addition, Zhou et al. (2008) presented one of the first overviews of the research conducted until that moment at the ISMAR conference and its predecessors. Although the research focus was on AR technologies, it also pointed out the significance of usability evaluation. The authors reported that a small number of collaborative AR prototypes were starting to emerge, but few had been evaluated in formal user studies. The authors also highlighted how the role of different displays would affect collaboration in the future and how the location of the task affected user behaviors in terms of verbal and non-verbal communication. Since collaboration and evaluation were not one of the focus of the survey, no further detail was provided [42].

In the same way, Bai et al. (2012) conducted an analytic review on usability evaluation at ISMAR. The authors suggested that while the design of usable systems were the main focus of collaborative AR research to that point, an increase in evaluation research was emerging. They also stated that measurements of particular interest in collaborative AR systems may include explicit communication (e.g., spoken and gestural messages), ease of collaboration and information gathering (e.g., basic awareness, eye gaze). The authors also reported that subjective answers may be collected via questionnaire and that direct observation was used to extract objective results. Moreover, signs of discomfort and enjoyment during collaboration were also taken into account by researchers [29].

Billinghurst et al. (2015) published a survey on AR, in which almost 50 years of research and development in the field were summarized.

The authors state that in Collaborative AR studies, besides the standard subjective measures, process measures may be more important than quantitative outcome measures. Process measures are typically gathered by transcribing interaction between users, like speech or gestures and performing a conversational analysis. Measures that have been found to be significantly relevant include: frequency of conversational turns, duration of overlapping speech, number of questions, number of interruptions, turn completions and dialog length, among others. Besides, gesture and non-verbal behaviors can also be analyzed for characteristic features. The survey acknowledges that there have been very few user studies with collaborative AR environments and almost none that examined communication process measures [12].

Table 1
Summary of evaluation surveys addressing Collaborative Augmented Reality (2008–2019).

Year & Authors	Pub.	# Pubs. analyzed	Aspects of Collaboration	Main outcomes
2008 - Duenser et al.	[41]	10	n/a	Studies that evaluated collaboration between users using AR were quite underrepresented. Only 10 papers were reported, which were divided according to the study type in formal and informal studies.
2008 - Zhou et al.	[42]	not specified	n/a	A small number of examples of collaborative AR prototypes were starting to emerge, but few had been evaluated in formal user studies.
2012 - Bai et al.	[29]	9	Communication, Awareness	An increase in measurements of particular interest in AR collaborative systems included explicit communication (e.g., spoken and gestural messages), ease of collaboration and information gathering (basic awareness, eye gaze).
2015 - Billinghurst et al.	[12]	not specified	Communication	Besides the standard subjective measures, process measures may be more important than quantitative outcome measures. Process measures are typically gathered by transcribing interaction between users, like speech or gestures and performing a conversational analysis. In this context, very few studies have examined communication process measures.
2018 - Kim et al.	[6]	not specified	n/a	A reduce but increasing number of publications explicitly focused on ways to improve collaboration using AR. A mixture of qualitative and quantitative experimental measures were used, such as performance time and accuracy (quantitative), and subjective questionnaires (qualitative).
2018 - Dey et al.	[43]	12	n/a	Need to conduct more user studies regarding collaboration using AR, more use of field studies, and the use of a wider range of evaluation methods. There is an urge to improve the reporting quality of user studies, and education of researchers on how to conduct good AR user studies.
2019 - Ens et al.	[14]	110	Time, Space, Symmetry, Artificiality, Focus, Scenario	Review of the history of collaborative MR systems, and investigation on how common taxonomies and frameworks in CSCW and MR research could be applied to such systems. The authors emphasize that MR systems have been facing significant engineering hurdles and have only recently started to mature to focused on the nuances of supporting collaboration.
2019 - Belen et al.	[25]	259	Task, Awareness, Presence, Social factors	A total of 112 papers studied how MR affects the sense of presence and the perception of social awareness, situational awareness and task awareness during collaboration. A considerable amount of research studied how collaboration reduces cognitive workload through MR environments. 55 papers were categorized under user perception and cognition studies.

Then again, Kim et al. (2018) revisited the trends presented at ISMAR conferences. According to their review, user evaluation and feedback has become one of the main categories for research presented at ISMAR, with 16.4% of publications reporting evaluation being conducted, showing a significant increase when compared to Zhou et al. 5.8% [42]. The authors extended Zhou et al. list of emerging research, including interactive collaborative systems for multiple remote or co-located users. A mixture of qualitative and quantitative experimental measures were used in studies that addressed collaboration, such as performance time and accuracy (quantitative), and subjective questionnaires (qualitative) [6].

Dey et al. (2018) conducted a systematic review of AR usability studies. A total of 291 papers have been reviewed. Among other things, over the years, there were few collaborative user studies, mostly directed towards remote collaboration. The authors reported 12 papers, in a total of 15 studies associated to the collaboration application area. One noticeable feature was the fact that there were no pilot studies reported, which is an area for potential improvement. Also, a reduced number (3 out of 15) of field studies was reported and all except one were performed indoors. Furthermore, a within-subjects design was used by 14 out of 15 studies, since these require fewer participants to achieve adequate statistical significance, with only 12 participants being recruited per study. Besides, roughly one-third of the participants were females in all studies. Hence, participant populations are dominated by mostly young, educated, male participants, suggesting that the field could benefit from more diversity. A majority of the studies, 8 out of 15 collected both objective (quantitative) and subjective (qualitative) data, while 5 studies were only based on subjective data, and 2 studies were based on only objective data. Aside from subjective feedback or ratings, task completion time and error/accuracy were also extensively used. Curiously, the NASA TLX was only used by one study. This analysis suggest the need of more user studies regarding collaboration using AR, particularly more field studies, and the use of a wider range of evaluation methods [43].

Although not strictly a survey, Kim et al. (2018) proposed a questionnaire including aspects regarding overall collaboration, namely the level of enjoyment and mental stress in communication with the partner, and whether collaboration was effective or not [7]. Moreover, the questionnaire included questions about who (presence of others — users' feeling of togetherness with the collaborating partner), what (users' activities — effectiveness in sending and receiving messages) and where (location of activities — whether seeing work space properly and asking the level of having a same focus with a partner). The questionnaire was based on previous work by Gutwin and Greenberg (1999), which suggested three types of experimental measurements are necessary to assess collaboration: product, process, and satisfaction. Product measures focus on assessing collaboration outcomes in terms of efficiency (e.g., task completion time) or quality (e.g., accuracy). Process measures assess user communication and patterns of collaboration and can be obtained by system log data, observation, and video/audio analysis. Satisfaction measures are adequate to assess participants' subjective opinions on the quality of their collaboration and can be collected through interviews and questionnaires [44].

More recently, Ens et al. (2019) revisited collaboration through MR. A total of 110 papers employing MR technology and motivated by challenges in collaborative scenarios was reviewed, showing a rise in the number of papers published from 2012 and onward. The authors emphasize that MR systems have been facing significant engineering hurdles, being limited by the contemporary capabilities of technology, and have only recently started to mature to the point where researchers can focus squarely on the human concerns that underlie communication and collaboration, instead of focusing on creating the enabling technology. The vast majority of papers analyzed (106, or 95%) focused on synchronous collaboration. Moreover, 30 papers (27%) worked on a co-located setting, while 75 papers (68%) worked on a remote setting, and 6 papers (5%) support both settings. In the early years (up to 2005), most research addressed co-located work. Then, the paradigm changed, and from 2006 forward most work tackled remote collaboration. In addition, 45 papers (41%) focus on symmetric collaboration, while 63 (57%) on asymmetric, and 2 (2%) supported both types. The review states that existing methods are not sufficient to characterize how collaboration occurs. Finally, it also emphasizes the need to deepen the understanding of collaborative work through more user studies [14].

Finally, Belen et al. (2019) performed a systematic review of the current state of collaborative MR technologies published from 2013 to 2018. A total of 259 papers have been classified based on their application areas, types of display devices used, collaboration setups, user interaction and experience aspects. Regarding the collaboration setups used, 129 papers (50%) report works that used a remote setup, 103 papers (40%) used a collocated setup, and 27 (10%) used both settings. The type of user interaction and user experience were categorized, resulting in 55 papers categorized under user perception and cognition studies, which aim to lessen cognitive workload for task understanding and completion time and increase users' perceptual (e.g., situational, social, and task) awareness and presence. Besides, a total of 112 papers studied how MR affects the sense of presence and the perception of social awareness, situational awareness and task awareness during collaboration. There was also a considerable amount of research on how collaboration reduces cognitive workload through MR environments. This review also showed that user interaction in a collaborative MR environment is an essential topic that requires further investigation [25].

2.2. Summary

Research is evolving from solving technical issues using AR and MR, towards more meaningful studies on collaboration. We were able to understand that evaluation is frequently done using single-user methods, which are not always applicable to groupware collaborative solutions. To clarify, by single-user methods, we are referring to the methodologies used in the collaborative studies. For example, focusing more on technological aspects of the solution being used than in the collaborative process; including tasks with low complexity that do not elicit real collaboration among participants; using only performance measures like task completion time and error/accuracy data, while other important dimensions are ignored; collecting participant data based only on standard practices with fixed answers, applying scales, questionnaires (e.g., System Usability Scale (SUS), NASA Task Load Index (TLX), among others), which are not thought for collaborative scenarios, thus ignoring detail on crucial aspects of collaboration.

The majority of papers mentioned in the surveys informed on the tasks, types of devices used (although not specific to on-site or remote users), evaluation design, evaluation methods and number of participants, but lack detail on the participants' role, if participants knew each other previous to the study, their previous experience with Virtual Reality (VR), AR or MR solutions, description on the experimental context, among other factors of collaboration. However, our review highlights some limitations included in previous surveys, namely the absence of information regarding specific characteristics of the collaborative context.

These characteristics are important since collaboration may occur at many levels and depends on several factors that may impact directly the collaborative outcomes [40]. Contextual information helps inform the conditions in which the collaborative effort took place. Without comprehending the contextual information, it becomes difficult to assess the important variables related to the collaborative process, which means the results and findings reported may be misleading or of limited value in these scenarios, thus being an important subject to improve the characterization of the collaborative process.

Hence, these aspects have an important impact on how the studies must be prepared and how they were conducted, influencing situation understanding, team-members communication, task performance, and even how AR-based tools were used among team-members, among others. Therefore, it is important to conduct thorough collaborative studies, allowing to retrieve the necessary amount of data for more comprehensive analysis that helps provide a perspective on the different factors of collaboration supported by AR.

To sum up, the use of AR-based multi-site solutions creates challenges to the contextualization of the actions of each user and the problems/barriers they may face. Therefore, having a grasp of those aspects is paramount to ensure characterization is genuine. By doing so, researchers may be able to better assess a wide range of information, namely individual and team personalities, motivations, performances, behaviors, who completed the tasks and who provided instructions, how was the communication process, details of the surrounding environments, as well as duration and type of interactions with the collaborative technology, among other aspects when analyzing data and establishing conclusions.

3. Method and overview of recent literature

To understand to what extent user evaluation is currently being reported covering collaborative AR and Mixed Reality (MR) research, we conducted an analysis of existing works through a systematic review.

This section presents the research methods employed to carry out the review process, which was divided into: the search, i.e., describing how the collection of publications was performed and the review, i.e., explaining the process employed to ensure that the papers follow our review criterion.

What differentiates our review from other surveys described in the previous section is the fact that we focus exclusively on evaluation and user studies in remote scenarios mediated by AR/MR to comprehend how the collaborative process has been captured and reported, rather than addressing the technology that made collaboration possible, which was the nucleus of the two only surveys that dedicated their efforts to the subject of Collaborative AR/MR, while the remaining ones are rather general in application scenario, although also addressing more technological aspects of AR/MR. Besides, by identifying relevant aspects that are missing from existing surveys regarding evaluation and user studies, we are able to include them in our analysis, leading to a discussion in which we critically analyze the field in light of the BRETAM model, thus providing a clearer understanding of how the characterization of the collaborative process has been achieved, which lead to the proposal of a roadmap of relevant research topics, aiming to help the community move the field forward.

3.1. Augmented reality vs mixed reality

While older papers used the term remote collaboration supported by AR, more recent efforts described in literature are beginning to replace the term AR by MR. Next, we elaborate on the meaning of MR, and why this sudden change has started to emerge. Many researchers see MR as a synonym for AR [45]. Some consider MR a superset of AR, i.e., a real-world object can interact with a virtual one in real-time to assist individuals in practical scenarios [46–48]. Yet, others consider MR distinct from AR in the sense that MR enables walking into, and manipulating a scene, whereas AR does not, i.e., there is a separation of the real and virtual world content, which may lead to lower user immersion [49].

Although MR is increasingly gaining in popularity and relevance, the research community is still far from a shared understanding of what MR actually constitutes. Speicher et al. (2019) highlights that currently, there is no single definition for MR, since this concept can be considered different things for different individuals. In their survey, six partially competing notions were identified based on literature analysis and experts' responses. Nevertheless, there is no universally agreed on, one-size-fits-all definition of MR. Moreover, the authors state that it is highly unrealistic to expect one single definition may appear in the future, which means discussions about MR become increasingly difficult. Therefore, it is extremely important to be clear and consistent in terminology while communicating one's understanding of MR in order to avoid confusion and ensure constructive discussion [45].

Among the most important applications of MR are collaborative solutions, that may be used as decision-making tools for daily life problems [14,49]. In this context, Speicher et al. (2019) suggested that MR can be considered as a type of collaboration that describes the interaction between physically separated users exploring AR and VR [45]. This definition includes mapping of the environment of an on-site AR collaborator, i.e., capturing more dimensional information about the local scene, which is reconstructed in VR for the remote collaborator [45,50] and so provides unique capabilities to achieve a common goal, e.g., improved communication cues for more efficient and easier collaboration [8,46,51,52]. Given the aforementioned panorama, we decided to include both terms in our analysis.

3.2. Search process

Our review was made as inclusive as possible. We collected papers from the Scopus database (since it covers most top journals and conferences on Collaborative AR) using the search terms:

> ("Augmented Reality" OR "Mixed Reality")
> AND
> ("Remote Collaboration" OR "Remote Cooperation" OR "Remote Assistance" OR "Remote Guidance" OR "Distributed Collaboration")
> AND
> ("User Evaluation" OR "User Study" OR "User Experiment")

The search for the terms was made in the Title, Abstract, and Keywords fields. All search results published in conferences and journals between 2000 and 2020 were taken into consideration. Only publications in the English language were considered as this is the current 'lingua franca' of the academic research.

3.3. Analysis process

We obtained a total of 64 publications. The search results were analyzed individually to identify whether or not it supported evaluation of remote scenarios supported by solutions using MR or AR. Only 42 publications satisfied the defined criteria. We started by filtering the initial collection of publications to meet our objectives. We removed articles that were incorrectly selected in the search process (false positives) and identified only those articles that included user evaluation. The reviews of each paper focused on the following attributes (Tables 2 and 3): application areas and keywords; type of collaboration; type of task; types of devices used (regarding on-site and remote users); type of study; type of data collected; evaluation design; evaluation methods; number of participants (number of female participants); participant role; participants' familiarity with each other; previous experience with AR/VR/MR; experimental context description; adaptation period provided; study average duration (min); recording of audio and video.

3.4. Validity limitations

A considerable amount of effort was invested on the selection and review process. Although the Scopus bibliographic database has been used to cover a wide range of publication venues and topics, there may be limitations with the described method.

The search terms used might be limiting, as other papers could have used different keywords to describe "Remote Collaboration", "Augmented Reality", "Mixed Reality" or "Evaluation". Therefore, it remains likely that there are papers which may have not been included in this review.

3.5. Results

Next, a high-level overview of the reviewed papers is provided (Tables 2 and 3), following a similar structure as the one used by Dey et al. (2018) in their systematic review [43], which is extended to include relevant aspects missing from the surveys analyzed in the previous section, such as collaboration details, task type, study type, data type, study design, evaluation methods, participants characteristics, experimental context, adaptation period, and duration.

3.5.1. User studies categorization

The papers (Tables 2 and 3) have the following distribution by application areas: assistance (25 papers, 59.5%); assembly (11 papers, 26.2%); co-design (3 papers, 7.1%); social presence (2 paper, 4.8%); education (1 paper, 2.4%), as presented in the orange bubbles in Fig. 1.

Regarding the collaboration details, 30 papers (71.4%) explored collaboration using a synchronous hierarchy approach, i.e., each member has a specific function or expertise and all team members are present and could act in real-time, while 11 papers (26.2%) studied synchronous parallel approach, where all elements have the same level of expertise and could act in real time and only 1 paper (2.4%) studied asynchronous parallel approach, i.e., all elements have the same level of expertise in which collaboration would take place at different times, as shown in the dark blue bubbles in Fig. 1.

Table 2
Summary of User studies on Remote Collaboration using AR or MR — Part 1. Legend: S — Subjective; O — Objective; HHD — Handheld Device; HMD — Head Mounted Display.

ID	Pub.	Year	Application area	Collaboration details	Task type	Devices used (On-site User)	Devices used (Remote User)	Study type	Data type	Study design
1	[53]	2020	Assembly	Hierarchy - Synchronous	Lego Brick Assembly	Projector, External Camera	HMD, Hand Tracker	Formal	S	Within-subjects
2	[54]	2020	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, Controllers, 360° camera	HMD, Controllers	Formal	S	Between-subjects
3	[55]	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360° camera	HMD, Controllers, Hand Tracker	Formal	O + S	–
4	[56]	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation, Puzzle Assembly	See-through HMD, 360° camera	HMD, Controllers	Formal	O + S	–
5	[57]	2019	Assembly	Hierarchy - Synchronous	Lego Brick Assembly	See-through HMD, Depth Sensors	HMD, Hand Tracker	Formal	O + S	Within-subjects
6	[58]	2019	Co-Design	Parallel - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD	See-through HMD,	Formal	S	Within-subjects
7	[59]	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation, Lego Brick Assembly	See-through HMD, 360° camera	HMD, Hand Tracker	Formal	O + S	–
8	[60]	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360° camera	HMD, Controllers	Formal	O + S	–
9	[61]	2019	Social Presence	Parallel - Synchronous	Puzzle Assembly	See-through HMD	HMD, Controllers	Formal	S	Within-subjects
10	[62]	2019	Assembly	Hierarchy - Synchronous	Lego Brick Assembly	Projector, Camera	HMD, Hand Tracker	Formal	O + S	Within-subjects
11	[63]	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360° camera	HMD, Hand Tracker	Formal	O + S	Within-subjects
12	[52]	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, Hand Tracker	HMD, Hand Tracker	Formal	O + S	Between-subjects
13	[64]	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, Hand Tracker	See-through HMD, Hand Tracker	Formal	O + S	Within-subjects
14	[6]	2018	Assembly	Parallel - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Mouse and Keyboard	Formal	O + S	Within-subjects
15	[65]	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360° camera	HMD, Hand Tracker	Formal	O + S	Within-subjects
16	[66]	2018	Assembly	Parallel - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
17	[67]	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HMD, Hand Tracker	HMD, Hand Tracker	Informal, Formal	S	Between-subjects
18	[22]	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HHD	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
19	[68]	2018	Assembly	Hierarchy - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Hand Tracker	Formal	O + S	Between-subjects
20	[69]	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, External Camera	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
21	[70]	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD	HMD, Controllers	Formal	O + S	Within-subjects
22	[71]	2018	Assistance	Parallel - Synchronous	Navigation, Object Selection and Manipulation	HHD	HHD	Formal	O + S	–
23	[72]	2018	Assistance	Parallel - Synchronous	Navigation, Object Selection and Manipulation	HMD, Controllers, Hand Tracker	HMD, Controllers, Hand Tracker	Formal	O + S	–
24	[73]	2018	Assembly	Parallel - Synchronous	Puzzle Assembly	Projector, External Camera	Computer, Gaze Tracker	Informal, Formal	O + S	Within-subjects
25	[74]	2017	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360° camera	HMD, Controllers, Hand Tracker	Formal	O + S	–
26	[75]	2017	Assembly	Hierarchy - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Gaze Tracker	Formal	S	Between-subjects
27	[76]	2017	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, External Camera, Body Tracker	Projector, Optitrack Capture Tracker	Informal	S	Within-subjects
28	[77]	2016	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, Hand Tracker	HMD, Controllers	Informal	O + S	Between-subjects
29	[78]	2015	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation, Lego Brick Assembly	Projector, External Camera	Computer, Mouse and Keyboard	Informal, Formal	O + S	Within-subjects
30	[79]	2015	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD	Computer, Mouse and Keyboard	–	O + S	Between-subjects
31	[80]	2015	Assembly	Parallel - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
32	[81]	2014	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD	Computer, Mouse and Keyboard	–	O + S	Between-subjects
33	[82]	2014	Assembly	Parallel - Synchronous	Puzzle Assembly	HHD or See-through HMD	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
34	[83]	2014	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HHD	Computer with Touch screen	Informal, Field	S	–
35	[84]	2014	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HHD	Computer, Mouse and Keyboard	Informal, Field	O + S	Within-subjects
36	[85]	2013	Assembly	Hierarchy - Synchronous	Puzzle Assembly	Monitor, External Camera	HMD	Formal	O + S	–
37	[86]	2013	Co-Design	Parallel - Synchronous	Navigation, Object Selection and Manipulation	HHD	HHD	Formal	O + S	–
38	[87]	2012	Co-Design	Parallel - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, External Camera	See-through HMD, external camera	Informal	S	Between-subjects
39	[88]	2012	Assistance	Hierarchy - Synchronous	Airplane Cockpit	HHD	Computer, Mouse and Keyboard	Formal	O + S	Within-subjects
40	[89]	2007	Education	Parallel - Asynchronous	Navigation, Object Selection and Manipulation	Computer, External Camera	Computer, External Camera, Gaze Tracker	Informal	S	Between-subjects
41	[90]	2006	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HMD, External Tracker	HMD, External Tracker	Formal	O	–
42	[91]	2004	Social Presence	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	Computer, Mouse and Keyboard	Computer, Mouse and Keyboard	Formal	S	Within-subjects

Table 3
Summary of User studies in Remote Collaboration using AR or MR — Part 2.

ID	Pub.	Evaluation methods	# Participants (# Females)	Participant role	Participants knew each other	Previous experience with AR/VR/MR	Description experimental context	Adaptation period	Duration (min)	Recording audio and video
1	[53]	Questionnaires, Interview	34 (11)	On-site or Remote	Yes and No	–	–	Yes	55	–
2	[54]	Questionnaires, User Preference	40 (-)	On-site or Remote	–	Yes	Yes	Yes	40	–
3	[55]	Task Performance, Questionnaires, User Preference	32 (8)	On-site or Remote	–	Yes	–	–	–	–
4	[56]	Task Performance, Questionnaires, User Preference	10 (1)	On-site or Remote	–	Yes	Yes	Yes	–	–
5	[57]	Task Performance, Questionnaires, User Preference	10 (4)	On-site or Remote	–	–	–	–	–	–
6	[58]	Questionnaires, User Preference	8 (4)	On-site or Remote	–	–	Yes	–	30	–
7	[59]	Task Performance, Questionnaires, User Preference	14 (-)	On-site	–	–	Yes	–	70	Yes
8	[60]	Task Performance, Questionnaires, User Preference	24 (5)	On-site	–	Yes	Yes	–	90	–
9	[61]	Questionnaires, Interview	48 (24)	On-site	Yes	Yes and No	Yes	–	–	–
10	[62]	Task Performance, Questionnaires, User Preference	13 (5)	On-site or Remote	–	No	–	Yes	–	–
11	[63]	Task Performance, Questionnaires, User Preference	12 (3)	On-site and Remote	Yes	Yes	Yes	–	–	–
12	[52]	Task Performance, Questionnaires, Interview	32 (9)	On-site and Remote	Yes	Yes and No	Yes	Yes	120	–
13	[64]	Task Performance, Questionnaires	20 (5)	On-site or Remote	No	–	Yes	Yes	35	–
14	[6]	Task Performance, Questionnaires, User Preference	24 (7)	On-site or Remote	Yes	–	–	–	–	Yes
15	[65]	Task Performance, Questionnaires, User Preference	8 (2)	On-site	–	Yes	Yes	Yes	–	–
16	[66]	Questionnaires, Interview	24 (4)	On-site or Remote	Yes	–	Yes	Yes	–	–
17	[67]	Questionnaires, User Preference	38 (23)	On-site or Remote	–	–	Yes	Yes	30	–
18	[22]	Task Performance, Questionnaires	30 (4)	On-site	–	–	–	–	–	
19	[68]	Task Performance, Questionnaires, User Preference	10 (0)	On-site	–	–	–	–	–	–
20	[69]	Task Performance, Questionnaires, User Preference	8 (4)	On-site or Remote	–	–	Yes	–	60	Yes
21	[70]	Task Performance, Questionnaires, User Preference	16 (5)	On-site or Remote	–	Yes	–	–	–	–
22	[71]	Task Performance, Questionnaires, User Preference	40 (-)	On-site or Remote	–	–	Yes	–	–	–
23	[72]	Task Performance, Questionnaires, User Preference	28 (-)	On-site and Remote	–	Yes	–	–	–	–
24	[73]	Task Performance, Questionnaires, User Preference	24 (16)	On-site or Remote	–	–	Yes	Yes	–	–
25	[74]	Task Performance, Questionnaires, User Preference	8 (-)	On-site and Remote	–	–	–	–	–	–
26	[75]	Questionnaires, User Preference	8 (2)	On-site or Remote	Yes	–	–	–	–	–
27	[76]	Questionnaires, User Preference	8 (-)	Remote	–	–	–	–	–	–
28	[77]	Task Performance, Questionnaires	10 (-)	On-site	–	–	–	–	–	–
29	[78]	Task Performance, Questionnaires	13 (-) + 24 (-)	On-site or Remote	–	–	–	–	–	–
30	[79]	Task Performance, Questionnaires, User Preference	36 (15)	On-site or Remote	–	–	–	–	–	Yes
31	[80]	Task Performance, Questionnaires, Interview	24 (7)	Remote	–	–	Yes	Yes	70	–
32	[81]	Task Performance, Questionnaires, User Preference	–	Remote	–	–	–	–	–	Yes
33	[82]	Task Performance, Questionnaires, User Preference	24 (7)	On-site or Remote	Yes	–	Yes	Yes	90	–
34	[83]	Questionnaires, User Preference	25 (-) + 11 (5)	–	–	–	–	–	50	–
35	[84]	Task Performance, Questionnaires, User Preference	20 (-) + 60 (29)	On-site	–	–	–	–	–	–
36	[85]	Task Performance, Questionnaires, User Preference	14 (-)	On-site and Remote	–	–	Yes	–	–	–
37	[86]	Task Performance, Questionnaires, User Preference	36 (-)	On-site or Remote	–	–	Yes	–	–	–
38	[87]	Interview	5 (-)	On-site	–	–	–	–	–	Yes
39	[88]	Task Performance, Questionnaires	48 (21)	On-site	Yes and No	–	Yes	Yes	–	–
40	[89]	Questionnaires, Interview	9 (3)	On-site	–	–	–	Yes	20	–
41	[90]	Task Performance	12 (2)	On-site or Remote	–	–	–	–	–	–
42	[91]	Questionnaires, User Preference	27 (8)	On-site or Remote	–	–	Yes	–	–	–

Fig. 1. Overview of the main results from the recent literature review on evaluation and AR-supported Remote Collaboration. In the first level are the categories considered for the systematic review, raging among the participants, application areas, collaboration details, study characteristics, task details, adaptation period and evaluation methods. Then, in the outer ring, the detailed topics of interest for each category are presented, respectively. For each, the number of publications covering it is illustrated, following the literature review analysis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.5.2. Study design

As shown in Table 2, 16 papers (38.7%) used a within-subjects design, while 15 papers (35.7%) used a between-subjects design. There were no mentions of a mixed-factorial design. In addition, 11 (26.2%) papers did not mention the method used, as illustrated in the green bubbles in Fig. 1.

3.5.3. Study type

We found that most papers (33, 78.6%) were formal user studies. On the opposite, 7 papers (16.6%) reported conducting informal studies. Only 2 papers (4.8%) conducted user studies in the field, which shows a lack of experimentation in real-world conditions, as exhibit in the green bubbles in Fig. 1.

3.5.4. Task type

As expected, most papers (26 out of 42, 61.9%) explored navigation, object selection and manipulation, forcing participants to communicate and use collaborative tools to provide indications to achieve a concrete goal. Additionally, 12 papers (28.6%) focused on assembly tasks using Lego bricks, or puzzles like tangram, pentominoes, origami, among others. Only 1 paper (2.4%) reported the use of an airplane cockpit as case study, as presented in the red bubbles in Fig. 1. This shows that there is an opportunity for conducting more user studies exploring different, more complex case studies, or even combinations of different types.

Moreover, just 14 papers (33.33%) claim to have provided an adaptation period before the performance of the tasks, as shown in the purple bubbles in Fig. 1. Finally, the bulk of the user studies were conducted in an indoor environment, but only 21 papers (50%) described the experimental context, although no clear pattern emerged.

3.5.5. Evaluation methods and data type

In terms of data type, 30 papers (71.4%) collected subjective and objective data, 11 papers (26.2%) collected only subjective data, and just 1 (2.4%) only objective data.

Concerning the evaluation methods, we found that the most popular method is filling out questionnaires (40 papers, 95.2%), followed by assessing task performance (31 papers, 73.8%) with error/accuracy measures and task completion time. Then, user preference (28 papers, 66.7%) and finally interviews (5 papers, 11.9%), as illustrated in the light blue bubbles in Fig. 1. Note that many papers used more than one evaluation method, so the percentages sum to more than 100%.

Another essential point: only 13 papers (31%) mentioned the average duration of the user study (58.5 min). Some papers mentioned the duration of the task, but no clear information on the collaboration process is provided, like dialog length, frequency of conversational turns, among others. Besides, none of the papers report to have conducted gesture or non-verbal behaviors analysis. This is supported by the lack of audio or video recording, since only 6 papers (14.3%) acknowledge to store this type of data.

Table 4
Overview of common approaches and what is missing regarding the evaluation process of remote collaboration mediated by AR.

Common:	•synchronous hierarchy collaboration •within-subjects design •formal user studies •navigation, selection, manipulation and assembly tasks •focus on technological aspects or interaction mechanisms of the collaborative AR solution •subjective and objective data collection •use of single-user questionnaires, task performance and user preferences assessment •young participants from universities •participants act as on-site or remote team-members
Missing:	•conduct outdoor and field studies •explore complex/adequate tasks •contemplate failure situations •provide an adaptation period •address participants relationships, knowledge and motivations •better description of the collaborative process supported by AR •reporting of study average duration •data collection on dialog turns, interaction types, main features and visual complexity •contextualized information on the team, task, environment and collaborative tool •improve existing frameworks •use of video, audio recordings, and post-task interviews

3.5.6. Participants

Our review of the participants shows that the number of participants involved in the analyzed studies ranged from 5 to 48, with an average of 21. Also, a total of 31 out of 42 papers (73.4%) reported involving female participants in their experiments, with the ratio of female participants to male participants being 47.6% of total participants in those 31 papers. Hence, most of the studies were run with young participants, mostly university students, rather than a more representative cross section of the population.

Equally important, 23 papers (54.8%) stated that participants would perform the role of the on-site or remote user during the studies. Moreover, in 5 papers (11.9%) the participants would perform the on-site and remote role. 11 papers (26.2%) only allowed the participants to perform the on-site user, while 3 papers (7.1%) only allowed to perform the remote role. In these cases, the counterpart would be performed by a monitor, as presented in the brown bubbles in Fig. 1.

Most papers, 32 out of 42 (76.2%) made no mention if participants knew each other, with only 9 clearly stating that information. Likewise, the same percentage did not mention any type of previous experience the participants might have with AR or MR systems.

3.6. Summary

Our review (Table 4 and Fig. 1) shows that the dominant type of collaboration is based on the hierarchy approach focused on synchronous communication between participants. Also, that assistance and assembly are the main areas of application, exploring navigation, selection and manipulation tasks in indoor environments, during approximately one hour.

On average, studies involved 21 participants, mostly young university students. Moreover, ruffly half of the papers reported that the participants would perform the role of the on-site or remote user during the studies. Besides, most papers lack information regarding if participants knew each other prior to the study and if they had previous experience with MR systems.

The majority of the studies conducted are formal studies, collecting objective and subjective data using evaluation methods like questionnaires, task performance and user preferences in that order respectively. As for collaborative measures, most works focus on effectiveness, only checking if participants were able to accomplish a given task collaboratively. Moreover, the evaluation design is distributed between within-subjects and between-subjects.

Besides, interviews are not used often, as is also the case of recording audio and video during the studies. In addition, half of the times the experimental context is not described and only one third of the times studies referred the existence of an adaptation period. It is important to report this last fact, as it can affect the way the collaboration process was performed between collaborators, i.e., those that had an opportunity to use, adapt and comprehend the technology that helped create a shared understanding prior to the tasks will easily interact better with their respective counterpart, when compared to the ones that have only done the adjustment process during the task itself.

Another observation is that single-user evaluation methods are applied to collaborative tasks, which mainly focus on the comparison of technological aspects or interaction mechanisms based on rather simpler procedures. We argue that collaborative tasks must be difficult and long enough to encourage interaction between collaborators and for the AR-based solution being used to provide enough contribution. In general, tasks can benefit from deliberate drawbacks, and constraints, i.e., incorrect, contradictory, vague or missing information, to force more complicated situations and elicit collaboration. For example, suggest the use of an object which does not exist in the environment of the other collaborator or suggest removing a red cable, which is green in the other collaborator context. Such situations help introduce different levels of complexity, which go beyond the standard approaches used, and elicit more realistic real-life situations where the surroundings are not always perfect. Likewise, multiple procedures may be applied to an evaluation, while also exploring different levels of complexity, contextual changes in the surroundings environments, as well as stress conditions.

4. Critical analysis

This section describes the main limitations hindering a better understanding regarding how AR supports collaborative work in remote scenarios. Analysis was mostly based in the results from the literature review process, complemented by meetings with domain experts, and authors' own experience creating and conducting evaluation studies in this domain [Refs omitted for review purposes]. The contributions presented in this paper were conducted in the scope of a larger multidisciplinary research line, with a total of nine individuals with several years of expertise (minimal of 6 years, and a maximum of 40 years of experience) in the areas of Human–Computer Interaction (HCI), Virtual and Augmented Reality (VR/AR), Information Visualization (IV), Multimodal Interaction (MMI), as well as remote collaboration in several scenarios of application. To this effect, face-to-face and remote meetings were conducted, as well as focus group and brainstorm sessions (sometimes with different combinations of experts according to their availability) over several months. To conclude the section, a global assessment of the field maturity is attempted, followed by a critical analysis on how that may affect the road ahead.

4.1. Main limitations

As was aforementioned, the characterization and evaluation of the collaborative process in remote scenarios using AR-based solutions have been reported mainly using single-user methods focusing on technological aspects, thus lacking information and focus on the important dimensions of collaboration. As a result the following main limitations can be identified.

Limitation 1: partial evaluation

According to Merino et al. *"designing appropriate evaluations that examine MR/AR is challenging, and suitable guidance to design and conduct evaluations of MR/AR are largely missing"* [37]. This fact is further evident in scenarios of remote collaboration, since the logistics associated with carrying out evaluations is even more demanding due to a significant number of variables that must be considered. The existence of two or more collaborators makes it more difficult to evaluate the solution as a whole, given that it requires to perform multiple evaluations at the same time and that validation from all users is required [40]. As a consequence, there is a clear lack in addressing crucial aspects of collaboration like how was the relation and communication of the collaborators during the tasks (only 10 out of 42 papers reported such information and just 6 recorded audio or video during the studies), whether they had previous experience with AR/VR/MR technologies and were able to use the available solutions to their full potential (a topic just mentioned by 11 out of 42 papers), how the available information was used to support the accomplishment of the tasks, among other aspects.

In this context, trying to apply conventional evaluation techniques to collaborative settings, without adapting them can lead to an incomplete vision of the process of collaboration and in turn to dubious results, falling short to retrieve the necessary amount of data for more comprehensive evaluations and characterizations of the collaborative process which may lead to an incomplete vision of the process of collaboration. Given the complex environments and situations collaborators may encounter, such methods alone provide insufficient information and rarely are good indicators for improving distributed solutions [31,38,92,93].

Limitation 2: lack of contextual information

Remote collaboration represents high levels of data by involving different types of distributed collaborators, tasks and in encompassing dynamical environments with contextual data. Dey et al. revealed that *"work needs to be done towards making AR-based remote collaboration akin to the real world with not only shared understanding of the task but also shared understanding of the other collaborators emotional and physiological states"* [43]. Moreover, Ratcliffe et al. suggested that *"remote settings introduce additional uncontrolled variables that need to be considered by researchers, such as potential unknown distractions, trust in participants and their motivation, and issues with remote environmental spaces"* [94]. However, our analysis shows that half of the papers analyzed (21 out of 42) did not described the experimental context of collaborators, and that 76.2% (32 out of 42) did not report participants knowledge of each other. The same percentage of papers did not mention previous experience with AR or MR technologies, as illustrated in Table 3. By doing this, evaluation scenarios disregard information such as contextual or user related data, obtaining only superficial results.

Limitation 3: failure situations are not contemplated

Bai et al. stated that: *"as deeper insight is obtained into the affordances of AR collaboration, more complex activities should be supported"* [29]. This is also corroborated by Ens et al. which highlighted that *"as new capabilities emerge, (...) we expect to see this trend continue, with an initial focus on perfecting the systems, followed by deeper explorations of collaboration"* [14]. Furthermore, this is also supported by our analysis from the selected data set, which shows that failure situations were not taken into account by any study. For example, in the case of failure to achieve the intended goals of the collaborative process, how can we understand what went wrong? Was it caused by problems in participants communication, by too much augmented information being displayed, by the actions of a particular collaborator that did not followed correctly some indications, or was it caused by an error in the AR-based solution being used?

Limitation 4: lack of theories and guidelines

Literature shows an absence of rules, guidelines and theories to guide the characterization of the collaborative process using solutions mediated by AR. For example, Dey et al. suggests that *"opportunities for increased user studies in collaboration, more use of field studies, and a wider range of evaluation methods"* [43]. Moreover, Ens et al. reported that *"MR systems faced significant engineering hurdles, and have only recently started catching up to provide new theories and lessons for collaboration"* [14]. A better evaluation strategy is required by researchers and developers to obtain a comprehensive description, given the challenges involved in evaluating many aspects that may influence the way collaboration occurs, e.g., relations between individuals, their interconnection as a team and how the use of AR/MR technologies affected the accomplishment of the tasks in relation to the collaborative effort.

Limitation 5: minimal support in existing frameworks

The constraints and challenges identified may change according to the maturity of the solution being used, the goal of the evaluation, the participants individual and group characteristics, among other parameters. In this context, existing frameworks are not sufficiently well suited to describe how collaboration mediated by AR/MR technologies happens, thus ignoring detail on crucial aspects of collaboration [7,14,29,36,39,43].

For example, Bai et al. emphasized that *"it can be hard to isolate the factors that are specifically relevant to collaboration"* [29]. Likewise, Ens et al. outlined that *"frameworks for describing groupware and MR systems are not sufficient to characterize how collaboration occurs through this new medium"* [14]. In addition, Ratcliffe et al. communicate that *"the infrastructure for collecting and storing this (mass) of XR data remotely is currently not fully implemented, and we are not aware of any end-to-end standardized framework"* [94]. Therefore, integration of proper characterization and evaluation methods and guidelines, covering different contexts of use and tasks, running in its intended (real or simulated) environment are of paramount importance.

Limitation 6: limited reporting of outcomes

There is now an opportunity to convince researchers to better document their work, and help improve evaluations and characterizations that are, in our view, a bottleneck in this research area. Currently, researchers struggle to analyze the state of the art, since much information on existing publications lack detail on the collaborative process as previously demonstrated. This may happen since most of the research efforts have been devoted on creating the enabling technology.

4.2. Maturity of the field

To put in perspective the evolution of the field, as well as consider current limitations, this section concludes with the analysis of the status of the area according to the BRETAM model (Fig. 2) [95]. This model has been considered useful for the introduction of new knowledge, technology or products and adopted in several scenarios, including for example, in a multimodal interaction review [96].

According to the current panorama reported in this publication, we argue that it is possible to situate the field of remote collaboration mediated by AR between the **Replication** and **Empiricism** phases of the BRETAM model as illustrated in Fig. 2. We argue that remote collaboration mediated by AR has already passed the **Breakthrough** phase, which means research institutions worldwide can replicate the basic concepts, as demonstrated by the last few decades of research [25].

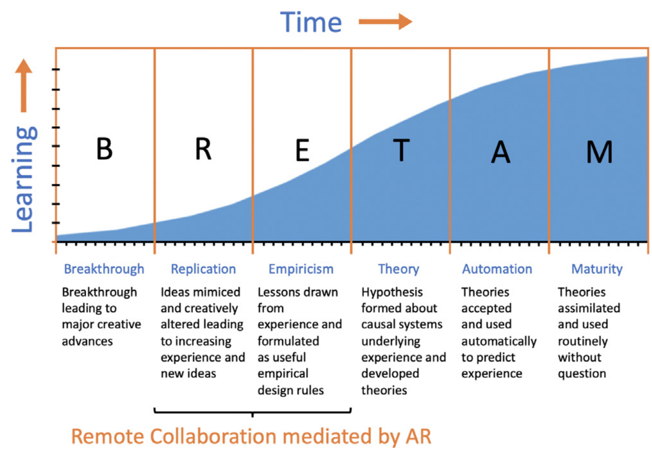

Fig. 2. Positioning of Remote Collaboration mediated by AR between the Replication and Empiricism phases of the BRETAM model. Inspired by [95].

The **Replication** and **Empiricism** phases on the other hand imply increased ideas to generate enough experience, leading to empirical design rules. As such, these phases seem adequate to the overall panorama described in this publication, reinforcing the need to deepen the understanding and characterization of the collaborative process through methods, frameworks, guidelines and various user studies.

In our view, remote collaboration mediated by AR has still not reached the **Theory** phase as it requires enough empirical experience to model the basis of success and failure, which cannot be performed without proper methods for the characterization and evaluation of the collaborative process [14,25].

Likewise, the **Automation** phase was also rejected, which implies automation of the scientific data-gathering and analysis, since existing systems are still limited by the contextual and multi-user data they are able to collect, thus not being sufficient to characterize how collaboration occurs [14].

As such, without fulfilling the previous phases, the field cannot be positioned into the **Maturity** phase, i.e., turn to cost reduction and quality improvements in what describes a mature technology [95].

5. Charting out a roadmap for the characterization and evaluation of the collaborative process

According to what was said in the critical analysis, it is important to address the main limitations to make the field achieve the Theory, Automation and Maturity phases of the BRETAM model [95]. Aiming at contributing to that, in this section we propose a first roadmap, to deal with the most pressing issues (Fig. 3), composed by five key topics:

- definition of dimensions of collaboration to face the partial characterization of the collaborative process;
- systematization of perspectives based on the acquired knowledge of the field, facing the lack of theories and guidelines;
- creation of new paradigms, architectures and frameworks to answer the limited support to development and evaluation of existing ones;
- enhanced support for data gathering, leading to better design, development and evaluation with distributed users supported by AR;
- new and better outcomes from the evaluation to support the assessment, leading to the creation of new theories, as well as improve the lack of contextual information.

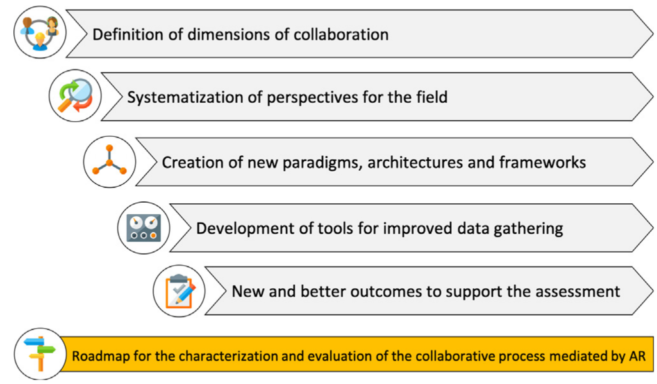

Fig. 3. Roadmap overview of the main topics that should be addressed regarding remote collaboration mediated by AR to make the field achieve the Theory, Automation and Maturity phases of the BRETAM model. Inspired by [97].

5.1. Definition of dimensions of collaboration

First, it is important to identify dimensions that need to be taken into consideration when performing the characterization of the collaborative process. In practical terms, given a concrete application context and a problem, the research community is still not able to provide an overall definition of the collaborative AR system that addresses it. Although there are works that have presented some dimensions of collaboration, existing efforts are mostly oriented towards technology. As the field matures, it is normal new proposals emerge to address new aspects related to collaboration. A comprehensive set of dimensions must be defined to more thoroughly classify and discuss the contributions of the collaborative work effort, not only addressing the technological features being used, but also encompassing the characteristics of the context.

For example, Ens et al. stated the following: *"While somewhat useful, the dimensions we used are fairly technical, and focus mainly on mechanical aspects of the system or properties of the underlying technologies. (...) Perhaps additional dimensions with a greater focus on user experience would better allow for capturing the essence of collaborative scenarios*" [14]. Therefore, some of the existing dimensions might still not reflect the full scope of some categories by encompassing all possibilities. Therefore, this effort cannot be intended as a closed work, but should, instead, be taken as the grounds that might enable the community to elaborate, expand, and refine the field.

This may be achieved by analyzing the literature regarding collaborative work supported by AR, in particular, existing categorization efforts [13,14,25,45,98–101]. Another possibility is to adopt a conceptual-to-empirical methodology by using a participatory design process, i.e., actively involving stakeholders in focus group and brainstorming sections. This entails going beyond Collaborative AR literature, considering other domains (e.g., CSCW, Groupware, Telerehabilitation, Remote Medicine, among others) that may be relevant to characterize the collaborative effort, to identify which dimensions should be taken into account when we move from asking what existing systems can do, to understanding what they would be able to do in particular contexts, i.e., the value of AR to the collaborative process.

5.2. Systematization of perspectives for the field

Ens et al. report that when considering if it is possible to clearly describe distinct categories of collaborative MR research based on the existing dimensions, the answer is *"to some extent, yes, however the result is not wholly satisfying (...) these dimensions* *do not suffice to describe all scenarios"* [14]. Therefore, another area of research that needs to be addressed given the lack of theories and guidelines [14,43], is the need to bring new dimensions forward into conceptual models, guidelines, taxonomies and ontologies, that might foster harmonization of perspectives for the field, thus creating a common ground for systematization and discussion [100,102].

Through these, it would be possible to structure the characterization of the collaboration process, which can form the basis for analysis and comparison, fostering a more detailed understanding of the field, and in turn ensure that the research adds to the body of knowledge and provides enough context and evidence to enable a transparent account [103] and transferability [104]. These can also work as a knowledge repository for evaluation, allowing researchers to observe and compare a variety of results inside the same domain and make considerations and conclusions about specific nuances of collaboration. For example, the proposal of human-centered approaches, i.e., focusing on collaboration, instead of the technology, might bring forward a perspective that is not rapidly deprecated with the advancements of technology [105].

To create conceptual models and taxonomies, it is important to ensure the dimensions of collaboration contain categories and characteristics that are mutually exclusive and collectively exhaustive [106,107]; Moreover, a detailed explanation of these objects of interest must be included, following, for example, a similar approach to the one used by Zollmann et al. [108]. It is also relevant to include discussion and refinement over several iterations with domain experts, to verify if the established dimensions, categories and characteristics are well defined, need to be merged, or if new ones can be identified [106]. Regarding ontologies, literature shows that its design is considered a creative process and no two ontologies by different individuals would be the same, since the applications of the ontology and the designer's understanding of the domain will undoubtedly affect the ontology design choices [109,110]. As such, one possible strategy is to adapt existing ontologies when they exist, or as an alternative when this is not the case, define and populate a new ontology considering relevant dimensions of collaboration as the core classes and establish their relations with each other based on the targeted application of the ontology.

5.3. Creation of new paradigms, architectures and frameworks

According to Merino et al. *"as MR/AR technologies become more mature, questions that involve human aspects will gain focus in MR/AR research. Consequently, we expect that future MR/AR papers will elaborate on human-centered evaluations that involve not only the analysis of user performance and user experience, but also the analysis of other scenarios, like understanding the role of MR/AR in working places and in communication and collaboration"* [37]. However, there is no standard methodology for characterization and evaluation, specifically tailored to assess how remote collaboration occurs through AR/MR technologies. In this vein, without the appropriate paradigms, methods and mechanisms, the research community does not accumulate enough experience to improve collaboration between distributed collaborators [7,14, 29,36,37,39,43,92].

Currently, there is too much focus on post-task evaluation. New paradigms must also consider continuous assessment, i.e., giving proper relevance to evaluation conducted during the accomplishment of open challenges, instead of pre-defined tasks, which fail short to mimic real scenarios of remote collaboration.

As such, architectures and frameworks capable of supporting the new paradigm(s) must be created, to assist researchers conducting future user studies, while eliciting more characterization of the collaboration process in remote scenarios. Such frameworks must include support for:

- defining the **evaluation scope** for individual and collective assessment by properly identifying which dimensions of collaboration will be evaluated;
- detailing **collaborative challenges** to be performed, including specification of the users minimum level of knowledge, definition of each collaborator activity, as well as definition of the procedures;
- defining the **experimental setup and design**, ensuring each dimension is defined in terms of the necessary variables and how they should be measured according to specific techniques;
- conducting **data gathering** through the use of a distributed evaluation tools focusing in the dimensions proposed specifically for remote collaboration;
- performing **data analysis**, including inspection of what happened during the tasks, to understand how the collaboration process occurred over time.

5.4. Development of tools for improved data gathering

According to Marques et al. *"it is essential the existence of holistic evaluation strategies that monitor the use and performance of the proposed solutions regarding the team, its members, and the technology, allowing adequate characterization and report of collaborative processes"* [36]. To achieve this, the operationalization of data gathering should also deserve its own line of work due to its importance. It is paramount to conduct thorough collaborative user studies to provide new perspectives [14,36–38,111]. A better evaluation process can be supported by improved data collection and data visualization tools [92,112]. In this context, it is necessary to collect, process and analyze a multiplicity of data, e.g., context, history, user related information like actions, emotional state, as well as the results of processing the various components of the data gathering tools, aiming at obtaining a more comprehensive understanding.

To accomplish this, tools must be designed and developed to allow researchers to run multiple evaluations at different locations simultaneously, following a distributed paradigm [36]. In this process, researchers should be able define measures, custom logging and register interesting events they detect, which can be later reviewed in post-task analysis, adapting and extending, for example the works by Pereira et al. [93,113,114]. Likewise, the following factors are crucial and must be taken in account to better understand the real impact of each aspect in the collaborative effort: team, collaborative tasks, context and AR-based solution [36]. These factors can help portrait the conditions in which collaborators performed a given action, received information or requested assistance. In addition, they can be used to assert uncommon situations or identify patterns that can lead to new understanding of a given artifact, as well as identify new research questions. Therefore, such tools are essential to help researchers when performing judgment over evaluation results.

5.5. New and better outcomes to support the assessment

A better characterization of the collaborative process coupled with improved and specific evaluation tools and methods will provide ground to improve how research is reported. Thus, increasing the awareness of researchers about the different dimensions of collaboration and elicit better reporting, as researchers understand the need to improve how they describe the nuances associated to the collaborative effort of their work. Currently, in most cases, data relevant to characterize the collaborative context is not reported.

To elaborate, most works focus only on individual performance, on the technological aspects of the AR-based solution or in quantifying effectiveness of tasks. It is important to consider a wide range of information, namely individual and team personalities, motivations, performances, behaviors, who completed the tasks and who provided instructions, how was the communication process, as well as duration and type of interactions with the collaborative AR-based technology, among other aspects when analyzing data and establishing conclusions. The reporting process must also integrate the context in which the collaborative effort took place, thus allowing the creation of a better understanding of the surrounding conditions, while contributing to support replication of such context, if they are relevant to other researchers, in future studies. Moreover, a complete definition of the data used to substantiate the usefulness of the results reported must be included, as well as the measures used, how was the data computed, based on what criteria, etc. This is essential to move towards replication and interpretability across contributions in the field.

A more systematic reporting can, in turn, lead to a community setting that enables easier communication, understanding, reflection, comparison, refining, as well as building on existing research and foster harmonization of perspectives for the field.

Furthermore, researchers can also compare their outcomes, as this is also a good opportunity for reflecting and refining. It is important to use what is learned during the studies and identify aspects which did not go according to what was expected or select additional ones which may improve on existing guidelines for future user studies.

6. Conclusions and future work

Collaborative AR solutions can be powerful tools for analysis, discussion and support of complex problems and situations in remote scenarios. By bringing different and sometimes opposing points of view together, such solutions can lead to new insights, innovative ideas, and interesting artifacts.

However, most research efforts have been devoted to creating the enabling technology for supporting the design and development of such solutions. Hence, the characterization and evaluation of the collaborative process is an essential, but a very difficult endeavor nowadays.

This paper describes a critical analysis supported by surveys that addressed evaluation and user studies in scenarios of remote collaboration mediated by AR. In addition, a literature review on works ranging from 2000 to 2020 is also presented. Based on the limitations and challenges identified, we argue that remote collaboration mediated by AR is currently between the Replication and Empiricism phases of the BRETAM model. To contribute to an advance to Theory, Automation and Maturity phases, based in the critical analysis, we propose a roadmap for important research actions that need to be addressed to facilitate and elicit more characterization of the collaboration process using AR-based solutions in the future.

Work is being continued through the creation of a conceptual model and taxonomy, as well as an initial architecture and framework aligned with the proposed roadmap. These can form the basis for a common ground, as well as the development of a framework for researchers who want to follow best practices in designing their own collaborative AR user studies in remote scenarios.

CRediT authorship contribution statement

Bernardo Marques: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **António Teixeira:** Conceptualization, Methodology, Writing – original draft, Writing – review, Visualization. **Samuel Silva:** Methodology, Investigation, Writing – original draft, Writing – review & editing. **João Alves:** Validation, Formal analysis, Investigation, Data curation, Writing – original draft. **Paulo Dias:** Conceptualization, Investigation, Data curation, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition. **Beatriz Sousa Santos:** Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Supervision Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank the reviewers for their thoughtful comments and suggestions towards improving on an earlier version of this manuscript. To everyone involved in case studies, and discussion groups, thanks for your time and expertise. This research was developed in the scope of the Ph.D. grant [SFRH/BD/143276/2019], funded by FCT - Foundation for Science and Technology. It was also supported by IEETA - Institute of Electronics and Informatics Engineering of Aveiro, funded by National Funds through FCT, in the context of the project [UID/CEC/00127/2019].

References

[1] Alem L, Tecchia F, Huang W. HandsOnVideo: Towards a gesture based mobile AR system for remote collaboration. Recent Trends of Mobile Collaborative Augmented Reality Systems. 2011, p. 135–48,
[2] Johnson S, Gibson M, Mutlu B. Handheld or handsfree? Remote collaboration via lightweight head-mounted displays and handheld devices. In: Proceedings of the 18th ACM conference on computer supported cooperative work & social computing; 2015, p. 1825–36.
[3] Lukosch S, Billinghurst M, Alem L, Kiyokawa K. Collaboration in augmented reality. In: Computer Supported Cooperative Work, CSCW 2015. 24, 2015, p. 515–25.
[4] Schneider M, Rambach J, Stricker D. Augmented reality based on edge computing using the example of remote live support. In: 2017 IEEE international conference on industrial technology; 2017, p. 1277–82.
[5] Thomas PJ. CSCW Requirements and Evaluation. Berlin, Heidelberg: Springer-Verlag; 1996.
[6] Kim K, Billinghurst M, Bruder G, Duh HB, Welch GF. Revisiting trends in augmented reality research: A review of the 2nd decade of ISMAR (2008-2017). IEEE Trans Vis Comput Graphics 2018;24:2947–62.
[7] Kim S, Billinghurst M, Lee G. The effect of collaboration styles and view independence on video-mediated remote collaboration. Comput Support Coop Work: CSCW: An Int J 2018;27(3–6):569–607.
[8] Kim S, Lee G, Billinghurst M, Huang W. The combination of visual communication cues in mixed reality remote collaboration. J Multimodal User Interf 2020;14(4):321–35.
[9] Kim S, Billinghurst M, Kim K. Multimodal interfaces and communication cues for remote collaboration. J Multimodal User Interf 2020;14(4):313–9.
[10] Arias E, Eden H, Fischer G, Gorman A, Scharff E. Transcending the individual human mind-creating shared understanding through collaborative design. ACM Trans Comput-Human Inter 2000;7:84–113.
[11] Grudin J, Poltrock S. The Encyclopedia of Human-Computer Interaction. The Interaction Design Foundation; 2013.
[12] Billinghurst M, Clark A, Lee G. A survey of augmented reality. Found Trends Human–Comput Interact 2015;8:73–272.
[13] Jalo H, Pirkkalainen H, Torro O, Kärkkäinen H, Puhto J, Kankaanpää T. How can collaborative augmented reality support operative work in the facility management industry?. In: Proceedings of the 10th international joint conference on knowledge discovery, knowledge engineering and knowledge management; 2018. p. 41–51.
[14] Ens B, Lanir J, Tang A, Bateman S, Lee G, Piumsomboon T, Billinghurst M. Revisiting collaboration through mixed reality: The evolution of groupware. Int J Human-Comput Stud 2019;131:81–98.

[15] Bruno F, Barbieri L, Marino E, Muzzupappa M, D'Oriano L, Colacino B. An augmented reality tool to detect and annotate design variations in an industry 4.0 approach. Int J Adv Manuf Technol 2019;105(1):875–87.
[16] Ong SK, Yuan ML, Nee AYC. Augmented reality applications in manufacturing: A survey. Int J Prod Res 2008;46(10):2707–42.
[17] Wang X, Ong SK, Nee AYC. A comprehensive survey of augmented reality assembly research. Adv Manuf 2016;4(1):1–22.
[18] Palmarini R, Erkoyuncu JA, Roy R, Torabmostaedi H. A systematic review of augmented reality applications in maintenance. Robot Comput-Integr Manuf 2018;49:215–28.
[19] Hall M, McMahon CA, Bermell-Garcia P, Johansson A, Ravindranath R. Capturing synchronous collaborative design activities: A state-of-the-art technology review. In: Proceedings of international design conference, DESIGN 2018; 2018. p. 347–58.
[20] Billinghurst M. Grand challenges for augmented reality. Front Virtual Reality 2021;2:1–4.
[21] Altug Y, Mahdy AM. A perspective on distributed and collaborative augmented reality. Int J Recent Trends Human Comput Interact (IJHCI) 2016;7:23–41.
[22] Choi S, Kim M, Lee J. Situation-dependent remote AR collaborations: Image-based collaboration using a 3D perspective map and live video-based collaboration with a synchronized VR mode. Comput Ind 2018;101:51–66.
[23] Bottani E, Vignali G. Augmented reality technology in the manufacturing industry: A review of the last decade. IISE Trans 2019;51(3):284–310.
[24] van Lopik K, Sinclair M, Sharpe R, Conway P, West A. Developing augmented reality capabilities for industry 4.0 small enterprises: Lessons learnt from a content authoring case study. Comput Ind 2020;117.
[25] Belen RAJ, Nguyen H, Filonik D, Favero DD, Bednarz T. A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. AIMS Electron Electr Eng 2019;3:181.
[26] Oda O, Elvezio C, Sukan M, Feiner S, Tversky B. Virtual replicas for remote assistance in virtual and augmented reality. In: Proceedings of the 28th annual acm symposium on user interface software & technology - UIST '15; 2015. p. 405–15.
[27] Elvezio C, Sukan M, Oda O, Feiner S, Tversky B. Remote collaboration in AR and VR using virtual replicas. ACM SIGGRAPH 2017 2017;1–2.
[28] Barroso Ja, Fonseca L, Marques B, Dias P, Sousa BS. Remote collaboration using mixed reality: Exploring a shared model approach through different interaction methods. In: Proceedings of european conference on computer-supported cooperative work, ecscw 2020 posters; 2020. p. 1–6.
[29] Bai Z, Blackwell AF. Analytic review of usability evaluation in ISMAR. Interact Comput 2012;24(6):450–60.
[30] Zillner J, Mendez E, Wagner D. Augmented reality remote collaboration with dense reconstruction. In: Adjunct proceedings - 2018 IEEE international symposium on mixed and augmented reality, ismar-adjunct 2018; 2018. p. 38–9.
[31] Neale DC, Carroll JM, Rosson MB. Evaluating computer-supported cooperative work: Models and frameworks. In: Proceedings of the 2004 ACM conference on computer supported cooperative work, CSCW '04; 2004. p. 112–21.
[32] de Souza Cardoso LF, Mariano FCMQ, Zorzal ER. A survey of industrial augmented reality. Comput Ind Eng 2020;139:106159.
[33] Röltgen D, Dumitrescu R. Classification of industrial augmented reality use cases. Procedia CIRP 2020;91:93–100.
[34] Fernández del Amo I, Erkoyuncu JA, Roy R, Wilding S. Augmented reality in maintenance: An information-centred design framework. Proc Manuf 2018;19:148–55.
[35] Jetter J, Eimecke J, Rese A. Augmented reality tools for industrial applications: What are potential key performance indicators and who benefits? Comput Hum Behav 2018;87:18–33.
[36] Marques B, Teixeira A, Silva S, Alves Ja, Dias P, Santos BS. A conceptual model for data collection and analysis for AR-based remote collaboration evaluation. In: IEEE international symposium on mixed and augmented reality, ISMAR; 2020.
[37] Merino L, Schwarzl M, Kraus M, Sedlmair M, Schmalstieg D, Weiskopf D. Evaluating mixed and augmented reality: A systematic literature review (2009–2019). In: IEEE international symposium on mixed and augmented reality, ISMAR; 2020. p. 438–51.
[38] Hamadache K, Lancieri L. Strategies and taxonomy, tailoring your CSCW evaluation. In: Groupware: Design, Implementation, and Use. 2009, p. 206–21.
[39] Antunes P, Herskovic V, Ochoa SF, Pino JA. Reviewing the quality of awareness support in collaborative applications. J Syst Softw 2014;89:146–69.
[40] Patel H, Pettitt M, Wilson JR. Factors of collaborative working: A framework for a collaboration model. Applied Ergon 2012;43(1):1–26.
[41] Duenser A, Grasset R, Billinghurst M. A survey of evaluation techniques used in augmented reality studies. In: SIGGRAPH 2008. 2008.
[42] Feng Zhou, Duh HB, Billinghurst M. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In: International symposium on mixed and augmented reality; 2008. p. 193–202.
[43] Dey A, Billinghurst M, Lindeman RW, Swan JE. A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. Front Robot AI 2018;5:37.
[44] Gutwin C, Greenberg S. The effects of workspace awareness support on the usability of real-time distributed groupware. ACM Trans Comput-Hum Interact 1999;6(3):243–81.
[45] Speicher M, Hall BD, Nebeling M. What is mixed reality? In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19, 2019, p. 1–15.
[46] Kato H, Billinghurst M. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In: Proceedings 2nd IEEE and ACM international workshop on augmented reality (IWAR'99); 1999. p. 85–94.
[47] Milgram P, Takemura H, Utsumi A, Kishino F. Augmented reality: A class of displays on the reality-virtuality continuum. Photon Indust Appl 1994;282–92.
[48] Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Trans Inform Syst 1994;77(12):1321–9.
[49] Rokhsaritalemi S, Sadeghi-Niaraki A, Choi S-M. A review on mixed reality: Current trends, challenges and prospects. Appl Sci 2020;10(2).
[50] Bai H, Sasikumar P, Yang J, Billinghurst M. A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing. Conf Human Factors Comput Syst - Proc 2020;1–13.
[51] Masai K, Kunze K, Sugimoto M, Billinghurst M. Empathy glasses. In: Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems, CHI EA '16; 2016. p. 1257–63.
[52] Piumsomboon T, Dey A, Ens B, Lee G, Billinghurst M. The effects of sharing awareness cues in collaborative mixed reality. Front Robot AI 2019;6.
[53] Wang P, Bai X, Billinghurst M, Zhang S, Han D, Sun M, Wang Z, Lv H, Han S. Haptic feedback helps me? A VR-SAR remote collaborative system with tangible interaction. Int J Human-Comput Interact 2020.
[54] Rhee T, Thompson S, Medeiros D, Dos Anjos R, Chalmers A. Augmented virtual teleportation for high-fidelity telecollaboration. IEEE Trans Vis Comput Graphics 2020.
[55] Teo T, Lee G, Billinghurst M, Adcock M. Investigating the use of different visual cues to improve social presence within a 360 mixed reality remote collaboration. In: Proceedings - VRCAI 2019: 17th ACM siggraph international conference on virtual-reality continuum and its applications in industry; 2019.
[56] Teo T, Hayati A, Lee G, Billinghurst M, Adcock M. A technique for mixed reality remote collaboration using 360 panoramas in 3D reconstructed scenes. In: Proceedings of the ACM symposium on virtual reality software and technology, VRST; 2019.
[57] Sasikumar P, Gao L, Bai H, Billinghurst M. Wearable remotefusion: A mixed reality remote collaboration system with local eye gaze and remote hand gesture sharing. In: Adjunct Proceedings of the 2019 IEEE international symposium on mixed and augmented reality, ismar-adjunct 2019; 2019. p. 393–4.
[58] Mahmood T, Fulmer W, Mungoli N, Huang J, Lu A. Improving information sharing and collaborative analysis for remote geospatial visualization using mixed reality. In: Proceedings - 2019 IEEE international symposium on mixed and augmented reality, ISMAR 2019; 2019. p. 236–47.
[59] Teo T, Lawrence L, Lee G, Billinghurst M, Adcock M. Mixed reality remote collaboration combining 360 video and 3D reconstruction. Conf Human Factors Comput Syst 2019.
[60] Piumsomboon T, Lee G, Irlitti A, Ens B, Thomas B, Billinghurst M. On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. Conf Human Factors Comput Syst 2019.
[61] Yoon B, Kim H-I, Lee G, Billinqhurst M, Woo W. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In: 26th IEEE conference on virtual reality and 3d user interfaces, vr 2019 - proceedings; 2019. p. 547–56.
[62] Wang P, Zhang S, Billinghurst M, Bai X, He W, Wang S, Sun M, Zhang X. A comprehensive survey of AR/MR-based co-design in manufacturing. Eng Comput 2019.
[63] Lee G, Teo T, Kim S, Billinghurst M. A user study on MR remote collaboration using live 360 video. In: Proceedings of the 2018 ieee international symposium on mixed and augmented reality, ISMAR 2018; 2019. p. 153–64.
[64] Waldow K, Fuhrmann A, Grünvogel S. Investigating the effect of embodied visualization in remote collaborative augmented reality. Lecture Notes in Comput Sci 2019;11883 LNCS:246–62.
[65] Teo T, Lee G, Billinghurst M, Adcock M. Hand gestures and visual annotation in live 360 panorama-based mixed reality remote collaboration. ACM Int Conf Proc Ser 2018;406–10.

[66] Kim S, Billinghurst M, Lee C, Lee G. Using freeze frame and visual notifications in an annotation drawing interface for remote collaboration. KSII Trans Internet Inform Syst 2018;12(12):6034–56.
[67] Congdon B, Wang T, Steed A. Merging environments for shared spaces in mixed reality. In: Proceedings of the ACM symposium on virtual reality software and technology, VRST; 2018.
[68] Yamada S, Chandrasiri N. Evaluation of hand gesture annotation in remote collaboration using augmented reality. In: 25th IEEE conference on virtual reality and 3d user interfaces, vr 2018 - proceedings; 2018. p. 727–8.
[69] Günther S, Avrahami D, Kratz S, Mühlhäuser M. Exploring audio, visual, and tactile cues for synchronous remote assistance. ACM Int Conf Proc Ser 2018;339–44.
[70] Piumsomboon T, Lee G, Hart J, Ens B, Lindeman R, Thomas B, Billinghurst M. Mini-me: An adaptive avatar for mixed reality remote collaboration. Conf Human Factors Comput Syst 2018;2018-April.
[71] Ryskeldiev B, Cohen M, Herder J. Stream space: Pervasive mixed reality telepresence for remote collaboration on mobile devices. J Inform Proc 2018;26:177–85.
[72] Hoppe A, Reeb R, van de Camp F, Stiefelhagen R. Interaction of distant and local users in a collaborative virtual environment. Lecture Notes in Comput Sci 2018;10909 LNCS:328–37.
[73] Akkil D, Isokoski P. Comparison of gaze and mouse pointers for video-based collaborative physical task. Interact Comput 2018;30(6): 524–42.
[74] Lee G, Teo T, Kim S, Billinghurst M. Mixed reality collaboration through sharing a live panorama. In: SIGGRAPH Asia 2017 Mobile Graphics and Interactive Applications. 2017.
[75] Lee G, Kim S, Lee Y, Dey A, Piumsomboon T, Norman M, Billinghurst M. Mutually shared gaze in augmented video conference. In: Adjunct proceedings of the 2017 ieee international symposium on mixed and augmented reality, ismar-adjunct 2017; 2017. p. 79–80.
[76] Komiyama R, Miyaki T, Rekimoto J. Jackin space: Designing a seamless transition between first and third person view for effective telepresence collaborations. ACM Int Conf Proc Ser 2017.
[77] Chenechal M, Duval T, Gouranton V, Royan J, Arnaldi B. Vishnu: Virtual immersive support for HelpiNg users an interaction paradigm for collaborative remote guiding in mixed reality. In: 2016 IEEE 3rd vr international workshop on collaborative virtual environments, 3dcve 2016; 2016. p. 9–12.
[78] Gurevich P, Lanir J, Cohen B. Design and implementation of TeleAdvisor: a projection-based augmented reality system for remote collaboration. Comput Support Coop Work: CSCW: An Int J 2015;24(6):527–62.
[79] Tait M, Billinghurst M. The effect of view independence in a collaborative AR system. Comput Support Coop Work: CSCW: An Int J 2015;24(6):563–89.
[80] Kim S, Lee G, Ha S, Sakata N, Billinghurst M. Automatically freezing live video for annotation during remote collaboration. Conf Human Factors Comput Syst 2015;18:1669–74.
[81] Tait M, Billinghurst M. View independence in remote collaboration using AR. In: ISMAR 2014 - IEEE international symposium on mixed and augmented reality - science and technology 2014, proceedings; 2014. p. 309–10.
[82] Kim S, Lee G, Sakata N, Billinghurst M. Improving co-presence with augmented visual communication cues for sharing experience through video conference. In: ISMAR 2014 - IEEE international symposium on mixed and augmented reality - science and technology 2014, proceedings; 2014. p. 83–92.
[83] Gauglitz S, Nuernberger B, Turk M, Höllerer T. In touch with the remote world: Remote collaboration with augmented reality drawings and virtual navigation. In: Proceedings of the ACM symposium on virtual reality software and technology, vrst; 2014. p. 197–205.
[84] Gauglitz S, Nuernberger B, Turk M, Höllerer T. World-stabilized annotations and virtual scene navigation for remote collaboration. In: UIST 2014 - Proceedings of the 27th annual acm symposium on user interface software and technology; 2014. p. 449–60.
[85] Huang W, Alem L, Tecchia F. HandsIn3D: Supporting remote guidance with immersive virtual environments. Lecture Notes in Comput Sci 2013;8117 LNCS(PART 1):70–7.
[86] Pece F, Steptoe W, Wanner F, Julier S, Weyrich T, Kautz J, Steed A. PanoInserts: Mobile spatial teleconferencing. Conf Human Factors Comput Syst 2013;1319–28.
[87] Poppe E, Brown R, Johnson D, Recker J. Preliminary evaluation of an augmented reality collaborative process modelling system. In: Proceedings of the 2012 international conference on cyberworlds; 2012. p. 77–84.
[88] Gauglitz S, Lee C, Turk M, Höllerer T. Integrating the physical environment into mobile remote collaboration. In: MobileHCI'12 - proceedings of the 14th international conference on human computer interaction with mobile devices and services; 2012. p. 241–50.
[89] Barakonyi I, Prendinger H, Schmalstieg D, Ishizuka M. Cascading hand and eye movement for augmented reality videoconferencing. In: IEEE symposium on 3D user interfaces, 3dui 2007; 2007. p. 71–78.
[90] Bannai Y, Tamaki H, Suzuki Y, Shigeno H, Okada K. A tangible user interface for remote collaboration system using mixed reality. Lecture Notes in Comput Sci 2006;4282:143–54.
[91] Regenbrecht H, Lum T, Kohler P, Ott C, Wagner M, Wilke W, Mueller E. Using augmented virtuality for remote collaboration. Presence: Teleoperat Virtual Environ 2004;13(3):338–54.
[92] Araujo RM, Santoro FM, Borges MRS. A conceptual framework for designing and conducting groupware evaluations. Int J Comput Appl Technol 2004;19(3):139–50.
[93] Pereira C, Teixeira A, e Silva MO. Live evaluation within ambient assisted living scenarios. In: Proceedings of the 7th international conference on pervasive technologies related to assistive environments, PETRA 2014; 2014.
[94] Ratcliffe J, Soave F, Bryan-Kinns N, Tokarchuk L, Farkhatdinov I. Extended reality (XR) remote research: a survey of drawbacks and opportunities. In: CHI Conference on Human Factors in Computing Systems. 2021, p. 1–13.
[95] Gaines BR. Modeling and forecasting the information sciences. Inform Sci 1991;57–58:3–22, Information Sciences-Past, Present, and Future.
[96] Lalanne D, Nigay L, Palanque p, Robinson P, Vanderdonckt J, Ladry J-Fc. Fusion engines for multimodal input: A survey. In: Proceedings of the 2009 International Conference on Multimodal Interfaces. ICMI-MLMI '09, New York, NY, USA: Association for Computing Machinery; 2009, p. 153–60.
[97] Teixeira A. A critical analysis of speech-based interaction in healthcare robots: Making a case for the increased use of speech in medical and assistive robots. In: Speech and Automata in Healthcare : Voice-Controlled Medical and Surgical Robots - Chapter 1. 2014, p. 29.
[98] Billinghurst M, Kato H. Collaborative augmented reality. Commun ACM 2003;45.
[99] Wang X, Dunston PS. Groupware concepts for augmented reality mediated human-to-human collaboration. In: Joint international conference on computing and decision making in civil and building engineering; 2006. p. 1836–42.
[100] Brockmann T, Krueger N, Stieglitz S, Bohlsen I. A framework for collaborative augmented reality applications. In: Proceedings of the nineteenth americas conference on information systems; 2013, p. 1–10.
[101] Sereno M, Wang X, Besancon L, Mcguffin MJ, Isenberg T. Collaborative work in augmented reality: A survey. IEEE Trans Vis Comput Graphics 2020;1–20.
[102] Collazos CA, Gutiérrez FL, Gallardo J, Ortega M, Fardoun HM, Molina AI. Descriptive theory of awareness for groupware development. J Ambient Intell Humaniz Comput 2019;10(12):4789–818.
[103] Talkad Sukumar P, Avellino I, Remy C, DeVito MA, Dillahunt TR, McGrenere J, Wilson ML. Transparency in qualitative research: Increasing fairness in the chi review process. In: Extended abstracts of the 2020 chi conference on human factors in computing systems; 2020. p. 1–6.
[104] Meyer M, Dykes J. Criteria for rigor in visualization design study. IEEE Trans Vis Comput Graphics 2019;26(1):87–97.
[105] Augstein M, Neumayr T. A human-centered taxonomy of interaction modalities and devices. Interact Comput 2019;31:27–58.
[106] Nickerson RC, Varshney U, Muntermann J. A method for taxonomy development and its application in information systems. Eur J Inform Syst 2013;22:336–59.
[107] Teruel MA, Navarro E, López-Jaquero V, Montero F, González P. A comprehensive framework for modeling requirements of CSCW systems. J Softw: Evolu Process 2017;29(5):e1858.
[108] Zollmann S, Grasset R, Langlotz T, Lo WH, Mori S, Regenbrecht H. Visualization techniques in augmented reality: A taxonomy, methods and patterns. IEEE Trans Vis Comput Graphics 2020;1–20.
[109] Chandrasekaran B, Josephson JR, Benjamins VR. What are ontologies, and why do we need them? IEEE Intell Syst Appl 1999;14(1):20–6.
[110] Noy NF, McGuinness DL. Ontology development 101: A guide to creating your first ontology. Stanford Knowl Syst Lab Tech Rep 2001;15(2):1–25.
[111] Herskovic V, Pino JA, Ochoa SF, Antunes P. Evaluation methods for groupware systems. In: Haake JM, Ochoa SF, Cechich A, editors. Groupware: Design, Implementation, and Use. Springer, Berlin, Heidelberg; 2007, p. 328–36.
[112] de Araujo RM, Santoro FM, Borges MRS. The CSCW lab ontology for groupware evaluation. In: 8th international conference on computer supported cooperative work in design, Vol. 2; 200. p. 148–53.
[113] Pereira C, Almeida N, Martins AI, Silva S, Rosa AF, Oliveira e Silva M, Teixeira A. Evaluation of complex distributed multimodal applications: Evaluating a TeleRehabilitation system when it really matters. Lecture Notes in Comput Sci (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2015;146–57.
[114] Pereira C, Teixeira A, Oliveira e Silva M. Dynamic evaluation for reactive scenarios. (Ph.D. thesis), University of Aveiro (MAPi); 2016, p. 1–196.