Universidade de Aveiro
2021

Bernardo José
Santos Marques

Conceitos e Métodos para apoio ao Desenvolvimento e Avaliação de Colaboração Remota utilizando Realidade Aumentada

Concepts and Methods to support the Development and Evaluation of Remote Collaboration using Augmented Reality

Universidade de Aveiro
2021

Bernardo José Santos Marques

Conceitos e Métodos para apoio ao Desenvolvimento e Avaliação de Colaboração Remota utilizando Realidade Aumentada

Concepts and Methods to support the Development and Evaluation of Remote Collaboration using Augmented Reality

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Informática, realizada sob a orientação científica do Doutor Paulo Miguel de Jesus Dias, Professor auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro e da Doutora Maria Beatriz Alves de Sousa Santos, Professora associada com agregação do Departamento de Eletrónica Telecomunicações e Informática da Universidade de Aveiro.

The research presented in this thesis was developed in the scope of the PhD grant [SFRH/BD/143276/2019], funded by the FCT - Foundation for Science and Technology (FCT) and the European Social Fund (FSE) under the III community support framework.

To my parents, Júlia and Francisco.

o júri / the jury

presidente / president	Professor Doutor João Carlos Matias Celestino Gomes da Rocha, Professor Catedrático do Departamento de Química da Universidade de Aveiro
vogais / examiners committee	Professor Doutor Maximino Esteves Correia Bessa, Professor Associado com Agregação do Departamento de Engenharias da Universidade de Trás-os-Montes e Alto Douro
	Professor Doutor António Fernando Vasconcelos Cunha Castro Coelho, Professor Associado com Agregação da Faculdade de Engenharia da Universidade do Porto
	Professora Doutora Maria Beatriz Duarte Pereira do Carmo, Professora Auxiliar da Faculdade de Ciências da Universidade de Lisboa
	Professor Doutor Joaquim João Estrela Ribeiro Silvestre Madeira, Professor Auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro
	Professor Doutor Paulo Miguel de Jesus Dias (*Orientador*), Professor Auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro

acknowledgements

I take this opportunity to acknowledge the essential support of all individuals, research projects and institutions that were involved directly and indirectly in this thesis. To all, I express my sincere gratitude. It was a great pleasure meet and work together with all of you. However, I would like to give some special thanks.

First, to my mentors, Professor Paulo Dias and Professor Beatriz Sousa Santos for accepting the challenge of working on the subject of remote collaboration and for their human characteristics. Moreover, for their availability and scientific guidance. Also, for including me in every new opportunity and challenge our research group faced, e.g., participating in teaching the classes of IHC - *Human-Computer Interaction*, as well as RVA - *Virtual and Augmented Reality,* advising students of master's dissertations and projects, help creating new research projects, and the inclusion to review conference publications and books. Furthermore, for encouraging me to be a Student Volunteer and for supporting my travels to international conferences, and my tuition fees before I obtained an FCT grant. Also, for the patience and trust in my judgment over these years ("*deixo ao critério do candidato*").

Second, to Professors António Teixeira and Samuel Silva for their outstanding support, for caring, motivating and sharing their knowhow with me. I was able to learn so much about research projects, how to question everything and always aim to achieve more. Your enthusiasm and passion are second to none. The two of you complement each other mindsets. You are capable of quickly pick up any subjects, while generating the most amazing ideas to organize and improve them. You really inspire me.

Third, to the amazing colleagues from Lab 1.10 that helped creating an extraordinary working environment: Mário Rodrigues, Catarina Oliveira, Nuno Almeida, Ana Martins, Ana Rocha, Ana Rita Valente, Luciana Pereira, Fábio Barros, David Ferreira, Diogo Cunha, Tiago Henriques and Maksym Ketsmur, to whom I built amazing relationships over the years. Your presence was essential through this journey. Thank you all for integrating me in your core, for supporting my sense of humor, my constant jokes and for participating in all diverse team-building events.

To Professor Carlos Ferreira for sharing his experience, words of wisdom and for his constant availability and contribution during numerous case studies and brainstorm sessions.

Moreover, thanks to all students I had the opportunity to collaborate with over the years, in particular João Alves (*you still owe me a chocolate croissant*), Tiago Araújo (‘*vai dar certo’*), Nuno Martins, Tiago Madeira, Miguel Neves, Hélder Serra, João Barroso, Fábio Ferreira, Inês Cruz, Renato Cabral, Anna Andreikanich, Rafael Esteves, Raphael Carvalho and all PEI students I helped supervise.

To everyone from the Institute of Electronics and Informatics Engineering of Aveiro (IEETA), funded by National Funds through the FCT in the context of the project [UIDB/00127/2020] who supported the research presented in this document and allowed my participation in interesting conferences and workshops. In particular, Armando Pinho, André Zúquete, Nuno Lau, as well as Sónia Gouveia, Rui Lebre, Anabela Viegas, Isabel Martins, Sónia Brandão, Sofia Ribeiro, and ‘Senhor Zé’.

Equally important, to Professor José Maria Fernandes, Doctor Susana Brás, Professor Ilídio Oliveira and Doctor Raquel Sebastião, who introduced me to the meaning of ‘*doing research*’. Although our paths have diverged, I owe you a special gratitude for encouraging me to pursue a Ph.D. program. Thanks for being such staggering human beings over the years. I really admire all of you.

Additionally, to Professor Paulo Jorge Ferreira who supports me in every way since my Master’s Degree. It has been a pleasure learning from you over the years. Thanks for all your advice and for making our university a better place for the academic community.

To all domain experts and participants involved in brainstorming sessions and case studies, thanks for your time and expertise. In particular, António Rocha, Pedro Cruz, Emanuel Fonseca and António Valente who helped keep our connection to the industry sector through the Smart Green Homes (SGH) Project [POCI-01-0247-FEDER-007678], a co-promotion between the University of Aveiro and Bosch Termotecnologia S.A., financed by Portugal 2020 under the Competitiveness and Internationalization Operational Program and by the European Regional Development Fund.

To my everlasting friends Michael and Diana. Our friendship continues to flourish over the years, now, with our little Tomás. Dear godson, one day I will tell you all about how your parents met and how lucky you are to have them by your side. Treasure them as you grow, because they will fight every day to give you every opportunity in the world. I love you all, you are like family to me. Thanks for being always present and for giving me the gift of being a Godfather. Also, for making me smile so big, laugh so hard, and always dream of how amazing life can be.

In addition, to Andreia Santos who has been by my side, making my days more colorful during this astonishing endeavor. Thanks for your dedication, concern and support through the years. I really appreciate having you as the designated remote expert during my research. It was an amazing experience being able to share my path with you. Nothing would be the same without your gorgeous smile and your curious big blue eyes to elevate my spirit.

Finally, I thank the most extraordinary human beings I know, my parents Júlia and Francisco, for being present in all moments of my life, for their unconditional love and support. I'm grateful for the constant reminder of the rewards that come from pursuing our goals. Also, for all the dedication and hard work put into my education, both as an individual with strong values and in terms of academic level. You manage to build everything we have from scratch, and you keep fighting every day for a better future. You sacrificed everything so that I could achieve my dreams. What a great gift is having such a wonderful family with me every day. You're my role models. I love you both very much.

' *After all, no regrets!* '

Surround yourself with people who challenge you,
teach you and push you to be your best.
Bill Gates

palavras-chave

Colaboração Remota, Realidade Aumentada, Trabalho Cooperativo Apoiado por Computador, Manutenção, Processo Colaborativo, Avaliação, Caracterização.

resumo

Colaboração Remota utilizando Realidade Aumentada (RA) apresenta um enorme potencial para estabelecer um entendimento comum em cenários onde membros de uma equipa fisicamente distribuídos precisam de atingir um objetivo comum. No entanto, a maioria dos esforços de investigação tem-se focado nos aspetos tecnológicos, em fazer experiências e propor métodos para apoiar seu desenvolvimento. À medida que a área evolui, a avaliação e caracterização do processo colaborativo tornam-se um esforço essencial, mas difícil, para compreender as contribuições da RA.

Nesta dissertação, realizámos uma análise crítica para identificar as principais limitações e oportunidades da área, ao mesmo tempo em que situámos a sua maturidade e propomos um mapa com direções de investigação importantes. De seguida, foi adotada uma metodologia de Design Centrado no Humano, envolvendo parceiros industriais de forma a compreender como a RA poderia responder às suas necessidades em manutenção remota. Estes resultados foram combinados com métodos da literatura num protótipo de RA e a sua avaliação foi realizada com um caso de estudo. Ficou então clara a necessidade de realizar uma reflexão profunda para melhor compreender as dimensões que influenciam e devem ser consideradas na RA Colaborativa. Foram então propostos um modelo conceptual e uma taxonomia centrada no ser humano para promover a sistematização de perspetivas. Com base no modelo proposto, foi desenvolvido um *framework* de avaliação para recolha e análise de dados contextualizados, permitindo apoiar o desenho e a realização de avaliações distribuídas de forma mais informada e completa. Para instanciar esta visão, o *CAPTURE toolkit* foi criado, fornecendo uma perspetiva adicional com base em dimensões de colaboração e medidas predefinidas para obter dados *in situ*, que podem ser analisados utilizando o painel de visualização integrado. O *toolkit* permitiu avaliar com sucesso vários colaboradores durante a realização de tarefas de manutenção remota apoiada por RA, permitindo mostrar a sua versatilidade e potencial em obter uma caracterização abrangente do valor acrescentado da RA em situações da vida real. Sendo assim, estabelece-se como uma solução genérica, potencialmente aplicável a uma gama diversificada de cenários colaborativos.

keywords

Remote Collaboration, Augmented Reality, Computer Supported Cooperative Work, Maintenance Context, Collaborative Process, Evaluation, Characterization.

abstract

Remote Collaboration using Augmented Reality (AR) shows great potential to establish a common ground in physically distributed scenarios where team-members need to achieve a shared goal. However, most research efforts in this field have been devoted to experiment with the enabling technology and propose methods to support its development. As the field evolves, evaluation and characterization of the collaborative process become an essential, but difficult endeavor, to better understand the contributions of AR.

In this thesis, we conducted a critical analysis to identify the main limitations and opportunities of the field, while situating its maturity and proposing a roadmap of important research actions. Next, a human-centered design methodology was adopted, involving industrial partners to probe how AR could support their needs during remote maintenance. These outcomes were combined with literature methods into an AR-prototype and its evaluation was performed through a user study. From this, it became clear the necessity to perform a deep reflection in order to better understand the dimensions that influence and must/should be considered in Collaborative AR. Hence, a conceptual model and a human-centered taxonomy were proposed to foster systematization of perspectives. Based on the model proposed, an evaluation framework for contextualized data gathering and analysis was developed, allowing support the design and performance of distributed evaluations in a more informed and complete manner. To instantiate this vision, the CAPTURE toolkit was created, providing an additional perspective based on selected dimensions of collaboration and pre-defined measurements to obtain “in situ” data about them, which can be analyzed using an integrated visualization dashboard. The toolkit successfully supported evaluations of several team-members during tasks of remote maintenance mediated by AR. Thus, showing its versatility and potential in eliciting a comprehensive characterization of the added value of AR in real-life situations, establishing itself as a general-purpose solution, potentially applicable to a wider range of collaborative scenarios.

I. Contents

II. List of Figures

III. List of Tables

IV. List of Acronyms

2D – Two Dimensional

3D – Three Dimensional

AR – Augmented Reality

BARRT – collaBorative Augmented Reality for Remote supporT

CAPTURE – Contextual dAta Platform for remoTe aUgmented Reality Evaluation

CSCW – Computer-Supported Cooperative Work

ET – Exposure Therapy

GS – Google Scholar

HCI – Human-Computer Interaction

HCD – Human-Centered Design

HHD – Handheld Displays

HMD – Head Mounted Displays

HUD – Head Up Display

IV – Information Visualization

MMI – Multi Modal Interaction

MR – Mixed Reality

R&D – Research and Development

UN – United Nations

VR – Virtual Reality

WoS – Web of Science

1 Introduction

Coming together is the beginning.
Keeping together is progress.
Working together is success.

Henry Ford

Collaboration can be described as "*the process of joint and interdependent activities between co-located or remote collaborators, performed to achieve a common goal*" (S. Kim, Billinghurst, & Lee, 2018; S. Kim, Billinghurst, et al., 2020; S. Kim, Lee, et al., 2020). Collaboration is essential in many situations, as is the case of industrial, medical, and educational domains, among others (Johson et al., 2015; Lukosch et al., 2015; Schneider et al., 2017; P. Wang et al., 2021).

Collaborative scenarios mediated by technology have evolved from simple co-located activities to more complex use cases of remote collaboration, involving several team-members with different experiences, expertise and multidisciplinary backgrounds, distributed geographically around the world. Therefore, the design and development of collaborative tools 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 (Arias et al., 2000; Lukosch et al., 2015; Schneider et al., 2017).

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 among distributed collaborators (Grudin & Poltrock, 2011, 1997; Ishii et al., 1994).

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 (Billinghurst et al., 2015; Ens et al., 2019; Grudin & Poltrock, 2011; Jalo et al., 2018).

Remote AR-based solutions are well suited for overlying responsive computer-generated information on top of the real-world environment combining the advantages of virtual environments and the possibility for seamless interaction with the real-world objects and other collaborators (Altug & Mahdy, 2016; Bottani & Vignali, 2019; K. Kim et al., 2018; P. Wang et al., 2021). Collaborative AR helps team-members establish a shared understanding, similar to their perception of the physical space, serving as basis for situation mapping, transmission of ideas, identification of issues, thus making assumptions and beliefs visible (Hall et al., 2018; K. Kim et al., 2018; Ong et al., 2008; X. Wang et al., 2016).

In this section, we outline the main challenges associated to the topic of interest of this thesis, Remote Collaboration mediated by AR, to substantiate the motivation of our research. Subsequently, the thesis statement and main research objectives are introduced. Then, the main contributions and published results are summarized. Last, the outline is presented.

1.1 Context

The research conducted aligns with the field of CSCW, an area in the study of Human-Computer Interaction (HCI) that examines "how people work together in groups and how groupware technologies can support collaboration" (Grudin & Poltrock, 2011, 1997; Ishii et al., 1994).

Besides, the topic of this thesis is associated with the Ph.D. Candidate own experience and motivations. In parallel with his master's degree, He worked as a part-time electrician for an Industrial automation company at Renault CACIA, S.A., Portugal. Throughout his daily tasks, He was responsible for technical tasks regarding automation, networks and electrical equipment, being confronted with problems which sometimes required assistance from experts. Nevertheless, such help was not always available, due to localization or communication constrains.

In addition, along this thesis, we have also been involved in a research project addressing challenges related to the use of AR in different areas of application. The Smart Green Homes project[1] [POCI-01-0247-FEDER-007678], a co-promotion between Bosch Termotecnologia S.A. and the University of Aveiro focused, among other subjects, on remote maintenance through the use of shared AR-based annotations, allowing a remote expert to assist an on-site technician while performing tasks. The participation in this project enabled us to experiment several technologies, while collaborating with experts from distinct areas of application, which helped identify challenges and provided the context for the user studies described in this thesis.

[1] - ua.pt/pt/smartgreenhomes/ [Accessed: 31-Mar-2021]

1.2 Motivation and Challenges

The use of AR-based solutions to assist in scenarios of remote collaboration is object of interest by many businesses worldwide (de Belen et al., 2019; Ens et al., 2019; K. Kim et al., 2018; S. Kim, Billinghurst, et al., 2020). It elicits more efficient collaboration, improves knowledge transfer and may minimize by 50% the need for expertise individuals to travel abroad. For example, Scope AR[2] reports the use of such solutions can contribute to boost response time above 90%, reduce errors, downtime by 50% and training time up to 85%, resulting in considerable savings (projected overall service expense savings up to 2.4M$) for the businesses that profit from the advantages of such solutions (additional details can be found in section 2.3).

Several studies described in literature have focused on the creation of AR-based prototypes using virtual annotations to augment the physical environment on top of images or live video scenes, such as drawings, pointers, gaze, gestures or arrows (Choi et al., 2018; S. Kim, Billinghurst, Lee, et al., 2018; Lukosch et al., 2015). As an alternative, some recent studies start to explore the use of virtual replicas (Barroso et al., 2020; Elvezio et al., 2017), as well as reconstructions of the physical environment (H. Bai et al., 2020; Zillner et al., 2018), although these required the existence of 3D models and additional hardware, which may limit their adoption in some scenarios of application. By creating a common ground environment, these technological approaches can enhance alertness, awareness, and understanding of the situation, allowing interaction between team-members at geographically dispersed locations (Neale et al., 2004).

Nevertheless, although some AR-based prototypes already exist, most of the research efforts, so far, have been devoted to explore and evolve the enabling technology, as well as propose novel methods to support its design and development (Ens et al., 2019; Marques, Teixeira, et al., 2020; Merino et al., 2020), i.e., create small proofs of concept to support remote collaboration based on exploring hardware devices available on the market. However, as aforementioned, these types of distributed activities imply substantial levels of complexity, which most proof of concepts still fall short to address. With the growing number of prototypes, the only way to have usable, realistic and impactful solutions is to focus on the nuances of supporting the collaborative effort, i.e., comprehend how collaboration occurs through this new medium and how AR may help achieve more effective collaboration in such scenarios. To achieve this, the characterization and evaluation of the collaborative process become an essential but difficult endeavor (Antunes et al., 2014; Ens et al., 2019; Hamadache & Lancieri, 2009; Marques, Teixeira, et al., 2020; Merino et al., 2020). These are essential to ensure the reporting process integrate the context in which the collaborative effort took place, thus allowing to create a better understanding of the information being provided, i.e., the impact of the various variables being considered, leading to the development of better AR-based tools capable of improve the collaborative work effort.

[2] - scopear.com [Accessed: 31-Mar-2021]

Nonetheless, given the focus on design and development, even when evaluation is performed, it is frequently done using single-user methods (R. Belen et al., 2019; Dey, Billinghurst, Lindeman, & Ii, 2018; Ens et al., 2019; Marques et al., 2021; Marques, Silva, Dias, et al., 2021a; Marques, Teixeira, et al., 2020), focusing only on the performance of one collaborator, i.e., on-site, or remote, on the technology being used or in quantifying task effectiveness, while important dimensions of the collaborative process are ignored. This means evaluation usually does not include interaction, and communication among the team members, and is not conducted in distributed scenarios, as should be the case to establish experimental conditions closer to real scenarios, a possible reason for the lack of contextual information on the experimental setup (Billinghurst et al., 2015; Dey, Billinghurst, Lindeman, & Swan, 2018; S. Kim, Billinghurst, & Lee, 2018).

Remote collaboration scenarios are intrinsically multifaceted: many aspects may affect the way teams collaborate. In addition, current frameworks are not adapted to characterize how collaboration occurs (R. Belen et al., 2019; Ens et al., 2019; Marques, Silva, Dias, et al., 2021c; Marques, Teixeira, et al., 2020; Merino et al., 2020; Ratcliffe et al., 2021), falling short to retrieve the necessary data for more comprehensive analysis, thus requiring an eclectic perspective.

As such, it is possible to identify the following main challenges that must be addressed regarding AR-supported remote collaboration:

- **Move beyond single-user evaluations**: The logistics associated with carrying out evaluation in remote scenarios is demanding. 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;
- **Contextualize what is happening:** Without contextual information, it becomes difficult to assess the variables that influence the collaborative process. It is important to construct evaluation methodologies that address/measure all parts involved in the collaborative process, allowing researchers to better characterize the collaborative context as a whole;
- **Improve existing frameworks:** Current frameworks are not sufficiently well suited to characterize how collaboration using AR occurs. Thus, the integration of data gathering, visualization and analysis tools is of paramount importance to explore the plethora of resulting data, foster insights, and help make sense of use patterns and specific events;
- **Select complex/adequate collaborative tasks:** Current approaches often rely on simple tasks, which minimize the need for proper collaboration between distributed team-members. We argue that collaborative tasks must be complex enough in difficulty and duration to encourage a real interaction between collaborators.

1.3 Thesis Statement and Objectives

Considering the challenges associated with evaluating AR-supported remote collaboration, and how the state-of-the-art is still lacking proper answers, we consider that a novel approach is essential, since evaluation needs to consider numerous aspects that may influence the collaborative process, many of those not only concerning AR, but also, e.g., the nature of the task and its context. To contribute to a more in-depth knowledge, it is paramount to understand where the field stands and how well it can address collaborative work using AR. In this context, we present the **research question** we aim to tackle in this thesis:

> *How can we improve the characterization and evaluation of the collaborative process in remote scenarios to better understand the contributions of AR to the work effort?*

In this regard, trying to apply conventional evaluation techniques to scenarios of remote collaboration without adapting them can lead to dubious results that may be misleading or of limited value. As such, without the appropriate probing mechanisms, the research community might not accumulate enough knowhow and experience to build better solutions and improve AR-based collaboration among distributed team-members. A better characterization of the collaborative process can lead to an additional perspective on the nuances of collaboration, and in turn, provide researchers with the possibility to determine the reasons for the success (or failure) of the collaborative effort, which may result in localized improvements in the AR-based solutions. Thus, we present the **thesis statement** of this research:

> *Eliciting more comprehensive characterizations of the collaborative process in scenarios of remote collaboration mediated by AR can be achieved through holistic strategies and more structured evaluation methods by exploiting contextualized approaches like improved data gathering regarding characteristics that may directly impact the collaboration outcomes.*

In this vein, the main objective of this thesis is to provide researchers with methods and tools to obtain an additional perspective on the added value AR may bring to the collaborative effort when compared to existing tools, for example based on video chat, through more structured evaluations within scenarios of remote collaboration. To achieve this, it is important to create a common ground for systematization based on the proposal of dimensions of collaboration, conceptual models and taxonomies. It is also important to design and develop data gathering frameworks, allowing researchers to analyze the implications of an improved characterization of the collaborative process, as well as extract better conclusions.

To pursue this objectives, the work carried out on this thesis adopted an **Engineering Design Methodology** (Dieter & Schmidt, 2013), which can be summarized in the following stages: 1) after defining the problem, the work starts by researching, observing and analysing current state of the art solutions, followed by the specification of the requirements; 2) conceptualization of a general prototype comprising paradigm, methodology, model, and architecture; 3) implementation of its core components; 4) development, deployment and test of a proof of concept. This process is repeated until a good degree of satisfaction is attained based on the extent to which the proposed solutions address the identified objectives.

To finish, the research described in this document may contribute to several *sustainability and development goals* proposed by the United Nations (UN)[3], such as:

- **Goal 8** - Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all;
- **Goal 9** - Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation;
- **Goal 10** - Reduce inequality within and among countries;
- **Goal 11** - Make cities and human settlements inclusive, safe, resilient and sustainable;
- **Goal 12** - Ensure sustainable consumption and production patterns;

1.4 Main Contributions and Published results

The relevance of the research described in this thesis to the state of the art, led to 13 scientific publications in the field of CSCW and Collaborative AR accepted in peer-reviewed journals, and conferences. In the following paragraphs a brief overview of some selected publications, deemed representative of the main contributions is presented.

Analysing the maturity associated with remote collaboration using AR and how it addressed the characterization of the collaborative process was essential to define a roadmap of important research actions and guide the research conducted in this thesis. The main outcomes, relevant for this context, were published in:

1. **Marques, B.**, Teixeira, A., Silva, S., Alves, J., Dias, P. & Santos, B. S. A critical analysis on remote collaboration mediated by Augmented Reality: making a case for improved characterization and evaluation of the collaborative process. In Computer & Graphics, 1-17, 2021.

   Main topic: discusses the maturity of remote collaboration mediated by AR through a critical analysis of current literature and proposes a road map of important research actions, as presented in Chapter 2.

[3] - sdgs.un.org/goals [Accessed: 21- Jul-2021]

One important objective of this thesis was understanding how AR could assist the collaborative effort in scenarios of remote collaboration. Also, the creation of prototypes for exploring different AR-based approaches to address requirements obtained from a user-centered methodology with partners from the industry sector and conducting initial evaluations of the collaborative process. In this vein, several works were published.

2. **Marques, B.,** Silva, S., Alves, J., Rocha, A., Dias, P., & Santos, B. S.. Remote Collaboration in Maintenance Contexts using Augmented Reality: Insights from a Participatory Process. In International Journal on Interactive Design and Manufacturing IJIDeM*,*1-21, 2021.

   Main topic: probes how AR can support remote maintenance through a user-centered design methodology with domain experts from the industry sector. Also, identifies a set of requirements and materializes them through the design and creation of a collaborative prototype based on AR annotations. Additionally, presents and discusses insights from a case study on a maintenance context, as described in Chapter 3.

3. **Marques, B.,** Silva, S., Rocha, A., Dias, P. & Santos, B. S.. Remote Asynchronous Collaboration in Maintenance scenarios using Augmented Reality and Annotations. In IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE VR, 567-568, 2021.

   Main topic: presents a pilot user study between distributed team-members to evaluate asynchronous remote collaboration while using AR and annotations, as illustrated in Chapter 3.

4. **Marques, B.,** Silva, S., Rocha, A., Santos, A., Ferreira, C., Dias, P., & Santos, B. S.. Exploring Remote Augmented Reality as a 2D Authoring Tool for Creation of Guiding Instruction. In International Conference on Graphics and Interaction, ICGI, 2021.

   Main topic: describes the use of an AR-based authoring tool to create instructions for aiding in scenarios of remote maintenance, as reported in Chapter 3.

5. Madeira, T., **Marques, B.,** Alves, J., Dias, P. & Santos, B. S.. Exploring annotations and hand tracking in Augmented Reality for remote collaboration. In Advances in Intelligent Systems and Computing, 83-89, 2020.

   Main topic: describes how different types of AR-based annotations can be controlled through hand-tracking to assist on-site collaborators during remote assistance, as presented in Chapter 3.

6. **Marques, B.,** Dias, P., Rocha, A., Alves, J. & Santos, B. S.. An Augmented Reality Framework for Supporting Technicians during Maintenance Procedures. In International Conference on Graphics and Interaction, ICGI, 2019.

   Main topic: introduces an initial AR-based prototype to assist on-site technicians during maintenance procedures, and discusses its potential for remote scenarios, which was used later in a focus group to elicit discussion regarding the role of AR in such scenarios, as reported in Chapter 3.

When addressing systematization of perspectives for the field of Collaborative AR, identification of dimensions of collaboration, the proposal of a conceptual model and a human-centered taxonomy are important topics, enabling harmonization of perspectives for the field, creating a common ground for understanding and discussion. In this regard, we highlight:

7. **Marques, B.**, Silva, S., Alves, J., Araújo, T. Dias, P. & Santos, B. S. A Conceptual Model and Taxonomy for Collaborative Augmented Reality. In IEEE Transactions on Visualization and Computer Graphics, TVCG, 1-21, 2021.

   Main topic: performs an analysis of the different dimensions that should be taken into account when analysing the contributions of AR to the collaborative work effort. Moreover, it proposes a conceptual model and a human-centered taxonomy for Collaborative AR, as described in Chapter 4.

8. **Marques, B.**, Araújo, T. Silva, S., Alves, J., Dias, P. & Santos, B. S.. Visually exploring a Collaborative Augmented Reality Taxonomy. In International Conference on Information Visualization, IV 1–6, 2021.

   Main topic: presents a first effort towards the creation of an interactive visualization tool for exploration and analysis of collaborative AR research, based on a set of dimensions of collaboration, as illustrated in Chapter 4.

Evaluation in the context of remote collaboration mediated by AR was one key topic to be tackled in this thesis, in order to support researchers in conducting evaluations in a more structured manner and in turn elicit a more comprehensive characterization of the collaborative process. The different contributions accomplishing this point were published in:

9. **Marques, B.**, Teixeira, A., Silva, S., Dias, P. & Santos, B. S. A vision for contextualized evaluation of remote collaboration supported by AR. In Computer & Graphics, 1-13, 2021.

   Main topic: outlines a conceptual framework and the CAPTURE evaluation toolkit to help researchers conduct evaluations in a more structured manner in remote scenarios mediated by AR. It also reports a user study comparing two distinct tools instrumented with the toolkit and demonstrates its usefulness/versatility, as described in Chapter 5.

10. **Marques, B.**, Silva, S., Dias, P. & Santos, B. S.. A Toolkit to Facilitate Evaluation and Characterization of the Collaborative Process in Scenarios of Remote Assistance Supported by AR. In IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 336-337, 2021.

    Main topic: presents a toolkit to instrument AR-based tools via visual editors, enabling rapid data collection and filtering during distributed evaluations, as reported in Chapter 5.

11. **Marques, B.**, Silva, S., Dias, P. & Santos, B. S. Evaluating Augmented Reality based Remote Collaboration: a contextualized approach. In Human-Automation Interaction: Mobile Computing, Springer International Publishing, 1-14, 2021.

    Main topic: describes a novel conceptual framework to support researchers performing contextualized evaluations regarding scenarios of remote collaboration mediated by AR, as presented in Chapter 5.

12. **Marques, B.**, Silva, S., Dias, P. & Santos, B. S.. An Ontology for Evaluation of Remote Collaboration using Augmented Reality. In European Conference on Computer-Supported Cooperative Work, ECSCW, 1-8, 2021.

    Main topic: proposes an ontology describing relations among dimensions of collaboration and the main concepts of the evaluation process to guide researchers in designing and conducting their evaluations, as reported in Chapter 5.

13. **Marques, B.**, Teixeira, A., Silva, S., Alves, J., Dias, P. & Santos, B. S.. A Conceptual Model for Data Collection and Analysis for AR-based Remote Collaboration Evaluation. In IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 1-2, 2020.

    Main topic: details a conceptual model for data collection and analysis to support evaluation of the collaborative process in scenarios of remote collaboration mediated by AR, as presented in Chapter 5.

Additional contributions were also made in other research areas that although not directly associated with the thesis objectives. These were fundamental to its success, as they contributed to elicit discussion and decisions taken during this research. From these, we highlight the following publications: in Remote Collaboration using a 3D shared Model approach (Barroso et al., 2020; Marques et al., 2021); in Co-located and Remote Interaction (Alves et al., 2018; Alves, Marques, Neves, et al., 2019); in manipulation of AR content (Marques, Alves, et al., 2020; Marques et al., 2019); in Industrial AR (Alves et al., 2020; Alves, Marques, Ferreira, et al., 2021; Alves, Marques, Dias, et al., 2021; Alves, Marques, Oliveira, et al., 2019; Marques et al., 2018).

Finally, it is important to note that the development of this thesis was initiated within the context of the research line LD5 of the Smart Green Homes project, completed with a high level of success, to which the work described in this thesis contributed to multiple reports, posters, videos, and live presentations in multiple public events through the years, which were continuously critically analyzed by a committee of external evaluators from the National Innovation Agency (ANI).

1.5 Outline

This thesis is organized in six chapters. **Chapter 1** introduced the motivation, main challenges and objectives for this work. Following this, the outline of the document is structured as follows.

In **Chapter 2**, we start by defining concepts and characteristics relevant to understand and discuss the domain of Collaboration. Next, we present the state-of-the-art regarding the use of AR for remote collaboration, highlight trends, considerations and open challenges. Then, we present a critical analysis in which we discuss the maturity of the field and propose a roadmap of important research actions that may help address how to improve the characterization and evaluation of the collaboration process supported by remote AR moving forward.

Chapter 3 identifies a set of requirements and describes the rationale behind them, obtained through a user-centered approach with domain experts, to be considered when developing collaborative solutions using AR for remote scenarios. Afterwards, we describe an initial user study to evaluate collaborative aspects of a method using AR annotations. The latter also allows us to understand the difficulties associated with evaluating remote collaboration mediated by AR.

Chapter 4 describes the research methodology adopted to understand how collaborative work is accomplished and how it can affect the design of collaborative solutions using AR based on: 1) what does it take to address the question at hand, i.e., which dimensions of collaboration need to be considered, and 2) how existing research is tackling each of these dimensions. After identifying such dimensions, we bring them forward into a conceptual model and a taxonomy, thus creating a common ground for systematization and discussion.

Next, in **Chapter 5** we address the challenges in evaluating collaborative AR-solutions for remote scenarios. We propose a methodological framework and apply this vision through an evaluation toolkit, by instrumenting two different tools in a remote maintenance case study. Then, we discuss the attained results of a contextualized data gathering.

Lastly, in **Chapter 6** we present our conclusions to this thesis. We discuss the main results of the research conducted and point out possible directions for future work.

2 Related Work and Critical Analysis on AR-supported Remote Collaboration

Challenges is what makes life interesting and
overcoming them is what makes life meaningful.

Joshua J. Marine

In this chapter, the main concepts and characteristics associated with Collaboration are explained. Next, the concept of Augmented Reality (AR) is also introduced including general areas of application and in particular the role of AR for supporting scenarios of remote collaboration. Also, important trends and open challenges for AR-supported remote collaboration are summarized. After a general introduction to the main topics of interest to the thesis, remote collaboration is analyzed through a systematic review on how characterization and evaluation of the collaborative process has been conducted. A literature review from 2000 to 2020 was performed on collaborative studies to provide a high-level overview of the field and identify strengths and weaknesses of existing methods. Based on the analysis, we describe the main challenges regarding the evaluation of 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 important aspects to support AR-based remote collaboration.

2.1 Collaboration and Computer Supported Cooperative Work

Several activities in life cannot be solely resolved by a single individual due to the rapidly growing in terms of scale, complexity and interdisciplinarity, prompting the request for collaborative assistance towards common goals from other individuals with a more comprehensive or technical expertise.

The term *collaboration* has received multiple definitions in the literature, being often considered a synonym to the term *cooperation* (Gervasi et al., 2020). However, it is important to emphasize that these terms may have different meanings as highlighted by the work of (Kozar, 2010), which summarizes the definitions given by different authors. Cooperation can be defined as "*working together to accomplish shared goals*" according to (K. A. Smith, 1995), while collaboration can be described as "*working in a group of two or more to achieve a common goal, while respecting each individual's contribution to the whole*" following the definition by (McInnerney & Roberts, 2004). Besides, cooperative work can be defined as a task that is accomplished by dividing it among collaborators, where "*each person is responsible for a portion of the problem*" and collaborative work as "*the mutual engagement of participants in a coordinated effort to solve the problem together*" as defined by (Roschelle & Teasley, 1995).

These definitions show that cooperation focuses on collaborators working together to achieve a final goal that can be obtained even if all team-members do their assigned tasks separately and join their results in the end. Collaboration requires knowledge sharing, implying direct interaction among individuals, i.e., negotiation, discussion, and consideration of other perspectives (Kozar, 2010). Hence, collaboration can be considered a more complex form of interaction when compared to cooperation, since it requires the fulfilment of additional conditions (Gervasi et al., 2020).

Collaborative actions may occur in different conditions: co-located, i.e., in the same place (face-to-face collaboration), or remote collaboration: between distributed team-members (Kim, Billinghurst, et al. 2020). This activity is essential to combine different and sometimes opposing points of view together in order to establish a shared understanding towards common goals among different stakeholders, which can lead to new insights, innovative ideas, and interesting artefacts (R. A. J. de Belen et al., 2019; Ludwig et al., 2021). In turn, these advantages may generate increased profit, cost reduction and improved decision making (Patel et al., 2012).

The field of Computer-Supported Cooperative Work (CSCW) has long been concerned with understanding and designing solutions to support collaboration, resulting in theories that have influenced the design of collaborative technologies. Groupware tools have been created to support interaction between multiple collaborators (Ens et al., 2019; Grudin & Poltrock, 2011; X. Wang & Dunston, 2006), which can be defined as: "*computer systems that support groups of people engaged in a common task (or goal) and that provide an interface to a shared environment*" (Ellis et al., 1991). The goal of such tools is to assist a team of individuals in communicating, collaborating, and coordinating their activities towards a common purpose.

Supporting CSCW advocates a user-centered design philosophy, in which the development and evolution of interactive systems is tightly coupled with systematic evaluation and user studies. Besides, a categorization for CSCW systems based on time and spatial location was proposed (Johansen, 1988). According to the time-space matrix (Figure 1), CSCW systems can be organized depending on when collaborators work together at the same or different times (synchronous vs. asynchronous collaboration) and the physical arrangement of the workplace (at the same location or in different places). The time-space matrix organization is still a cornerstone of the categorization of software tools for the collaborative activity (Ens et al., 2019). Moreover, two essential elements that CSCW solutions need to address to support collaborative work have emerged: 1) enable mutual awareness about the collaborators in the workspace, as well as about the tasks they are performing; and 2) understand and articulate how information is used to support the collaborative effort (Ens et al., 2019; Nguyen & Duval, 2015).

	Synchronous (same time)	Asynchronous (different time)
Co-located (same place)	Face-to-face interactions	Continuous task
Remote (different place)	Remote interactions	Communication & Coordination

Figure 1 – CSCW Time-space matrix. Adapted from (Johansen, 1988).

The first studies associated to collaboration focused on conversation and information exchange between two co-located workers working on an assembly task. Collaborators started by exchanging information related to objects that needed to be handled, as well as instructions for performing specific instructions and confirmation of achieved goals. Later, further observation studies delve into the use of sketches and writing, showing that such approach helped to support the collaborative process while improving information exchange (Aschenbrenner et al., 2018).

Moving forward into remote settings, there are five essential factors (Figure 2) that must be considered (S. Kim, Billinghurst, et al., 2020):

- **Task:** a multi-person activity towards a shared goal that collaborators try to complete together. It often involves manipulating physical objects located in the on-site collaborator environment (S. Kim et al., 2014).
- **On-site Collaborator:** a person located in the physical environment where the task and the task objects are situated. To fulfil the task objectives, the on-site collaborator requests assistance from a remote collaborator, who might have expert knowledge on the subject (Lukosch et al., 2015).
- **Remote Collaborator:** a person located in a remote environment who communicates with the on-site collaborator for situation understanding and providing guidance towards the completion of the shared task (Lukosch et al., 2015).
- **Communication:** an activity that collaborators perform to share their thoughts and intent while completing the task together. To establish a common ground with the goal of completing the task successfully, collaborators may use different communication cues, e.g., verbal and nonverbal behaviors (S. Kim et al., 2019).

- **Tool:** a technological system used to establish a *common ground*, show the state of the task and the activity between physically distributed collaborators. The affordances associated with the collaborative tool determine the general user experience, as well as the communication and collaboration performance (S. Kim, Billinghurst, Lee, et al., 2018).

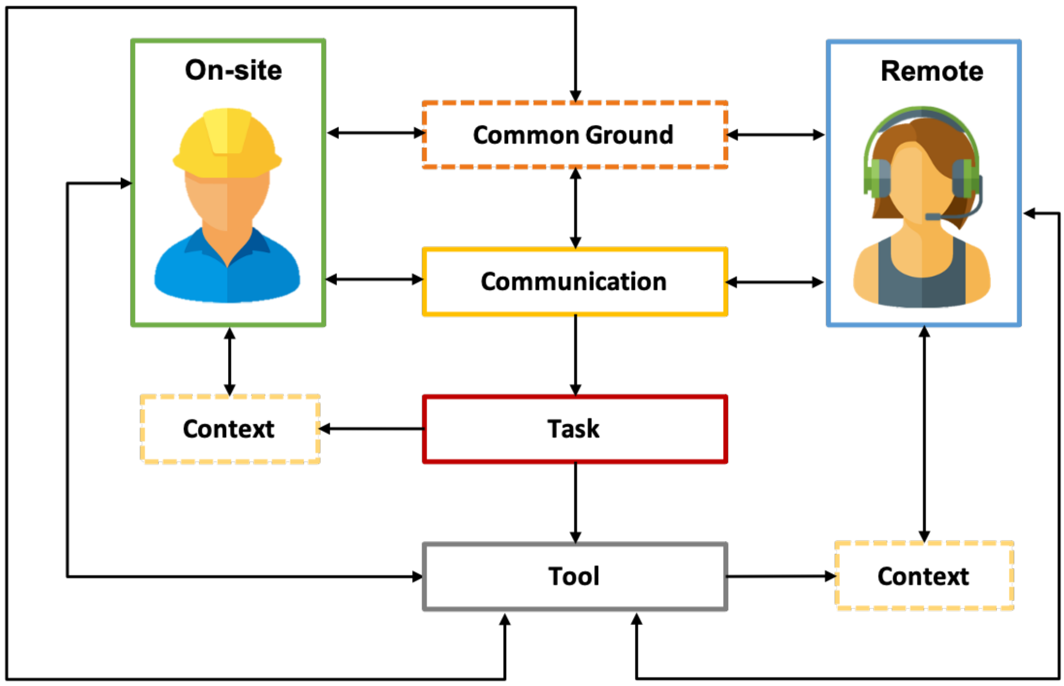

Figure 2 – Conceptual model for essential factors that must be considered regarding remote collaboration.

In this context, it is important to consider how the on-site collaborators can capture and transmit a view of their surroundings. CSCW research showed the potential of video systems to address such scenarios of collaboration, providing evidence that shared visual context is essential for remote collaborative work (Aschenbrenner et al., 2018; Billinghurst et al., 2017). Usually, an on-site person solving task problems does not know what information to share with a remote expert when explaining his/her situation context. In this case, the remote expert may have difficulty to recognize the problem correctly just by seeing images or video scenes. As such, sharing video-based information alone limits proper communication and interaction between collaborators due to the lack of information to promote proper situation understanding, as well as inaccurately guided instructions (Choi et al., 2018; S. Kim, Billinghurst, Lee, et al., 2018; Teo, Lawrence, et al., 2019). In this vein, collaborators often work around these by engaging in complex verbal negotiation, which although it helps communicate their intended directions, affects the conversational grounding and in turn the collaborative performance and situation awareness (Fakourfar et al., 2016; P. Wang, Zhang, Bai, Billinghurst, Zhang, et al., 2019). To overcome these limitations, it is essential to combine visual information with other interaction metaphors such as AR, which can superimpose virtual content on physical artifacts (Billinghurst et al., 2017; Choi et al., 2018).

2.2 Augmented Reality

Solutions using AR can be described as a "*human machine interaction tool that overlays computer-generated information on the real-world environment. The information display and image overlay are context-sensitive, which means that they depend on the observed objects*" (Ong et al., 2008). These definitions do not limit AR to a specific technology, such as Head Mounted Displays (HMDs), allowing other technologies to be used, as long as they retain the essential characteristics of AR.

The use of AR provides new forms of visualization of information and interaction, by adding new content, including 3D models, animations, sound, images and video aligned with the physical world. Nevertheless, AR is not limited to sight, as it may apply to other senses like hearing, touch and smell (Krevelen & Poelman, 2010). In short, any AR application should satisfy three characteristics to provide an augmented experience to the user: 1 - combine real and virtual worlds to let the user perceive them at the same time; 2 - give impression of coherence between, by aligning real and virtual objects; 3 - be interactive in real-time (Azuma, 1997).

The concept of AR is closely related to Virtual Reality (VR), since AR evolved as a variation of VR. On the left side of Figure 3 lies the real environment, composed solely by real objects. On the right is virtual reality, constituted exclusively by virtual objects. The area between these two extremes is defined as Mixed Reality (MR). AR lies between the left and the middle of the continuum. Both VR and AR aim at immersing the user, although using different approaches to accomplish this goal. While VR offers a digital recreation of a real-life environment, AR uses computer-generated technology to blend virtual reality and real life. AR displays virtual elements as an overlay to the real world, allowing to interacting with them in real-time (Milgram et al., 1994), meaning that from a technical perspective, AR is clearly more challenging (Sandor et al., 2015). Currently, the AR concept is evolving into Pervasive AR: an AR experience that is continuous in space and in time (based on the knowledge of the position of the device or user within the environment), being aware of and responsive to the user's context (Grubert & Zollmann, 2017).

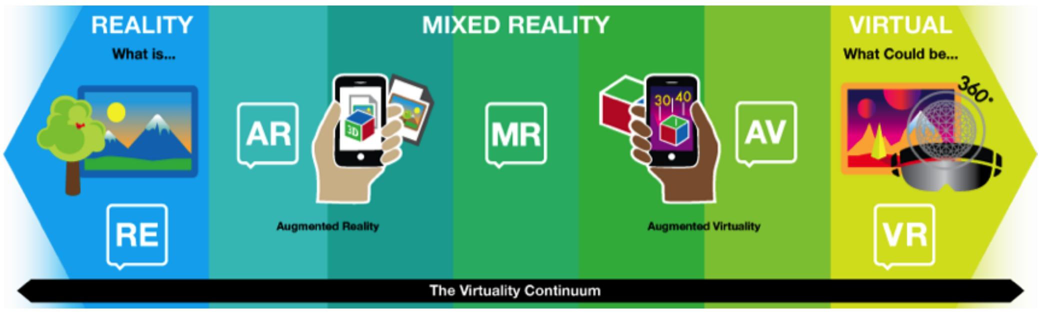

Figure 3 – Representation of Milgram et al. (1994) Reality-Virtuality Continuum: RE - Real World environment; AR - Augmented Reality; MR - Mixed Reality; AV - Augmented Virtuality; VR - Virtual Reality.

Nowadays, AR is considered an interdisciplinary area, transcending boundaries between concepts, traditionally not associated with AR, such as Optics, Computer Graphics, Computer Vision, Human-Computer Interaction, among others. Moreover, to guide the rise of new AR-based solutions, input from other areas might also be required, like ethics, art, philosophy, and social sciences (Sandor et al., 2015).

This technology benefits all areas where on-site and real-time 3D visualization is required. The first AR-based solutions appeared confined to research in the military, industrial and medical fields (Zhou et al., 2008). Yet, through the years, open-source and commercial AR systems began to emerge and helped to expand the use of this technology. Nowadays, AR technology is used in several commercial/real-case scenarios, such as visualization, industry, robotics, education, tourism, marketing, and entertainment, among others (K. Kim et al., 2018; Schmalstieg & Höllerer, 2016).

Numerous studies have proven that AR has the potential to improve human cognition and reduce errors in different areas of application (Li et al., 2017; Neumann & Majoros, 1998). Henderson et al. showed that AR reduced by 50% the time needed to identify maintenance steps (Henderson et al., 2011). Also, Richardson et al. showed AR can also reduce errors by up to 85% for complex assembly procedures (Richardson et al., 2014).

Recently, AR is also turning into a commercially viable technology, proven to have the potential to improve human cognition and reduce errors (Li et al., 2017), with the potential to be used in various practical day-to-day areas of application, such as museums, tourism, advertising, education, industry, among others (K. Kim et al., 2018; Schmalstieg & Höllerer, 2016).

Figure 4 shows the Global AR Industry Landscape for 2019, according to the Venture Reality Fund[4]. Over the years, this landscape has grown to incorporate different application scenarios, as well as new tools and platforms. Subsequently, more and more brands are working to create the necessary infrastructure required for exploring AR to its full potential. Goldman Sachs Global Investment Research predicts that, by 2025, VR and AR applications will reach a value of 80 billion dollars a year, 35 billion dollars in software (Figure 5) and 45 billion dollars in hardware[5] (Bellini et al., 2016), which means the current landscape is expected to continue to grow even more.

[4] - thevrfund.com [Accessed: 31-Mar-2021]
[5] - tinyurl.com/potentialApplicationsVRAR [Accessed: 31-Mar-2021]

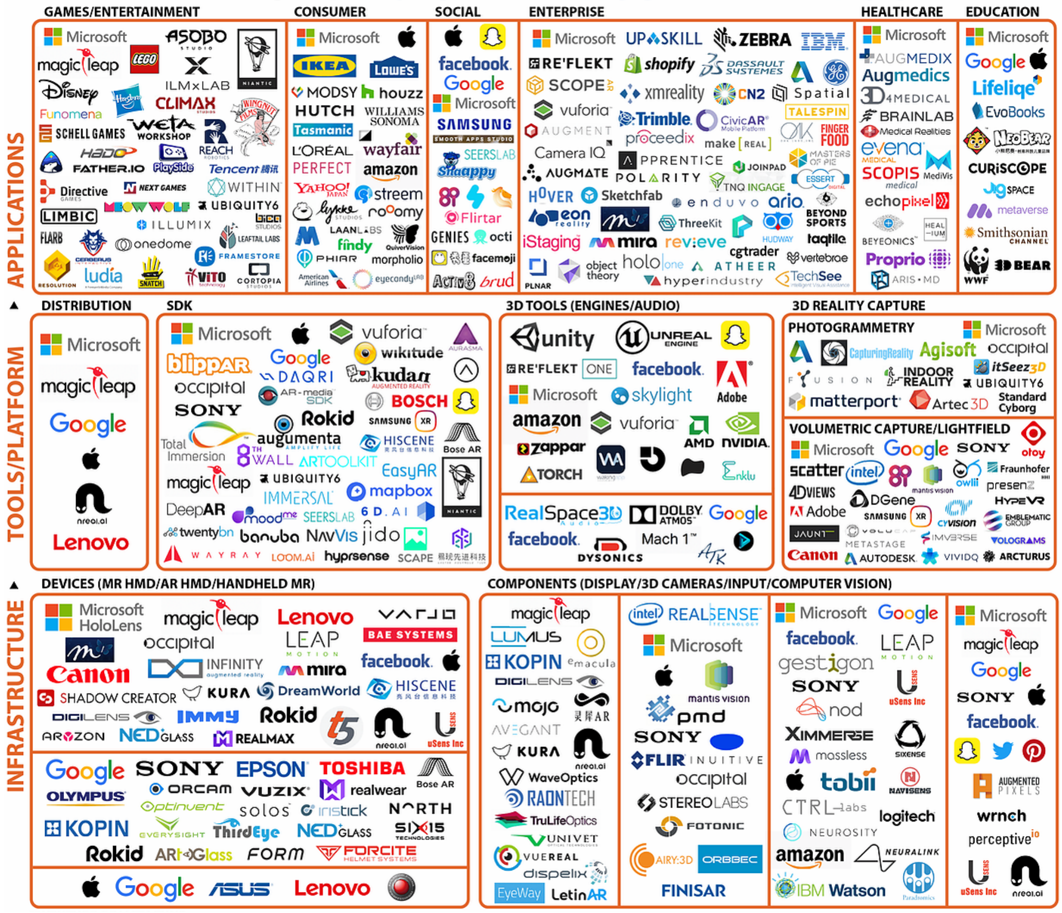

Figure 4 – Global AR Industry Landscape for 2019, according to the Venture Reality Fund[3].

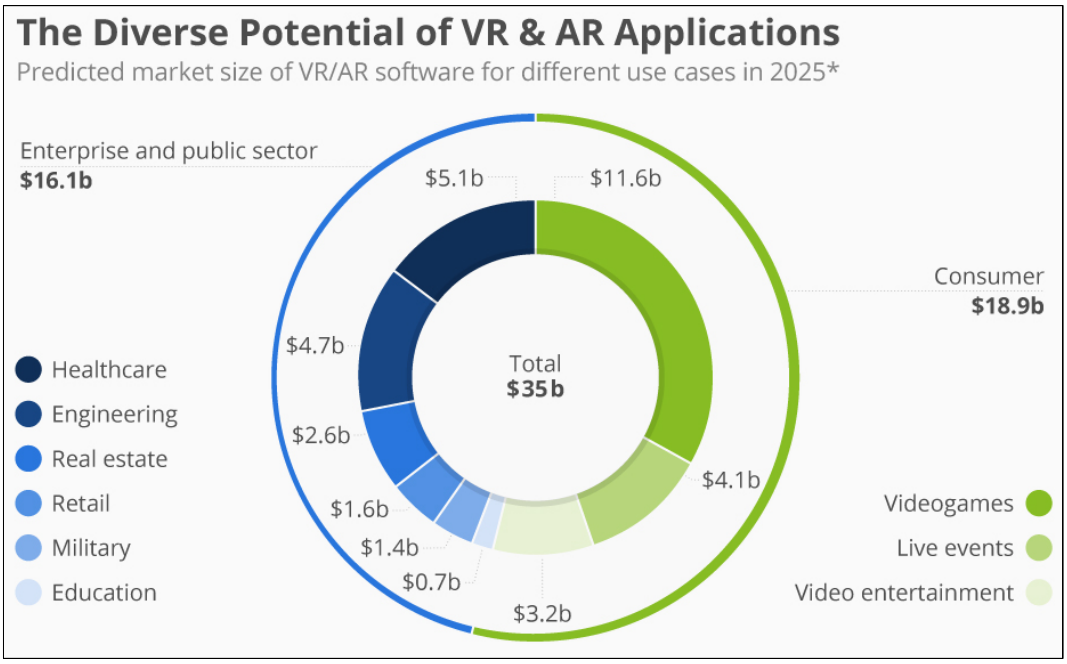

Figure 5 – The Diverse Potential of VR & AR Applications – Predicted market size of VR/AR software for different use cases in 2025. Source: Bellini et al., 2016 and The Statistics Portal[4].

Next, some general applications of AR are described. Starting by the **medical** field, where AR can be used for visualization, enabling doctors to visualize 3D medical data on the patient's body in real time (Figure 6), improving doctors focus on the task at hand, since they do not look away to find information associated with the medical procedure. Another relevant application is in training and simulation, allowing students to train complex surgical procedures in a controlled and safe environment (Barsom et al., 2016).

Figure 6 – AR to support doctors during a guided surgery. Source: Barsom, Graafland, & Schijven, 2016.

Moreover, it is also possible to combine the use of AR serious games and wearables in **Exposure Therapy** (ET) for anxiety disorders (Figure 7) to modulate the response of individuals to different AR stimulus exposure, which may be of great value for diagnostic and treatment purposes in anxiety disorders, namely specific phobia (Brás et al., 2018).

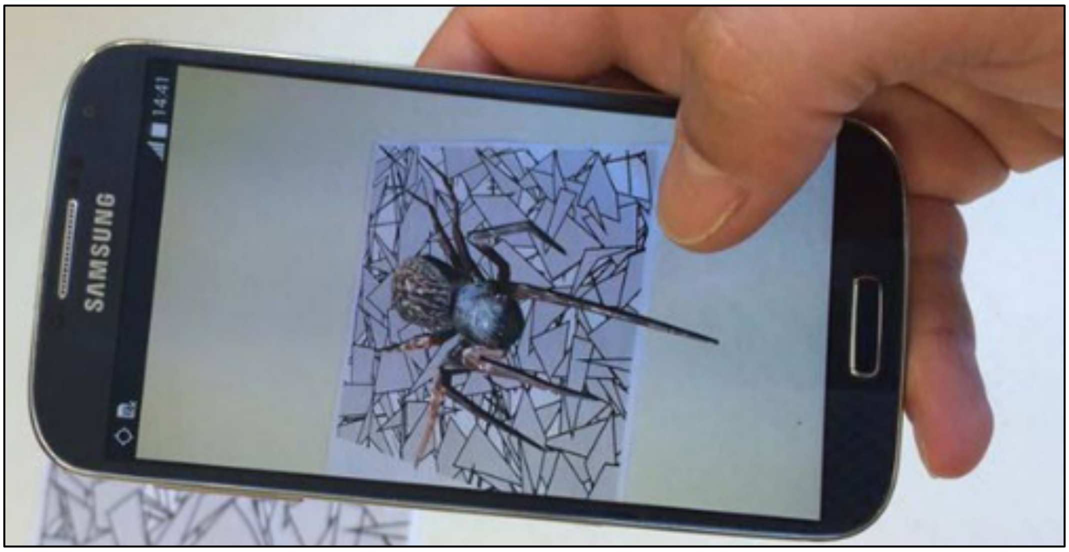

Figure 7 – Participant using the smartphone as a magnifying glass. Source: Brás et al., 2018.

In **entertainment and gaming**, there is a huge market for AR (Pope, 2018), not only due to the visibility of games but also because of the **advertising potential**, promoting brands and products in an interactive manner (Mekni & Lemieux, 2014).

AR can also be used as a virtual guide in **tourism, cultural heritage and museums tours** (Han et al., 2017). Maps that use AR tips are able to show information regarding places of interest. Museums try to incorporate these new technologies in their exhibits to be more appealing to visitors and improve the overall experience (Figure 8) (Miyashita et al., 2008).

Figure 8 – Visitor exploring a map using an AR guiding system. Source: Miyashita et al., 2008.

Then again, AR can also be used as a **navigation** tool to aid with wayfinding in indoor and outdoor spaces (Figure 9). Plus, it can also be used in car navigation for driver safety, using Head-Up Displays (HUDs) to send notifications about warnings, speed limits, etc. (Bhorkar, 2017).

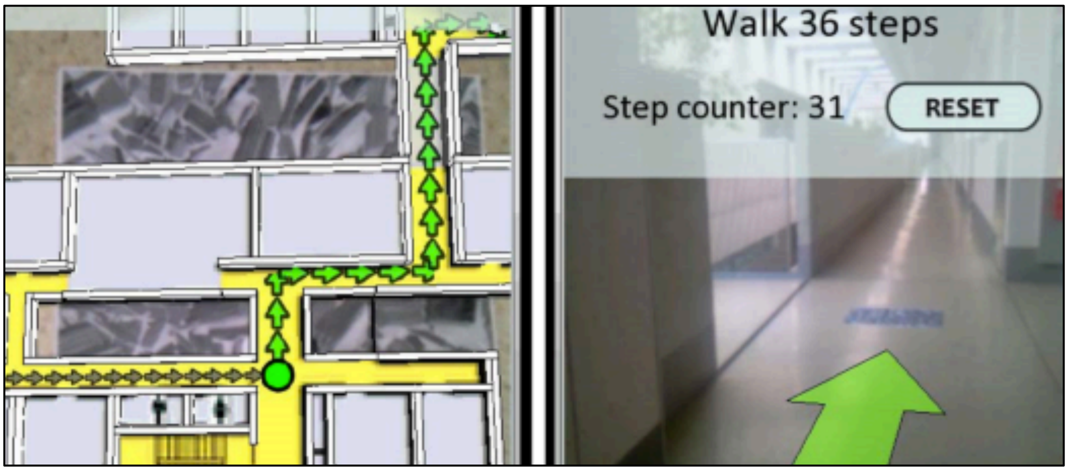

Figure 9 – AR used as a navigation tool to help find a path in indoor environment. Source: Bhorkar, 2017.

In **marketing and sales**, AR can show virtual models of watches, clothes, shoes, etc. on the human body before a buy is made (Carmigniani & Furht, 2011). Also, AR can help to preview the final 3D build of a LEGO set or even show the final look of an architecture project (Mekni & Lemieux, 2014).

As for **education**, AR can be used as a complement to improve teaching and learning quality in a classroom with more information and interaction, increasing student engagement and focus (Al-Azawi & Shakkah, 2018; Bacca et al., 2014).

Since the origins of AR, the **industrial sector** has been one of AR most prominent application domains. Nowadays, with the rise of Industry 4.0 and digital revolution, the potential to use AR is gaining interest as to provide workers with real-time information, aiming to improve decision-making and work procedures (Jetter et al., 2018; Rüßmann et al., 2015). AR-based solutions can present significant potential to improve several industrial sectors, providing better customer services, and an increased engineering and manufacturing quality (Campbell et al., 2017, 2018). Studies show that AR might help to perform complex operations more efficiently by improving cognitive knowledge (Fiorentino et al., 2014). Also, AR can contribute to increasing workers motivation and interest, resulting in less error rate and faster time completion (Radkowski et al., 2015; Sanna et al., 2015).

It is possible to identify at least five major areas of application in the industry sector (Pace et al., 2018; Paelke, 2014): human-robot collaboration, training operations, product inspection and monitoring operations, maintenance and assembly repair tasks. Human-robot collaboration enables the use of AR-based interfaces to interact with industrial robots and visualize relevant information. As such, it is possible to improve operators awareness of the system, while also creating a safer environment, by understanding in advance the robots intentions, movements and forces (Mekni & Lemieux, 2014). In training, users are able to use AR to improve their skills before moving to real-life scenarios (Ong et al., 2008). In product inspection and monitor operations, technicians can use powerful and versatile AR system to notice discrepancies of items and highlight any errors (Pace et al., 2018). In maintenance and assembly tasks, AR can be used to improve productivity and to assist technicians with specific problems, while reducing errors and time duration (Figure 10). Moreover, unexpected failures can be (highly) disturbing, which means technicians face unfamiliar incidents, requiring specific knowhow and additional information. Recently, there is a growing interest for AR systems to connect on-site technicians to remote experts, in order to help provide clues regarding unexpected problems (Martinetti et al., 2017; Palmarini et al., 2018; Quint et al., 2017).

Figure 10 – Maintenance technician using AR for task guidance. Source: Aromaa, Aaltonen, Kaasinen, Elo, & Parkkinen, 2016.

2.3 Remote Collaboration mediated by Augmented Reality

Even though AR shows great promise and is commonly used to enhance a single user perception of reality, one of the areas where it can be most useful, perhaps AR greatest potential, is to assist in collaborative scenarios (Billinghurst et al., 2015; Dey, Billinghurst, Lindeman, & Ii, 2018; Lukosch et al., 2015; Schmalstieg & Höllerer, 2016). The widespread of mobile devices allows individuals to access AR through a personal perspective and share and interact with other users through collaborative interfaces. Hence, researchers from the CSCW domain have been investigating the use of AR to provide collaborators with a shared understanding over the last three decades (Gergle et al., 2013). The concept of Collaborative AR can be described as a system where: "*multiple collaborators share the same augmented environment locally or remotely*" (H. T. Regenbrecht et al., 2002), "*which enables knowledge transfer between them*" (Jalo et al., 2018). Moreover, the augmentation of the real environment of one collaborator occurs through the actions of other collaborators and does not merely rely on information previously stored (X. Wang & Dunston, 2006).

Co-located AR solutions can be used to elicit the performance of specific tasks between a group of users, allowing interaction with shared virtual content as naturally as with physical objects while maintaining important natural face-to-face communication cues (Lukosch et al., 2015). While early work focused mainly on co-located scenarios, there is a great interest for remote scenarios as technological limitations are being overcome (Ens et al., 2019; Hall et al., 2018).

Even though remote collaboration using AR started roughly at the same time as co-located collaboration, only in recent years has a growing interest emerged as literature shows. AR-based solutions can be used to empower specific situations in which individuals may require knowhow and additional information from professionals unavailable on-site (Gurevich et al., 2015;

Schmalstieg & Höllerer, 2016; Teo, Lawrence, et al., 2019). In this vein, remote experts can use AR-based solutions regardless of their localization to guide on-site collaborators, providing real-time spatial information, highlighting specific areas of interest, or sharing situated information associated with relevant objects in the on-site physical environment (Hall et al., 2018; Lamberti et al., 2014; Palmarini et al., 2018).

Over the years, AR has been explored to augment video streams, shared virtual replicas or reconstructions of a physical environment, covering multiple areas of remote collaboration like: Entertainment and Education, Healthcare, Crime Scene Investigation, Training, Maintenance, and others (R. Belen et al., 2019; R. A. J. de Belen et al., 2019; K. Kim et al., 2018; S. Kim, Billinghurst, et al., 2020).

To illustrate, in **Entertaining and Education**, AR can be used to solve jigsaw puzzles in a collaborative manner. For example, (S. Kim et al., 2014) explored visual communication cues to improve the experience of distributed team members that needed to assemble a Tangram puzzle collaboratively (Figure 11). The on-site collaborator could capture a live video of the task space using a HMD or a mobile device. Then, communication was achieved through an augmented pointer, spatial annotations or drawings on the shared view. A control study was conducted with 24 participants to investigate if augmented visual cues could improve the feeling of connectedness. Results showed both pointing and spatial annotations could improve the shared experience in terms of feeling more connected, being together, and understanding the remote collaborator. Results also suggested that pointing is quicker and easier than spatial annotations, which require more time to be created. Later, another work followed this approach by introducing eye gaze in both directions, between remote team-members, aiming to understand how this affected collaboration and communication. Results from a formal study with eight participants showed that sharing gaze significantly improved awareness of both users, hence improving collaboration when compared to the sharing of eye gaze from one user only or no sharing (Lee et al., 2017).

Figure 11 – Assembly of jigsaw puzzles though remote AR. Source: (S. Kim et al., 2014).

As for **Healthcare**, an on-site doctor with little expertise might need to conduct an operation on a patient with the help from a distant expert surgeon. For example, (S. Wang et al., 2017) explored the use of see-through HMD and hand gestures to facilitate remote medical interaction during complex medical procedures. The remote hands hand gestures are captured using a dedicated sensor which displays a virtual representation in the AR space of the on-site doctor (Figure 12). As part of a pilot user study, twelve novice medical trainees were guided by an expert through the proposed solution during a simulated ultrasound exploration in a trauma scenario. The study allowed to collect insights to improve the proposed solution when compared to a more traditional multi-camera telemedicine alternative.

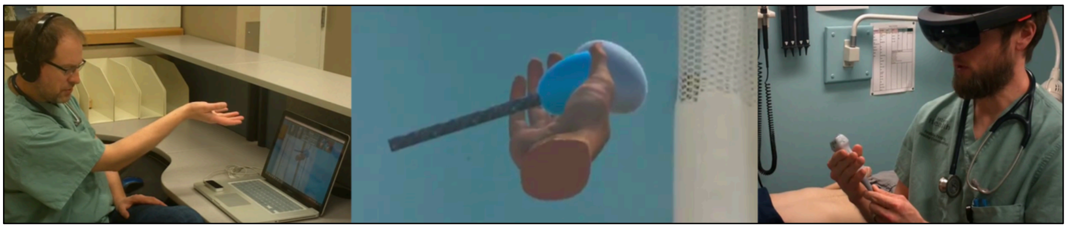

Figure 12 – Remote telemedicine using a 3D shared model AR platform. Source: (S. Wang et al., 2017).

Regarding **Crime Scene scenarios**, an on-site investigator can request help from a forensics expert in a remote place. An example is the research by (Datcu et al., 2016), which allows a remote expert to provide visual and audio cues whenever the on-site investigator requires assistance (Figure 13). Besides, the on-site investigator used a mobile device attached to the wrist and a video camera placed on his shoulder (providing the view from the environment), while the remote experts used a laptop computer to provide assistance through visual cues, e.g., AR annotations like arrows and drawings. The proposed solution was evaluated with three experienced forensic investigators and allowed to conclude that the on-site investigator division of attention between mobile devices and the real environment impacted the situational awareness. However, using mobile devices for visualization and limited interaction from time to time was considered feasible, in particular when compared to earlier experiments using HMDs, which might be more suitable in scenarios that require constant attention and interaction with AR information.

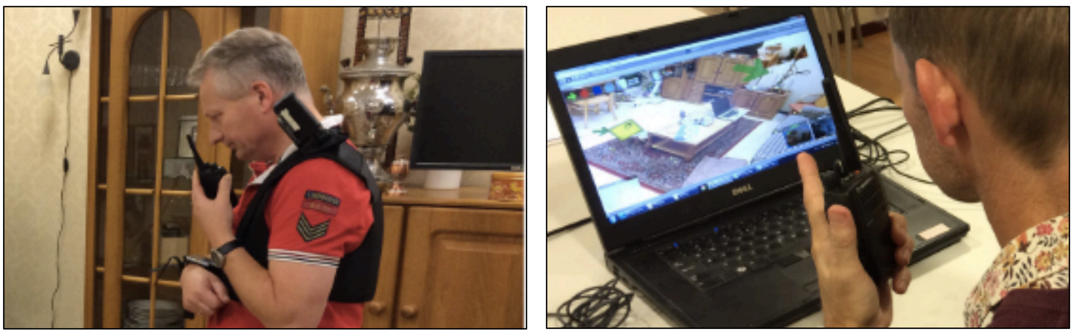

Figure 13 – AR for Distributed Collaborative Crime Scene Investigation. Source: (Datcu et al., 2016).

Training, Guidance and Maintenance can also benefit from AR in scenarios od Industry 4.0. In **Training**, it is possible to use 3D shared models related to the local physical context, i.e., take advantage of pre-existing virtual replicas (or digital twins in industry 4.0) to provide guidance between distributed team-members using a richer common ground. For example, (Elvezio et al., 2017; Oda et al., 2015) proposed a solution for training an on-site technician while performing an intervention in an aircraft combustion engine. The remote expert has access to a virtual replica of the physical object, that he/she can manipulate while providing training instructions to the on-site trainee (Figure 14).

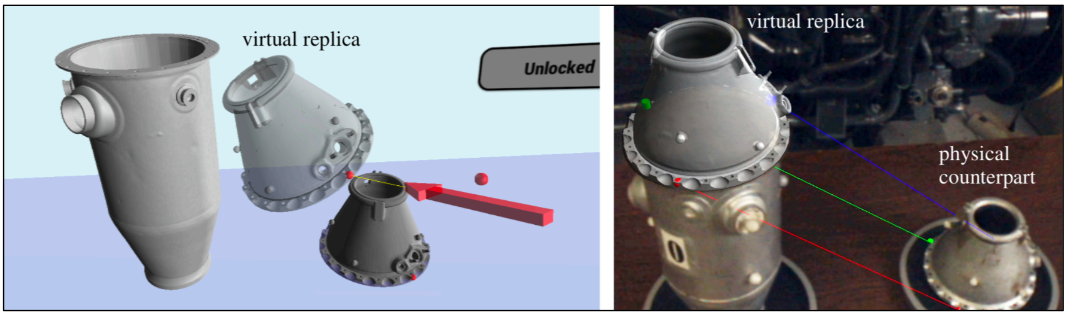

Figure 14 – Remote training using AR virtual replicas. Source: (Elvezio et al., 2017; Oda et al., 2015).

Another possibility is high-fidelity, dense scene reconstruction for precise AR-based remote **Guidance**. An on-site worker in need of assistance can use the HMD sensors to generate a 3D mesh of the surrounding physical environment and stream it to a remote expert (Zillner et al., 2018). Then, the expert can explore the reconstructed virtual model in six degrees of freedom as well as segment-colored objects from the mesh. Plus, AR-based annotations (e.g., text, image, drawing) can be placed in the model to convey precise instructions (Figure 15). This extensive use of dense scene reconstruction provides a unique interaction experience for remote work with AR HMD glasses.

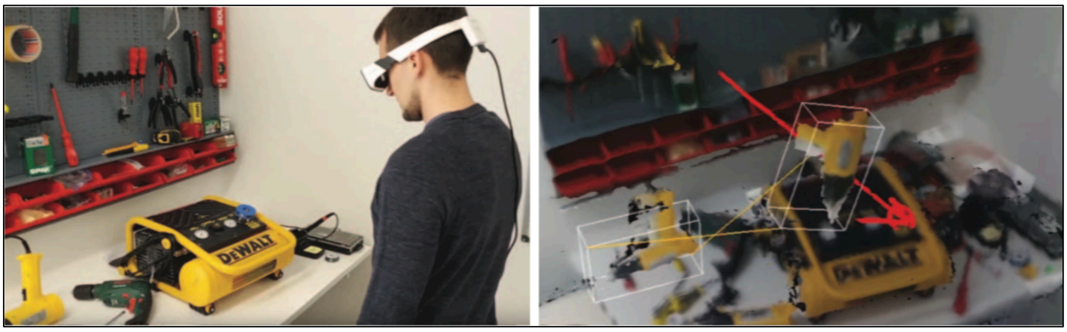

Figure 15 – Remote guidance using dense reconstruction and AR annotations. Source: (Zillner et al., 2018).

In **Maintenance**, an on-site novice collaborator might require support from a remote expert to perform a repair, as reported by (Masoni et al., 2017). The authors proposed a solution based on off-the-shelf mobile devices and a desktop computer, which allowed to connect a skilled user with an untrained worker (Figure 16). The on-site collaborator could take a photo of the environment, use it as a marker and share it with the remote expert for context understanding. Then, the photo could be annotated by the expert using AR-based visual cues based on common operations in maintenance: unscrew, screw, indications, warning, disassemble and assemble, as well as sketches or notes. A control study with partners from the industrial sector showed the potential of AR-technology as a tool for industry scenarios.

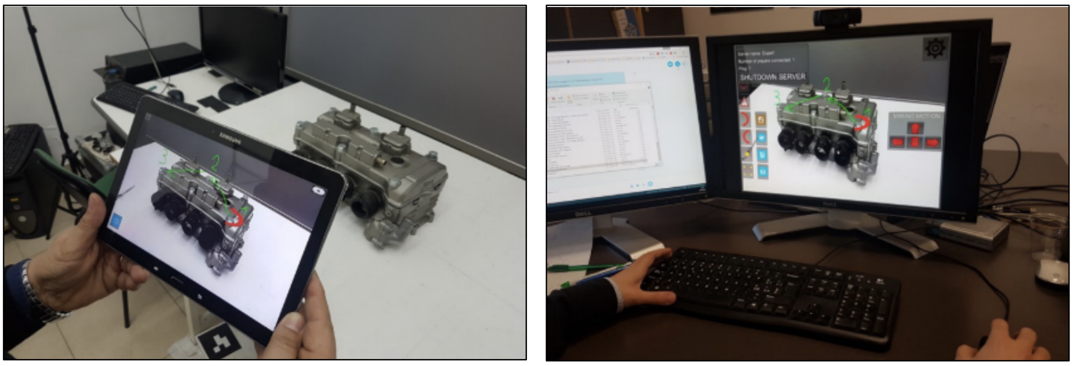

Figure 16 – Collaborative AR for remote maintenance. Source: (Masoni et al., 2017).

Although research in AR-supported remote collaboration is still in its infancy, it has the potential to support effective knowledge transfer between distributed collaborators allowing them to interact with each other in a context-sensitive manner that may result in significant benefits (Jalo et al. 2018; K. Kim et al. 2018). By creating a common ground environment, AR-based solutions can support information exchange and improve collaborators situational awareness, enhance alertness, and understanding of the situation through different types of communication aids (e.g., pointers, annotations, gaze, hand gestures, among others). Hence, enhance a scene as it is captured by an on-site collaborator and provide real-time spatial information about objects, events and areas of interest (Ens et al., 2019; Gurevich et al., 2015; S. Kim, Billinghurst, Lee, et al., 2018; Palmarini et al., 2018).

2.3.1 Augmented Reality vs Mixed Reality

While the research described in this thesis uses the expression ***remote collaboration supported by AR***, some recent efforts described in literature are beginning to replace the term AR by MR (H. Bai et al., 2020; Ens et al., 2019; S. Kim, Lee, et al., 2020; Piumsomboon, Dey, et al., 2019).

But what exactly is MR, and why this sudden change? Many researchers see MR as a synonym for AR (Speicher et al., 2019). Some consider MR a superset of AR in terms of a mix of real and virtual objects that are presented together on a single display (Billinghurst & Kato, 1999; Milgram et al., 1994; Milgram & Kishino, 1994), i.e., a real-world object can interact with a virtual one in real-time to assist individuals in practical scenarios. 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 (Rokhsaritalemi et al., 2020).

While MR is increasingly gaining popularity and relevance, being recently considered as one of the top 10 ranked ICT technologies (Rokhsaritalemi et al., 2020), the research community is still far from a clear definition of what MR actually constitutes. As reported in a recent survey by Speicher et al. (2019), currently, there is no single definition for MR, since this concept can be considered different things as its understanding is based on one's context, e.g., perceptions of MR, technological capabilities and design practices. 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 (Speicher et al., 2019).

Among the most important applications of MR are collaborative tools due to the recent advances of commodity technology, e.g., availability of new AR/VR HMDs and Handheld Displays (HHDs) platforms (Ens et al., 2019; Nebeling et al., 2020) that may be used as decision-making tools for daily life problems (Rokhsaritalemi et al., 2020). 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. 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 (H. Bai et al., 2020; G. A. Lee, Teo, et al., 2017; Speicher et al., 2019) and so provides unique capabilities to achieve a common goal, e.g., improved communication cues for more efficient and easier collaboration (Billinghurst & Kato, 1999; S. Kim, Lee, et al., 2020; Masai et al., 2016; Piumsomboon, Dey, et al., 2019).

To clarify the terminology used in this thesis, since the research reported does not focus on scenarios where remote users are supported by VR to obtain a reconstructed environment of the on-site collaborator, we decided to use the expression ***remote collaboration using AR***. However, the contributions described in this thesis are broad enough to also support how the collaborative process can be reported in scenarios supported by MR-technology, following the definition described above (Speicher et al., 2019).

2.3.2 Commercial Tools

Recently, commercial AR tools have emerged to assist in scenarios of remote collaboration with companies like ScopeAR[6], XMReality[7], ViewAR[8], PTC Vuforia[9], Vsight[10], Re'flekt[11], Upskill[12], Glartek[13], or Microsoft remote assistant[14], among others, promising increase in productivity, cost, downtime and error reduction, as well as lower training time.

Evaluating these tools is a challenging endeavor, since most, if not all of them require paid subscriptions in order to have access to the full range of features. Nevertheless, during the pandemic crisis caused by COVID-19, some of these companies made available trial versions of their tools for free. This exception allowed us to test the Vuforia Chalk (Figure 17) and XMReality tools, while others did not answer our request for a demo trial.

Figure 17 – Example of Vuforia Chalk Annotations features - Toyota use case. Source: (PTC, 2020).

We were able to understand that such tools mainly focus on synchronous collaboration between a remote expert and an on-site worker in need of assistance, being developed mainly for industry scenarios, e.g., diagnostic, maintenance, inspection, and others. These tools aim to increase response capacity. To do so, they provide different instances of the tool, according to the role of the person using it, i.e., desktop version through browsers for local experts and handheld or smart glasses for on-site workers, which means different features are available during the collaborative process. To assist in the decision-making process, existing tools offer video, audio, chat, AR annotations and file transfer. Regarding AR annotations, it is possible to point, place arrows, share text bubbles, as well as drawing or creation of highlights on top of real-world objects. In

[6] - scopear.com [Accessed: 31-Mar-2021]
[7] - xmreality.com [Accessed: 31-Mar-2021]
[8] - viewar.com [Accessed: 31-Mar-2021]
[9] - ptc.com/en/products/vuforia/vuforia-chalk [Accessed: 31-Mar-2021]
[10] - vsight.io [Accessed: 31-Mar-2021]
[11] - re-flekt.com [Accessed: 31-Mar-2021]
[12] - upskill.io [Accessed: 31-Mar-2021]
[13] - glartek.com [Accessed: 31-Mar-2021]
[14] - dynamics.microsoft.com/en-gb/mixed-reality/remote-assist/ [Accessed: 31-Mar-2021]

addition, the XMReality tool also enables sharing the remote expert hands to suggest specific actions during the assistance process.

Figure 18 – On-site technicians receiving hand gestures from a remote collaborator through XMReality. Source: (XMReality, 2020)

Although some of these companies suggest they are able to provide AR 3D models in their tools, this is mostly applied to pre-defined on-site use-cases, instead of scenarios of remote collaboration, probably because these last often occur in dynamic environments, which limit the amount of models commercial tools can provide. Thus, the wager on AR annotations, since they are easier to implement and to use, is an advantage to dynamic scenarios. In this context, it seems that such tools are better suited for simple tasks, and that they still need to tackle many research challenges before they can be consider truly useful in complex realistic tasks for different areas of application.

For example, much of the above-mentioned features are only available to the remote expert. The on-site worker still lacks identical features to express their thoughts while communicating and requesting assistance. Also, the use of Multi Modal Interaction (MMI), i.e., voice recognition, hand gestures, gaze, and others are not yet contemplated. To elaborate, interaction methods are still limited, relying on traditional mouse and keyboard on the remote side and on touch for the on-site worker. Plus, if the on-site worker uses smart glasses, interaction is minimal, thus being limited to capture the context of the problem and to visualization of AR information on top of the real-world.

We argue that in such cases it is difficult to collaborate, and what actually occurs is a situation of guidance, where the on-site worker only follows instructions, which, without the presence of the remote expert, would be similar to a scenario of following step-by-step instructions. Lastly, although tracking technologies have been improving in recent years, a common problem with AR annotations is related to the incorrect anchoring to an object that might occur if the on-site worker changes the viewpoint of the shared view when the remote experts is creating annotations.

From the analysis of existing marketing videos of other companies, which follow the trends previously presented, it appears that these challenges are common to the vast majority, if not all, of current commercial tools, possibly because they are still in an early phase of design and development and companies are still struggling with engineering hurdles. Consequently, it is not possible to identify any features in existing commercial tools that may help with the characterization of the collaborative process, an important topic to understand how AR-based solutions can support target stakeholders, and in what conditions, as well as comprehend how collaborative performance evolves over time. Since most of these companies address industry related use cases, this subject may prove useful to quantify the impact of such tools in industrial processes in the long-term.

2.3.3 How Remote AR can help Businesses during the Pandemic

The recent pandemic caused by COVID-19 has been generating an unprecedented impact across the business landscape all over the world. During this period, several countries have been implementing various forms of lockdown, e.g., communities were put into quarantine and were restricted from traveling and even stepping outside, severely limiting how business is conducted, and in some cases even leading to the full stop of company's activity. Next, some of the main impacts of COVID-19 are described[15]:

- Traditional face-to-face interactions are affected by social distancing measures which limit the size of groups;
- The ability to conduct business, manage effective team operations, and share knowledge where it is needed is being diminished by the inability to travel and prevalence of expertise individuals, working from home.
- Organization's ability to continue operations 'as before' is limited by fewer on-site workers due to illness, self-isolation or financial restrictions;
- Train new or existing workers on product and processes becomes a challenging endeavor due to the lack of classroom hands-on training.
- Supply chains are affected by interruptions in the shop floor, thus requiring more flexible processes to ensure product continuity;
- Workers avoid surfaces and objects that may have been touched by others due to the potential virus transmission.

To help address these limitations, research and development of enhanced tools is of paramount importance, although any lingering transformation of what we do at a distance will rely on convincing and accessible forms of remote collaboration (Matthews et al., 2021). In fact, remote collaboration gained new significance given the pandemic circumstances since it represents an important artifact/alternative to overcome many limitations in the most effective and fastest way possible.

Therefore, businesses need to adapt to this new reality by becoming more flexible and agile to embrace new and innovative technologies. For example, the pandemic has increased use cases

[15] - thearea.org/covid-19-how-augmented-reality-is-helping-mitigate-business-impact/ [Accessed: 30-Apr-2021]

like remote assistance and training[16] given that employees are less willing to be in close contact with each other. Moreover, travel restrictions prevent the most specialized individuals from being present in the field, while various issues may arise[17]. Solutions exploring AR can play an important role in mitigating a number of COVID-19 restrictions, while offering opportunities to provide long-term improvements.

Remote collaboration using AR is reaching several areas of applications, including technical support, inspection processes, training and maintenance repairing, which are of extremely usefulness in a wide range of industries[18]. The use of such technologies can help mitigate impacts of travel reduction, reduced staffing, as well as social distancing. Furthermore, it is expected a reduction in errors, downtime, training time, and all other forms of cost and loss by eliminating the need for mass gatherings while keeping the workforce connected, improving efficiency, and data transfer, thus disseminating knowledge that challenges physical boundaries while also minimizing the need for expertise individuals to travel abroad.

Next, some opinions of individuals currently handling such technologies in the daily activities of their workforces are presented to illustrate the role AR-based solutions for remote collaboration are having during this challenging times. For example, Sarah Reynolds, Vice President of Marketing at PTC emphasizes that[19]: "*as organizations look to maintain business continuity in this new normal, they are embracing AR to address travel restrictions, social distancing measures, and other challenges impacting their front-line workers' ability to go on-site and operate, maintain, and repair machines of all kinds (…) improve the clarity, precision, and accuracy of their communication and collaboration*". Additionally, Stegan Goeris, Process Consulting for Manufacturing Digital Business at Henkel[20] reports the following regarding the use of a specific AR-based tool for remote collaboration: "*exchange between employees is essential. Vuforia Chalk has enabled us to promote virtual collaboration between employees on the production floor and employees working remotely. Despite the current situation, sharing of expertise has not been sacrificed and we can continue to drive knowledge sharing*". Jaume Carreras, Project Manager for digital transformation at the Laundry & Home Care production facility also advocates that "*Vuforia Chalk has provided us with an efficient alternative to physical meetings at the plants in times of border closures and travels restrictions (…) The direct exchange and immediate feedback made possible via the remote assistance tool ensures efficient virtual collaboration*".

Hence, the wide potential of such technologies is quite favorable for further use by businesses since remote collaboration becomes feasible and extremely important to facilitate the way the workforce performs its tasks in the current panorama. It is likely that the ongoing crisis may have

---

16 - valmet.com/media/articles/services/remote-services-prove-valuable-during-the-covid-19-pandemic/ [Accessed: 02-Jun-2021]
17 - vsight.io/the-covid-19-field-service-challenge-ensuring-business-continuity-with-ar-remote-assistance/ [Accessed: 30-Apr-2021]
18 - vsight.io/ar-remote-support-keeping-people-connected-during-covid-19/ [Accessed: 30-Apr-2021]
19 - thearea.org/covid-19-how-augmented-reality-is-helping-mitigate-business-impact/ [Accessed: 30-Apr-2021]
20 - ptc.com/en/blogs/corporate/henkel-uses-vuforia-chalk-real-time-remote-assistance-covid-19 [Accessed: 30-Apr-2021]

an impact in the months ahead and beyond, as the transition to the 'new normal' is achieved. As such, successfully adopting AR-based Remote Collaboration appears as a step towards ensuring business continuity during this global pandemic and beyond.

2.3.4 Summary

Despite the amount of research on remote AR-technologies, there are still several gaps that must be tackled by researchers and practitioners before these technologies can be used efficiently in a daily manner. To summarize, it is possible to identify the following research opportunities regarding AR-supported remote collaboration:

- Improve usability of collaborative AR-based user interfaces (S. Kim, Billinghurst, et al., 2020; Lukosch et al., 2015);
- Address remote collaboration between multiple teams and in asynchronous scenarios (R. A. J. de Belen et al., 2019; Irlitti et al., 2017)
- Develop effective tools to empowered distributed users to interact with each other and with the environment (K. Kim et al., 2018; S. Kim, Billinghurst, et al., 2020);
- Resort to multimodal approaches, as well as different interaction devices to present augmented information efficiently, without interfering with the users' task (S. Kim, Billinghurst, et al., 2020; Lukosch et al., 2015);
- Assess how to maintain focus of the users to certain events and parts of the environment (Lukosch et al., 2015; P. Wang, Zhang, Bai, Billinghurst, He, et al., 2019);
- Evaluate which interaction paradigm is more effective for communication between users, and in which scenarios (K. Kim et al., 2018; S. Kim, Billinghurst, et al., 2020);
- Create novel methodologies to evaluate solutions in this context (Andreas & Billinghurst, 2011; R. A. J. de Belen et al., 2019; Dey, Billinghurst, Lindeman, & Ii, 2018; Ens et al., 2019).
- Conduct more user studies on collaborative AR systems (Andreas & Billinghurst, 2011; Dey, Billinghurst, Lindeman, & Ii, 2018; K. Kim et al., 2018; Teo, Lawrence, et al., 2019);
- Explore human factors to understand perception, cognition, and behavior towards the use of AR-technology (S. Kim, Billinghurst, et al., 2020);

Literature shows that, so far, most of the research efforts have been devoted to exploring and evolving the AR technology for scenarios of remote collaboration. Nevertheless, as well observed by (Merino et al., 2020), future works on MR and AR will elaborate on human-centered evaluations involving not only the analysis of user experience and performance, but also understanding the role of such technologies in working places, in communication and in collaboration. As Collaborative AR research starts to focus on the nuances of supporting the collaborative effort, characterization and evaluation of the collaborative process become of paramount importance (Antunes et al., 2014; Ens et al., 2019; Hamadache & Lancieri, 2009; Marques, Teixeira, et al., 2020; Merino et al., 2020) to ensure the quality and relevance of the growing number of prototypes by assessing different aspects of collaboration itself (Antunes et al., 2014; Dey, Billinghurst, Lindeman, & Swan, 2018; Hamadache & Lancieri, 2009; S. Kim, Billinghurst, & Lee, 2018; Marques, Teixeira, et al., 2020; Merino et al., 2020; Neale et al., 2004).

The research presented in this thesis delves into these subjects, which are of the utmost importance in order to improve the characterization of the collaborative process, which in turn can better inform how AR contributes to the collaborative work effort.

2.4 Overview of Existing Evaluation Efforts of Collaborative AR

To understand to what extent user evaluation is currently being reported regarding collaborative AR and collaborative MR research, we conducted an analysis of existing works based on two phases (Marques et al., 2021). First, we started by understanding what has been reported through the analysis of existing survey papers. Second, we identify relevant aspects that are potentially missing from these surveys and based on this analysis, we re-visit recent literature through a systematic review to see if this is not reported or if it has not been considered relevant information for describing the collaborative process. Thus, this section starts by analysing previous surveys, and then 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.

2.4.1 Previous Surveys including User Evaluation information

This section analyzes existing survey papers addressing evaluation in Collaborative AR, which are summarized in Table 1. The 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. While this is the case, the two last entries of the list 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.

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 (Zhou et al., 2008).

In addition, Duenser et al. (2011) reported on user evaluation techniques used in AR research. Back 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 (Andreas & Billinghurst, 2011).

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 was 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 (Z. Bai & Blackwell, 2012).

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 (Billinghurst et al., 2015).

Authors & Year	# pubs. analyzed	Aspects of Collaboration	Main Outcomes
(Zhou et al., 2008)	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.
(Andreas & Billinghurst, 2011)	10	n/a	Studies that evaluated collaboration between users using AR were quite under-represented. Only 10 papers were reported, which were divided according to the study type in formal and informal studies.
(Z. Bai & Blackwell, 2012)	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).
(Billinghurst et al., 2015)	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.
(K. Kim et al. 2018)	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).
(Dey et al. 2018)	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.
(Ens et al., 2019)	110	Time, Space, Symmetry, Artificiality, Focus, Scenario	Proposal of a questionnaire about overall collaboration, namely the level of enjoyment, the level of mental stress in communication with partner, and whether collaboration went well or not. It also included product measures focus on assessing collaboration outcomes in terms of efficiency or quality, process measures to assess user communication and identify patterns of collaboration and satisfaction measures.
(R. Belen et al., 2019)	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.

Table 1 – Summary of Surveys addressing User Evaluation and Collaborative AR (2008-2019).

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% (Zhou et al., 2008). 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) (K. Kim et al., 2018).

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 and mostly directed toward 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 on average. 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 and subjective, as well as quantitative and qualitative, while 5 studies were only based on subjective data, and 2 studies 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 suggests the need of more user studies regarding collaboration using AR, particularly more field studies, and the use of a wider range of evaluation methods (Dey, Billinghurst, Lindeman, & Ii, 2018).

More recently, Ens et al. (2019) revisited collaboration through MR, taking into account the evolution of groupware. 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 various user studies (Ens et al., 2019).

Finally, Belen et al. (2019) performed a systematic review of the current state of collaborative MR technologies including works 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 co-located 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 (R. Belen et al., 2019).

Summary

Research is evolving from solving technical issues using AR and MR solutions, towards more meaningful human-centered studies on collaboration. We were able to understand that, even when evaluation is performed, it 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) or AR solutions, description on the experimental setup, 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 (Patel et al., 2012). 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.

2.4.2 Method and Overview of Recent Literature

To understand to what extent user evaluation is currently being reported regarding 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 (Gaines, 1991), 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.

Search Process

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

- “Augmented Reality” OR “Mixed Reality”;
- “Remote Collaboration” OR "Remote Cooperation" OR "Remote Assistance" OR "Remote Guidance" OR "Distributed Collaboration";
- "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. Besides, the decision to use Scopus and not considered other alternatives, e.g., Google Scholar was based on the conclusions reported by (Martín-Martín et al., 2018), in which the authors investigated 2,448,055 citations to 2,299 English-language highly cited documents from 252 subject categories published in 2006, comparing Google Scholar (GS), the Web of Science (*WoS*), and Scopus, where the authors reported the following: “*About half (48%-65%, depending on the area) of GS unique citations are not from journals but are theses/dissertations, books or book chapters, conference proceedings, unpublished materials (such as preprints), and other document types. These unique citations are primarily written in English, although a significant minority (19%-38% depending on the area) are in other languages. The scientific impact of these unique citations themselves is, on average, much lower than that of citations also found by WoS or Scopus, suggesting that the GS coverage advantage is mostly for low impact documents. Taken together, these results suggest caution if using GS instead of WoS or Scopus for citation evaluations. Without evidence, it cannot be assumed that the higher citation counts of GS are always superior to those of WoS and Scopus, since it is possible that the inclusion of lower quality citing documents reduces the extent to which citation counts reflect scholarly impact.”*

Analysis Process

Regarding the analysis of reviews, 9 publications were identified focusing on the pre-establish attributes. Moreover, in regard to the systematic review, we obtained a total of 64 publications. Then, 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: 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 VR or AR; experimental setup description; adaptation period provided; study average duration (min); recording of audio and video.

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.

Results

Next, a high-level overview of the reviewed papers is provided (Table 2, Table 3, Table 4 and Table 5) following a similar structure as the one used by (Dey, Billinghurst, Lindeman, & Ii, 2018) in their systematic review, which is extended to include relevant aspects missing from the surveys analyzed in the previous section, such as collaboration type, task type, study type, data type, study design, evaluation methods, participants characteristics, experimental setup, adaptation period, and duration.

User Studies Categorization

The papers (Table 2 and Table 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 (1 paper, 2.4%); education (1 paper, 2.4%); tele-presence (1 paper, 2.4%), as presented in the orange bubbles in Figure 19. Regarding the collaboration type, 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 Figure 19.

Study type

We found that most papers (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 Figure 19.

Study Design

As shown in Table 4 and Table 5,16 papers (38.7%) used a within-subjects experimental design, while 15 papers (35.7%) used a between-subjects design. There were no mentions of a mixed-factorial design. In addition, 11 papers (26.2%) did not mention the method used, as illustrated in the green bubbles in Figure 19.

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 Figure 19. 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 Figure 19. Finally, the bulk of the user studies were conducted in an indoor environment, but only 21 papers (50%) described the experimental setup, although no clear pattern emerged.

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 Figure 19. 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 dialogue 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.

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 Figure 19. 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.

Summary

Our review (Table 6 and Figure 19) shows that most studies focus exclusively on the performance of one collaborator, i.e., on-site, or remote. This means evaluation usually does not include interaction, and communication details, being the focus given to the technological aspects of the solution being used, as well as to quantifying the effectiveness in completing the tasks, which mostly lack difficulty and diversity. In this context, 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, roughly 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.

ID	Pub	Year	Application areas	Collaboration Details	Task type	Devices used (On-site User)	Devices used (Remote User)	Study type	Data Type	Study Design
1	(P. Wang et al., 2020)	2020	Assembly	Hierarchy - Synchronous	Lego Brick Assembly	Projector, External Camera	HMD, Hand Tracker	Formal	S	Within-subjects
2	(Rhee et al., 2020)	2020	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, Controllers, 360º camera	HMD, Controllers	Formal	S	Between-subjects
3	(Teo, Lee, et al., 2019)	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360º camera	HMD, Controllers, Hand Tracker	Formal	O + S	—
4	(Teo, Hayati, et al., 2019)	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation, Puzzle Assembly	See-through HMD, 360º camera	HMD, Controllers	Formal	O + S	—
5	(Sasikumar et al., 2019)	2019	Assembly	Hierarchy - Synchronous	Lego Brick Assembly	See-through HMD, Depth Sensors	HMD, Hand Tracker	Formal	O + S	Within-subjects
6	(Mahmood et al., 2019)	2019	Co-Design	Parallel - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD	See-through HMD,	Formal	S	Within-subjects
7	(Teo, Lawrence, et al., 2019)	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation, Lego Brick Assembly	See-through HMD, 360º camera	HMD, Hand Tracker	Formal	O + S	—
8	(Piumsomboon, Lee, et al., 2019)	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360º camera	HMD, Controllers	Formal	O + S	—
9	(Yoon et al., 2019)	2019	Social Presence	Parallel - Synchronous	Puzzle Assembly	See-through HMD	HMD, Controllers	Formal	S	Within-subjects
10	(P. Wang, Zhang, Billinghurst, et al., 2019)	2019	Assembly	Hierarchy - Synchronous	Lego Brick Assembly	Projector, Camera	HMD, Hand Tracker	Formal	O + S	Within-subjects
11	(G. A. Lee et al., 2019)	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360º camera	HMD, Hand Tracker	Formal	O + S	Within-subjects
12	(Piumsomboon, Dey, et al., 2019)	2019	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, Hand Tracker	HMD, Hand Tracker	Formal	O + S	Between-subjects
13	(Waldow et al., 2019)	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	(K. Kim et al., 2018)	2018	Assembly	Parallel - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Mouse and Keyboard	Formal	O + S	Within-subjects
15	(Teo et al., 2018)	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360º camera	HMD, Hand Tracker	Formal	O + S	Within-subjects
16	(S. Kim, Billinghurst, Lee, et al., 2018)	2018	Assembly	Parallel - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
17	(Congdon et al., 2018)	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HMD, Hand Tracker	HMD, Hand Tracker	Informal, Formal	S	Between-subjects
18	(Choi et al., 2018)	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HHD	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
19	(Yamada & Chandrasiri, 2018)	2018	Assembly	Hierarchy - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Hand Tracker	Formal	O + S	Between-subjects
20	(Günther et al., 2018)	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, External Camera	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
21	(Piumsomboon et al., 2018)	2018	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD	HMD, Controllers	Formal	O + S	Within-subjects

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

ID	Pub	Year	Application areas	Collaboration Details	Task type	Devices used (On-site User)	Devices used (Remote User)	Study type	Data Type	Study Design
22	(Ryskeldiev et al., 2018)	2018	Assistance	Parallel - Synchronous	Navigation, Object Selection and Manipulation	HHD	HHD	Formal	O + S	—
23	(Hoppe et al., 2018)	2018	Assistance	Parallel - Synchronous	Navigation, Object Selection and Manipulation	HMD, Controllers, Hand Tracker	HMD, Controllers, Hand Tracker	Formal	O + S	—
24	(Akkil & Isokoski, 2018)	2018	Assembly	Parallel - Synchronous	Puzzle Assembly	Projector, External Camera	Computer, Gaze Tracker	Informal, Formal	O + S	Within-subjects
25	(G. A. Lee, Teo, et al., 2017)	2017	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, 360º camera	HMD, Controllers, Hand Tracker	Formal	O + S	—
26	(G. A. Lee, Kim, et al., 2017)	2017	Assembly	Hierarchy - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Gaze Tracker	Formal	S	Between-subjects
27	(Komiyama et al., 2017)	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	(Chenechal et al., 2016)	2016	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD, Hand Tracker	HMD, Controllers	Informal	O + S	Between-subjects
29	(Gurevich et al., 2015)	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	(Tait & Billinghurst, 2015)	2015	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD	Computer, Mouse and Keyboard	—	O + S	Between-subjects
31	(S. Kim et al., 2015)	2015	Assembly	Parallel - Synchronous	Puzzle Assembly	See-through HMD, External Camera	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
32	(Tait & Billinghurst, 2014)	2014	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	See-through HMD	Computer, Mouse and Keyboard	—	O + S	Between-subjects
33	(S. Kim et al., 2014)	2014	Assembly	Parallel - Synchronous	Puzzle Assembly	HHD or See-through HMD	Computer, Mouse and Keyboard	Formal	O + S	Between-subjects
34	(Gauglitz et al., 2014a)	2014	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HHD	Computer with Touch screen	Informal, Field	S	—
35	(Gauglitz et al., 2014b)	2014	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HHD	Computer, Mouse and Keyboard	Informal, Field	O + S	Within-subjects
36	(Huang et al., 2013)	2013	Assembly	Hierarchy - Synchronous	Puzzle Assembly	Monitor, External Camera	HMD	Formal	O + S	—
37	(Pece et al., 2013)	2013	Co-Design	Parallel - Synchronous	Navigation, Object Selection and Manipulation	HHD	HHD	Formal	O + S	—
38	(Poppe et al., 2012)	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	(Gauglitz et al., 2012)	2012	Assistance	Hierarchy - Synchronous	Airplane Cockpit	HHD	Computer, Mouse and Keyboard	Formal	O + S	Within-subjects
40	(Barakonyi et al., 2007)	2007	Education	Parallel - Asynchronous	Navigation, Object Selection and Manipulation	Computer, External Camera	Computer, External Camera, Gaze Tracker	Informal	S	Between-subjects
41	(Bannai et al., 2006)	2006	Assistance	Hierarchy - Synchronous	Navigation, Object Selection and Manipulation	HMD, External Tracker	HMD, External Tracker	Formal	O	—
42	(H. Regenbrecht et al., 2004)	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 1 - II. Legend: S- Subjective; O- Objective; HHD- Handheld Device; HMD- Head Mounted Display.

ID	Pub	Evaluation Methods	# Participants (# Females)	Participant Role	Participants knew each other	Previous experience with VR/AR/MR	Description experimental context	Adaptation period	Duration (min)	Recording audio and video
1	(P. Wang et al., 2020)	Questionnaires, Interview	34 (11)	On-site or Remote	Yes and No	—	—	Yes	55	—
2	(Rhee et al., 2020)	Questionnaires, User Preference	40 (-)	On-site or Remote	—	Yes	Yes	Yes	40	—
3	(Teo, Lee, et al., 2019)	Task Performance, Questionnaires, User Preference	32 (8)	On-site or Remote	—	Yes	—	—	—	—
4	(Teo, Hayati, et al., 2019)	Task Performance, Questionnaires, User Preference	10 (1)	On-site or Remote	—	Yes	Yes	Yes	—	—
5	(Sasikumar et al., 2019)	Task Performance, Questionnaires, User Preference	10 (4)	On-site or Remote	—	—	—	—	—	—
6	(Mahmood et al., 2019)	Questionnaires, User Preference	8 (4)	On-site or Remote	—	—	Yes	—	30	—
7	(Teo, Lawrence, et al., 2019)	Task Performance, Questionnaires, User Preference	14 (-)	On-site	—	—	Yes	—	70	Yes
8	(Piumsomboon, Lee, et al., 2019)	Task Performance, Questionnaires, User Preference	24 (5)	On-site	—	Yes	Yes	—	90	—
9	(Yoon et al., 2019)	Questionnaires, Interview	48 (24)	On-site	Yes	Yes and No	Yes	—	—	—
10	(P. Wang, Zhang, Billinghurst, et al., 2019)	Task Performance, Questionnaires, User Preference	13 (5)	On-site or Remote	—	No	—	Yes	—	—
11	(G. A. Lee et al., 2019)	Task Performance, Questionnaires, User Preference	12 (3)	On-site and Remote	Yes	Yes	Yes	—	—	—
12	(Piumsomboon, Dey, et al., 2019)	Task Performance, Questionnaires, Interview	32 (9)	On-site and Remote	Yes	Yes and No	Yes	Yes	120	—
13	(Waldow et al., 2019)	Task Performance, Questionnaires	20 (5)	On-site or Remote	No	—	Yes	Yes	35	—
14	(K. Kim et al., 2018)	Task Performance, Questionnaires, User Preference	24 (7)	On-site or Remote	Yes	—	—	—	—	Yes
15	(Teo et al., 2018)	Task Performance, Questionnaires, User Preference	8 (2)	On-site	—	Yes	Yes	Yes	—	—
16	(S. Kim, Billinghurst, Lee, et al., 2018)	Questionnaires, Interview	24 (4)	On-site or Remote	Yes	—	Yes	Yes	—	—
17	(Congdon et al., 2018)	Questionnaires, User Preference	38 (23)	On-site or Remote	—	—	Yes	Yes	30	—
18	(Choi et al., 2018)	Task Performance, Questionnaires	30 (4)	On-site	—	—	—	—	—	
19	(Yamada & Chandrasiri, 2018)	Task Performance, Questionnaires, User Preference	10 (0)	On-site	—	—	—	—	—	—
20	(Günther et al., 2018)	Task Performance, Questionnaires, User Preference	8 (4)	On-site or Remote	—	—	Yes	—	60	Yes
21	(Piumsomboon et al., 2018)	Task Performance, Questionnaires, User Preference	16 (5)	On-site or Remote	—	Yes	—	—	—	—

Table 4 – Summary of User studies in Remote Collaboration using AR or MR - Part 2 – I.

ID	Pub	Evaluation Methods	# Participants (# Females)	Participant Role	Participants knew each other	Previous experience with VR/AR/MR	Description experimental context	Adaptation period	Duration (min)	Recording audio and video
22	(Ryskeldiev et al., 2018)	Task Performance, Questionnaires, User Preference	40 (-)	On-site or Remote	—	—	Yes	—	—	—
23	(Hoppe et al., 2018)	Task Performance, Questionnaires, User Preference	28 (-)	On-site and Remote	—	Yes	—	—	—	—
24	(Akkil & Isokoski, 2018)	Task Performance, Questionnaires, User Preference	24 (16)	On-site or Remote	—	—	Yes	Yes	—	—
25	(G. A. Lee, Teo, et al., 2017)	Task Performance, Questionnaires, User Preference	8 (-)	On-site and Remote	—	—	—	—	—	—
26	(G. A. Lee, Kim, et al., 2017)	Questionnaires, User Preference	8 (2)	On-site or Remote	Yes	—	—	—	—	—
27	(Komiyama et al., 2017)	Questionnaires, User Preference	8 (-)	Remote	—	—	—	—	—	—
28	(Chenechal et al., 2016)	Task Performance, Questionnaires	10 (-)	On-site	—	—	—	—	—	—
29	(Gurevich et al., 2015)	Task Performance, Questionnaires	13 (-) + 24 (-)	On-site or Remote	—	—	—	—	—	—
30	(Tait & Billinghurst, 2015)	Task Performance, Questionnaires, User Preference	36 (15)	On-site or Remote	—	—	—	—	—	Yes
31	(S. Kim et al., 2015)	Task Performance, Questionnaires, Interview	24 (7)	Remote	—	—	Yes	Yes	70	—
32	(Tait & Billinghurst, 2014)	Task Performance, Questionnaires, User Preference	—	Remote	—	—	—	—	—	Yes
33	(S. Kim et al., 2014)	Task Performance, Questionnaires, User Preference	24 (7)	On-site or Remote	Yes	—	Yes	Yes	90	—
34	(Gauglitz et al., 2014a)	Questionnaires, User Preference	25 (-) + 11 (5)	—	—	—	—	—	50	—
35	(Gauglitz et al., 2014b)	Task Performance, Questionnaires, User Preference	20 (-) + 60 (29)	On-site	—	—	—	—	—	—
36	(Huang et al., 2013)	Task Performance, Questionnaires, User Preference	14 (-)	On-site and Remote	—	—	Yes	—	—	—
37	(Pece et al., 2013)	Task Performance, Questionnaires, User Preference	36 (-)	On-site or Remote	—	—	Yes	—	—	—
38	(Poppe et al., 2012)	Interview	5 (-)	On-site	—	—	—	—	—	Yes
39	(Gauglitz et al., 2012)	Task Performance, Questionnaires	48 (21)	On-site	Yes and No	—	Yes	Yes	—	—
40	(Barakonyi et al., 2007)	Questionnaires, Interview	9 (3)	On-site	—	—	—	Yes	20	—
41	(Bannai et al., 2006)	Task Performance	12 (2)	On-site or Remote	—	—	—	—	—	—
42	(H. Regenbrecht et al., 2004)	Questionnaires, User Preference	27 (8)	On-site or Remote	—	—	Yes	—	—	—

Table 5 – Summary of User studies in Remote Collaboration using AR or MR - Part 2 - II.

Common approaches:

- 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

What is missing:

- conduct outdoor and field studies
- explore complex/adequate tasks
- contemplate failure situations
- provide an adaptation/training period
- address participants relationships, knowledge and motivations
- better description of the collaborative process supported by AR
- reporting of study average duration
- data collection on dialogue 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

Table 6 – Overview of common approaches and missing steps regarding the evaluation process of remote collaboration mediated by AR.

The evaluation design is distributed between within-subjects and between-subjects. The majority of the studies conducted are formal studies, collecting objective and subjective data at the end of the tasks using standard practices namely fixed answers, scale-based questionnaires (e.g., System Usability Scale (SUS), NASA Task Load Index (TLX), among others) or direct observation. Only a reduced set of studies include measurements collected during the collaborative process (e.g., duration and error/accuracy), as well as explicit communication (e.g., spoken messages or gestural cues), ease of collaboration and information gathering (e.g., basic awareness, eye gaze). Although the collection of a reduced subset of contextual and behavioral data can already contribute to the understanding of the evaluation outcomes, the collection of more data is often not considered or hindered due to the complexity it entails regarding acquisition, processing and analysis.

Interviews and audio/video recordings are seldom used during the studies and in roughly half of the studies the experimental context is not described. Only one third of the times studies referred the existence of an adaptation period. The lack of adaptation/training periods is an important factor as it affects directly the collaboration process, i.e., those that had an opportunity to use, train and comprehend the technology prior to performing the tasks will interact much better with their respective counterpart when compared to the users that have only done the adjustment process during the task realization itself.

Figure 19 – Overview of the main results from the recent literature review on User studies 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.

Another observation is that single-user evaluation methods are often applied directly to collaborative tasks resulting in the comparison of technological aspects or interaction mechanisms based on rather simpler procedures. We argue that collaborative tasks must be complex and with an adequate duration to encourage significant interaction between collaborators and that many studies involving short-term simple tasks (Lego assemblies) are insufficient to evaluate correctly the collaboration effectiveness. In this line, tasks may also include deliberate drawbacks and constraints, i.e., incorrect, contradictory, vague or missing information, to encourage more complicated situations and elicit collaboration. Simple examples could be suggesting the use of an object which does not exist in the environment of the other collaborator or ask to remove a red cable, which is green in the other collaborator context. Such situations introduce different levels of complexity, resulting in more realistic real-life situations where the environment is complex, the information might not be available and unknown events can occur frequently.

2.5 Critical Analysis

This section describes the main limitations hindering a better understanding regarding how AR may support collaborative work in remote scenarios. This analysis is mostly based in the literature review complemented by meetings with domain experts, and the author's own experience creating prototypes and conducting user studies (Alves et al., 2018; Alves, Marques, Neves, et al., 2019; Madeira et al., 2020; Marques et al., 2021; Marques, Alves, et al., 2020; Marques, Silva, Dias, et al., 2021c; Marques, Teixeira, et al., 2020).

The analysis was 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 30 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.

2.5.1 Main limitations

Based on our literature review, meeting with experts and own experience, the following limitations were 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*" (Merino et al., 2020).

The existence of two or more collaborators makes it more difficult to evaluate the solution since it requires to perform multiple evaluations in parallel. The logistics associated with evaluations in remote scenarios are demanding since a significant number of variables must be considered (Patel et al., 2012). As a consequence, crucial aspects of collaboration are often left apart: how was the interaction 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 were able to use the AR-based solutions to their full potential, how the available information was used to complete successfully the tasks, among other aspects (11 out of 42 papers).

In this context, trying to apply conventional evaluation techniques to collaborative settings without adapting them can lead to a reduced vision of the process of collaboration and in turn to dubious results. Given the complex environments and situations collaborators may encounter, conventional single user methods alone provide insufficient information and rarely are good indicators for improving distributed solutions (Araujo et al., 2004; Hamadache & Lancieri, 2009; Neale et al., 2004; Pereira et al., 2014).

Limitation 2: Lack of contextual information

Remote collaboration represents high levels of data by involving different types of distributed collaborators, on common tasks and in encompassing dynamical environments with contextual data. Dey et al. revealed that "*work needs to be done toward 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*" (Dey, Billinghurst, Lindeman, & Ii, 2018). 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*" (Ratcliffe et al., 2021). 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.

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*" (Z. Bai & Blackwell, 2012). 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*" (Piumsomboon, Dey, et al., 2019). 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 due to problems in participants communication, too much augmented information being displayed, the actions of a particular collaborator that did not follow correctly 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* (Dey, Billinghurst, Lindeman, & Ii, 2018). 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*" (Ens et al., 2019). 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 affected the accomplishment of the tasks in relation to the collaborative effort.

Limitation 5: Limited support in existing Frameworks for evaluation

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 crucial aspects of collaboration (Antunes et al., 2014; Z. Bai & Blackwell, 2012; Dey, Billinghurst, Lindeman, & Ii, 2018; Ens et al., 2019; S. Kim, Billinghurst, & Lee, 2018; Marques, Teixeira, et al., 2020). For example, Bai et al. emphasized that "*it can be hard to isolate the factors that are specifically relevant to collaboration*" (Z. Bai & Blackwell, 2012). 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*" (Ens et al., 2019). 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*" (Ratcliffe et al., 2021).Therefore, integration of evaluation strategies, covering different contexts of use and complex tasks, running in its intended (real or simulated) environments is 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 of AR Collaborative systems 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, which may happen since most research efforts have been focused on creating the enabling technology.

2.5.2 Maturity of the field

To put in perspective the evolution of the field, this section concludes with an analysis of the status of the area according to the BRETAM model (Figure 20) (Gaines, 1991). 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 (Lalanne et al., 2009).

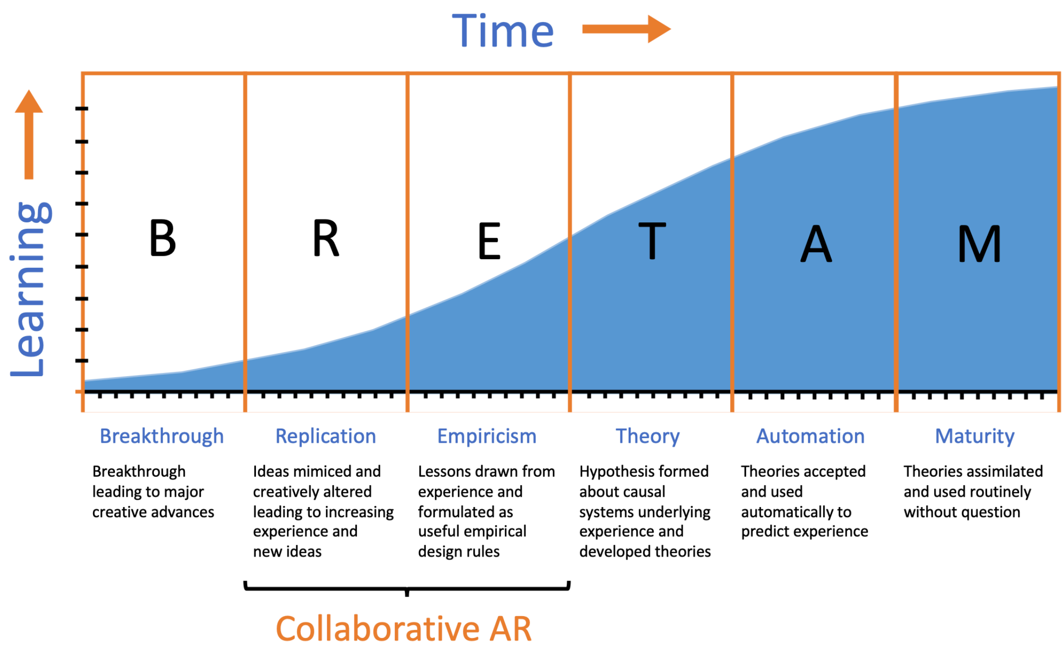

Figure 20 – Remote Collaboration mediated by AR positioning between the Replication and Empiricism phases of the BRETAM model. Inspired by (Gaines, 1991).

According to the current panorama, it is possible to situate remote collaboration mediated by AR between the **Replication** and **Empiricism** phases of the BRETAM model as illustrated in Figure 20. We argue that the field 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 (R. Belen et al., 2019). 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, Collaborative 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 (R. Belen et al., 2019; Ens et al., 2019). 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 (Ens et al., 2019). 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 (Gaines, 1991).

2.6 Roadmap for the characterization and evaluation of the collaborative process

According to the critical analysis, it is important to address the main limitations to carry the field to the Theory, Automation and Maturity phases of the BRETAM model (Gaines, 1991). With this objective in mind, this section we propose a roadmap to deal with the most pressing issues (Figure 21), composed by five key topics:

- definition of dimensions of collaboration to face the partial characterization of the collaborative process;
- systematization of perspectives on the acquired knowledge of the field, facing the lack of theories and guidelines;
- creation of new paradigms, architectures and frameworks to answer to the limited support of development and evaluation;
- 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.

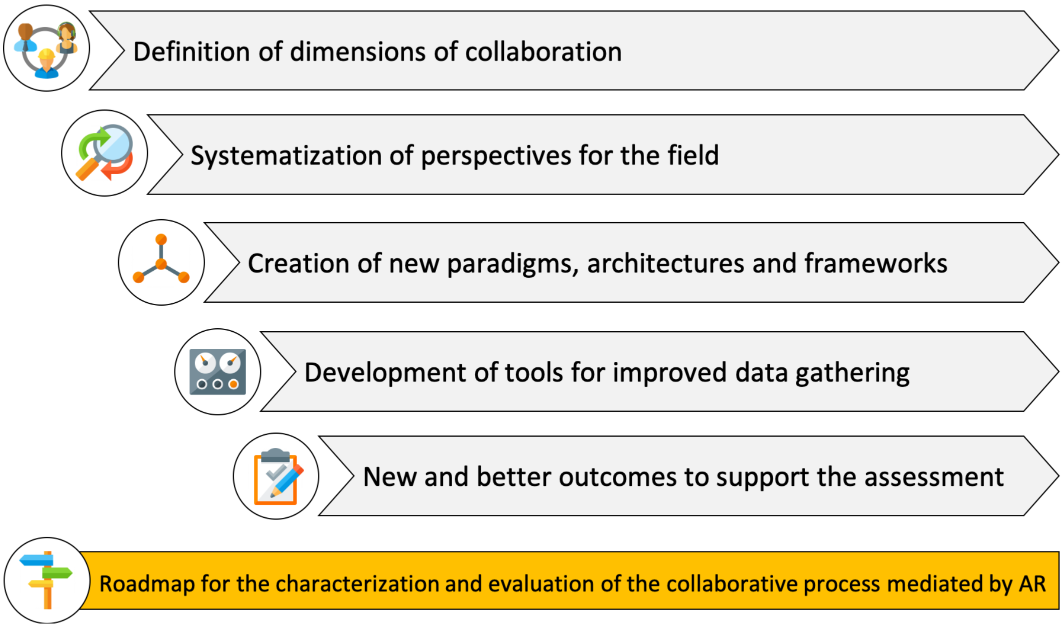

Figure 21 – Roadmap overview of the main topics that should be addressed regarding remote collaboration mediated by AR to make the field achieve the Theory, Automation and Maturity phases of the BRETAM model. Inspired by (Teixeira, 2014).

2.6.1 Definition of dimensions of Collaboration

First, it is important to identify the 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, new proposals must emerge to address new aspects related to collaboration. A comprehensive set of dimensions must be defined to classify more thoroughly 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*" (Ens et al., 2019). Therefore, the existing dimensions might still not reflect the full scope of some categories. 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 analysing the literature regarding collaborative work supported by AR, in particular, existing categorization efforts (Billinghurst et al., 2003; Brockmann et al., 2013; R. A. J. de Belen et al., 2019; Ens et al., 2019; Jalo et al., 2018; Sereno et al., 2020; Speicher et al., 2019; X. Wang & Dunston, 2006). 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 and to identify which dimensions should be considered 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.

2.6.2 Systematization of perspectives for the field

Ens et al report that when considering if it is possible to clearly describe categories of collaborative MR research based on 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*" (Ens et al., 2019).Therefore, another area of research to be addressed given the lack of theories and guidelines (Dey, Billinghurst, Lindeman, & Ii, 2018; Ens et al., 2019) is the need to bring these dimensions forward into conceptual models, guidelines, taxonomies and ontologies, that foster harmonization of perspectives for the field, thus creating a common ground for systematization and discussion (Brockmann et al., 2013; Collazos et al., 2019).

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. This aspect is paramount to ensure that the research adds to the body of knowledge and provides enough context and evidence to enable a transparent account (Sukumar et al., 2020) and transferability (Meyer & Dykes, 2019). These methods 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 (Augstein & Neumayr, 2019).

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 (Nickerson et al., 2013; Teruel et al., 2017). 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., 2020). It is also relevant to include application, 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 (Nickerson et al., 2013). Regarding the creation of 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 (Chandrasekaran et al., 1999; Noy & McGuinness, 2001). As such, one possible strategy is to adapt existing ontologies when they exist, or as an alternative 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.

2.6.3 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*" (Merino et al., 2020). As literature shows, there is no standard methodology for characterization and evaluation, specifically tailored to assess how remote collaboration occurs through AR/MR technologies (Ens et al., 2019). Without the appropriate paradigms, methods and mechanisms, the research community might not accumulate enough experience to improve collaboration between distributed collaborators (Antunes et al., 2014; Araujo et al., 2004; Z. Bai & Blackwell, 2012; Dey, Billinghurst, Lindeman, & Ii, 2018; Ens et al., 2019; S. Kim, Billinghurst, & Lee, 2018; Marques, Teixeira, et al., 2020; Merino et al., 2020; Ratcliffe et al., 2021).

Currently, the main focus is 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 fall short to mimic real scenarios of remote collaboration. As such, architectures and frameworks capable of supporting this new paradigm(s) must be created, to assist researchers conducting future user studies. Such frameworks must include support for:

- **evaluation scope** for individual and collective assessment by properly identifying which dimensions of collaboration will be evaluated;
- **collaborative challenges** to be performed, including specification of the user's minimum level of knowledge, definition of each collaborator activity, as well as definition of the procedures;
- **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;
- **data gathering** through the use of a distributed evaluation tool focusing on the dimensions proposed specifically for remote collaboration;
- **data analysis**, including inspection of what happened during the tasks, to understand how the collaboration process occurred over time.

2.6.4 Development of tools for improved data gathering

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 (Ens et al., 2019; Hamadache & Lancieri, 2009; Herskovic et al., 2007; Marques, Teixeira, et al., 2020; Merino et al., 2020). A better evaluation process must be supported by improved data collection and data visualization tools (Araujo et al., 2004, 2003), 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 run multiple evaluations at different locations simultaneously, following a distributed paradigm (Marques, Teixeira, et al., 2020). In this process, researchers should be able to define metrics, 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. (Pereira et al., 2015, 2016, 2014). 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, surrounding context and AR-based solution (Marques, Teixeira, et al., 2020). 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 in the evaluation of remote collaboration scenarios to help researchers when performing judgment over evaluation results.

2.6.5 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 may improve on existing guidelines for future user studies.

2.7 Summary

Collaborative AR solutions can be powerful tools for analysis, discussion and support of complex situations in scenarios of remote collaboration. By bringing different and sometimes opposing points of view together, such solutions can lead to new insights, innovative ideas, and interesting artefacts, while supporting sessions of collaboration between distributed team-members. However, most efforts have been devoted to creating the enabling technology for supporting the design and development of such solutions while the characterization and evaluation of the collaborative process has been left apart, although it is a crucial, very difficult endeavor.

This chapter presents a critical analysis supported by surveys on collaborative user studies mediated by AR, as well as a literature review on works ranging from 2000 to 2020. 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, 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. In the next chapters, we describe in detail our efforts towards the concretization of some of these research actions towards improving evaluation and characterization of the collaborative process that ultimately will boost the development and effectiveness of AR mediated collaboration.

3 Initial approach to Remote Collaboration using AR

You should never view your challenges as a disadvantage. Instead, it's important for you to understand that your experience facing and overcoming adversity is actually one of your biggest advantages.

Michelle Obama

In this chapter, a general introduction to the role of human operators and Augmented Reality (AR) for remote collaboration in the industry sector are presented. Due to an industrial collaboration under the SGH project, we focus on understanding the use of AR for remote maintenance through a literature analysis, and given the access to domain experts, through a participatory process, including motivations and identification of requirements for a specific case study. Based on these, the main features of an AR-based prototype are described, which was used during a user study to evaluate its use on a real use-case scenario. Finally, the main results are reported and discussed, and a summary table of the main outcomes of this chapter are presented.

3.1 Human operators and AR-based remote collaboration in the industry sector

Human presence can be considered essential to complete specific procedures through knowledge and experience sharing in the context of Industry 4.0, while also contributing to address unplanned situations and accidents as they occur (Hofmann & Rüsch, 2017; Kong et al., 2019; Liao et al., 2017). As Industry 4.0 takes shape, human operators experience an increased complexity of their everyday practices, compelling them to manage a variety of manual operations, be highly flexible in a very dynamic working environments, as well as learn from remote experts when additional knowhow not available on-site is required (Bottani & Vignali, 2019; Geng et al., 2020; Limbu et al., 2018; Longo et al., 2017; Lukosch et al., 2015). Thus, ensuring the conditions to support remote collaboration, the process of joint and interdependent activities performed to achieve a common goal, is of paramount importance for the fourth industrial revolution (Ens et al., 2019; Geng et al., 2020; S. Kim, Billinghurst, et al., 2020; S. Kim, Billinghurst, Lee, et al., 2018; Masood & Egger, 2020), in particular, in the field of training, assembly, quality control, repair or maintenance (Johson et al., 2015; K. Kim et al., 2018; Lukosch et al., 2015; Piumsomboon, Dey, et al., 2019)

One of the most promising innovation accelerators to support these needs is AR, being considered a key pillar of Industry 4.0 to facilitate the digitization of the manufacturing sector, with potential for a higher level of efficiency by speeding up the entire production chain (Bruno et al., 2019; de Souza Cardoso et al., 2020; Egger & Masood, 2020; Hernandez-de-Menendez et al., 2020; Masood & Egger, 2019, 2020; Quandt et al., 2018; Röltgen & Dumitrescu, 2020; Tzimas et al., 2019). Solutions using AR have been explored to provide a common ground environment among distributed collaborators, i.e., serve as a basis for situation mapping, making assumptions and beliefs visible, since it allows overlying responsive computer-generated information on top of the real-world environments, combining the advantages of virtual environments and the possibility for seamless interaction with the real-world objects and other collaborators (Bottani & Vignali, 2019; Bottecchia et al., 2009; Bruno et al., 2019; Choi et al., 2018; Egger & Masood, 2020; Geng et al., 2020; Hall et al., 2018; Masood & Egger, 2020; Palmarini et al., 2018; van Lopik et al., 2020; X. Wang et al., 2016). It is expected that AR will improve efficiency and accuracy of the performed tasks by enhancing the perception of the shared understanding (Choi et al., 2018; de Souza Cardoso et al., 2020; K. Kim et al., 2018; van Lopik et al., 2020), as well as collaboration times, knowledge retention, increased problem context and awareness (Fernández Del Amo et al., 2018; Jetter et al., 2018; Ong et al., 2008; Röltgen & Dumitrescu, 2020; X. Wang et al., 2016).

However, there is still plenty of room to investigate collaborative studies (R. Belen et al., 2019; Billinghurst, 2021; Dey, Billinghurst, Lindeman, & Swan, 2018; Ens et al., 2019). In particular, real implementations in industrial scenarios are challenging, since most of the research, so far, have been performed under controlled settings adopting simple tasks as proofs-of-concept with, as main objective to showcase what the technology can achieve rather than about its level of integration as part of a solution for a specific problem (Ens et al., 2019; Masood & Egger, 2019). While this is the case, there is also a lack of insights into how human operators use current AR-based solutions and the type of challenges they face in real industrial environments (Ashtari et al., 2020; Flotyński, 2021). We especially lack an understanding of motivations, needs, and barriers of the targeted users.

This landscape opens up the space for academia and industry to work side-by-side in obtaining an overview of existing challenges that need to be overcome, like the creation of AR-based solutions that meet the needs of human operators (Ardito et al., 2020; Egger & Masood, 2020; Flotyński, 2021), bring domain experts into the proposal and validation of such technology due to the value of their knowledge about the problem and the workflows, which may lead to increase the adoption of such technologies by a larger audience, who might not be experts in AR technology. Therefore, understanding domain experts needs to be better integrated with the design and development processes (Ferrari et al., 2016; Hofmann & Rüsch, 2017; Kong et al., 2019; Liao et al., 2017) by intertwining human expectations and practices, as well as spaces and digital artifacts into cohesive interaction solutions for Industry 4.0 (Ferrari et al., 2016). It is important to contribute to support research that places AR in close relation with the collaborative contexts it aims to address and reflect on the extent of its contributions. The design focus must evolve and move from technology deployment to devise how the technology can augment human capacities as individuals or members of a team (Ferrari et al., 2016). Moreover, it is paramount to ensure that the research adds to the body of knowledge and provides enough context and evidence to enable a transparent account (Sukumar et al., 2020) and transferability (Meyer & Dykes, 2019), thus contributing to support the wide scope of challenges concerning Industry 4.0 and facilitate the digitization of the manufacturing sector.

Accomplishing these goals is not without its challenges. One of the most prominent aspects that needs to be tackled is the integration of domain experts in the process of designing and developing AR-based solutions, particularly in regard to understand how collaborative work mediated by AR can be accomplished in for remote maintenance. The work presented here includes the following main contributions: 1) a proposal for addressing these challenges and explore its applicability and suitability by working with domain experts to assess the extent to which AR-supported collaboration might be useful and contribute to remote maintenance; 2) identification of a list of relevant aspects, which can be the starting point for further evolving the design of new collaborative solutions and propose AR-based prototypes to assist in scenarios of remote maintenance; 3) a discussion about the lessons learned from a case study, which can be applied to other remote settings to facilitate the digitization of the industry sector.

3.2 Related work on AR for remote maintenance

Maintenance can be defined as an elaborated combination of activities that occur during the life cycle of an equipment, to return it or keep it in a state where it can perform the required function. It aims to ensure equipment performance, reduction of downtime and minimize disruption of production schedules. With the increasing complexity of industrial facilities due to the rise of Industry 4.0, maintenance processes play an extremely important role, improving competitiveness and contributing to sustainable development in Industry. Maintenance is a core activity of the production life-cycle, accounting for as much as 60 to 70% of its total costs (Mourtzis et al., 2017). Therefore, the provision of the right information to the right professional, with the right quality and in time is critical to increase efficiency (Fernández del Amo et al., 2018; Fiorentino et al., 2014; Zhu et al., 2014).

Unfortunately, some issues cannot be easily fixed by on-site technicians alone, and an in-depth analysis with experts is required. However, skilled specialists are usually in short supply due to the time required for these individuals to obtain such expertise. Moreover, such kind of intervention can be expensive and sometimes requires travels of the experts to reach the location where the maintenance tasks must be performed. As such, remote collaboration using AR among off-site experts and on-site technicians is a prominent topic in current research (Egger and Masood 2020) for dealing with the increasingly complex maintenance procedures.

For more than two decades, the field of CSCW has been concerned with designing solutions to support remote maintenance (Grudin & Poltrock, 2011; Lukosch et al., 2015), sometimes referred to as “collaborative maintenance” or “remote assistance” (Palmarini et al., 2018).

The most common solution is the use of video conference systems, which are widely available and easily accessible (K. Kim et al., 2018; S. Kim, Billinghurst, & Lee, 2018). Unfortunately, with this technology, collaborators are limited to passively watching video feeds with no means for interaction with the remote physical environment (Gauglitz et al., 2014b). Such systems only allow assistance through verbal cues or hand gestures in response to a visual feed (S. Kim, Billinghurst, Lee, et al., 2018; Ranatunga et al., 2013). Another constraint of video conference systems is the limited ability to reference areas of interest or specific objects on the environment, i.e., it can become ambiguous or vague, leading to confusion and errors, since video conferencing is not suitable for conveying spatial information (R. Belen et al., 2019; S. Kim et al., 2015; S. Kim, Billinghurst, Lee, et al., 2018). Because these systems do not support the same level of awareness as co-located collaboration, professionals tend to adopt time consuming, complex verbal negotiations to communicate their intended directions and achieve a common goal (Fakourfar et al., 2016; S. Kim, Billinghurst, & Lee, 2018).

As an alternative to video conferencing, AR has been investigated to combine knowledge between distributed professionals (Ens et al., 2019; Jalo et al., 2018; Liao et al., 2017). The concept of Collaborative AR can be described as an AR system where: ”multiple users share the same augmented environment locally or remotely and which enables knowledge transfer between different users” (Jalo et al., 2018).

AR-based solutions can be used in situations where knowhow and additional information from professionals unavailable on-site is required (Gurevich et al., 2015; Lukosch et al., 2015; Ong et al., 2008; Teo, Lawrence, et al., 2019). Remote professionals can add augmented visual communication cues to enhance a scene as it is captured by an on-site professional and provide real-time spatial information about objects, events and areas of interest (Fakourfar et al., 2016; Gurevich et al., 2015; S. Kim et al., 2019; S. Kim, Billinghurst, Lee, et al., 2018). By creating a common ground environment, such solutions can provide a shared understanding, i.e., enhance alertness and awareness, improve the overall (level of) understanding of the working situation, as well as contribute faster and more accurate task completion (Bottani & Vignali, 2019; Choi et al., 2018; Ens et al., 2019; Hall et al., 2018; S. Kim, Billinghurst, Lee, et al., 2018; Palmarini et al., 2018).

A number of studies have explored different methods to improve mutual work understanding, task efficiency, and information sharing (S. Kim, Billinghurst, & Lee, 2018; Lukosch et al., 2015). Most systems rely on virtual annotations to augment 2D images or live video with drawings, pointers, or pre-defined shapes (e.g., arrows, circles, and others) (Choi et al., 2018; S. Kim, Billinghurst, Lee, et al., 2018; Lorenz et al., 2018). Annotations are an essential interaction method in daily life, being used to summarize and highlight important elements of the physical environment or to add reminders, explanations or messages for others. A step further in the virtualization of annotations has been achieved thanks to the development of AR technology, as it is a powerful way of offering users more information about the real-world surrounding them (García-Pereira. et al., 2020). An example was proposed by (Masoni et al., 2017) based on off-the-shelf mobile devices and a desktop computer to remotely connect a remote expert with an unskilled worker performing maintenance procedures on an internal combustion engine of a car. The local user could capture a picture of the environment, use it as a visual marker and share it with the remote expert. Then, the skilled remote collaborator can annotate the received photo on the desktop computer, selecting what kind of feedback to send (based on common operations in maintenance: unscrew, screw, indications, warning, disassemble and assemble), sketches, and notes.

To complement, (Aschauer et al., 2021) proposed a solution using video stream sharing in a similar setup as previous approaches, in which freezing and unfreezing functions were integrated with the annotation's features. Touching the video freezes the live stream and provides drawing features. Afterwards, the solution switches back to the live video view, which shows the annotations created in the environment. Besides, voice and message chat were also available for communication between collaborators.

To provide on-site technicians with a hands-free approach while conducting maintenance procedures, annotations may also be visualized using a see-through HMD, as described in (Madeira et al., 2020). After capturing the on-site technician context and share it with a remote expert to provide assistance, hand tracking can be used to manipulate the annotations, enabling the adjustment of their position and scale in the real-world according to the context, thus enriching the on-site professional experience and improving visualization of instructions.

Another approach consists in using 3D shared models related to the on-site worker context, i.e., take advantage of pre-existing virtual objects, also known as virtual replicas (or digital twins in industry 4.0 terminology) to provide assistance among distributed collaborators. For example, (Oda et al., 2015) presented a solution for guiding an on-site worker during interventions in an aircraft combustion engine. The remote expert has access to a virtual replica of the physical object, that he/she can manipulate and add annotations to, thus providing situated instructions. This approach enables high accuracy since it involves 3D representations when compared with the traditional image-based 2D approaches. Nevertheless, it needs to be adapted to each new context, since 3D models must exist for each new situation requiring each relevant physical object in the on-site worker environment to be modelled and tracked to be used in the expert's virtual environment.

The use of 3D shared models has also been explored to assist in robotic scenarios, as described by (Mourtzis et al., 2017). The solution focuses on a cloud implementation to facilitate communication between on-site technicians and remote experts by sharing maintenance instructions based on pre-existing 3D models. The remote expert uses CAD models of the products in the cloud database that allows the technicians to interact with the shared models. After the instructions are created, technicians are notified, downloads them and can proceed with the maintenance task.

Although current literature reports initial efforts towards the creation of AR-based prototypes, these efforts still rely on exploring how current technology can be directly applied to remote maintenance. However, in order to properly support the challenges faced by human operators in such tasks, it is paramount to understand how AR technology can better assist them, which means specialists must be included in the design and development of new cohesive solutions for such scenarios. In this line of thought, domain experts may contribute to the process of providing context and increasingly realistic requirements to challenge and assess the capacities of AR technology in responding to real collaborative scenarios.

3.3 Understanding AR Remote Maintenance through a participatory process

In this section we present the methodology adopted to identify the needs from an Industrial context regarding remote collaboration and discuss our findings derived from the focus group in light of the relevant literature. We select a maintenance context due to its impact on work methodologies and benefited from an on-going collaboration with partners from the industry sector. The work reported here contributes to demonstrate the importance of applying such approaches, as described in the next sections.

3.3.1 Methodology

To understand how collaborative work is accomplished and how it can affect the design of solutions using AR, we propose a methodology, comprising four steps as shown in Figure 22. Step 1 consists in the identification of industrial needs characterized by a desire for knowledge sharing between experts and on-site technicians (Gattullo et al., 2019). In this context, we capitalized on a framework of tools and features for AR-supported collaboration, which resulted from the experience of our research group in creating and testing different technologies (Marques et al., 2018; Marques, Dias, Rocha, et al., 2019) and methods, mostly proposed in the scope of user-centered design approaches, over the years. This was harnessed to create storyboards on the possibilities of AR resources and their potential use in collaborative contexts. Using a strategy focused on an existing framework enabled a very low-resource approach to the creation of tangible concretizations of some of the concepts and features in discussion allowing to materialize ideas and providing a common language among all individuals involved in the discussion, i.e., researchers on AR and collaboration and experts in maintenance in remote scenarios. Step 2 requires adaptation and integration of the defined requirements into the maintenance prototype, thus providing it with collaborative capabilities. Step 3 implies the creation of the necessary architecture to support interaction with the shared context. Step 4 enables iterative refinement of the prototype through evaluation with different targeted audiences.

3.3.2 Focus Group with Domain Experts

Among the available elicitation techniques, focus groups and interviews with stakeholders are considered among the most effective for knowledge transfer (Ceyssens et al., 2020; Ferrari et al., 2016). In this context, we establish a user-centered methodology through participatory design, i.e., by actively involving stakeholders in the design process (Barnum, 2010; Jacko, 2012; Krauß et al., 2021) to understand how AR could be leveraged for remote collaboration, and how it can address professionals' expectations to ensure our work meets their needs and is usable. We conducted a focus group with eight domain experts (Figure 23) to collect qualitative data (Ferrari et al., 2016; Jerald, 2015; Mani et al., 2016; Plummer-D'Amato, 2008; J. A. Smith, 2015).

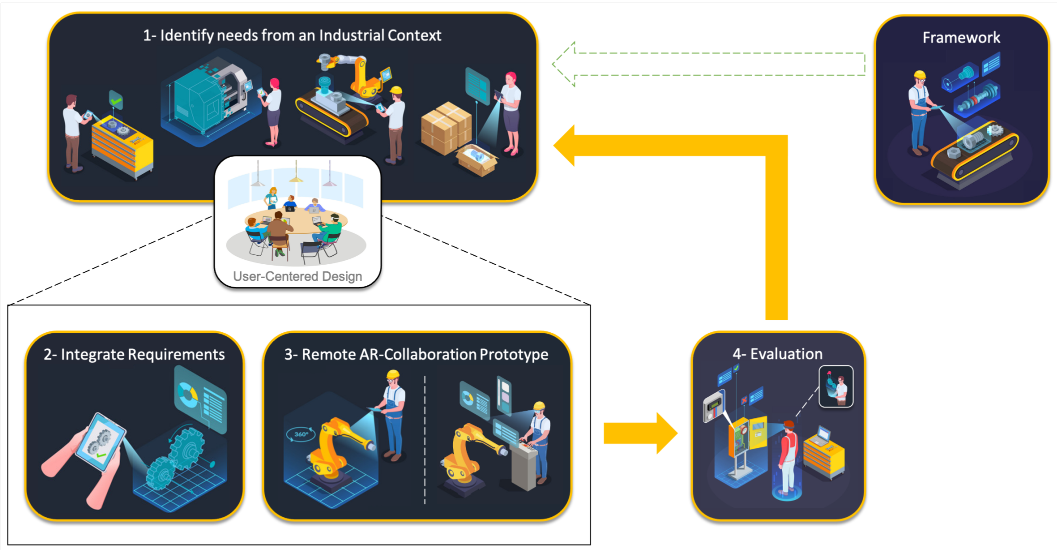

Figure 22 – Methodology adopted to bring domain experts into the understanding of how collaborative work is accomplished in an Industry context and how it may affect the design of collaborative solutions using AR. 1- A focus group was conducted to identify the needs from an Industrial context based on a framework in which tangible artifacts supporting the creation and discussion of storyboards were used; 2- this effort led to the creation of a set of requirements of relevant features suggested by the domain experts; 3- these requirements were fulfilled through the creation of a remote AR-based prototype for remote scenarios; 4- last, an evaluation was conducted following a set of tasks identified as relevant in maintenance contexts.

#	Occupation	Years of experience	Experience with collaboration tools	Experience with AR tools
1	Project Manager – Smart Factories, Automotive after market	6	Yes	Yes
	Previous: Software Developer	10		
2	Project Manager – Smart Homes, Smart Factories, IoT	25	Yes	Yes
3	Remote Support and Technical Instructor	20	Yes	No
4	Designer and UI Developer	6	Yes	Yes
5	Software Tester, Quality Assurance	7	No	No
6	Associate Professor and Researcher in HCI, VR, AR, IV	30	Yes	Yes
7	PhD Research Fellow in CSCW, VR, AR, HCI	5	Yes	Yes
	Previous: Industrial maintenance technician	9		
8	PhD Research Fellow in 3D reconstruction, AR, HCI	3	Yes	Yes

Figure 23 – Profile of the participants of the focus group session, including: project managers, technicians for remote support, UI designer, software tester and quality assurance engineers, as well as an associated professor and two PhD research fellows.

3.3.3 User motivations and context of collaboration

The focus group allowed to identify user motivations and context of collaboration (Figure 24) to discern what drives the different stakeholders (Miaskiewicz & Kozar, 2011; Tullis & Albert, 2013):

- **Technical Instructor**

  Motivation: Elicit performance of tasks between co-located workers, favoring the acquisition of knowledge, new skills, and the development of attitudes appropriated for the professional context.
  Technological literacy: medium to high.

- **On-site Technician**

  Motivation: Conduct maintenance and repairs on facility or domestic equipment's. They often require knowhow and additional information from professionals unavailable on-site.
  Technological literacy: low to medium.

- **Remote Expert**

  Motivation: Ensure assistance to on-site technicians using different mechanisms according to the complexity of the problem.
  Technological literacy: medium to high.

Example of questions used during the focus group to elicit discussion
Can you tell us your experience regarding collaboration in your daily tasks? What kind is more common? Co-located/Remote? Have you experience both? Can you elaborate? What type do you use more? Synchronous/Asynchronous?
Do you frequently use Remote Collaboration? What is the periodicity? In what scenarios/environments/tasks do you use it? What is the locus of knowledge? What is the Team size/ Role structure? What is the movement and size of workspace? What is the Complexity of physical task?
Can you elaborate on the type of teams? Always with the same remote expert? What is the average duration of the task? What is the average duration of your interaction with the remote technician? What tools do you currently use? What is the desired communication method? Verbal, textual, graphical, gestural, other?
What is the biggest challenge? Can you elaborate? How could the current tools be improved?
Are you familiar with the concept of AR? Can you elaborate? Do you believe it can be used in your daily tasks? How about as a tool for Remote-Collaboration? In what type of tasks do you consider it can be most useful?

Figure 24 – Example of questions used during the focus group to elicit discussion among the participants, which focused on understanding how collaboration is achieved.

Two contexts of collaboration were identified: co-located and remote collaboration (Figure 25). We focused on remote collaboration, this being our main interest and the scenario our partners had more experience and potential applicability:

- **Co-located collaboration:**

  Conducted in training situations between technicians and an instructor to promote the acquisition of new skills. Usually performed two months per year using text, images, videos, etc.

- **Remote collaboration:**

  A remote expert assists on-site technicians facing unfamiliar problems that require additional know-how. The size of the workspace is constant and focused on a specific equipment. In some cases, the technician must move around the physical environment (e.g., due to electrical connections). Three types of complexity tasks were identified:

  - **Simple issues:** Collaborators use synchronous communications through voice calls on handheld devices to help with simple procedures, e.g., locate a certain component in an equipment, which usually takes between 2 min and 2 hours;
  - **Moderate issues:** Collaborators share text and images in an asynchronous way, since the procedures require understanding the physical environment with more detail, e.g., installing a filter in a new equipment. The remote expert needs to use a graphic editor tool on a computer to create annotations based on drawings. Later, the on-site technician receives these instructions via email. This type of collaboration is frequently used when voice calls are insufficient to reach a solution, taking between 10 min to 90 min;

- **Complex issues:** Collaborators use synchronous sharing of text, images, video, and annotations, since the complexity of the procedures demands constant supervision and assistance, e.g., replacing an electronic board in an existing equipment, ensuring all connections are properly handed. A commercial tool[21] is used, in which the remote expert uses a computer, while the on-site technician a handheld device. Communication usually takes between 45 min and 120 min.

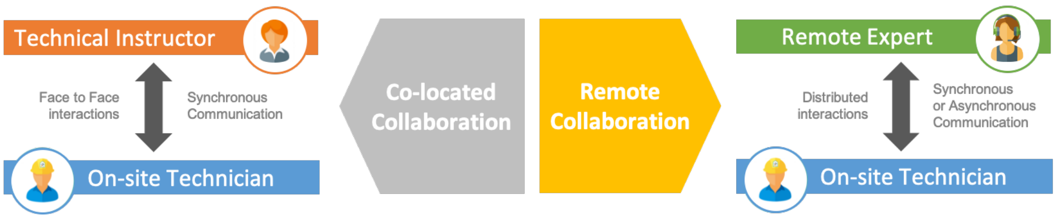

Figure 25 – Context of use obtained from a focus group session. Left: co-located collaboration using synchronous communication among a technical instructor and multiple on-site technicians. Right: remote collaboration using synchronous and asynchronous communication between an on-site technician and a remote expert.

3.3.4 Reflections on AR-based remote maintenance

Using AR technology in a maintenance context was introduced and discussed through storyboards and videos from our previous research, including mechanisms for visualization of components, presentation of step-by-step instructions, use of digital documentation, among others (Marques et al., 2018; Marques, Dias, Rocha, et al., 2019). This was considered extremely important to provide a visual overview of AR features that can be extended to remote collaboration.

Participants found it relevant to visualize situated AR-based content aligned with the real-world environment. They recognized AR can contribute to a better understanding of where to perform a given action. Displaying annotations on top of a region of interest was highlighted since currently they are unable to do so.

Two constraints were raised. First, the existence of large amounts of information in moderate and complex tasks, since these may involve several procedures, that last for significant amounts of time. The amount of information combined with the lack of means to view it aligned with the regions of interest creates confusion and periods of discussion between professionals, while trying to understand which information was created by who, as well as the order in which to consume it. Second, share step-by-step content could help minimize the problem of the amount of visual content, as well as serve as basis for re-visiting annotations created for a specific problem,

[21] - https://sightcall.com/visual-support/ [Accessed: 31-Mar-2021]

at a later time. Therefore, when a similar problem occurs, existing content may be re-used, saving time and authoring effort. Such an approach would be useful to create a kind of AR documentation that might be used with or without a remote collaborator.

When questioned about the use of other types of content, e.g., 3D models, participants stated their line of equipment’s feature more than 150 models, with thousands of individual components, which may hamper the modelling process. This large variety of models could affect the performance of a collaborative solution since the technicians face multiple contexts, in which internet connection may not be adequate to support sharing large amounts of data. Furthermore, they mention that the low technological literacy of their work force could entail training to understand how to handle such content. Given these limitations, they would give priority to a simpler, more generic solution for a wide range of scenarios.

Regarding hardware possibilities, participants emphasize their technicians are constantly moving, requiring easy-to-carry handheld devices, while also enabling the augmentation of annotations. Although this type of device requires the technician to place the device on a surface to perform a task, they reported that on-site technicians consider this as more natural, when compared to the use of a hands-free approach through HMDs, that would require additional training and adaptation periods as well as a significant investment in Hardware (HMDs, laptops, etc.). Regarding the remote experts, they consider a computer as the ideal device, but also find relevant having a handheld device for situations in which they work abroad, e.g., being present inside factories or warehouses.

Another relevant finding was the decision on whether AR is applied in remote scenarios depends on the complexity of the collaborative tasks. It would require a significant effort to create a solution, introduce it in the field (including training of professionals) and maintain it over time. In other words: “*applying AR to remote maintenance needs to be worth the effort*”.

3.3.5 Definition of requirements

From the feedback obtained, a set of requirements were outlined for the design of collaborative prototypes using AR, as illustrated in Table 7. The main purpose of sharing these requirements is to evidence the impact of the followed methodology in obtaining useful and detailed requisites that cover a wide range of features to support the collaborative effort.

Requirement	Description
Support for collaborative functions	The prototype should include standard built-in mechanisms providing functions such as:
• Capturing and sharing	Capture the real-world environment (on-site technician) or pause the video stream received (remote expert). Besides, it must support asynchronous (enhanced images) or synchronous (video) sharing of annotations.
• Communication	Support communication through audio, text, image(s), document(s) or video sharing. It could also be useful to represent the most common tasks through symbols and facilitate exchange of information among people of different languages and cultures without misunderstandings. In addition, it should not use voice-recognition, due to the potential disruption of noisy environment.
• Annotation whiteboard	Enable the use of different mechanisms to create annotation, following existing approaches as the use of drawing, notes, virtual objects and pointing to create instructions, highlight areas of interest, among others. It must not require 3D models, typically not available nor easily accessed.
• Sorting	Allow temporal ordering and clustering of information during its creation, especially important when an annotation with several instructions is received.
• Step-by-step instructions	A remote expert should be able to create step-by-step content to assist with a specific problem. Instead of sharing different enhanced images and repeating the process, he/she should be able to create multiple steps associated with a specific problem.
• Re-visit past actions	It should be possible to re-use annotations if the same failure occurs in another equipment, or with another team member. Plus, the annotations could also be a relevant source for manuals or documentation.
• Tracking and Augmentation	Must allow the visualization of annotations in standard 2D setting or as an augmentation of reality, i.e., situated annotations, as an additional layer of information, offering greater context.
• Awareness	A notification mechanism should exist to increase awareness, highlighting at each moment which collaborator is editing, or that new content is available, thus minimizing possible conflicts.
Modularity	Should be modular to support functions not anticipated that may arise. Functions must be easy to add, delete, or replace, since new technologies could lead to new approaches.
Scalability - handle multiple context	The prototype should comprehend multiple types of equipment's, ranging from small to large appliances.
Ease of learning and ease of use	Must be easy to learn, i.e., have a gentle learning curve with minimal impact on the collaborators workflow, aiming to reduce time and cost for flexible knowledge sharing between remote expert and novice technicians.
Persistence and Reliability	AR content cannot disappear or be corrupted. The prototype must ensure that they are available for re-use whenever a problem which was previously solved re-appears with a different team member.
Portability	The solution must guarantee multi-platform capabilities, being able to run on any operating system since team members may be exposed to a large variety of devices (e.g., handheld, desktop, laptop), with a different range of resources.
Performance	The architecture design of the prototype must pay special attention to satisfactory response time.

Table 7 – Functional and non-functional requirements for the creation of AR-based solutions.

3.3.6 Discussion

Based on the participatory process, it was possible to identify challenges and opportunities regarding remote collaboration in industrial scenarios, which otherwise would not be considered. In this section, we discuss these topics in light of the relevant literature, which suggest little research has been conducted on AR-based collaborative studies (Billinghurst, 2021; R. A. J. de Belen et al., 2019; Dey, Billinghurst, Lindeman, & Swan, 2018; Ens et al., 2019; Masood & Egger, 2019).

One possible limitation to a broad adoption of remote AR-based solutions is associated to the maturity level of such solutions, since most prototypes focus only in assisting specific situations, leading to proofs-of-concept limited to explore what current technology can achieve. As emphasized by the domain experts during the participatory process, this lack of adaptability to dynamic industrial scenarios may be one of the main reasons why existing remote AR-based solutions are not yet widely adopted by most companies. Despite the existence of AR-based prototypes, cohesive remote supporting tools for Industry 4.0 can only be better prepared to provide assistance by following user centered design methodologies, i.e., intertwining human motivations, expectations and practices, as well as knowing in advance the challenges these professionals face in real industrial environments. Therefore, the design focus must evolve and move from technology deployment to devise how the technology can augment human capacities.

Domain experts also suggest that although most studies reported in literature have focused on AR-based solutions for remote synchronous scenarios, it is also important to address remote asynchronous scenarios, in which collaborative actions take place at different times, since their off-site experts may not always be available at the exact moment assistance is required. Asynchronous scenarios present several research opportunities to complement existing AR-based solutions, namely study retention of produced information and its consumption at a later time, how multiple annotations are related and coexist within an environment to support a concrete set of actions, as well as the study of temporal sorting and clustering of information (Irlitti et al., 2017).

Another important topic is the creation of step-by-step annotations and re-use such type of instructions later in other collaborative sessions, if an identical task demands it, which may also be interesting for other scenarios of collaboration due to its clear advantages in time and resources. Furthermore, these features may also support content authoring, which remains a significant barrier to the wide spread use of AR in industrial scenarios (Bhattacharya & Winer, 2019; Fernández del Amo et al., 2018; Masood & Egger, 2020). The creation of step-by-step instructions may allow generation of documentation captured in the maintenance context, which can replace traditional manuals, without the need for content libraries or programming expertise.

Another essential point that stands out regards the use of 3D shared models. Although these solutions provide higher levels of detail for team-members, it was not considered the best alternative by the domain experts, mostly due to the fact that it requires the existence of digital twins for every maintenance scenario, which may not always be possible for companies that cover hundreds, sometimes millions of products. Likewise, it also requires a higher technological literacy to handle 3D aspects of said approach, which most workforces may not contemplate.

Equally important, despite the large speculation regarding the advantages of HMDs, the use of such devices requires careful analysis, since not all scenarios may benefit from their in-situ visualizations. The industrial partners stated that in the past they had surveyed part of their workforce on such topic, reporting that the majority of their on-site technicians preferred using traditional handheld devices, despite the obvious limitation of not offering a hands-free approach when conducting maintenance tasks. They argue that their price and required training does not make them the most suitable solution to the larger workforce's they represent.

Hence, combining the requirements derived from the participatory process with those arising from the literature provides a broad context that informs the contribution/significance of the work reported in this chapter, by intertwining human expectations and practices and digital artifacts into cohesive interaction solutions for Industry 4.0.

3.4 Prototype for AR-supported remote maintenance

Next, we describe **BARRT – collaBorative Augmented Reality for Remote supporT**, an effort towards the creation of an AR-based prototype for remote collaboration based on the requirements. The prototype aims to support scenarios that may require knowhow and additional information from professionals unavailable on-site, as is the case of maintenance scenarios. Therefore, it focuses on two types of users: on-site technicians and remote experts. Since on-site technicians are constantly moving, it seems adequate to equip them with easy-to-carry handheld devices (although the prototype also supports see-through HMD), while also enabling augmentation of annotations. Regarding the remote expert, the prototype supports multiple types of devices, including computers, interactive projectors, or handheld devices as illustrated in Figure 26 and Figure 27.

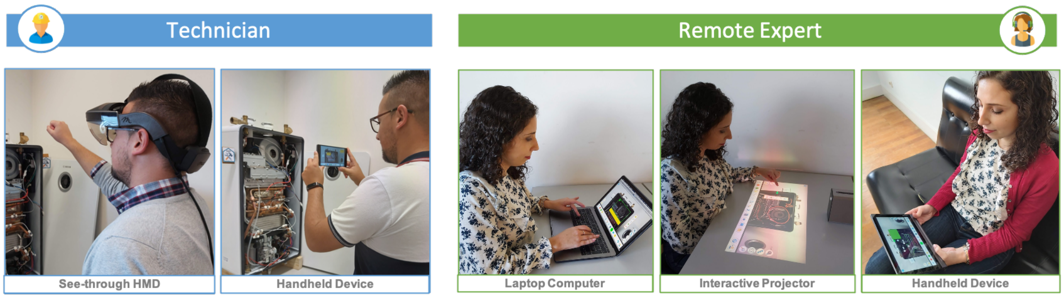

Figure 26 – Example of the multi-platform capabilities of the prototype. The on-site technician is able to use a handheld device or see-through HMD, while the remote expert can select between several devices: a computer, an interactive projector or a handheld device.

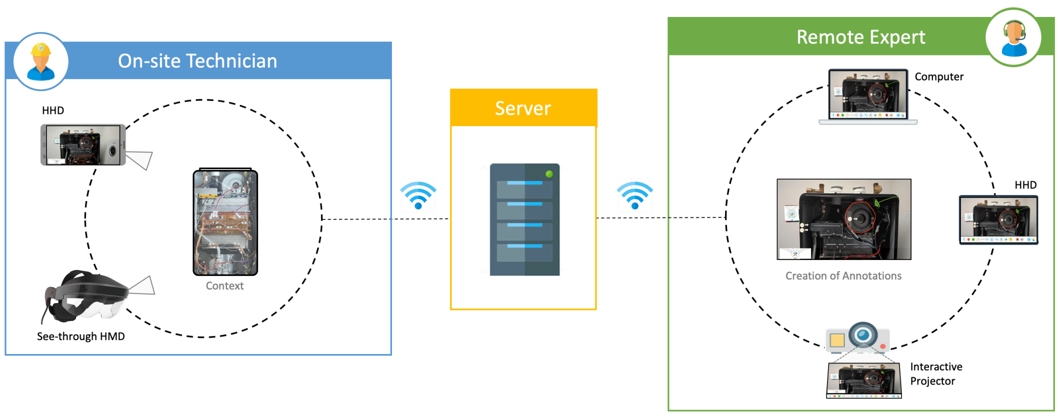

Figure 27 – Architecture overview to enable communication and interaction between distributed team-members. These are performed over Wi-Fi through specific calls to a centralized server responsible for storing and sharing the AR content accordingly (Madeira et al., 2020).

Figure 28 presents an overview of the prototype, which implements a subset of the requirements, in the spirit of an iterative user-centered approach. When facing unfamiliar problems, on-site technicians can point a handheld device at the situation that requires assistance and manually capture (freeze) its context. Then, using annotation mechanisms, he/she can edit the captured picture, creating layers of additional information to illustrate difficulties, identify specific areas of interest or indicate questions. Next, the enhanced picture is sent to the expert to provide a relevant illustration of the situation and enable the expert to suggest instructions accordingly using similar annotation features, plus some specific functions to facilitate the creation of content.

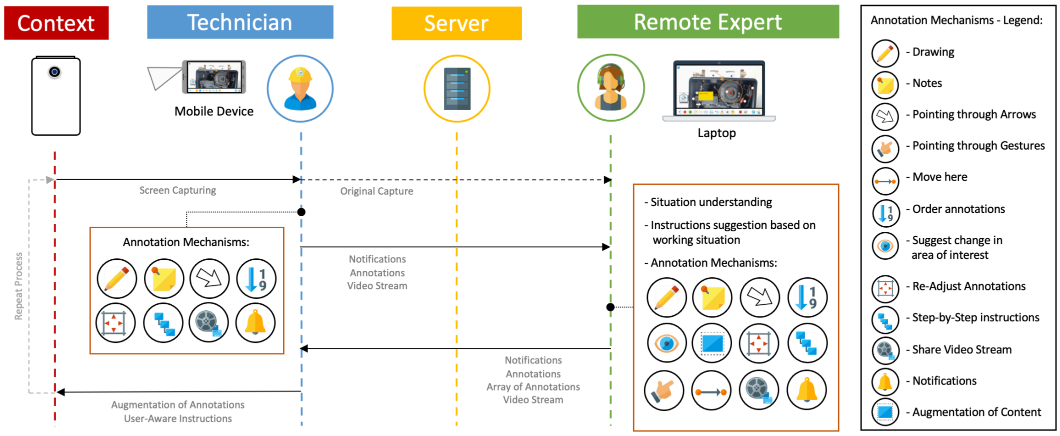

Figure 28 – Prototype Overview. Goal: Allow an on-site technician to capture the real world and use mechanisms to annotate it. Then, the content is shared with a remote expert for them to analyze and provide instructions (using identical mechanisms as those aforementioned). Finally, the technician can view the real world augmented with the instructions and perform an intervention Adapted from: (Marques et al., 2021).

Afterwards, the on-site technician receives the enhanced picture showing the annotations from the remote expert. Technicians can place a handheld device nearby and follow the instructions in a hands-free setting. At any time, the technician can pick up the device and perform an augmentation of the shared context, by re-aligning the annotations with the real world, thus receiving stabilized spatial information. Moreover, remote experts can receive video, manually freeze and annotate on the still video frame, rather than in a live feed, improving awareness and situation understanding. During this process, experts can generate content captured during real maintenance procedures and produce detailed documentation, which can be used for recording the procedure or reporting the current task progression. This process can be repeated iteratively until the task is successfully accomplished. Besides, audio communication is also available.

According to the team member role, the prototype provides a tailored set of functions, as illustrated in Figure 29. The use of shared images provides contextual information. Therefore, we added two mechanisms, one to **suggest capturing a specific region of interest**, thus improving awareness, and another for re-adjusting the shared image (e.g., rotate, scale, move).

Both collaborators can draw in different colors on top of the shared images. This enables them to highlight a specific component to be replaced by drawing distinct areas of interest or sketching an arrow. In addition, it is possible to add notes, such as relevant instructions i.e., important warnings or other contextual information.

Figure 29 – Example of the prototype functions associated to the on-site technician (left: drawing and notifications; augmentation of content; visualizing remote expert screen) and the remote expert (right: sorting annotations; pointing through 3D gestures; creation of step by step instructions). Adapted from: (Marques et al., 2021).

Pointing can be extremely important to address several aspects of remote collaboration. To address this, the prototype allows **pointing using 2D arrows**, generating virtual arrows on a required location of the captured/shared image. These can be selected and manipulated i.e., modify size, rotation, position. The remote expert can also **point using 3D gestures**, for example to illustrate how to perform an action (e.g., indicate where to plug a specific wire). This is only available when using a computer and an external sensor (e.g., Leap Motion) for hand recognition.

Sorting annotations allows sequential generation of IDs, providing temporal information on how annotations should be analyzed and consumed, facilitating understanding of problems involving several instructions.

The remote expert can create **step-by-step instructions**, particularly relevant in asynchronous collaboration scenarios, where team members may be unable to cooperate/communicate simultaneously.

Both collaborators can **re-use previous annotations** from other sessions/teams, since they can be an important source for documentation, also allowing to reduce the response time. As such, besides being used for collaboration, annotations can also be leveraged to minimize the need for expert assistance if similar situations happen in the future by providing some augmented documentation. Specific sets of annotation sequences created to address a maintenance task can be stored in the server. As such, if the same malfunction may come up, a possible solution can be re-used to instantly recall existing AR sequences.

On-site technicians can visualize an **augmentation** of annotations using the pictures captured as a marker, i.e., situated instructions on the real-world environment as an additional layer of spatial information.

Notifications also exist, e.g., images, text, and sound to enable awareness between collaborators. This is especially important in synchronous collaboration, avoiding possible conflicts. A preview of the annotations is presented before an image is shared and a confirmation panel is displayed, allowing validation before sending.

Video streaming can be relevant when combined with other features, e.g., hand gestures, providing a richer source of situation understanding, allowing an on-site technician to view the hands of a remote expert, while he/she explains how to perform a given action.

The prototype was developed using the Unity 3D game engine, based on C# scripts. To place the virtual content in the real-world environment, we used the Vuforia library. Communication between the different devices was performed over Wi-Fi through specific calls to a PHP server responsible for storing and sharing the enhanced content accordingly.

3.4.1 User Study

We conducted a user study to evaluate the viability of the BARRT prototype in a real remote setting, identify usability constrains, and understand participants satisfaction. As a case study, we focused on a typical remote maintenance scenario, where an on-site technician (using a handheld device) had to perform a set of maintenance procedures on an equipment, while requiring assistance from a remote expert (using a laptop computer) (Figure 30). We defined a set of synchronous and asynchronous tasks with the assistance of our partners from the Industry sector, which resulted from analysing the most common procedures their professionals face.

Tasks

Participants would act as on-site technician and as remote expert[22], while an experimenter was the respective counterpart. **On-site tasks**: capture the equipment context and request which component must be replaced and how. Then, perform the instructions provided using the augmented annotations displayed on top of the equipment. During this process, the experimenter (acting as the remote expert) would force multiple iterations through several completion stages, resulting in the need for collaboration to fulfil the task; **Remote tasks**: instruct the on-site participant on how to install a new filter and deal with several components in the process, while suggesting which tools to use from a large set of options. Also, create a step-by-step guide on how to replace a specific component of the boiler. During this process, the researcher would also force collaboration by asking how to handle multiple aspects associated to the task, including re-visiting some aspects of some instructions to enforce the re-use of annotations.

[22] - tinyurl.com/annotationToolOverview [Accessed: 12-Apr-2021]

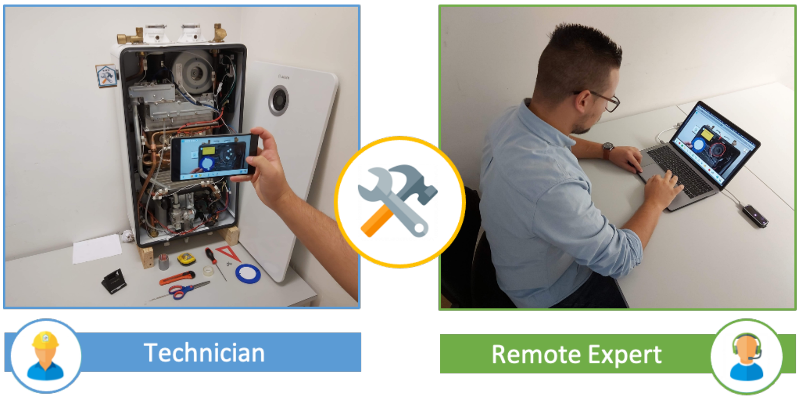

Figure 30 – Overall setup considered for the user study. The on-site participant (left) using a handheld device and the remote expert (right) using a laptop computer.

Procedure

Participants were instructed on the experimental setup, the tasks, and gave their informed consent. Afterwards, they were introduced to the BARRT prototype and a time for adaptation was provided, i.e., a training period to freely interact its functions. Then, the tasks were performed, while being observed by an experimenter who provided assistance, if necessary. After finishing, participants answered a post-experiment questionnaire.

Participants

Nine participants (3 female) performed the tasks and completed the post-experience questionnaire (although a sample of just 5 users is anticipated to find approximately 80% of usability issues (Nielsen & Landauer, 1993; Tullis & Albert, 2013)). For this stage of evaluation, we recruited participants from our University encompassing Faculty members, MSc and Ph.D. students, that had no prior experience with the defined case study but had experience with usability evaluation and collaborative tools (e.g., Skype, Team Viewer, etc.) in their daily activities, as well as in evaluating AR solutions.

Data Collection

Two types of data were collected. Task performance, namely the time needed to complete all procedures, logged in seconds by the device, and number of errors, logged by the device and an experimenter. The focus was to understand the time required to perform such tasks, and assess errors caused by communication issues or by malfunctions in our prototype; Participants' opinion was gathered through a post-task questionnaire that included demographic information and questions concerning collaborative aspects and through notes from a post-task interview to understand participants' opinion towards the collaborative process and to assess ease of use of the prototype features, as well as preferences.

Some examples of open-answer questions are illustrated in Figure 31. We decided to prioritize participant opinions at this stage and leave validated methods, such as the SUS, or NASA TLX for future studies with more experienced participants. The data collection was conducted under the guidelines of the Declaration of Helsinki.

Example of questions used in the post-task questionnaire
Do you believe AR can be used to improve collaborative work scenarios?
In what type of tasks do you consider it can be most useful?
To what kind of user (local or remote) do you think this features can be most advantageous? Why?
Were you able to collaborate with your counterpart?
Do the features of the prototype seem able to express my ideas properly?
Indicate which feature you believe could be most useful and why?
Describe other features which you consider relevant and were not presented.
Are there things you are dissatisfied with, that you would like to see changed?
The lack of video stream could be a limitation? Why?
Do you think using multiple visual communication cues would result in faster task completion time?
Do you think a notification mechanism is required to understand when the remote user is creating visual cues?
Please add additional comments you may consider relevant.

Figure 31 – Example of questions asked to the participants at the end of the study, during the post-task questionnaire.

3.4.2 Results and Discussion

All participants were able to collaborate using the AR-based features of BARRT. On average, each test lasted for 70 min (the tasks took 40 min to complete). They found relevant seeing AR-based annotations (Figure 32) and recognized it contributed to a better understanding of where to perform a given action, which facilitated communication and discussion. Moreover, they considered augmentation of content, drawing, creation of step-by-step and re-use of annotations as the most useful features and suggested the integration of voice-recognition into the prototype for command activation, a feature which was discarded as a priority by the domain experts, given the type of environments they usually face, which demonstrates that including the domain experts in the design and development processes helps to focus on the necessary functionalities to achieve collaborative work in multiple environments.

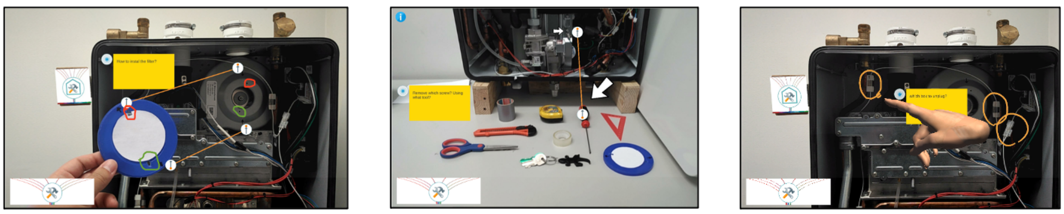

Figure 32 – Example of annotations created by the participants: how to install a new filter (left), suggest which tool to use (center) and identify which component must be unplugged (right).

Next, we present the main insights associated with each feature of the prototype. We chose to present issues and suggestions made by the participants, as well as possible solutions whenever they have already been implemented, highlighting the importance of using a user-centered approach to improve our prototype.

Participants enjoyed augmentation of content, e.g., annotations aligned with the real-world environment, recognizing it contributed to a better understanding of where to act and what to do. Participants pointed out that this feature requires the handheld device to be faced at the boiler to visualize the AR content, which might not be practical when performing some maintenance tasks that could require the use of both hands.

The possibility to freeze the video stream was also well received, since it gives more control to the remote expert. Most participants stated that although video enables sharing each step of the creation of the annotations, simple enhanced images would be enough to solve most simple collaborative problems. The only exception was the combination between video and the use of 3D gestures for pointing, which is much more useful with a video than with a still image.

Participants recognized they would use drawing often, being versatile to address most needs and suggested using different levels of line thickness. They also identified the need to display a preview of the annotations before sending them, which was already integrated into the prototype using a pop-up module (Figure 33).

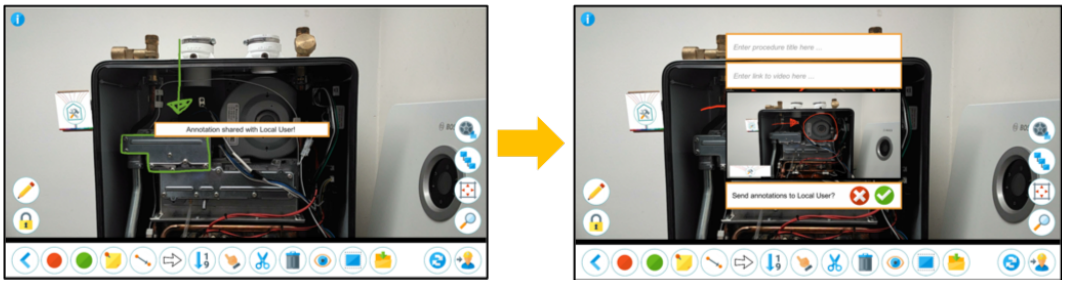

Figure 33 – Drawing: Interfaces before (left) and after (right) the inclusion of a mechanism to preview the annotations before being shared.

The use of notes was considered useful to share important messages, especially for the case of asynchronous communication conditions. Yet, participants highlighted longer text might not be practical to write or see on handheld devices.

The sorting function was considered important, as it fixes a problem which could become more relevant when a significant number of annotations exists. The possibility to select and re-adjust the order of specific annotations was also considered relevant.

Pointing through arrows was considered relevant to identify specific regions of interest. The enhancement of the selection and manipulation of this type of annotations was suggested in order to facilitate the creation of content. This was already integrated into the prototype using a pop-up module with shortcuts (e.g., rotate clock/counter-clockwise, scale and delete) (Figure 34). Participants also stated that the only reason they would use drawing instead of this feature, would be to create personalized arrows. Besides, it could be useful to have predefined shortcuts to other common shapes (e.g., circles, rectangles, etc.).

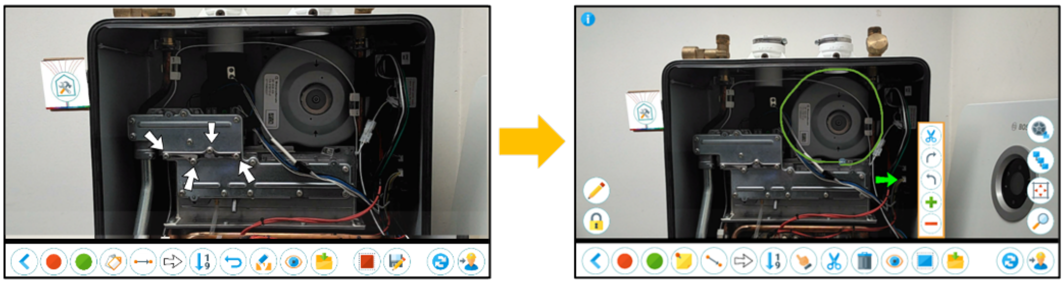

Figure 34 – Pointing through Arrows: Interfaces before (left) and after (right) the inclusion of a mechanism to facilitate selection and manipulation of virtual content.

Participants stated the step-by-step feature was useful and recognized its capacity to allow generating a set of simpler annotations, instead of larger ones with more visual content (Figure 35). Finally, re-visiting annotations created for a specific problem at a later time was considered interesting to help minimize the need for remote assistance in some cases.

Notifications were considered relevant to participants awareness during the collaboration process, in particular the use of sound to re-call attention for asynchronous situations, where the on-site team-member may be doing something else while waiting for the feedback of the remote expert.

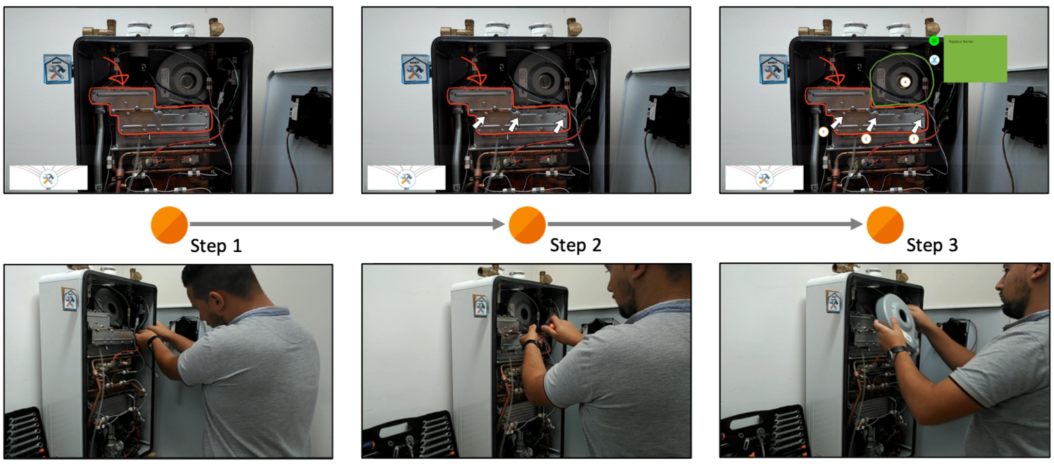

Figure 35 – Example of step-by-step instructions shared by the expert to assist in a maintenance task. Starting on the left, the on-site participant is provided with the identification of which component to remove through a red contour. Then, in the center, three arrows mark which screws must be removed. Finally, on the right, an order to do such activities is provided as well as identification to replace the boiler fan.

Issues identified

One of the main challenges identified is how to deliver contextualized information, i.e., how information can be shared without cluttering the users' field of view and without interfering with their task. This problem was referred to in some situations by the on-site participant when the remote annotations appeared in an intrusive way, thus occluding/cluttering important parts of the environment. A possible solution our prototype already supports as a result from the session of focus group discussion consists in using temporally situated data: step-by-step to create a full stack of operations showing only the relevant information at each step. Another challenge is ownership of virtual content. We must explore methods on how to present or discard information at a given moment, according to the collaborator's needs, aiming to support multiple team-members at once, e.g., 1 one-site technician and 3 remote experts.

Thus, the use of tangible artifacts to create a common language during the elicitation period with domain experts proves to be an advantage, which may indicate that research groups may capitalize on their work to create common ground for discussion with partners from the industry sector, aiming to better understand the needs of their workforce. The proposed methods and steps lead us to obtain new knowledge, as well as a first prototype that can be brought into a more realistic environment for evaluation, discussion, and refinement.

However, and as recognized by the community, experimental validation is often limited or absent when we address collaboration using AR technologies (R. A. J. de Belen et al., 2019; Dey, Billinghurst, Lindeman, & Swan, 2018; Ens et al., 2019; Kangsoo Kim et al., 2018; Marques, Teixeira, et al., 2020; Merino et al., 2020; van Lopik et al., 2020).

With the increasing interest in remote maintenance, it is imperative to develop better validation solutions and methodologies by/while trying to understand some of the following questions:

- Can we use/adapt evaluation methods from other domains?
- What tasks are relevant to evaluate these types of solutions, so that we can encompass the full complexity of the solution, its tasks and interaction capabilities?
- What aspects of collaboration should we consider?

Although this case study was conduct in simulated conditions, using real-life tasks, evaluation is even more relevant in real-life industrial environments, to understand if AR-supported solutions can be useful in such contexts, and in what conditions. This is challenging, as evaluation needs to provide measures that, in the long-term, also help decision makers have an increased quantification of its impact on industrial processes to inform the adoption of these technologies.

Table 8 presents a summary of the main insights of this work, following the work by (van Lopik et al., 2020), which presents a summary of issues and recommendations grouped according to known and emerging items regarding AR capabilities for industry 4.0, and considering that it is important that the community adopts more systematic methods to provide insights from the analysis conducted. We followed the same approach to address the main results of the case study we conducted, focusing on AR remote maintenance.

	Recommendations	Issues/Challenges
Known	• remote collaboration through AR can be a powerful tool for discussion and analysis of complex situations; • less intrusive devices are favoured so that on-site professionals can use both of their hands; • provide dependent and independent views of a shared environment; • retain collaborators actions through spatial information within leads to a more effective and efficient workflow.	• research has been limited by the capabilities of technology; • most studies don't conduct any type of evaluation of their solutions; • need to increase the ecological validity of evaluations; • lack of methods and guidelines to evaluate scenarios of remote collaboration mediated by AR; • existing frameworks are not sufficient to characterize how collaboration occurs through this new medium; • many aspects may affect the way teams collaborate, making it difficult for researchers to identify all variables related to the collaborative process.
Emergent	• include domain experts in the discussion of the suitability of AR for remote collaboration (due to the maturity of the field); • incorporate tangible artifacts to create a common understanding during identification of requirements; • minimal impact on the collaborator's workflow should be guaranteed by a gentle learning curve; • give priority to a simpler, more generic solution for a wide range of scenarios, which does not require 3D models; • annotation capabilities should be provided to on-site and remote collaborators; • allow temporal ordering and clustering of information during its creation; • notifications should be provided to increase collaborators awareness; • creation of step-by-step annotations should be considered in situations where large amounts of information must be exchanged; • cross referencing problems to reuse annotations established in a similar case; • customization should be available according to the collaborator profile; • provide on-site and remote workers with easy-to-carry handheld devices due to constant changes in the work environment; • conduct thorough collaborative user studies in real industrial scenarios with dynamic conditions to provide additional perspective.	• the decision on whether AR is applied in remote scenarios depends on the complexity of the collaborative tasks and the expected return; • voice-recognition should only be integrated in scenarios where potential disruption of noisy environments are out of the question; • the use of 3D models could affect the performance of a collaborative solutions, due to the number of models that may be required; • the existence of large amounts of information associated with complex tasks may affect worker's performance; • many scenarios include work forces with low technology literacy and are not prepared for significant investments, precluding the use of technologies such as HMD and, thus, eliciting novel approaches; • characterization and evaluation of the collaborative process is challenging in this multifaceted context; • absence of rules, guidelines and theories to guide the characterization of the collaborative process using solutions mediated by AR; • understand which tasks and measurements are relevant to evaluate these types of scenarios; • additional dimensions of collaboration with a greater focus on evaluation would allow better capturing the essence of these scenarios.

Table 8 – Summary of the main results and insights of the case study.

3.5 Summary

Remote maintenance relying on AR is complex, multidisciplinary and extremely relevant in Industry 4.0, since the expertise to solve a particular problem is often distributed among multiple remote professionals.

In this chapter, we set out to understand how collaborative work is accomplished and how it affects the design of solutions using AR technology to mitigate obstacles of remote scenarios. Designing AR-based solutions that intertwine human expectations and practices, as well as digital artifacts is a multifaceted process which relies on iterative and multidisciplinary approaches. This work gave us the opportunity to uncover insights on the real needs of the industry sector through the involvement of domain experts. The process of acquiring information about the scenario with such experienced individuals confirms the need for traceability, offering useful qualitative feedback on how to support remote collaboration. By merging these outcomes with literature methods into the BARRT prototype and performing its evaluation through user study, we were able to cover existing gaps of recent works. In this vein, we found that the AR-based prototype based on manual stabilized annotations and video sharing provides means for remote experts to collaborate with on-site professionals, regardless of their localization and time. While we must be prudent with generalizing our findings, we expect our results to be valuable for future reproduction in other domains of remote collaboration.

Besides, one of the more challenging areas we unveiled in this work is the necessity to develop a set of methods and processes to evaluate collaborative AR-based solutions. We felt this necessity when planning tests with the industry partners after the study. Therefore, this study is being expanded by investigating evaluation of remote collaboration using AR, to understand what should be taken into consideration to characterize the collaborative proves more effectively. Although the study presented was conducted in an environment purposely configured to be as realistic as possible, we recognize the need to perform field studies with domain experts to test our findings and validate our prototype in real design settings.

4 Conceptual Model and Taxonomy for Collaborative AR

You need to be aware of what others are doing, applaud their efforts, acknowledge their successes, and encourage them in their pursuits. When we all help one another, everybody wins.

Jim Stovall

In the past decade researchers have devoted their efforts to experiment with technology and mature methods for Collaborative Augmented Reality (AR) (Ens et al., 2019; Marques, Teixeira, et al., 2020; Merino et al., 2020). Now, is the time to understand where we stand and how well can we address the domain of collaborative work with AR. In this context, there are two aspects that need to be tackled to obtain a proper perspective:

(1) what does it take to address the question at hand, e.g., which dimensions need to be considered;

(2) how existing research is tackling each of these dimensions (i.e., how it is done).

In this regard, literature reviews can help us understand which research can provide an answer to a specific research question or problem (Collazos et al., 2019; Petersen & Stricker, 2015) and, particularly for areas with an already high level of maturity, existing work helps to identify enough of these two aspects. However, there are cases for which the literature addressing a particular research question does not provide enough information to understand if all relevant dimensions of the problem are being covered.

Therefore, performing a review without a first effort to identify these dimensions can provide limited insight, since it might precisely miss those aspects that, albeit important, are still not addressed by existing research. We are not only seeking what the community has achieved, but also if it fully addresses all aspects to solve domain-specific problems.

In this line of thought, we need to take a step back to have a wide perspective that goes outside the AR research boundaries to also encompass the context it strives to support collaboration. This entails going beyond Collaborative AR literature, considering other domains that may be relevant to characterize the collaborative effort, to identify which aspects (dimensions) should be taken into account when we move from asking what existing systems can do to understanding what they were able to do in particular contexts, i.e., the value of the solutions they provide.

The importance of such effort is, in our perspective, twofold: first, it allows gathering a structured insight on the defining dimensions of Collaborative AR, fostering a more detailed understanding of the field; and second, by doing so, contributes to support research that places the solutions in close relation with the collaborative context they address and reflects on the extent of its contributions (Marques et al., 2021).

As the Collaborative AR community, having matured technologies and methods (Ens et al., 2019), approaches domain experts to address their collaborative needs, this latter aspect is paramount to ensure that the research adds to the body of knowledge and provides enough context and evidence to enable a transparent account (Sukumar et al., 2020) and transferability (Meyer & Dykes, 2019).

To this end, it is important to materialize the devised dimensions of Collaborative AR into a conceptual framework and taxonomy that might foster a harmonization of perspectives for the field creating a common ground for systematization and discussion of past, present, and future works (Collazos et al., 2019). The proposal of a taxonomy should influence and improve how research is reported (Hadwin et al., 2006) by providing a structure and, in a way, a check-list to the defining characteristics that need to be clarified.

A more systematic reporting can, in turn, lead to a community setting that enables easier building on existing research. By gathering dimensions that both cover the collaborative context and the AR solution, the taxonomy may also foster going beyond the description of the methods and into the methodology (Sedlmair et al., 2012), i.e., how the research moved from the problem to the choice of the methods. Additionally, a taxonomy should also improve the awareness of researchers about different dimensions of the contexts they target. In this regard, the work presented here contributes to research on Collaborative AR by:

- providing an explicit consideration of the work in Collaborative AR in tight relation with several characteristics identified as defining collaborative work;
- performing an analysis of different dimensions to be considered when developing collaborative AR-based systems;
- proposing an extended human-centered taxonomy for the categorization of the main features of Collaborative AR stemming from the identified dimensions.

4.1 Categorization efforts

Throughout the years, several categorizations have been proposed for AR and collaborative technologies. The ability to draw inferences is a critical condition for a useful categorization. Taxonomies are a good example, allowing to structure the knowledge of a field, understand the relationships among concepts, analyze complex domains, and provide relevant input to the development of theories (Augstein & Neumayr, 2019). A literature survey in several disciplines was presented by (Nickerson et al., 2013), which discussed thoroughly the problem of taxonomy development. The authors proposed the following qualitative attributes for the creation of a useful taxonomy: it must be *concise, robust, comprehensive* and *extensible*. Likewise, a good taxonomy must also be *explanatory, and not descriptive*: it must contain dimensions and characteristics that do not describe in detail specific objects of interest, but rather provide useful explanations of their nature, allowing the taxonomy to be useful for several purposes.

4.1.1 Augmented Reality

Several taxonomies have been proposed in the field of AR, starting with the one by (Milgram et al., 1994), which performs a categorization based on the types of visual displays used. The taxonomies by (Mackay, 1998), (Suomela & Lehikoinen, 2004), (Lindeman & Noma, 2007), (Braz & Pereira, 2008), (Tönnis & Plecher, 2011), and (Hugues et al., 2011) are fairly general, not addressing any particular type of AR technique or area of application. These are summarized and organized by (Normand et al., 2012) into four different types:

- **Technique oriented** refers to taxonomies that group concepts related with the system environment knowledge, realistic representations, centricity of the type of display (egocentric or exocentric), congruency of control-display mapping and sense of presence.
- **User-centered** encompasses taxonomies that categorize stimuli based on the insertion point: the real world (in Spatial AR) or the virtual world (when the content is only visible through a device). Other taxonomies categorize AR application based on other users' properties: mobility (stationary/mobile), number of users and space (co-located/remote).
- **Interaction-centered** taxonomies focus on interaction aspects, such as the target of augmentation (user or physical object), input and output devices, system and persons, and connections between the system and the real world.
- **Information-centered** taxonomies focus on concepts related to the data available: model dimensionality (ranging from 0D to 3D), viewpoints (first or third person), temporality (continuous or discrete presentation of information), registration and referencing (objects that present information about other objects in the environment).

4.1.2 Collaborative Augmented Reality

Regarding Collaborative AR, some categorization efforts can be found in literature. Early work by Benford et al. proposed an interaction-centered taxonomy for classification of MR approaches according to the shared spaces based on three dimensions: *transportation* - the extent to which a group and objects leave behind their on-site space and enter into a new remote space in order to meet with others; *artificiality* - the extent to which a space is either synthetic or is based on the physical world; and *spatiality* - the level of support for fundamental physical spatial properties such as containment, topology, distance, orientation, and movement (Benford et al., 1998).

Billinghurst et al. defined the following characteristics as relevant for a Collaborative AR environment: *virtuality* - virtual objects with no direct relation with the real environment that can be seen and examined in AR; *augmentation* - virtual objects (e.g., annotations, visual guides, etc.), directly related to real objects existing in the scene; *cooperation* - possibility for multiple users to see each other and cooper- ate; *independence* - each user controls their own independent view; individuality - the displayed data/representation might be different for each user (Billinghurst & Kato, 2002).

Then, Wang et al. focused into the design of effective AR systems to mediate human-human collaboration for shared production tasks in the Construction sector. Three dimensions were identified for categorizing AR systems in such contexts: *mobility* - user's location and orientation, divided in small local area environments and large distances; *number of users* - single-user AR and multi-user AR systems; *space* - distance between users in a multi-user AR system, which can be classified as either collocated or distributed systems (X. Wang & Dunston, 2006).

In addition, Brockmann et al. provided a categorization for collaborative AR-applications, based on a literature review focused on six dimensions: *space, time, mobility, virtual content, role concept, and visualization hardware*. According to the authors, their user-centered taxonomy "shall support the user in identifying the most appropriate collaborative AR-application fitting the respective communication and collaboration scenario" (Brockmann et al., 2013).

In the same way, the research by Jalo et al. reported the following characteristics for the development of Collaborative AR systems for the Industry sector: *dimensions* - depending on whether the collaboration happens synchronously or asynchronously and whether the users are located in the same place or not; *stakeholders* - collaboration inside a company, between companies or between a company and its customers; *type* - depends on the number of participants and can be divided into: one-on-one, one-on-many and many-on-many categories; *functionalities* - visual digital information, such as text, pictures, videos and models; *device* - collaboration using AR can happen though the use of a multitude of hardware; *senses* - all human senses can be used in AR (Jalo et al., 2018).

Speicher et al. also present the notion of mixed reality as a combination of AR and VR users that are potentially physically separated. Although the authors emphasize the MR landscape is highly fragmented, a conceptual framework with seven dimensions was created to categorize MR applications in terms of *number of environments* - total of physical and virtual environments; *number of users* - users required for a certain type of activity; *level of immersion* - how immersed the user feels based on the digital content they perceive; *level of virtuality* - how much digital content the user perceives (whether or not restricted to a specific sense); *degree of interaction* - which can be divided into implicit (e.g., walking around a virtual object registered in space) and explicit (e.g., intentionally providing input to); *input* - refers to input besides explicit interaction, used to inform the experience, which can be anything sensors can track; *output* - considers output to one or more of the user's senses in order to change their perception (Speicher et al., 2019).

Another example is the work by Belen et al., who performed a systematic review of the current state of collaborative MR technologies, published from 2013 to 2018. This review presented a high-level overview of collaborative MR influence across several research disciplines. A total of 259 papers have been categorized based on their *application areas, types of displays used, collaboration setups, user interaction and user experience aspects* (R. Belen et al., 2019).

Ens et al. revisited collaboration through MR, taking into account the evolution of groupware. The authors review investigated how common taxonomies and frameworks in CSCW and MR research could be applied to such systems. A set of six dimensions were defined, namely: *time* and *space* - including the values synchronous/asynchronous and co-located/remote respectively; *symmetry* - whether collaborators have the same basic roles and capabilities (symmetric) or whether they have different roles or capabilities (asymmetric); *artificiality* - extent to which a space is based on the physical world or either synthetic, spanning between physical, mostly digital, or hybrid; *focus* - primary target of collaborative activity, which can be defined as environment, workspace, person and object; *scenario* - overall concept of a system according to the users and use case. 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 on the nuances of supporting collaboration (focus squarely on the human concerns that underlie communication and collaboration), instead of the need to focus on creating the enabling technology, that makes AR collaboration possible (Ens et al., 2019).

Finally, recent work by Sereno et al. presented a systematic survey of 65 papers along the dimensions of *space, time, role symmetry* (whether the roles of users are symmetric), *technology symmetry* (whether the hardware platforms of users are symmetric), *output and input modalities*. The authors derived design considerations for collaborative AR environments, and identified research topics to further investigate, such as the use of heterogeneous hardware and 3D data exploration. The survey also contemplated collaborative immersive analytics using AR technologies to provide an overview of the field for newcomers, researchers and domain experts. In fact, the effort of searching for a categorization in this recent work is evidence that the research community is trying to bring forward a systematic view over the literature. An aspect this manuscript tries to address (Sereno et al., 2020).

4.1.3 Summary

While existing categorization efforts focus on specific use cases or aspects of Collaborative AR, they do not intend to cover the complete landscape as described in this chapter.

Considering the reviewed literature, there are several aspects that deserve further attention. Most papers report the characteristics of the technologies developed to address collaborative efforts, which is a valuable contribution to the community but, to the best of our knowledge, there are no efforts to explicitly convey the concrete contexts that each solution is designed to serve, defining the characteristics of collaboration, e.g., the team, environment, and collaborative context.

Additionally, existing efforts are mostly oriented towards technology. A human-centered approach, i.e., focusing on the feature that needs to be provided to serve the collaborators, instead of the technology, might bring forward a perspective that is not rapidly deprecated with the advancements of technology (Augstein & Neumayr, 2019).

4.2 Methodology

To reach a first proposal of the dimensions characterizing Collaborative AR, we adopted a conceptual-to-empirical methodology, adapted from the work of (Nickerson et al., 2013) and partially inspired by the research method used by (Collazos et al., 2019).

The methodology followed for the creation of the taxonomy was based on a participatory design process (Halskov & Hansen, 2015), i.e., actively involving stakeholders in focus groups and brainstorm sessions (Barnum, 2010; Jacko, 2012). To this effect, we gathered a set of multidisciplinary experts 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), Visualization, Multimodal Interaction, as well as remote assistance and maintenance in industrial contexts, involved in various types of collaborative work. In total, 15 experts were involved, from these different areas, although, in many cases, not all at once. These individuals had various professions, e.g., PhD students, researchers, faculty members, as well as project managers, maintenance technicians, remote support and technical instructors, thus benefiting from on-going collaborations with partners from the Industry sector.

We conducted several face-to-face and remote meetings, focus group and brainstorm sessions (sometimes with different combinations of experts according to their availability) over several months. In this process, we used illustrative materials like storyboards, diagrams and videos of our own work in the field of Collaborative AR. In addition, multiple collaborative tools were used for discussion, analysis and brainstorm, e.g., Evernote, SimpleMind, NodeD, VoiceRecorder, Zoom, OneDrive, Microsoft Word and PowerPoint, Overleaf. A moderator facilitated the discussion using scripts to elicit richer discussion, and as the work progressed, different iterations of the conceptual model and taxonomy were used to generate debate and deliberation. During these sessions, we focused on the following phases:

- **Explore collaborative realities** of each individual and progressively introduce and discuss the subject of Collaborative AR through the use of different materials, e.g., images and videos of a collaborative AR-based solution being used to address a maintenance problem between two distributed collaborators;
- **Identification of relevant terms**, i.e., determine a set of objects of interest, based on the area of interest and the expected use of the taxonomy by the research community. In this process, several terms may be added, removed and renamed. Based on the experience of the individuals, an effort was made to identify the defining blocks of a collaborative effort supported by AR;
- **Conceptual Model definition** using the terms identified, which then elicited an analysis of the literature, not only about Collaborative AR, but covering the key aspects of the conceptual model to support their definition; Categorization of objects of interest, i.e., identification of reoccurring objects and conceptualization of components that may be appropriate to differentiate between those using a graphical process as suggested in (Collazos et al., 2019);
- **Creation of the Taxonomy**, i.e., form the initial dimensions of the taxonomy, following a similar approach to the one used by (Zollmann et al., 2020). Their definitions must be clarified and agreed upon. Moreover, each dimension contains categories and characteristics that must be mutually exclusive and collectively exhaustive (Nickerson et al., 2013);
- **Detailed explanation of all objects of interest** included in the taxonomy, and description of the main decisions and design alternatives related to them, as well as inclusion of the main related bibliographical references;
- **Application, discussion and iterative refinement** of the taxonomy to verify if the established dimensions, categories and characteristics are well defined, need to be merged, or if new ones could be identified (Nickerson et al., 2013), which resulted in several iterations to the initial taxonomy.

4.3 Conceptual Model for Collaborative AR

During the creation of the conceptual model for Collaborative AR, there was no particular concern with specific supporting technologies, but mostly with the steps required to accomplish it. In this regard, the first step consisted of a conceptualization for the single user AR scenario to establish a baseline. Our goal was to represent the collaboration nature and the most common tasks (which can be present in more than one scenario), the context (co-located or remote collaboration), the collaborative setups, i.e., necessary apparatus to capture and share AR, the collaborator role, the predominant interaction modalities, the level of engagement, and the multi-sensory context. Moreover, in the context of co-located scenarios, we had in consideration the collaborative setup and cooperation modalities for different levels of engagement with virtual and real objects. Likewise, for remote collaboration scenarios, other modalities can be foreseen, allowing to differentiate the role of each collaborator, according to abilities and prerogatives.

The boxes represent different key elements with arrows loosely indicating a flow between elements. Before looking into how collaboration is performed using AR, let's look into AR when used by a single user to identify the main elements for such a goal. In a typical scenario supported by an AR system (see Figure 36), users try to accomplish a task by interacting with the environment, while their senses may be provided with some contextual data. The contextual data is considered from the very beginning of the pipeline that is building the AR view, and it should be assumed that it goes through the different entities and can be used accordingly. Registration data is considered to enable the identification of points of interest in the capture content and providing the grounds for augmentation spatially (and temporally) aligned with the reality. An AR view is generated and presented to the user through some output presentation. Different levels of user actuation may be possible entailing, e.g., the ability to modify how the scene is augmented through compositing, i.e., which elements are visible, exploring particular aspects of the scene, or interacting with content through the interactive input.

From this single-user model, we evolve to a collaborative conceptual model (see Figure 37). When moving into the collaborative setting, several aspects are replicated for the different users involved, as further discussed below, but there are a few additional elements that take shape. First, a team is now involved with its characteristics, e.g., the number of elements, their profiles, and their location. For the sake of simplicity, the diagram only shows the different conceptual blocks for one local and one remote user in detail, despite having present additional team members. With the collaboration effort, emerges the need for communication among the team members considering the channels suited for the context and task, which may be affected by different time aspects. Within the field of CSCW, the term awareness can be defined as "an understanding of the activities of others, which provides a context for your own activity" (Dourish & Bellotti, 1992). Awareness relates to the knowledge one has of other team members' actions. Such knowledge is used to inform one's own action in a way that makes the whole team move forward in the collaborative effort. While awareness is associated to the knowledge of what is going on at a particular moment, the notion of common ground refers to the common understanding of joint goals, shared resources and the state of the task solving process (Cidota et al., 2016). More formally, it can be defined as "a state of mutual understanding among participants about the topic at hand" (Clark & Brennan, 1991). The existence of a mutual understanding among team-members is based on working vocabulary, practices and norms, that contribute to a sense of shared knowledge and awareness (Patel et al., 2012) and is of utmost importance. These allow team members to work together effectively, adjusting their activities as necessary through different shared context sources.

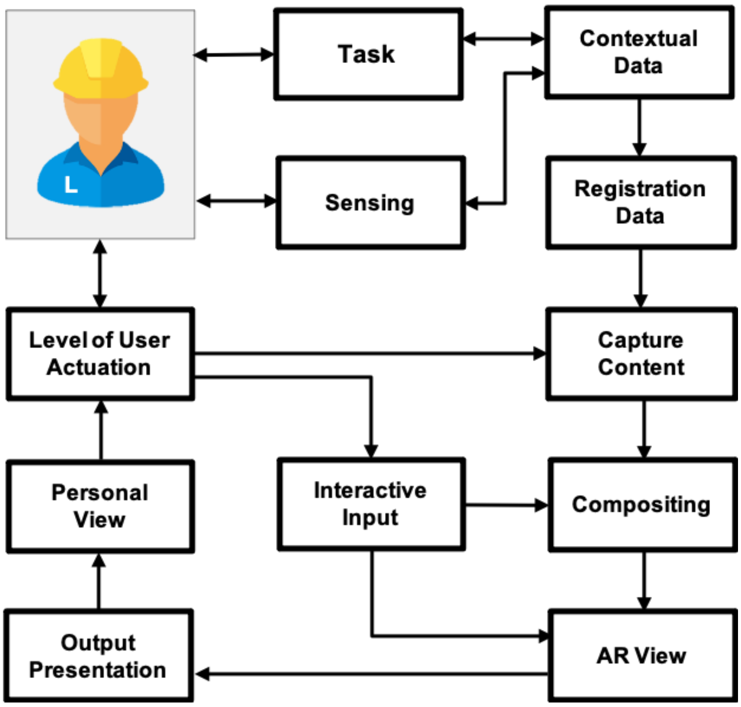

Figure 36 – Conceptual model illustrating the elements associated with how a single user interacts with AR to accomplish a given task.

The contextual data of the remote user is updated according to the capture content to present him/her an updated view of the local context data. On the remote side, the user can be provided with different levels of information regarding the local context, including, e.g., a video feed, a virtual scene or, even, tangible landmarks or a reproduction of the physical context. In the simplest situation, a view of the task setting is provided and can, if possible, ask the local user to provide different views. The view(s) available can also include augmented content whether sharing what other team members are seeing or adding information relevant for the remote member's function. Additionally, the remote user might have some control over the capture content and be able to select particular views, e.g., controlling a camera. The coexistence of more than one team member might also motivate the consideration of more than one view on each side: one that is shared, e.g., the remote user seeing the same view as the local user; and an additional view, e.g., for an overall analysis of the environment or selecting a different augmentation to explore additional information. These realities can be shared with more or less fidelity at the counterpart of each user: the local user can see aspects of the physical reality and vice-versa. Also, both local and remote user can interact and modify some attributes of the augmented elements in their counterpart view. Finally, each user might have a shared view (common between both) and a personal view not shared.

The conceptual diagram was continuously analyzed to identify potential grouping of blocks according to their overall purpose and interrelation. Finally, the terminology considered for each of the elements was harmonized, as best as possible and without affecting the overall concept, to adopt nomenclature already used in recent literature (Zollmann et al., 2020).

The outcomes were then considered as the basis for the proposed taxonomy, as presented in what follows. To elaborate, the team, time, task, communication and the level of user actuation dimensions were kept directly from the conceptual model as is. The scene capture and tracking dimension are the result of merging the registration data and the capture content from the conceptual model. Moreover, the shared context sources dimension is the result of adapting the collaborative context (including the local and remote contexts particularities like physical context, sensory context and sensing) in order to encompass more aspects like the human, environmental and collaborative factors. In turn, the input modalities dimension, as well as the output and augmentation were created based on the output presentation and interactive input. In addition, the research dimension resulted from the analysis and evaluation of the taxonomy itself, not being directly related to the conceptual model.

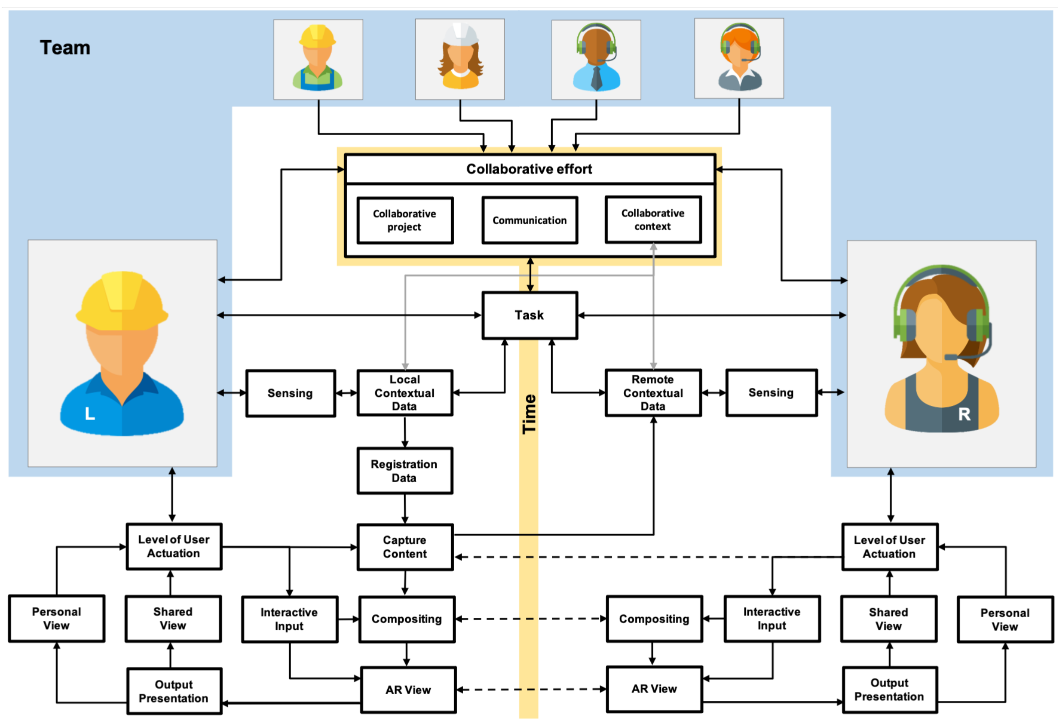

Figure 37 – Conceptual model for Collaborative AR. For the sake of simplicity, the diagram only shows the different conceptual blocks for one local and one remote user in detail. Dashed lines imply the existence of connections between elements, which are not mandatory, but may occur if needed during the collaborative effort.

4.4 Taxonomy for Collaborative AR

Our taxonomy aims not only to propose a first systematic approach to the more intrinsic (technological) characteristics of Collaborative AR, but also to put them in relation to key aspects of collaborative work (Figure 38). In what follows the different dimensions included in the taxonomy are presented and the considered categories and characteristics detailed and discussed. Additionally, for easier reference, a companion table is provided for each dimension.

Team

(*physical distribution*, *role structure*, *size*, *life-span*, *turnover*, *technology usage*, *multidisciplinarity*)

The characteristics of a team involved in a collaborative effort define much of how the tasks progress (Patel et al., 2012). First of all, the **physical distribution** of its members (Table 9), if they are in the same location (*collocated*), if they are all in remote locations (*distributed*) or a mix of these two cases (*mixed-presence* (Billinghurst et al., 2018)). This corresponds to one of the dimensions identified in the seminal work of (Johansen, 1988).

Another important aspect dictating how the teams need to work is their **role structure**. If the team is *functional*, it means that each member has a specific function or expertise, but if the team is *divisional*, all elements have the same level of expertise, but collaborate to divide the work (S. Kim et al., 2014), (Norman et al., 2019).

The **team's size** is also an important aspect to consider, since it can have impact on several aspects during design and run-time, for instance, in how to make it clear, for all, who is intervening or performing a certain task, at a certain time. At this stage, we distinguish between *two* and *three or more* elements (C. P. Lee & Paine, 2015).

Team **life-span** (or permanence), refers to the amount of time a team exists as so. If a team is assembled to tackle a particular task and, then disbanded, it is said to be *short-term*; however, if the team persists, over time, across multiple tasks it is classified as *long-term*.

Team **turnover** (C. P. Lee & Paine, 2015) refers to the amount of expected change in the elements intervening in the collaborative effort, i.e., how often team members leave and/or new team members are added, ranging between *low, intermediate or high*.

Moreover, **technology usage** includes the amount of effort devoted to the use of technology, i.e., how frequently technological solutions are used during the collaboration effort. Team members may use it *sporadically*, i.e., once in a while, not a common practice or *systematically*, i.e., technology has been established as one of the practices for collaboration and used often (Patel et al., 2012), (Stokols et al., 2008).

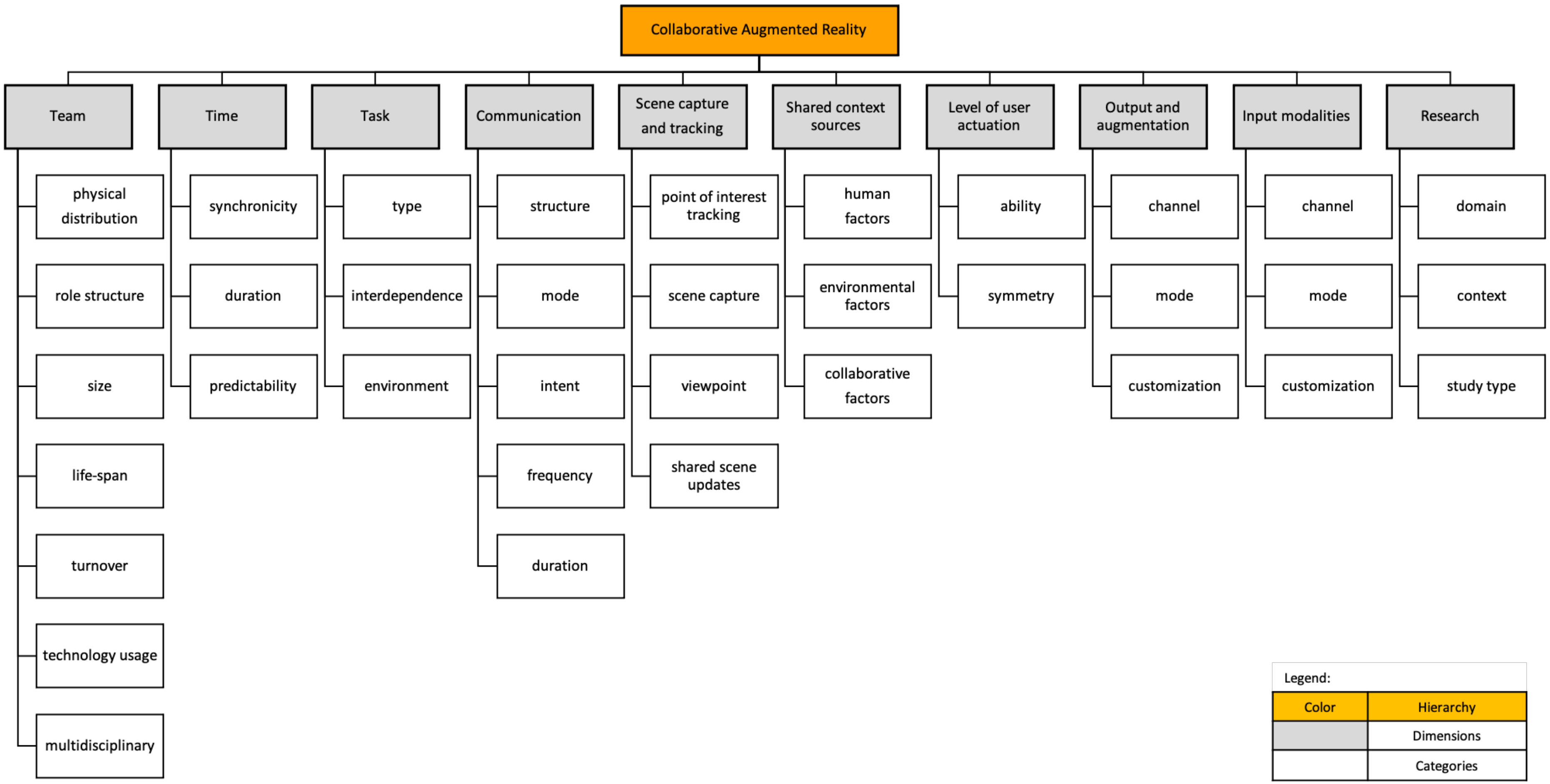

Figure 38 – Taxonomy including the different dimensions and categories identified for Collaborative AR categorization.

Finally, a **multidisciplinary** team, i.e., the presence of members with different backgrounds and perspectives over the task (analogous to the number of communities of practice discussed by (C. P. Lee & Paine, 2015) might pose particular challenges regarding how, e.g., a more elaborate context needs to be provided, communication is supported, or adaptation needs to be available to allow custom discipline specific augmentation.

Team	
physical distribution	
	co-located, distributed, mixed-presence
role structure	
	functional, divisional
size	
	two, three or more
life-span	
	short-term, long-term
turnover	
	low, intermediate, high
technology usage	
	sporadically, systematically
multidisciplinarity	
	yes, no

Table 9 – Team categories and characteristics.

Time

(*synchronicity, duration, predictability*)

This dimension groups characteristics that have to do with how the different elements of collaboration relate over time, considering (Table 10): **synchronicity**, i.e., if all team members are present and can act in real time (*synchronous*), or if collaborative actions, performed by different elements, take place at different times (*asynchronous*).

The mixed synchronicity refers to the fact that supporting both might be relevant, for some tasks; **duration**, refers to the time required for the collaboration effort, without interruption, to accomplish, e.g., a *short* (less than 30min), *intermediate* (between 30min and 90min), or *long* (more than 90min) task. This is important, since a certain setup might be adequate for short usage times, but be uncomfortable for longer periods; and **predictability** of the collaboration (Bolstad & Endsley, 2005) describes if it happens at well-defined *schedules* (predictable) or *unscheduled* (unpredictable) times.

Time	
synchronicity	
	synchronous, asynchronous, mixed
duration	
	short, intermediate, long
predictability	
	scheduled, unscheduled

Table 10 – Time categories and characteristics.

Task

(*type, interdependence, environment*)

The task is central in a collaborative effort and dictates much of the communication, information, and augmentation requirements (S. Kim, Lee, et al., 2020; Patel et al., 2012). Regarding the task **type**, it can be divided in (Table 11) (Wildman et al., 2012): 1) *management*, where someone assumes the supervision and coordination of others; 2) *advisory*, entailing professional support, e.g., providing expert advice; 3) *negotiation*, when two or more parties need to resolve conflicts and reach agreement; 4) *psycho-motor action*, referring to those tasks consisting of the manipulation of a machine or product involving elaborate movements and/or psychological processing, whether in physical or virtual reality. 5) *defined problem*, i.e., problems with well-defined answers, e.g., when a remote expert has the solution and provides instructions; 6) *ill-defined problem*, when no partner has an immediate solution and they, e.g., generate/share ideas or plans through brainstorming.

Interdependence describes how team member actions are influenced or limited by those from other members and can be (Wildman et al., 2012): 1) *pooled*, when each member can make their contribution independently from others, possibly asynchronously; 2) *sequential*, when the actions are performed in a well-established sequence of team member contributions; 3) reciprocal, when the different actions to accomplish the task are performed in sequence, but there is a continuous adjustment to how the task is progressing, entailing a certain level of unpredictability, e.g., choosing the kind of expertise required for particular situations. To clarify, sequential interdependence presumes a fixed and well defined sequence of steps with, typically, a precise definition of the team member involved in each, while in reciprocal interdependence, team members need to work in sequence, but there is a back and forth adjustment depending on how the task progresses and the expertise required, at each time; and 4) *intensive*, when all team members need to work simultaneously, i.e., synchronously to accomplish the task.

Finally, it is relevant to know if the task is performed in an **environment** located *indoor*, *outdoor*, or *mixed*, since this might impact on how the system is designed.

Task	
type	
	management, advisory, negotiation, psycho-motor, defined problem, ill-defined problem
interdependence	
	pooled, sequential, reciprocal, intensive
environment	
	indoor, outdoor, mixed

Table 11 – Task categories and characteristics.

Communication

(*structure, mode, intent, frequency, duration*)

The communication **structure** describes how the message flows inside the team (Patel et al., 2012). Inheriting from the work of (Wildman et al., 2012), three structures can be considered (Table 12): 1) *hub-and-wheel,* where all communication passes through one team element (e.g., a leader), and flows to others through him/her; 2) *chain*, where the message reaches each team element through a hierarchy; and 3) *star*, where every team member freely passes and receives information from others.

The communication **mode** (Ostergaard & Summers, 2009) characterizes what communicative elements are possible, such as *verbal, textual* (e.g., messaging), *graphical* (e.g., sketch), or *gestural* (e.g., hand gestures) elements.

Communication is also characterized by the **intent** (Ostergaard & Summers, 2009), (Yoshioka et al., 2001) that it serves: *inform, commit, guide, request, express, decide, propose, respond, and record*. The identification of intent is an important characteristic of communication. It emphasizes the existence of an explicit goal, marking a notable difference to other aspects of the collaboration dynamics which, although related to communication, are not explicit. For instance, using hand gestures to, explicitly, convey or complement a message (e.g., pointing to a specific area) is communication related (Huang et al., 2019), while sharing hands' position to contextualize how the task is being performed would not.

Frequency characterizes how often communication can (or needs to) occur to accomplish the task: *never, sometimes, and continuous*, and **duration**: e.g., *short* (less than 5 seconds), *intermediate* (between 5s and 5min) or *long* (more than 5min). Both these aspects are dependent (or might face challenges) on a number of other factors such as the type of task and team distribution. For instance, for particular tasks, frequent communication might be mandatory.

Communication	
structure	
	hub-and-wheel, chain, star
mode	
	verbal, textual, graphical, gestural
intent	
	inform, commit, guide, request, express, decide, propose, respond, record
frequency	
	never, sometimes, continuous
duration	
	short, intermediate, long

Table 12 – Communication categories and characteristics.

Following, some clarifications are presented regarding the difference between some categories of the previous dimensions, namely: "Time: duration", "Team: life-span", and "Communication: duration". While the Time: duration and Communication: duration may vary in similar characteristics, i.e., short, intermediate or long task/long term, Time duration refers to the total amount of time required to accomplish a given task, while Communication duration refers to the amount of time used for sharing information between team-members. Finally, the Team life-span refers to a different aspect, in particular, the amount of time a team exists with its members, either long-term, if the team persists, over time or short-term, if a team is assembled to tackle a particular task and disbanded at the end. So, for instance, a short-term team assembled to solve an emergency may perform a long task supported on small duration communications for, e.g., action synchronization.

Scene Capture and Tracking

(*point of interest tracking, apparatus, viewpoint, shared scene updates*)

A first aspect to consider is **point-of-interest tracking**, i.e., how the system knows where the relevant key features are (e.g., objects, location) enabling proper registration of the augmented content. In this regard, we identify three options (Table 13): 1) *computer vision* methods, resorting to marker or marker-less approaches; 2) *sensor* (e.g., electromagnetic, GPS (Global Positioning System), IMU (Inertial Measurement Unit)); or 3) *non-existent*. The latter option encompasses those situations for which augmentation is done without a direct connection with scene elements, e.g., instructions provided by remote teammates presented on the corner of an on-site technician field-of-view, while he/she performs a maintenance procedure.

The apparatus is the technological device used to support **scene capture** (and tracking) and can range from a simple *camera*, to more complex devices, such as *stereo cameras, depth cameras or 360 cameras* (S. Kim, Billinghurst, & Lee, 2018).

Finally, **viewpoint** refers to the nature of the views of the scene that are available. In this regard, we consider three alternatives: 1) *fixed*, e.g., from a fixed overhead camera; 2) *mobile*, dependent on the user, e.g., from a handheld device or a user mounted camera providing POV; and 3) *mobile, independent*, e.g., a camera mounted in a robotic arm that can be oriented to provide any particular view (S. Kim, Billinghurst, & Lee, 2018).

The **shared scene updates** are associated with how often the environment is updated. It may be 1) *static*, e.g., 360 images; 2) *dynamic*, e.g., 360 videos; or 3) *live*, e.g., 360 video streams.

Scene Capture and Tracking		
point-of-interest tracking	computer vision	
		marker, markerless
	sensor	
		electromagnetic (RFID), GPS, IMU
scene capture		
	camera, stereo camera, depth camera, 360 camera	
viewpoint		
	fixed, mobile dependent, mobile independent	
shared scene updates		
	static, dynamic, live	

Table 13 – Scene Capture and Tracking categories and characteristics.

Shared Context Sources

(*human, environmental, collaborative*)

Context-awareness is a field of research deserving strong attention in multiple areas. In this regard, it is important to distinguish two main purposes: 1) to provide the computational system with a context that might enable adaptive behaviors; and 2) provide the users with information that contextualizes the task they are involved in (Antunes et al., 2014; Fernández del Amo et al., 2018). Regarding collaboration and AR, two notable examples of a systematic (taxonomic) approach to the subject are the works by (Collazos et al., 2019), proposing a taxonomy for context information sources in the scope of their descriptive theory of awareness for groupware development. Here, we leave out the subject of discussing which context sources can be considered for providing adaptive system features (see (Grubert et al., 2017) for an in-depth discussion on system context-awareness). Instead, we focus on these context elements considering their potential importance to increase the awareness of team members regarding the collaborative context. Since system-side context-awareness has a parallel with collaborative context awareness (by sharing a range of context sources), we consider the nomenclature proposed by (Grubert et al., 2017) selecting those dimensions and characteristics with a more immediate relevance for collaborative scenarios and inherit, where deemed relevant, from (Collazos et al., 2019). To clarify, we do not consider as context those elements that arise from explicit communication by any team member (Collazos et al., 2019). Furthermore, it is important to note that context sharing is not only useful for remote collaboration, but can also be an important resource in collocated efforts (Olsson et al., 2020). In fact, the increasingly common multi-device ecologies often generate team and task awareness fragmentation (Fischer et al., 2018; Scott et al., 2015) that might be tackled by an explicit presentation of context elements.

Human factors, i.e., those pertaining team members and their performance encompass *personal, task-specific,* and *social aspects* Table 14). Personal human factors specifically relate to individual user characteristics or states including, e.g., age, abilities and knowledge, perceptual, cognitive (e.g., cognitive load) and affective states. Additionally, it can also encompass aspects directly specifying the user's context within the task, such as gaze orientation and focus (point-of-interest), availability, location, task progression, and motor activity. Social human factors account for the broader scope of interaction among people. The dimension of social interaction in collaborative efforts is highly relevant as it fosters improved learning, group formation and group dynamics (Kreijns et al., 2013). Considering the framework proposed by (Kreijns et al., 2013), social interaction depends on the systems sociability, the creation of a social space, and on social presence. *Presence* goes beyond the simple information regarding the location or availability of a team member (as included in the personal factors, above), as it entails a sense of someone being present and following what the person is doing (Collazos et al., 2019). The inclusion of such feature (e.g., through avatars (T. Y. Wang et al., 2019; Yoon et al., 2019)), is relevant since it potentially enables a remote collaboration experience and performance that is closer to what is possible in collocated work. Additionally, research hints that, as happens with collocated work, the sense of someone being present might have an influence on team member performance. For instance, (Miller et al., 2019) have shown how a sense of presence might have a social facilitation or inhibition effect depending on the difficulty of the task: having the sense of someone present, when performing a difficult task, lowered performance. The extent to which these elements are present depends on several aspects, such as the supported level of communication, but also greatly depends on how social cues are made available to team members.

Environmental factors concern everything about the *physical* and *digital* (i.e., augmented) environment that the user is experiencing. The physical environmental factors describe the characteristics of the place where the user is positioned including elements such as temperature, ambient noise and light intensity, and the spatial and geometric configuration of relevant artefacts (Collazos et al., 2019; Irlitti et al., 2019; Müller et al., 2017; Stokols et al., 2008). Digital environmental factors refer to elements that provide context about the characteristics of the AR environment and the features it provides. This is important in certain contexts, in which team members rely on different technological resources which should be known by other members (Collazos et al., 2019) is critical to guarantee success of collaboration teams (Stokols et al., 2008). For example, the amount of information each team member views, if the tracking/alignment mechanism is working properly, if all virtual elements are being rendered in a satisfactory manner, the availability of adequate infrastructure factors, like bandwidth for distance technology tools, state-of-the-art workstations or the availability of technical support.

Collaborative factors pertain information that provides a wide contextualization of the collaboration effort, further supporting, for instance, the conditions required for coordination, complementing communication, an aspect considered of the utmost relevance for collaborative work, by contributing to a shared team cognition and potentially driving anticipatory behavior and implicit coordination (D. Wang et al., 2020). The collaborative *action timeline* (Collazos et al., 2019) refers to information regarding the sequence of past relevant actions of different team members, and annotations or outcomes that would, for example, provide a reference procedure to solve a problem, a context of what has been attempted or performed, so far, or support auditing procedures. In turn, the *progression* of the collaborative effort refers to a less granular level of information than the timeline. While the latter can work similarly to a logbook, progression entails an additional level of detail providing team members with a runtime performance monitoring of each member of the team on their current task (including self-monitoring). This can be important to enable a team-level perspective of the ongoing work, serve to support team coordination, and help create some of the conditions required for the team to adjust to different phases of the tasks (Schmutz et al., 2015), e.g., informing when expert support may be required and, facilitating the articulation of individual actions with the collaborative efforts (She & Li, 2017).

Shared Context Sources		
human	personal	
		age, abilities and knowledge, perceptual, cognitive, affective state
	task-specific	
		gaze orientation and focus, availability, location, motor activity, social factors
		system sociability, social space, social presence
environmental		
	physical, digital	
collaborative		
	Action timeline, progression	

Table 14 – Shared Context Sources categories and characteristics.

Level of User Actuation

(*ability, symmetry*)

The user's actuation **ability** can range from *passive-view*, which can be on-site or remote, to *interacting/exploring*, e.g., manipulating content present in the scene, and to *sharing/creating*, e.g., adding annotations to the scene or new views or content that others can see (Table 15). According to Isenberg et al. (Isenberg et al., 2011), these are associated to the level of engagement and may apply to none, on-site and/or remote users. We chose 'actuation' to avoid a clash with 'engagement' also being used, in the literature, to refer to the amount of motivation and commitment a user is devoting to a task (Chittaro & Buttussi, 2015; Ens et al., 2019).

A user involved in AR-supported collaboration is also influenced by the level of **symmetry**, which represents if all parts have the same level of actuation: *symmetric*, i.e., whether collaborators have the same basic roles and capabilities; *asymmetric*, i.e., whether they have different roles or abilities (B. C. Kwon et al., 2017); or *fully asymmetric*, i.e., a remote user is equipped with the abilities that can help solve an onsite user's problem without any help being provided onsite (J. U. Kwon et al., 2019). The inclusion of full asymmetry is motivated by the passive role of the onsite user, which creates a context with specific challenges to address, beyond those of asymmetry, since the onus of action is on the remote user.

Level of User Actuation		
ability	passive-view	
		on-site, remote
	interact / explore	
		none, on-site, remote
	share / create	
		none, on-site, remote
symmetry		
	symmetric, asymmetric, fully asymmetric	

Table 15 – Level of User Actuation categories and characteristics.

Output & Augmentation

(*channel, mode, customization*)

We choose to have a level devoted to the sensory **channel** receiving the output rather than directly addressing the technological apparatus since this enables an easier grasp of which channels are specifically considered, to avoid uncertainty when the device might serve many channels (e.g., a tablet might provide visual, haptic, and auditory output). Additionally, centering the categories on the users, it should enable a more versatile categorization to encompass novel technologies and devices. In this line of thought, our proposal inherits from the detailed work of

(Augstein & Neumayr, 2019) proposing a human-centered taxonomy for interaction. The authors identify six modalities (sensory channels), related to human perception capabilities that we use to characterize output augmentation. Output and Augmentation can be performed through (Table 16): *vision*, including standalone self-appearance changing devices (e.g., monitor), their wearable alternatives (e.g., HMD), and external medium appearance changing devices, i.e., devices that can change the appearance of an external element (e.g., video projector); *audition*, considering airborne sound propagation (e.g., sound speakers), through a structure (e.g., bone), and possibly wearable (e.g., headphones); *touch* (Fernandes & Albuquerque, 2012), including tactility (i.e., devices that simulate being touched), haptics (i.e., devices that shift their physical properties, e.g., shape, temperature) and vibration, also considering wearable alternatives; *kinesthetics*, considering our senses of proprioception (i.e., body orientation and position), equilibrioception (i.e., body balance), and kinematics (i.e., acceleration); *olfaction* through a device located in the ambient or wearable (e.g., olfactometer (Narumi et al., 2011)); and *gustation*. Additionally, (Augstein & Neumayr, 2019) distinguish between the set of channels above, which entail a perception/action from one of the senses that is further processed by (or originates at) the brain, i.e., indirect processing, and those that directly deal with brain or muscle activity, i.e., direct processing. For the latter, the authors identify: *neural oscillation* and *galvanism*.

Output & Augmentation		
channel	vision	
		self, modifier, wearable
	audition	
		airborne, structure, wearable
	touch	
		tactility, haptics, vibration, wearable
	kinesthetics	
		proprioception, equilibrioception, kinematics
	olfaction	
		ambient, wearable
	gustation	
	neural oscillation	
	galvanism	
mode		
	unimodal, redundant, complementary	
customization		
	adaptable, adaptive, non-customizable	

Table 16 – Output & Augmentation categories and characteristics.

To explicitly convey if a system allows multimodal augmentation, i.e., through multiple channels, **mode** can be: *unimodal*, if only a channel is used, at each time, regardless of how many are available; *redundant*, and/or *complementary*, if multiple augmentation channels are used to reinforce or add information.

Finally, **customization** refers to the possibility of the user (*adaptable*) and/or the system (*adaptive*) to automatically choose or customize the most suited channels for output. It can also be *non-customizable*.

Input modalities

(*channel, mode, customization*)

For the input modalities, we adopt a similar rationale as the one adopted for the output modalities, considering a human-centered characterization aligned with the work of (Augstein & Neumayr, 2019), encompassing the following six **channels** (Table 17) related to human perception: *vision*, covering fixed (e.g., kinect) or wearable (e.g., eye traker glasses) devices that capture/process visual data; *audition*, including devices that capture airborne sound waves (e.g., microphone), through structural propagation in other materials (e.g., ear-bone microphone), and if they are wearable; *touch*, encompassing tactility (i.e., a device sits passively and is touched), haptics (manipulation of an explorable physical surface, e.g., braille keyboard) and vibration (a device sensing vibrations, e.g., tremors); *kinesthetics*, considering proprioception (i.e., position and orientation of the body), equilibrioception (i.e., body balance), and kinematics (i.e., acceleration); *olfaction; and gustation*. Additionally, two channels are considered to cover input though brain or dermal activity, i.e., *neural oscillation and galvanism*.

Mode refers to how the available modalities can be used to perform interaction. As for the output and augmentation dimension, we consider the options *unimodal*, when only one modality can be used, at once, and when these are explored together: *redundant*, when modalities can be used simultaneously to perform the same action, or *complementary*, when multiple modalities are used in sequence to provide different parts of an action (e.g., pointing to an annotation and saying ”delete”).

Finally, **customization** refers to the possibility of the user (*adaptable*) and/or the system to automatically choose (*adaptive*) the most suited channel. It can also be *non-customizable*.

Input modalities		
channel	vision	
		fixed, wearable
	audition	
		airborne, structure, wearable
	touch	
		tactility, haptics, vibration
	kinesthetics	
		proprioception, equilibrioception, kinematics
	olfaction	
	gustation	
	neural oscillation	
	galvanism	
mode		
	unimodal, redundant, complementary	
customization		
	adaptable, adaptive, non-customizable	

Table 17 – Input modalities categories and characteristics.

Research

(*domain, context, study type*)

The last dimension we have considered is devoted to research, allowing to clarify the maturity and detail of the collaborative work being reported (Table 18). In this context, the research **domain**, or topic is associated to the area of application, ranging between *medicine, industrial, education / training, architecture / construction, tourism / heritage, entertaining / gaming, among others* (Dey, Billinghurst, Lindeman, & Swan, 2018), (R. Belen et al., 2019).

According to the collaborative effort and the tasks being addressed, the research **context** may vary between *basic research*, i.e., the technologies and/or methods investigated are novel and have not matured, yet, to be usable in real scenarios, often considering dummy tasks as the case study (e.g., assembly of Lego blocks, tangram puzzles) and evaluation; and *applied research*, i.e., the technologies and methods are implemented in practice using problems inspired by real-world scenarios (e.g., industry related procedures), and an evaluation of the technique is conducted (Olalere, 2012), (Carvalho et al., 2017) for those scenarios.

There are various types of scientific studies (Dey, Billinghurst, Lindeman, & Swan, 2018). The choice of **study type** mainly depends on the research goal, and may vary between *pilot*, i.e., small-scale preliminary studies aimed to investigate crucial components of a main study; *informal*, i.e., studies aimed at getting more input, in a quicker manner, without following any structured method; *formal*, i.e., studies that follow structured methods to obtain measures; *field*, i.e., studies conducted outside a laboratory environment.

Research	
domain	
	industrial, education / training, architecture / construction, tourism / heritage, medicine, entertaining / gaming, among others
context	
	basic research, applied research
study type	
	pilot, informal, formal, field

Table 18 – Research categories and characteristics.

4.5 Critical analysis and refinement

To understand if the taxonomy can be applied to recent research for assessing how the reporting of the works has been conducted, along with a possible ambiguity of how some characteristics of the taxonomy might affect its use, we asked four experts to critically analyze our proposal. The group of selected experts (one female), aged from 27 to 62 years old, included PhD students, researchers and faculty members, sharing several years of expertise in the areas of HCI, VR/AR and Visualization. Moreover, they also had previous experience using tools for remote collaboration (e.g., Skype, Zoom, Microsoft Teams, Team Viewer, etc.), co-authored multiple publications, and participated in international projects on these subjects, over the years.

We choose to use an approach focused on utility demonstration, in which the experts were required to classify subject matter examples, i.e., publications selected by the experts in the field of Collaborative AR (Šmite et al., 2014; Usman et al., 2017). In this context, the experts were provided with instructions, in conjunction with the definitions above. In summary, they were asked to select at least two subject matter examples each and ensure that they could be clearly, concisely and thoroughly classified into the taxonomy, to verify if the established dimensions and categories were well defined, needed to be merged, or if new ones could be identified.

After a period of understanding, adjustment and use of the taxonomy, the experts reported that the taxonomy is globally straightforward to use and apply. Overall, most of the dimensions were easy to follow and understand, and in case of doubts, as reported by one of the experts, the description was generally enough to understand and systematically organize the subject matter examples with a significant level of confidence. The main difficulty reported was related to the **Task, Team and Shared Context Source** dimensions. In the examples analyzed, the source of the information could only be vaguely found in the evaluation sections, which was done resorting to simple tasks (e.g., involving Lego bricks), lacking information on these dimensions, which meant, they could not be easily mapped. According to all reviewers, this lack of information on the aforementioned dimensions meant they could not be easily mapped. The emphasis of the experts' feedback regarding this issue, particularly for such recent works, may hint that many of the works exploring Collaborative AR aim to address real life scenarios, but are not mature enough to be used in such cases. As such, many of the dimensions in the taxonomy incorporate categories and characteristics currently not properly reported in existing works, which are still focused on technology aspects of collaboration. Nevertheless, the experts believe that this makes the taxonomy even more interesting, since it opens questions in areas where most researchers are not yet focused, which may be interesting opportunities for the future. Moreover, one expert suggested feeling the need identify which side of the collaborative process is being described by this dimension. The expert proposed that the dimension was revisited, in order to include characteristics that allow to ease this gap.

Afterwards, we focused on identifying patterns regarding the experts' concerns, e.g., dimensions that could require further attention, structure of some categories, or characteristics missing some examples to better be understood. Also, how easy it was to use the taxonomy, what was their approach in case of doubt, and how they decided to proceed when some information could not be quickly identified in the publications analyzed.

Then, we carefully examined and addressed the main observations that were raised. This refinement process resulted in a new iteration of the taxonomy, which are reflected in the aforementioned proposal. More specifically, we performed updates to address doubts pertaining some dimensions, mostly tackled by adding examples of contextualization and improving descriptions. Examples of these changes include, for instance, clearer definitions of what was considered as short, intermediate or long duration in some dimensions, and a better definition of **Task** interdependence. **Level of User Actuation** was also improved to reflect which side of the collaboration (i.e., on-site or remote) was being reported.

Finally, the main difficulty that raised up was addressed by creating the **Research** dimension, devoted to clarifying the context, maturity and detail of the collaborative work.

4.6 Applying the taxonomy to collaborative AR works

To illustrate the use and utility of the taxonomy (Šmite et al., 2014; Usman et al., 2017), ten publications that explored different aspects of remote collaboration (e.g., collaborative systems, aspects being addressed and evaluated, among others) were selected and analyzed to create an illustrative data set (Aschenbrenner et al., 2018; Gupta et al., 2016; S. Kim et al., 2019; S. Kim, Billinghurst, & Lee, 2018; Obermair et al., 2020; Piumsomboon et al., 2017; Piumsomboon, Dey, et al., 2019; Teo, Hayati, et al., 2019; P. Wang, Zhang, Bai, Billinghurst, Zhang, et al., 2019). The publications were thoroughly classified into the taxonomy dimensions, categories and characteristics by the same group of experts that created the taxonomy, to reflect the full extent of the reported information, in each publication. These works[23] were selected from journals and conferences between 2016 and 2020, including the ACM Conference on Human Factors in Computing Systems (CHI), ACM Symposium on Virtual Reality Software and Technology (VRST), Supported Cooperative Work (CSCW), Frontiers in Robotics and AI, International Conference on Industrial Engineering and Applications, IEEE International Symposium on Mixed and Augmented Reality (ISMAR), International Journal of Advanced Manufacturing Technology, IEEE Transactions on Visualization and Computer Graphics (TVCG), and the Symposium on Mobile Graphics & Interactive Applications.

To foster insight on the examples analyzed, as well as the way they are classified using the proposed taxonomy we created a quick illustration through a visual representation based on a sunburst (as shown in Figure 39). We chose to use a sunburst diagram, thus allowing to visualize the hierarchical data of the taxonomy, depicted by three levels of concentric rings. Each ring corresponds to a level in the hierarchy, with the inner ring representing the root node associated to the proposed dimensions. The hierarchy moves outward from the center to represent categories in the center ring and the characteristics in the outer ring. The data of each publication is represented by slicing and dividing rings based on their hierarchical relationship to the parent slice. In addition, a graphical representation of data in the outer ring is depicted using color to highlight the number of publications that addressed each specific characteristic. Mapping the number of publications that address each characteristic to color helps getting an overall understanding on how they are classified using our taxonomy. This approach allows to comprehend which categories and characteristics get the most attention and identify existing patterns or gaps in a visual way.

The overview provided by Figure 39 shows that for the **Team**, only distributed collaboration cases were addressed by the selected publications. Likewise, they all focused on teams composed by two collaborators, with a short-term life-span. From these, 7 out of 10 teams were functional, 2 teams reported divisional aspects and the remaining publication did not report the type of role structure. Regarding technology usage, 7 publications acknowledge their teams had continuous use of collaborative tools. In addition, 2 publications reported low turnover, and 1 high turnover, with the others not reporting any information on this characteristic. Regarding multidisciplinarity, only 3 publications reported the presence of team members with diverse background.

[23] - tinyurl.com/datasetTaxonomyCollAR [Accessed: 12-Apr-2021]

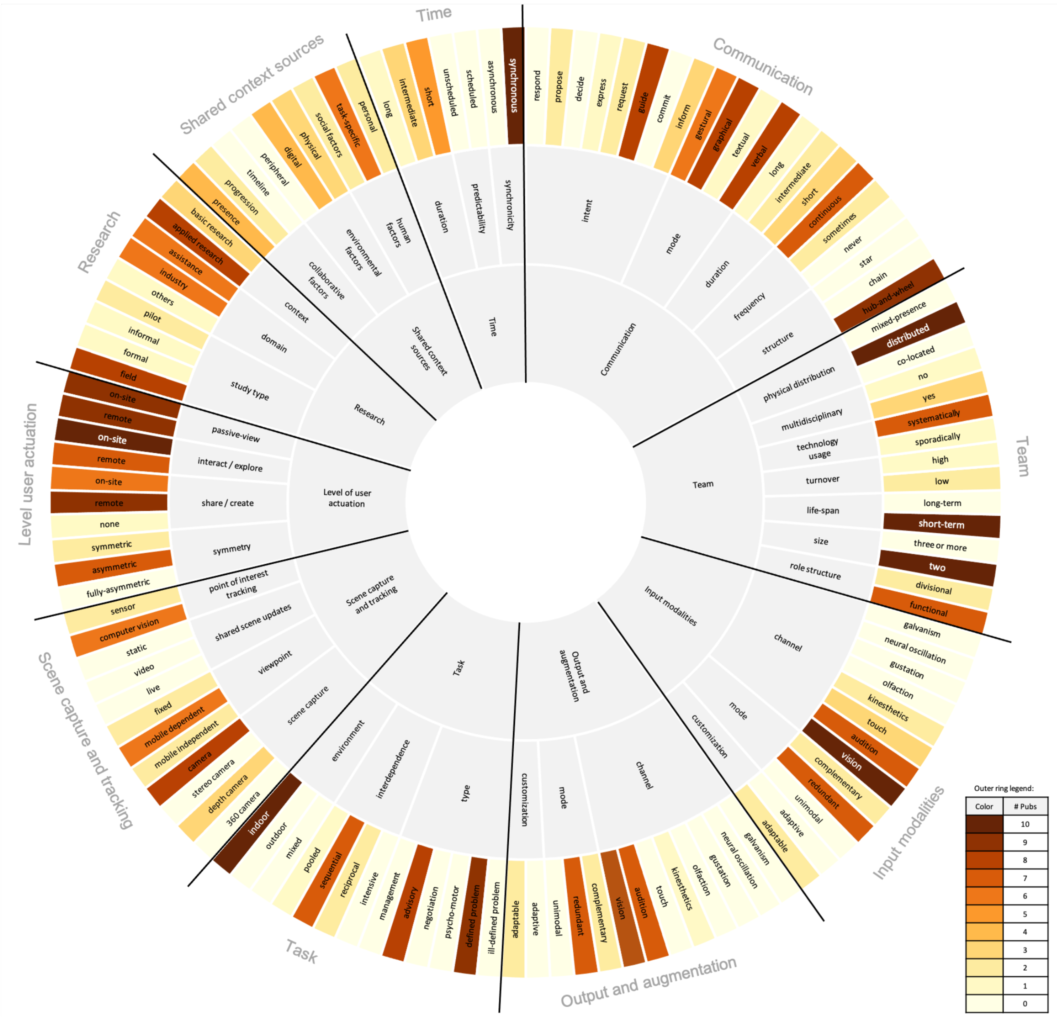

Figure 39 – Sunburst diagram displaying the hierarchical levels of the taxonomy: the inner ring represents the dimensions while categories and characteristics are showed as moving away from the center, respectively. The color scale shows the number of publications (out of a total number) addressing each characteristic. This example presents the results for the classification of ten publications: (Aschenbrenner et al., 2018; Gupta et al., 2016; S. Kim et al., 2019; S. Kim, Billinghurst, & Lee, 2018; Obermair et al., 2020; Piumsomboon et al., 2017; Piumsomboon, Dey, et al., 2019; Teo, Hayati, et al., 2019; Teo, Lawrence, et al., 2019; P. Wang, Zhang, Bai, Billinghurst, Zhang, et al., 2019).

Concerning **Time**, all publications focused on synchronous collaboration, with their efforts divided between short (5 publications) and intermediate (3 publications) periods of collaboration. There was no report regarding predictability.

As for the **Task**, only indoor environments (10 publications) were considered while addressing advisory (8 publications) and defined problems (9 publications). Sequential interdependence was explored by 7 publications, reciprocal by 2 publications and pooled by only 1 publication.

The preferred structured type associated to **Communication** was hub-and-wheel, being used by 9 publications. Moreover, 3 publications reported the communication duration was short, 2 publications intermediate, and only 1 publication long. In addition, 8 publications used verbal and graphical mode for communication, 6 explored gestural and 1 textual communication. Regarding communication frequency 7 publications reported it as continuous. Finally, the intent of communication was split among guide (8 publications), propose and request (2 publications) and inform (3 publications).

In respect to the **Scene capture and tracking**, 2 publications reported the use of sensors for tracking, while 6 publications used computer vision. Moreover, the majority used cameras for the scene capture (8 publications), and 3 resorted using a depth camera. Likewise, 6 publications used a mobile dependent viewpoint, 2 publication used mobile independent viewpoint and another 2 explored a fixed viewpoint.

Regarding the **Shared context sources**, the classification shows that 6 publications mention task-specific and 2 personal aspects, as well as social aspects in the human factors category. Besides, 3 publications report physical and 4 digital aspects in the environmental factors category. As for the collaborative factors, 4 publications mention presence and 2 progression.

Additionally, the **Level of user actuation** emphasizes that for the distributed teams, 9 publications reported on-site and remote team members had access to a passive-view of the task context. In the same way, all on-site collaborators could interact/explore, and only 5 publications report that the remote collaborator could do the same. In terms of share/create, the opposite situation occurs, with 9 publications reporting this ability being available to collaborators. Asymmetric level of user actuation was reported by 6 publications, while only 2 included mentions to symmetric possibilities.

For the **Output and Augmentation**, it is possible to observe that vision (10 publications), audition (7 publications) and kinesthetics (1 publication) were used as output channel. In addition, 7 publications explored a redundant approach and 2 complementary. Customization was only reported by 2 publications, using an adaptable approach.

In the same way, the **Input modalities** shows that vision (8 publications), audition (5 publications) touch (3 publications) and kinesthetics (1 publication) were used as input channel. The input mode was reported by 6 publications and focused on a redundant approach (5 publications) and complementary approach (1 publication). Regarding customization, only 1 publication explored an adaptable approach, suggesting that there is a lack of information in the selected papers or that this is not being addressed.

Regarding **Research**, 8 publications focused on basic research in the assistance domain, while 2 concentrated on applied research in the industry domain. Moreover, 8 publications described having conducted a formal study and reported on their results, while 1 publication reported an informal study and another a pilot study.

We further observe that some of categories and characteristics are not totally filled, due to lack of information being reported in the selected publications, as previously mentioned in the critical analysis and refinement by independent experts during their use of the taxonomy.

4.7 Discussion

While existing categorization efforts are suited for specific use cases or aspects of Collaborative AR, they do not cover the complete landscape of the field. Moreover, since different authors may use different notions when referring to the same aspects, depending on their context, it is important to make this context clear and provide a coherent common ground for systematization and discussion, thus fostering harmonization of perspectives, as well as reporting, and thus making comparative analyses easier.

4.7.1 Design and Validation

To explore this opportunity, we focused on a participatory design process and adopted a conceptual-to-empirical approach to understand the defining aspects of collaborative work supported by AR, i.e., what needs to be described about a work to provide a full account of its characteristics. As a result, our proposal is different from other efforts described in the literature, which use existing works as grounds and then, propose a taxonomy that encompasses them. These previous works have their merits and usefulness. However, such an approach implicitly assumes that existing research already covers all the different aspects required to fully address the problem. Therefore, the outcomes speak about where we are, but not necessarily if we are ticking all the important requirements and where should we go next research-wise, particularly if the field has not matured, as a whole. In the case of Collaborative AR, research has evolved tremendously regarding the supporting AR technologies, but at the onset of our work we argued that the field now needs to devote more efforts to understand how collaboration is being served in this context. To address this goal, we started by proposing and refining a conceptual model which then allowed the identification of ten dimensions that embody an extended human-centered taxonomy for the categorization of the main features of Collaborative AR. The work presented here proposes a set of dimensions that can be used to characterize collaborative AR not only addressing the technological features, but also encompassing the characteristics of the context they serve in the collaborative effort.

Considering the overall methodology adopted to reach the current stage of the taxonomy, it is important to note that its suitability to provide a structured view of Collaborative AR is not inferred solely from how well the four experts managed to classify ten recent articles. Although this is, naturally, a relevant outcome, and one could be tempted to increase the number of articles, the taxonomy is the result of an iterative participatory design method composed of several stages and it is the overall process that ensures its validation. In this context, the classification of ten articles works as one more refinement stage and the outcomes are presented here as further clarification on how to interpret the taxonomy given a set of recent research.

4.7.2 Utility and Impact

One of the purposes served by the adoption of the taxonomy is to increase the awareness of the research community regarding the characterization of the collaborative process, while also helping to identify those aspects that remain as research gaps. Additionally, the taxonomy provides a common ground to structure the different elements when conducting research. By increasing the awareness for the different dimensions, the taxonomy can foster additional transparency of the research, through better reporting, which, in turn, enables easier assessment by the community, fosters replicability, and the transferability of knowledge. By being provided with detailed information pertaining the different dimensions covered by a work, a researcher can assess the applicability and relevance of what is proposed to a new problem.

The range of the taxonomy results from the fact that Collaborative AR is an interdisciplinary area integrating different aspects from other research fields. Although in some particular situations one may advocate that simpler approaches as the ones described in section 4.1 may be used, the proposed taxonomy may also be used as a checklist of relevant dimensions to take into consideration, avoiding oversimplifying the collaborative process characterization. In this regard, nothing precludes researchers from considering a subset of the taxonomy, in particular cases, as the scope of their work. However, this will also put evidence on what is left out, and on the need to provide a rationale for it, based on the targeted research goals, along with a discussion of the contributions that is properly adjusted to the selected scope.

One essential point to note is that the proposed taxonomy is not intended as a closed work, but should, instead, be taken as the grounds that might enable the community to elaborate, expand, and refine it. Although some of the proposed dimensions might still not reflect the full scope of some categories by encompassing all possibilities, we consider that they create a clear enough organization to make itself evident where to insert new characteristics.

4.7.3 Identification of Novel Research Opportunities

The publications analyzed show how the taxonomy can be applied, yet, a larger use of the taxonomy is a paramount step expected to hint on interesting trends, identify research gaps (i.e., concrete areas which are not yet fully addressed or reported) and layout future directions in light of the proposed taxonomy. However, since, so far, most of the research efforts on Collaborative AR have been devoted to creating the enabling technology and proposing novel methods to support its design and development (Ens et al., 2019), it is expected that a majority of the data being reported corresponds to a subset of the dimensions covered by the taxonomy. As the field matures to focus on the nuances of supporting collaboration, the remaining dimensions will also flourish (e.g., the consideration of social aspects (Kreijns et al., 2013) as context sources) and accompany the growth already shown by the leading dimensions.

4.8 Visually exploring a Collaborative AR Taxonomy

To foster harmonization of perspectives and create a common ground for systematization and discussion, the design and development of tools that may allow to explore information associated with Collaborative AR in an interactive and visual way is of paramount importance. In this vein, we present a first effort towards the creation of an interactive visualization tool for exploration and analysis of collaborative AR research (Marques et al., 2021). In this section we motivate the visual design and describe its main characteristics. The contribute of using a visualization approach is to assist researchers to obtain an understanding of the field and how the dimensions relate with the literature, possibly informing further refinement of the taxonomy. Therefore, the goals we aimed to address were:

- better understand and systematize the collaboration dimensions included in the taxonomy, as well as their categories and characteristics;
- assess the number of papers addressing each dimension allowing to find gaps and research opportunities;
- select papers based on pre-defined criteria for analysis and comparison.

The proposed visualization[24] is based on a sunburst representing the hierarchy of dimensions, categories and characteristics in three levels of concentric rings (Figure 40 and Figure 41). This design allows a clear view of all dimensions and categories, while the number of papers do not overload the position as a visual encoding channel[25].

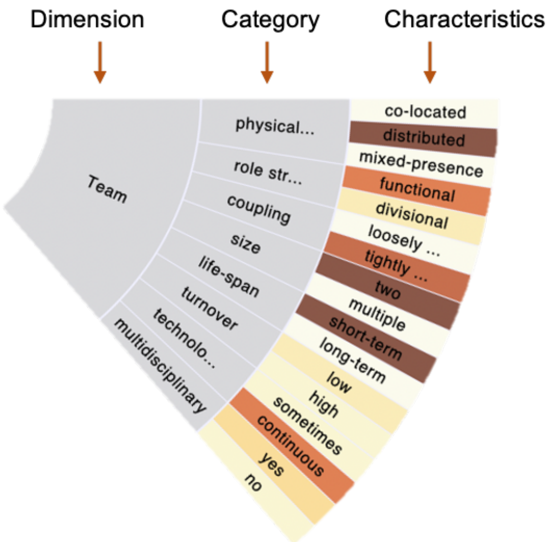

Figure 40 – Example of the visualization hierarchy associated with the dimensions, categories and characteristics of the taxonomy.

24 - tinyurl.com/visualizationTollTaxCollAR [Accessed: 12-Apr-2021]
25 - tinyurl.com/visualizationToolOverview [Accessed: 12-Apr-2021]

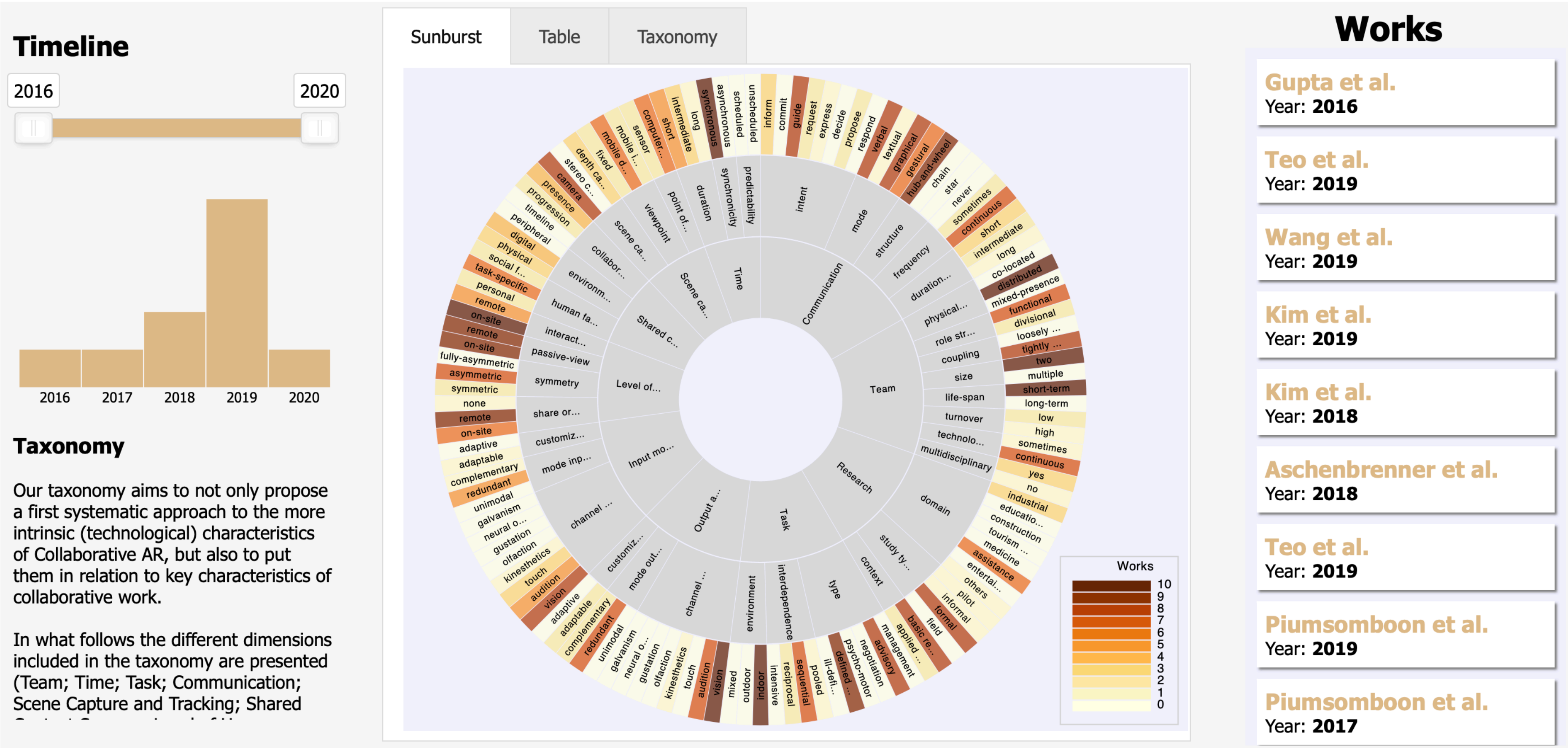

Figure 41 – Interactive visualization tool for analysis of different dimensions from a Collaborative Augmented Reality Taxonomy. On the left, the timeline slider to filter papers according to a time interval. In the center, the interactive sunburst visualization. On the right, the papers included in the data set used, which may be selected to preview the hierarchy of a specific paper. The data set includes papers by (Aschenbrenner et al., 2018; Gupta et al., 2016; S. Kim et al., 2019; S. Kim, Billinghurst, & Lee, 2018; Obermair et al., 2020; Piumsomboon et al., 2017; Piumsomboon, Dey, et al., 2019; Teo, Hayati, et al., 2019; P. Wang, Zhang, Bai, Billinghurst, Zhang, et al., 2019).

The sunburst visualization (Stasko et al., 2000) is enriched with a hierarchy navigation that can rearrange the hierarchy according to selection of an internal partition. The implemented visualization builds a visual idiom using the radial layout and hierarchy navigation of the sunburst. Using a partitioning method, it changes the original area from the aggregation by filling the area accordingly for each level, i.e., according to the dimensions, categories and characteristics addressed by the selected papers.

We decided to use this visualization design, given the taxonomy hierarchy it aims to address. Other approaches, for example, based on other visualization techniques of hierarchical data, such as treemaps present some limitations that our visualization overcomes for this particular scenario. To elaborate, a traditional tree structure brings attention to branches with many leafs, and it can hinder the navigation, besides it doesn't scale well with many nodes. Moreover, a treemap, even without the space filling algorithm, does not present a layout to highlight same level dimension, as it favors the values arrangement over the hierarchy alignment.

A radial layout also offers more space for the smaller nodes at the lower level of the hierarchies (Schulz et al., 2011; Woodburn et al., 2019), and this feature is vital for the Taxonomy hierarchy, as each characteristic at the leave nodes can hold meaningful information. The main change of the visual encoding channel of size from a sum to a fixed size is to prevent that the angle of the slices are too slim at the lower levels of the hierarchy, a known problem of radial designs (Woodburn et al., 2019). As such, a radial disposition of elements allows for the tilted disposition of labels, using the necessary space of each partition. In this context, the data of each paper is represented by slicing and dividing rings based on their hierarchical relationship to the parent sector.

The number of papers addressing each specific characteristic is represented in the outer ring through color mapping using a double-hue (Yellow to Brown) color scale. Mapping the number of papers that address each characteristic to color helps understand how the analyzed set of works is classified. This approach allows to understand which sectors (categories and characteristics) get the most attention and visually identify patterns or gaps.

The selection of a dimension rearranges the hierarchy to show only the selected dimension and its categories, partitioning the categories on the new space (Figure 42). In turn, selecting a category presents only the characteristics. Besides the interaction on the visual marks, the visual idiom also has contextual widgets to filter papers by year using a slide (Figure 42). A histogram with the distribution by year is available to guide the range selection.

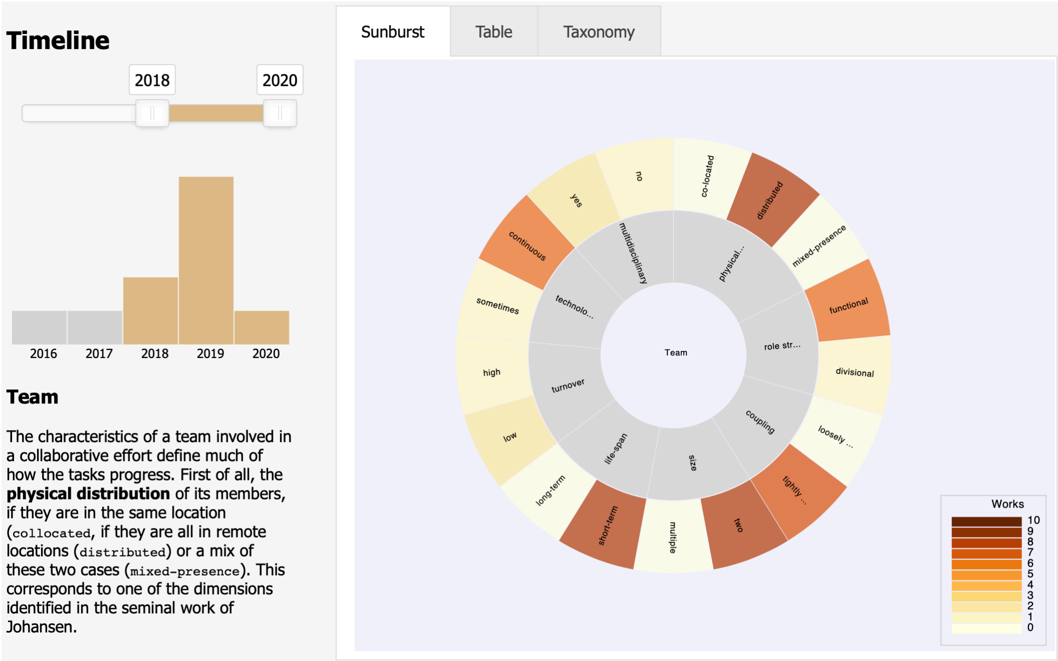

Figure 42 – Visualization displaying papers according to the Team dimension, ranging between 2018 to 2020.

Last, the cards on the right side of the visualization can be used for individual selection of specific papers highlighting the characteristics it embodies, as seen in Figure 43.

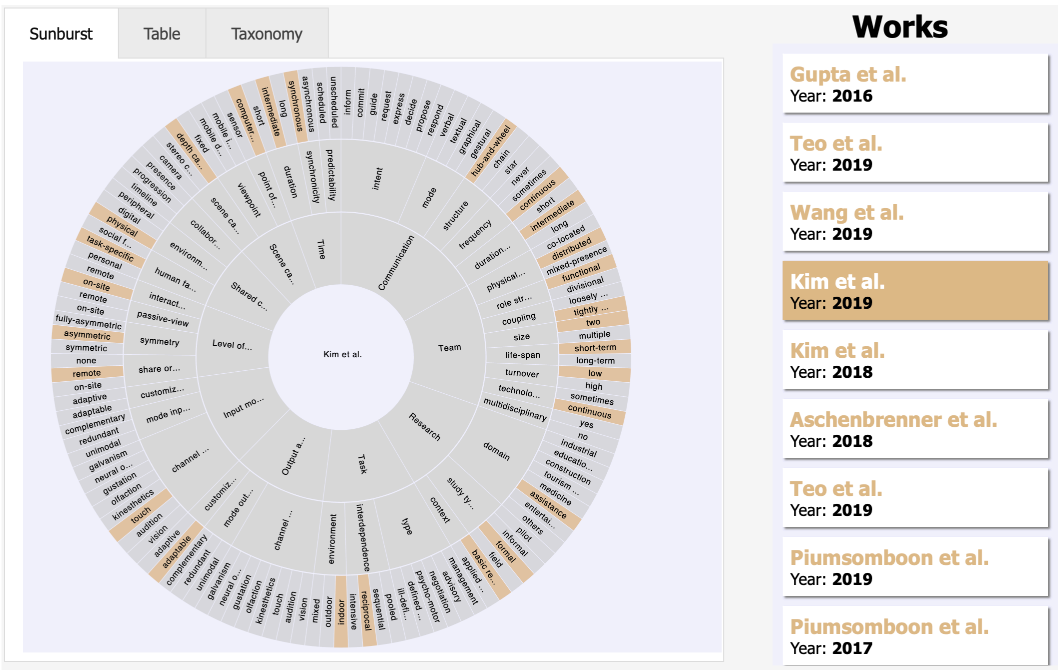

Figure 43 – Visualization of the dimensions, categories and characteristics of a specific paper (S. Kim et al., 2019)

Besides, the visualization design of the proposed tool is generic enough to be applied to taxonomies in other domains, as long as they follow a similar hierarchy (i.e., dimensions, categories and characteristics) to the one being used. Moreover, this information can be adapted over time, as the field matures, since the visualization is created dynamically based on the elements of hierarchy. As such, e.g., if new dimensions appear, as long as that information is included in the hierarchy description (e.g., text file containing all necessary elements of the taxonomy being addressed), the proposed tool can automatically adapt to support new content.

The visualization tool was developed using D3.js and web technologies, using a web server to host the application, as illustrated in Figure 44 fostering greater versatility of use and distribution among the research community.

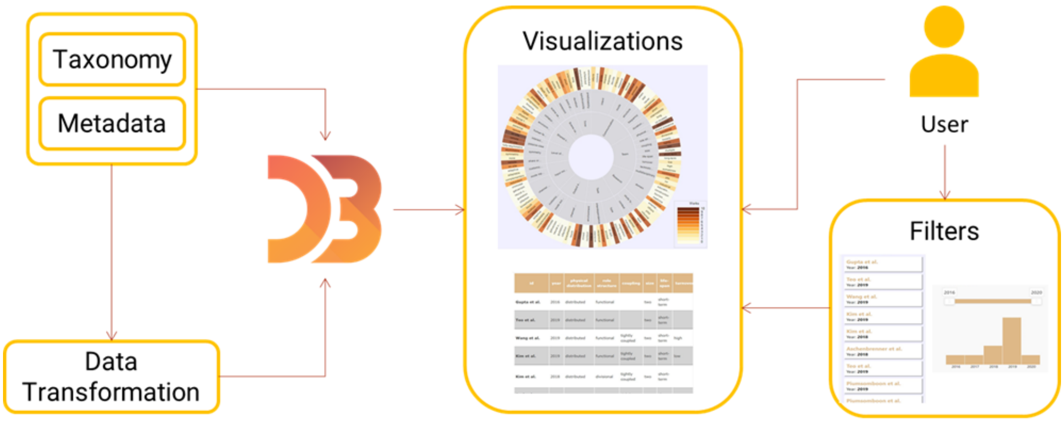

Figure 44 – System architecture and update process. The entire tool runs on a web browser, using D3.js to create and manage the visualizations. The filters update the visualizations on the go, as each filter can impact the number of visible slices.

4.8.1 User Study of the Taxonomy visualization tool

Next, we describe a user evaluation conducted in two universities from different countries. It was started through a use case showing how a set of tasks can be solved with the tool to gain insights into whether the hierarchy was easy to perceive and analyze, as well as understand if the representation used was simple and efficient to use. Then, a survey was conducted, where we asked participants to evaluate the usability and acceptance of the tool.

Dataset

Regarding the data set, we included the papers[26] described in section 4.6 (Aschenbrenner et al., 2018; Gupta et al., 2016; S. Kim et al., 2019; S. Kim, Billinghurst, & Lee, 2018; Obermair et al.,

[26] - tinyurl.com/datasetTaxonomyCollAR [Accessed: 31-Mar-2021]

2020; Piumsomboon et al., 2017; Piumsomboon, Dey, et al., 2019; Teo, Hayati, et al., 2019; Teo, Lawrence, et al., 2019; P. Wang, Zhang, Bai, Billinghurst, Zhang, et al., 2019), since these explored different aspects of remote collaboration and represent an illustrative repository of works between 2016 and 2020.

Tasks

Participants were asked to complete the following tasks, deemed relevant to understand if the visualization is useful to survey the selected papers, i.e., comprehend relationships among concepts, as well as infer trends and opportunities within the field:

- Check how the sunburst representation of the data evolves over the years;
- Evaluate which characteristics had more papers addressing them;
- Describe the characteristics of the *team* for a specific paper;
- Count how many papers reported *basic research* in the last 2 years;
- Identify opportunities for new research.

Measures

Participants' opinion was obtained through a post-task survey, including: 1- demographic information (e.g., age, gender, occupation, previous experience in the domains of visualization of information and AR, as well as with tools for remote collaboration); 2- SUS; 3- additional questions concerning the possibility to comment the characteristics of the visualization tool, suggest additional features or possible changes.

Study procedure

At the beginning of the study, participants were instructed about the experimental setup, the tasks and gave their informed consent. Then, they completed the tasks, while observed by a researcher who assisted them if they asked for help. A standard form was used to take notes (e.g., main difficulties, etc.). Immediately after completing the tasks, participants answered the post-study survey. During this process, all measures were followed to ensure a COVID-19 safe environment.

Participants

Forty participants (4 female) aged from 20 to 45 years old, performed the tasks and completed the post-study survey. Participants had various professions, e.g., Master and PhD Students, Researchers and Faculty members, as well as Software Engineers and Front-End Developers.

From these, 36 participants had previous experience in the domain of Information Visualizations and 25 in the field of AR. All participants had previous experience using tools for remote collaboration like Skype, Zoom, Team Viewer, among others.

4.8.2 Results and Discussion

All participants were able to use the interactive visualization tool to complete the tasks. The SUS score was 71.8, implying an above average usability, which can still be improved. In addition, answers to the post-task questionnaire show that the proposed visualization can be used to quickly relate the characteristics of each paper to the center and inner rings, thus allowing to understand how each individual paper is represented, which can be used to compare papers addressing similar research questions. The color mapping in the outer ring was considered useful by 35 out of 40 participants, although some stated that it was not intuitive at first to convey the number of existing papers, which must be revised in future iterations.

Conducting an in-depth analysis of the dimension, categories and characteristics of several papers at once generated mixed feelings between participants, with 16 out of 40 participants stating that identifying gaps and opportunities was not straightforward. To elaborate, one of the main challenges is the fact that moving between the different rings removes insight on the other ones, while affecting the notion of the full picture. In this context, participants suggested that the timeline may also require some changes, since it is not possible to filter among the provided dimensions, to understand how multiple papers addressed them over time, which can reveal other research gaps. Also, some future directions may be derived during the exploration of dimensions. One possible solution for these challenges is to extend the visualization using a Sankey diagram as (Ens et al., 2019), thus representing the evolution of each dimension along a given time period, without losing understanding of existing dependencies and correlations, either for a set of papers or for a single one.

Concerning future improvements, participants identified the need to expand the data set, in order to use a richer data sample that may provide additional challenges and insights during the analysis process. The illustrative data set being used was created to provide an initial case study for a first assessment of the proposed visualization, which although not representative of the field, can be used to assert most usability issues. Nevertheless, we plan to expand the data set with additional papers from top conferences and journals in the near future. By using the tool with a larger data set to understand the relationships among concepts, new gaps and trends may arise, which can help identify new research opportunities to move the field of Collaborative AR forward.

Furthermore, since the proposed tool aims to visually explore a taxonomy for Collaborative AR, it is important to consider how scalability may affect its design and performance, since taxonomies are not intended as a closed work, but should, instead, be taken as the grounds that might enable the community to elaborate, expand, and refine it. Although we consider the proposed tool provides an organization clear enough to insert new dimensions, categories and characteristics, we must be careful to ensure these last topics are properly addressed in future iterations, thus guarantying the exploration and analysis of information is not affected.

Finally, it is also important to continue to **explore interactive visualizations** to support the inclusion of new dimensions according to the evolution of the field. In this effort, we intend to share the tool with the research community, providing the ability to process data faster and properly explore, analyze and compare the characteristics of the collaborative effort mediated by AR as addressed in the literature.

5 A Vision for Evaluating Remote Collaboration using AR

We cannot solve our problems with the
same level of thinking that created them.

Albert Einstein

Given the challenges and constraints involved in evaluating the way collaboration occurs through Augmented Reality (AR) (see chapter 2), we argue it is paramount to address a set of important topics, namely: 1- **conduct more collaborative-centric evaluations**, i.e., move beyond usability testing, which fails to obtain a more comprehensive understanding of the work effort. Equally important, 2- **develop evaluation strategies that rely in more complete contextual data collection and more advanced visualization**, i.e., collect a richer data set to better understand how AR contributes to the collaborative process, in order to shape more effective collaboration.

Motivated by these, in this chapter, we start by presenting a knowledge-based ontology describing relations among dimensions of collaboration and the main concepts of the evaluation process. Then, we propose a conceptual framework for supporting researchers in conducting evaluations in a more structured manner with the goal of obtaining an additional perspective on several dimensions of collaboration. Then, we present the CAPTURE toolkit (Marques, Silva, Dias, et al., 2021b), a first instantiation towards the proposed vision to support evaluations in such scenarios, aiming to provide a strategy that monitors data concerning the level of collaboration, behaviour and performance of each intervening party, individual and as a team, as well as contextual data. To illustrate the advantages of the conceptual framework, the toolkit usefulness and versatility is demonstrated through a case study in a remote maintenance scenario, comparing two distinct methods: sharing of video and AR-based annotations. Then, the results obtained are discussed, and the added value of our proposal is summarized. Overall, the vision instantiated by the toolkit help allows to have an additional level of insights to better understand what happened, eliciting a more complete characterization of the collaborative work effort.

5.1 Ontology for Evaluation of Remote Collaboration using AR

Planning, designing, performing, and replicating an evaluation are demanding activities in remote scenarios mediated by AR. Given the difficulty to evaluate such scenarios, it is important to propose enhanced evaluation strategies to conduct thorough collaborative studies and provide an additional perspective on the different dimensions of collaboration. Thus, we presented an initial effort towards the creation of an ontology to guide researchers in designing and conducting their evaluations, aiming to generate an additional perspective on the nuances of collaboration. The proposed ontology establishes relations among the main dimensions of collaboration (proposed in chapter 4) and concepts of the evaluation process.

The goal of an ontology is to define a common vocabulary, i.e., set of terms for researchers who need to describe the facts of a given field. It captures the intrinsic conceptual structure of a field and usually covers classes that describe concepts of interest. In addition, ontologies may also share knowledge about reasoning strategies or problem-solving methods with others who have similar needs for knowledge representation, thereby eliminating the need for replicating the knowledge-analysis process (Chandrasekaran et al., 1999; Noy & McGuinness, 2001).

Literature shows that there is no correct methodology for developing an ontology, since its design is considered a creative process and no two ontologies by different individuals would be the same. The applications of the ontology and the designer's understanding of the domain will undoubtedly affect the ontology design choices (Chandrasekaran et al., 1999; Noy & McGuinness, 2001). To this effect, we performed brainstorming sessions (Jacko, 2012; Jerald, 2015) involving experts with several years of expertise in Human-Computer Interaction (HCI), Virtual and Augmented Reality (VR/AR), multimodal interaction, as well as remote collaboration.

The strategy chosen to define and populate the ontology was to consider the main dimensions of collaboration (e.g., team, time, task, communication, interaction, among others) as the core classes and associate them with other concepts that exist in common evaluation processes, like scope, design, setup, data, instruments, and others, through the reuse of an existing ontology for groupware evaluation (Araujo et al., 2004, 2003), as depicted in (Figure 45). The ontology aims to support a semantic knowledge base, which can be used to understand the scope of the evaluations of remote collaboration mediated by AR, how they were designed, their results and interpretations. More specifically, for registry how contextualized information can be used during the evaluation of the nuances of collaboration in scenarios where distributed team-members need to collaborate through AR to achieve a common goal.

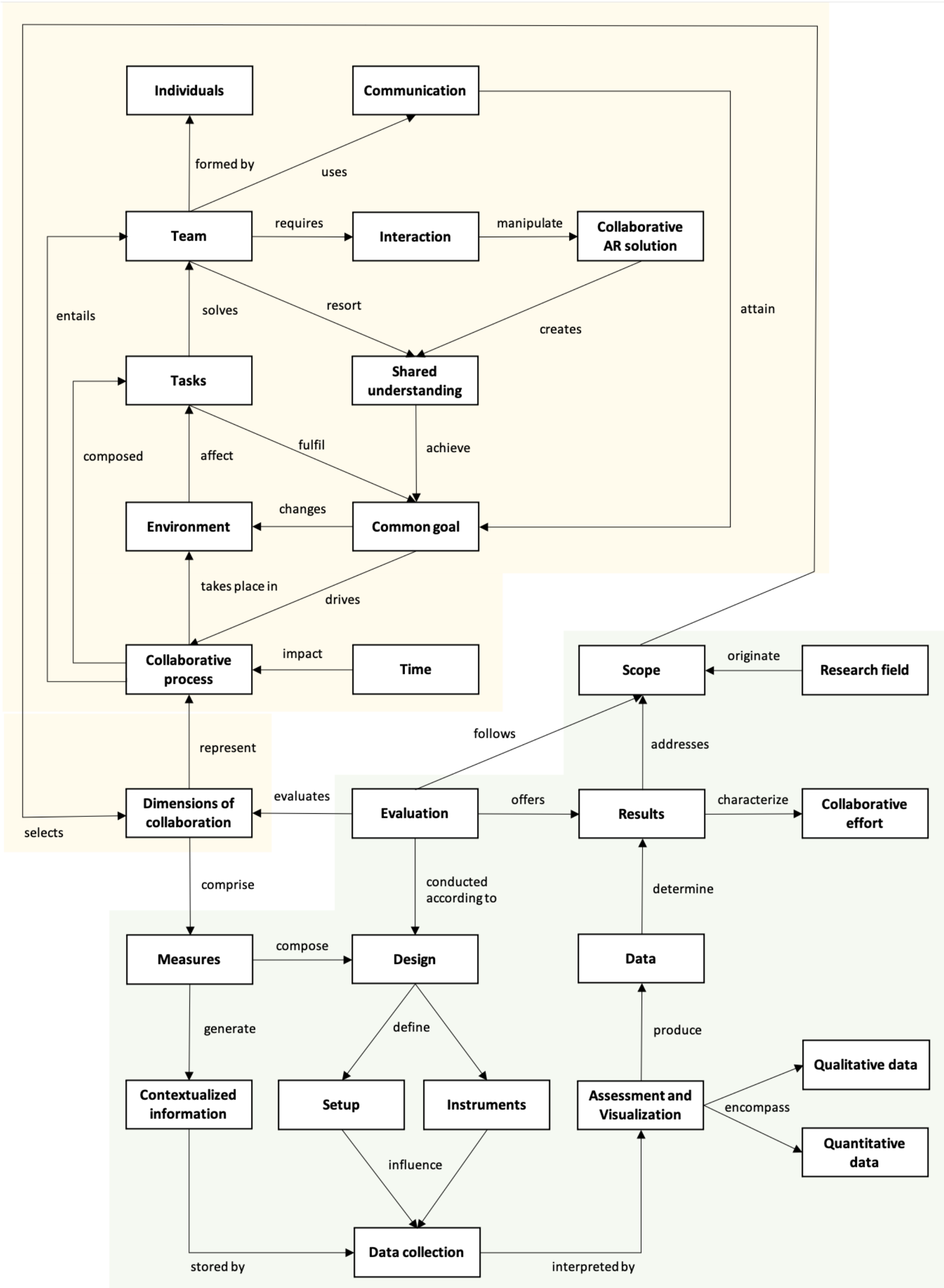

Figure 45 – Ontology for contextualized evaluation in scenarios of remote collaboration mediated by AR, which can be used to understand and guide the scope of the evaluations, how they were designed, their results and interpretations (Marques, Silva, Dias, et al., 2021c). In yellow: classes, properties and relations among dimensions of collaboration. In green: the main concepts of the evaluation process.

Evaluation is needed in order to address a specific scope generated within a research field. The classes, concepts and attributes in the ontology serve as a guideline for the evaluation design. In scenarios of remote collaboration, the collaborative process entails tasks that need to be fulfilled, time representing the synchronicity of the tasks, environments in which these tasks occur, and a team formed by distributed individuals. The team members must interact with each other through a collaborative AR solution, which serve as basis for situation mapping, thus creating a shared understanding. By communicating, the team members can analyze and discuss possible solutions to attain a common goal.

The first step for designing an evaluation in scenarios of remote collaboration is to identify the scope being addressed. Consequently, define which dimensions of collaboration make sense and are going to be evaluated, following the scope. Each evaluation dimension comprises pre-defined measures that can be chosen to compose the evaluation design as dependent/independent or secondary variables. The evaluation can be designed in terms of the setup that will be used by the team, as well as the instruments that can be used to collect data based on the selected dimensions.

Contextualized data are the expected outcomes of the evaluation and must be considered both quantitatively and qualitatively. The results of an evaluation comprise contextualized (qualitative and quantitative) data as the outcome of the collaboration process, which can be used to characterize the collaborative effort.

5.2 Conceptual Framework for AR-based Remote Collaboration

The area being addressed in this work is part of a complex phenomenon. To allow answering existing problems, it is necessary to systematize knowledge and perspectives, so that it can be applied transversely. For this, it is necessary the creation of evaluation frameworks, i.e., capitalize on the hierarchies and dimensions of collaboration from ontologies and taxonomies, as well as the development of tools that allow contextualizing the use of collaborative solutions.

Taking into account the challenges and needs identified in this thesis, Figure 46 structurally presents an evaluation framework of the collaborative process when using a given tool, with a proposal of several levels of information that must be considered for contextualization. In this effort, we argue that the evaluation process must be addressed by the research community, namely the definition of the evaluation purpose, as well as the team characteristics and the details of the collaborative tasks. Also, carefully establish the experimental setup and design. Equally important, explore contextualized data gathering and analysis, which requires the creation of novel tools. This last being the aspect this work further contributes. Next, we elaborate on these with more detail.

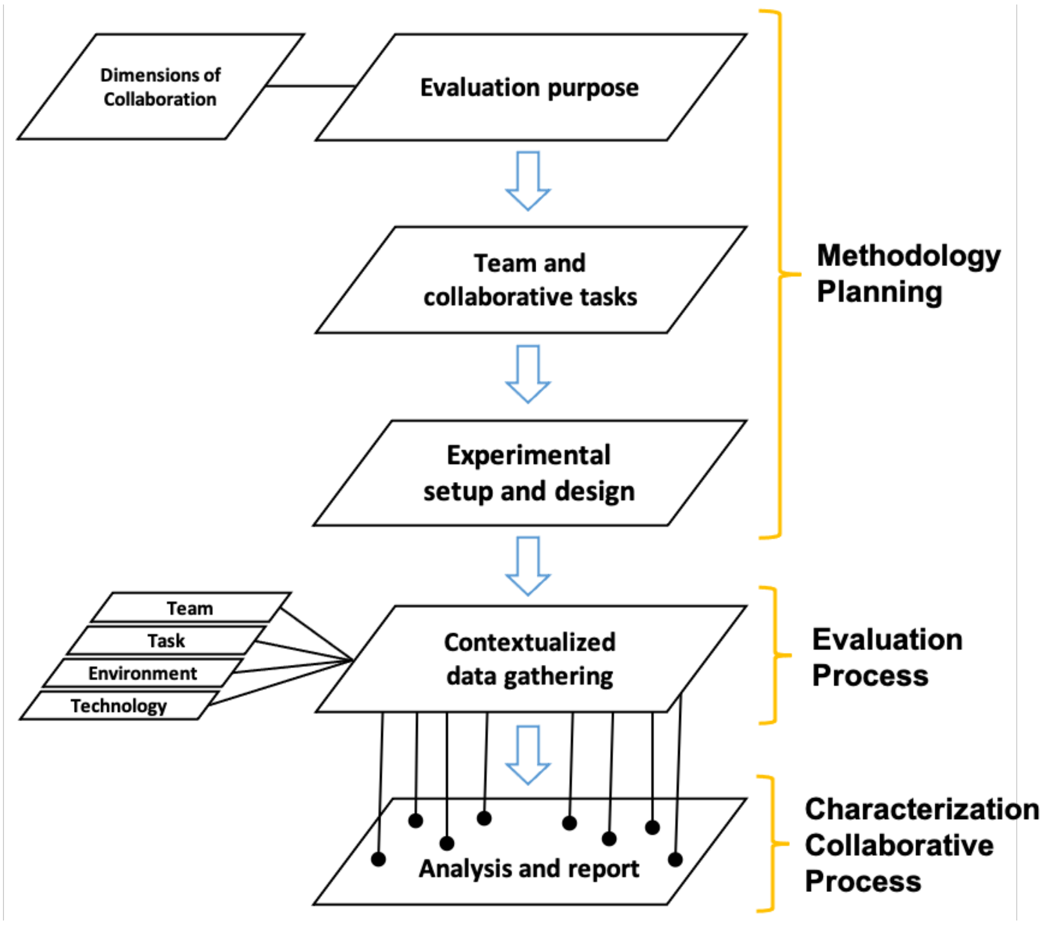

Figure 46 – Conceptual framework for helping researchers evaluate AR-remote collaboration in a more structured manner.

5.2.1 Evaluation Purpose

To begin, the scope of the study must be clearly defined, taking into account existing dimensions of collaboration to clarify what will be evaluated, so that relevant research questions are formulated in the design phase and answered in the evaluation analysis (Hamadache & Lancieri, 2009). As such, evaluation can be divided as follows: 1) technology centric associated with evaluation of prototypes or more mature systems, as well as comparison between systems with different approaches, e.g., different methods of tracking, interaction, etc.; 2) human centric focusing on design and human factors, e.g., performance, behaviour or emotions (S. Kim, Billinghurst, et al., 2020; Neale et al., 2004).

5.2.2 Team and Collaborative Tasks

It is also important to determine the team-members' characteristics, i.e., role structure, coupling level, life-span, technology literacy and multidisciplinarity. In this context, participants with different ages, perspectives, motivations, and multidisciplinary background should be considered, which might pose particular challenges, but lead to more relevant insights. Moreover, understanding of VR/AR, as well as remote tools is a benefit for the adaptation, thus removing the 'wow factor' that makes participants feel excitement or admiration towards such technologies. Besides, participants should only perform one role, , i.e., on-site or remote, so that they are only exposed to a set of tasks, concerns and responsibilities, unless the evaluation scope requires otherwise for context purposes.

Furthermore, the collaborative tasks goals must be clearly established including which team-members will be accountable for achieving at each completion stage. It is also important to consider the tasks, e.g., if they are performed indoor, outdoor, or mixed between the two; A balance must be kept between task complexity and duration. Tasks must be complex and long enough to encourage interaction through AR. However, longer tasks may cause fatigue or boredom, affecting the evaluation outcomes, as well as its ecological validity. Equally relevant, tasks can introduce deliberated drawbacks, i.e., misleading cues like incorrect, contradictory, vague or missing information, to force more complex situations and elicit collaboration. For example, suggest the use of an object which does not exist on the environment of the other collaborator or suggest remove 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. Additionally, tasks may elicit exploration of the physical workspace (i.e., environments that cannot be captured in a single camera frame), in particular, investigating different perspectives, rather than simply focusing on objects on a flat surface.

5.2.3 Experimental Setup and Design

Establishing the experimental setup and design is equally key. When considering prototypes, evaluation under laboratory settings should be used. Afterwards, when considering more mature solutions, evaluation should be made in the field, with real stakeholders and domain experts, moving beyond typical laboratory settings to increase the ecological validity of the evaluations. Regarding the environment, two separated rooms in the same/different building(s) should be used. Otherwise, participants must be separated by some kind of physical barrier when in the same room. Furthermore, an adaptation period must be provided so that participants can explore the technology possibilities before the tasks, individually and as a team. Besides, a proper amount of time must be defined for other aspects, e.g., presentation of the study, pre- and post-task questionnaires, team interview, and others.

5.2.4 Contextualized data gathering

As well observed by (Merino et al., 2020), future works on Mixed and Augmented Reality (MR/AR) will elaborate on human-centered evaluations involving not only the analysis of user experience and performance, but also understanding the role of such technologies in working places, in communication and in collaboration. In this scope, contextual information helps inform the conditions in which the collaboration took place. It can also be used for understanding interaction and communication changes, namely if the surroundings affected the way teams collaborate, in such a way that they needed to adapt it. Also, it helps portrait the conditions in which team-members performed a given action, received information or requested assistance, which can be used to assess uncommon situations or identify patterns that can lead to new understanding of a given artifact, as well as identify new research opportunities. Without comprehending contextual information, it becomes difficult to assess important variables related to the collaborative process, which means the findings reported may be misleading or of limited value. Hence, these aspects have an important impact on how the studies must be prepared and conducted, influencing situation understanding, team-members communication, performance, and usage of AR.

Literature on Computer-Supported Cooperative Work (CSCW) shows that a better evaluation process can be supported by improved data collection and data visualization tools (Araujo et al., 2004, 2003). In particular, 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, tasks, environment and collaborative tool** (Marques, Silva, Dias, et al., 2021a; Marques, Teixeira, et al., 2020). Through these, a wide range of information is provided when performing judgment over the results and establishing conclusions. To elaborate, contextual data can be materialized by a wide number of relevant facts from the surroundings, providing context to a person, team or event. The following aspects are crucial to any collaborative systems and must be taken in account when developing and evaluating remote collaboration to better understand the real impact of each aspect in the collaboration effort:

- **Team:** Aspects related to individual attributes like skills, abilities, individual personalities, emotional state, cognitive load, motivations and agendas. Equally important, consider homogeneity of abilities, knowledge of others, level of trust and performance throughout. For example, data gathering regarding team-members relation and communication during the execution of the tasks might significantly impact the understanding of the collaborative effort during the analysis process, informing if team-members are compatible or should be matched with other professionals in future iterations;

- **Task:** Aspects related to shared activities like duration and scope of the activities, type of collaboration (synchronous or asynchronous), amount of information required to conduct the activity, completion stages, order of steps, resources necessary. An illustration is registering how the available information was used to support the accomplishment of the tasks, like the nature of the material (physical or digital), the interdependencies with the work of other collaborators, or the number and type of interactions;

- **Environment:** Aspects related to contextual factors in the surrounding space, e.g., weather conditions, noise, dimensions, resources available, illumination, among others, which may not even be part of the immediate task or goal, but can still make a significant impact in the collaborator's performance; For instance, data collection about the amount of movement in the environment, the duration spent in each space, the tools available and the number of augmented annotations created can help understand the context of each space visited by the team-members to fulfill the common goal, which may inform how to prepare for future sessions of remote collaboration in these spaces, thus avoiding some surprises generated by its conditions;

- **Collaborative Tool:** Aspects related to the use of AR for supporting the remote collaboration process, including communication, interaction, multimodal features, as well as access virtual content used to build a common ground, e.g., identification of use patterns, visual complexity of augmented scenes. As an example, gathering data about the cognitive load of collaborators and the visual complexity of the augmented content being shared (e.g., level of awareness; perception; interest; and mental stress) can inform about the mental state of the team-members and help explain situations of confusion, missing information, as well as inform on necessary redesigns to the system in order to avoid overload of information in next session of remote collaboration.

In this vein, a conceptual model for contextualized data collection and analysis in scenarios of remote collaboration using AR was created (Marques, Teixeira, et al., 2020), as illustrated in Figure 47. It uses a distributed paradigm, that allows researchers to run multiple evaluations at different locations simultaneously. In this process, they may define measures, custom logging and register interesting events they detect, which can be later reviewed in post-task analysis. By gathering additional data, researchers can assess a wide range of aspects regarding several characteristics when establishing conclusions regarding the collaborative process. The proposed model makes available to researchers and developers a comprehensive description of relations between individuals, their interconnection as a team and how an AR-based solution can affect the accomplishment of the tasks and the collaborative process.

These factors can help portrait the conditions in which collaborators performed a given action, received information or requested assistance. In addition, it 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. Contextual data can also be used for understanding interaction and communications changes, namely if the surroundings affected the way the team collaborate, in such a way that they needed to adapt it, which is especially important during longer or challenging evaluation scenarios. Therefore, the gathering and analysis of contextual data is essential in the evaluation of scenarios of remote collaboration to help researchers when performing judgment over evaluation results.

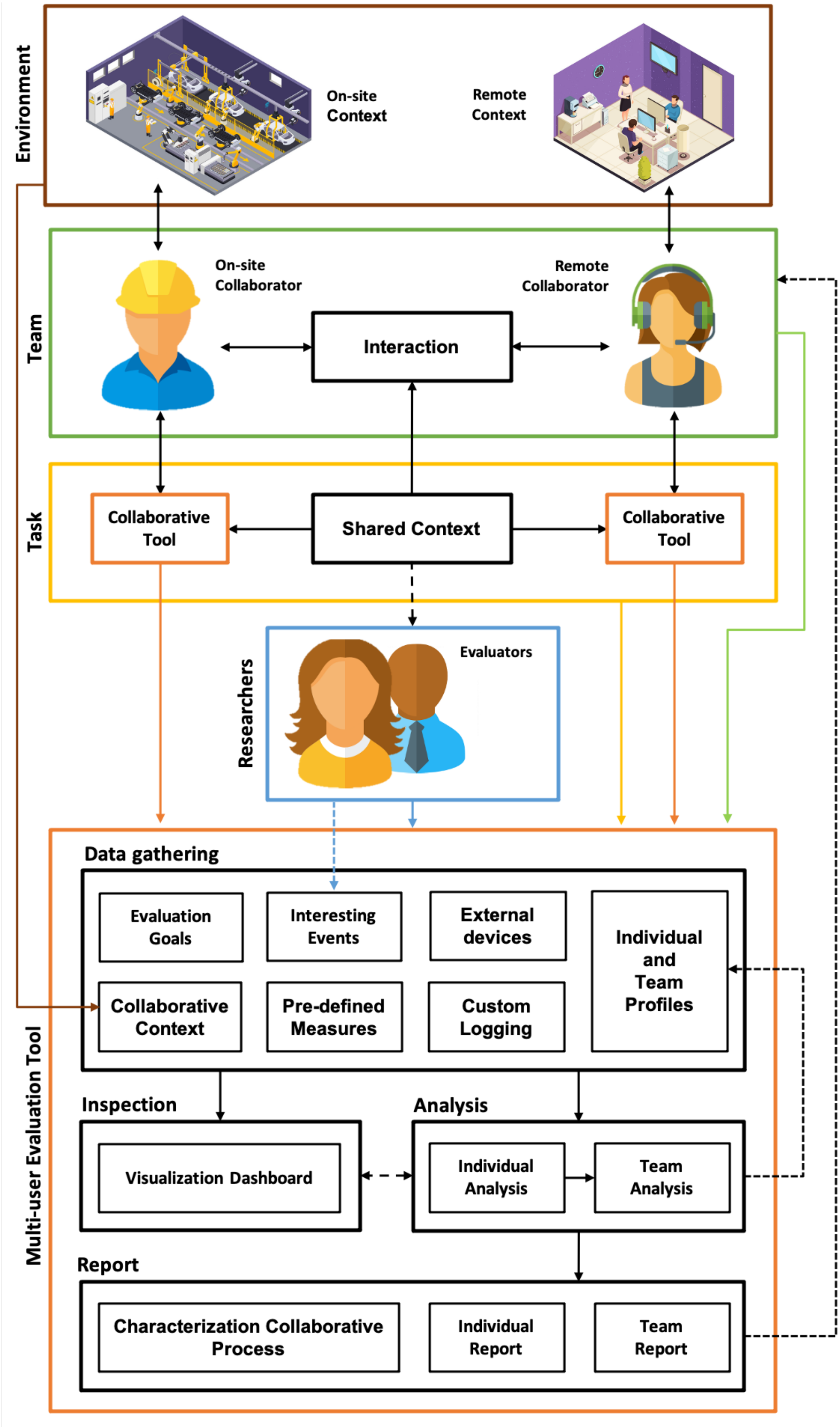

Figure 47 – Conceptual model for contextualized data gathering for remote collaboration using AR (Marques, Teixeira, et al., 2020).

Therefore, data collection while team-members collaborate, considering different forms of measure according to the evaluation goals is paramount and should include:

- **pre-task measures** like demographic questionnaires (e.g., age, gender, occupation, years of experience, etc.), information on participants background: if they knew each other, previous experience with VR/AR technologies and remote tools, and other aspects;
- **runtime measures** may comprise:
  - *performance metrics* including overall duration of specific events, e.g., when a task is started or completed; number and type of errors; number and type of interactions; frequency of using each feature of the tool, e.g., time spent and number of augmented content created; including when they are created, when a position and orientation changes, accuracy, etc., for a comprehensive understanding of what happens; screenshots of the enhanced content shared (for post-study analysis); custom logging, among others;
  - *behaviour measures* including conversational analysis (e.g., frequency of conversational turns, number of questions or interruptions, and dialog length, duration of overlapping speech, and others); physical movement around the environment; number and type of hand gestures; physiological variables and emotions; eye gaze, etc.;
  - *collaboration measures* including the level of effectiveness; perception; interest; engagement; awareness; togetherness; social interest; enjoyment/satisfaction; mental stress, or others also relevant;
  - *additional media:* researchers may collect audio (or video) and register *interesting events* including the type (e.g., guide, request, express, propose) and frequency of communication (e.g., never, sometimes, often, continuously), if the goals were accomplished, difficulties detected, if the participants requested assistance and how many types, among other relevant aspects.
- **post-task measures** can encompass:
  - register usability concerning the tools(s) used;
  - record *collaboration measures* including the level of effectiveness; perception; interest; engagement; awareness; togetherness; mental stress, etc.;
  - collect participants reactions, opinions and preferences through semi-structured interviews and questionnaires.

5.2.5 Analysis and Report

The use of more contextualized approaches will provide ground to improve how research is analyzed and reported. Hence, increasing the awareness of researchers about the different dimensions of collaboration and the need to improve how the nuances associated to the collaborative effort are described. In turn, a more systematic characterization can lead to a community setting that enables easier communication, understanding, reflection, comparison and refining, building on existing research while fostering harmonization of perspectives for the field. In this context, some noticeable recommendations are:

- researchers can profit from the outcomes generated to improve the level of detail provided in their reports;
- the collaborative context needs to be widely described, allowing the creation of a better understanding of the surrounding conditions, including relations between individuals, their interconnection as a team, how AR was used, the characteristics of the environment, etc.;
- the outcomes can help identify limitations and promising functionalities regarding AR, providing opportunities for future work at a technical level;
- the insights obtained may also lead to improvements in individual behaviour and team collaboration in specific procedures and tasks over longer periods of time.

5.3 CAPTURE – Toolkit for distributed evaluations using AR

Given the challenges in evaluating the way remote collaboration occurs, the absence of proper frameworks and tools, this section describes **CAPTURE - *Contextual dAta Platform for remoTe aUgmented Reality Evaluation***, a first instantiation towards addressing the vision previously described, in particular the need to include more contextual data in the evaluation of the collaborative process (Figure 48) following the conceptual model proposed above (Marques, Teixeira, et al., 2020). To inform the design and conceptualization, we conducted brainstorm sessions with domain experts sharing several years of expertise in HCI, VR/AR, Visualization and remote collaboration. Hence, the toolkit must support:

- data gathering at distributed locations simultaneously to run multiple evaluations;
- explicit input on different dimensions of collaboration, following the evaluation ontology for remote scenarios presented above;
- data collection regarding team interaction, custom logging and registration of interesting events according to the selected scenarios of remote collaboration;
- easy instrumentation into remote tools by providing ready to use scripts and Unity prefabs for non-experts in programming, i.e., each process can be configured via visual editors;
- modularity to ensure flexibility and adaptation to different evaluation goals;
- data storage and aggregation via a centralized server;
- post-task analysis through a visualization dashboard.

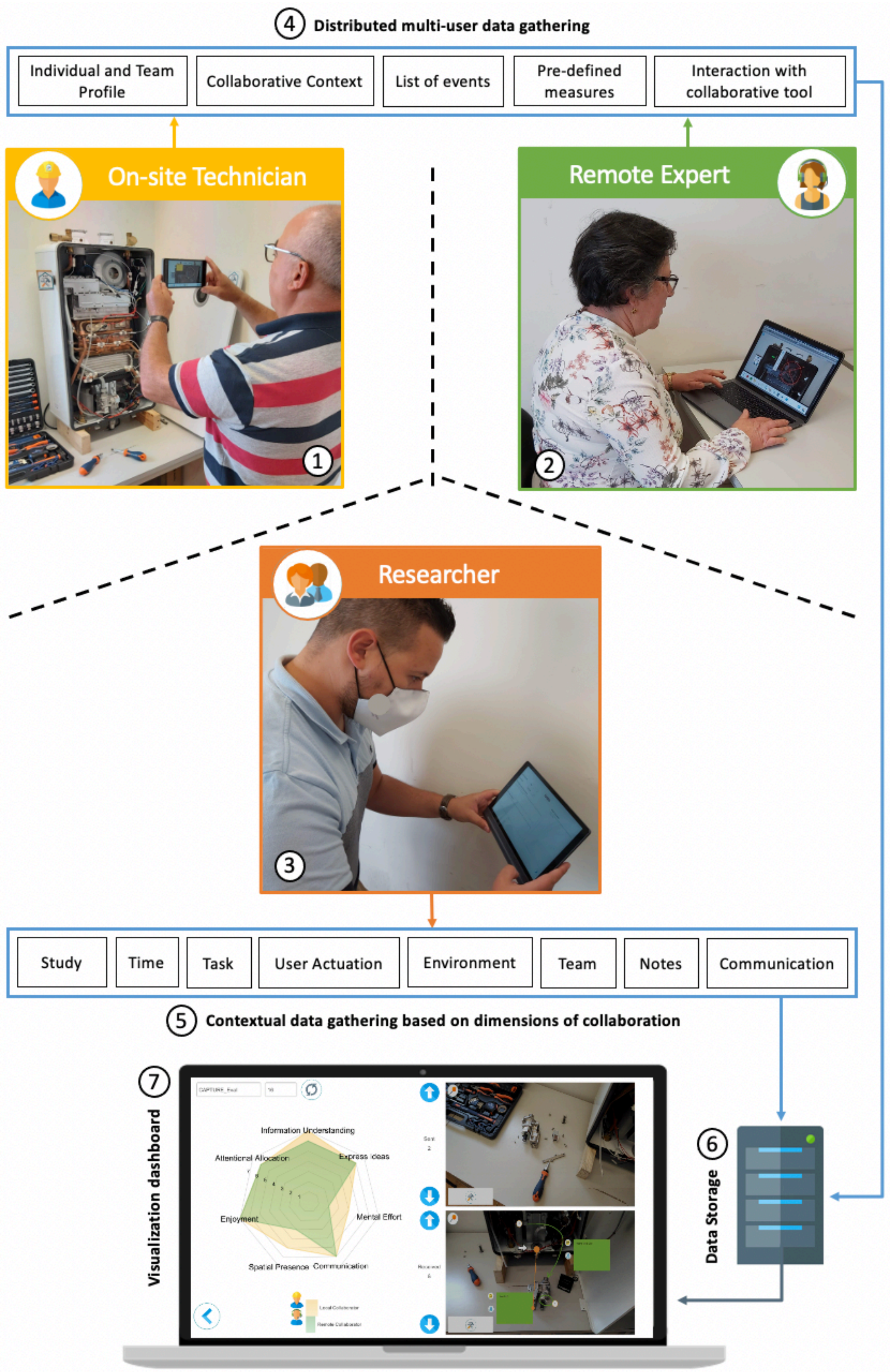

Figure 48 – Scenario of remote collaboration using an AR-based tool instrumented with the CAPTURE toolkit: 1- On-site technician requiring assistance; 2- Expert using AR to provide remote guidance; 3 - Researcher(s) following the evaluation process; 4- Distributed multi-user data gathering; 5- Contextual data collection based on existing dimensions of collaboration; 6- Evaluation data storage; 7- Visualization dashboard for analysis of the collaborative process.

To elaborate, for team-members, the CAPTURE toolkit provides native off-the-shelf modules to support explicit input and data gathering regarding (Figure 48 - 4):

- **individual/team profile**: demographic data, participants background, knowledge of other collaborators, emotional state (Izard, 2007), experience with AR and remote tools;
- **collaborative context**: details on the task and the environment, like the number of completion stages, resources available or the amount of persons, movement and noise;
- **list of events**: task duration, augmented content shared and received, other occurrences;
- **pre-defined measures**: characteristics associated to the collaborative process, including, but not limited to, easy to communicate or express ideas and the level of spatial presence, enjoyment, mental effort, information understanding, attention allocation or others (Figure 49 - top). Also, the Microsoft reaction card methodology (Barnum, 2010) to have a grasp on team-members reaction towards the tool used for shared understanding (Figure 49 - bottom);
- **interaction with the collaborative tool**: duration of the collaborative process and specific events, e.g., when creation of content is started or completed, number and type of interactions, frequency of using each feature, as well as captures of the augmented instructions being shared.

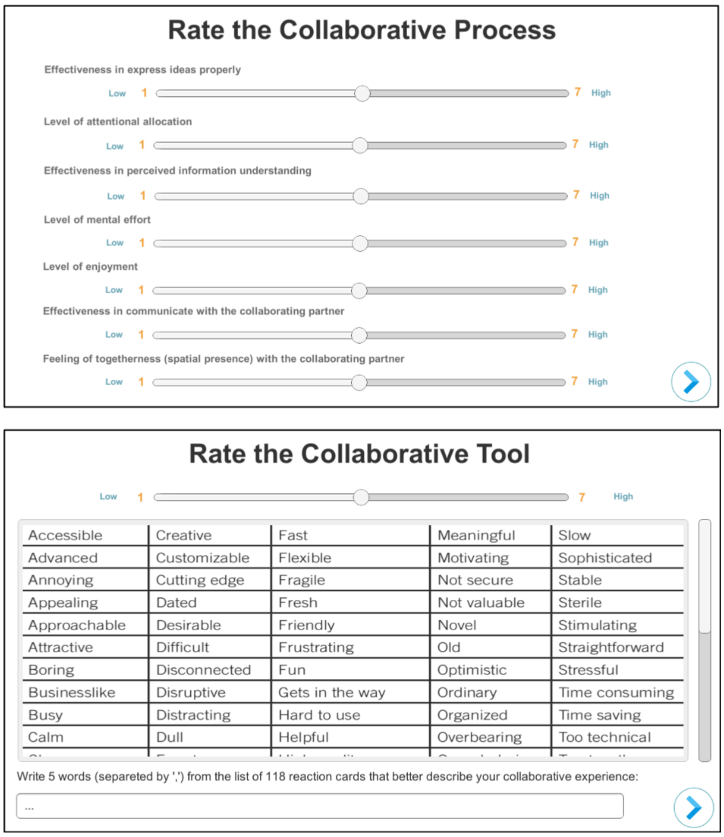

Figure 49 – CAPTURE toolkit - example of pre-defined screens associated with post-task measurements. Top - questionnaire regarding the collaboration process; Bottom - questionnaire regarding the collaborative tool.

Regarding pre-defined measures, it is important to clarify the aspects of collaboration proposed are the result of carefully surveying existing literature to create a list of important topics facing the lack of methodologies and frameworks, thus eliciting a better perspective into the collaborative work effort supported by AR. To elaborate, we took inspiration from the questionnaires used by (Gutwin & Greenberg, 1999; Huang et al., 2019; S. Kim et al., 2014; S. Kim, Billinghurst, Lee, et al., 2018; Nilsson et al., 2009), as well as the works by (Aschenbrenner et al., 2018; Fakourfar et al., 2016; Johson et al., 2015; S. Kim et al., 2019; S. Kim, Lee, et al., 2020; Patel et al., 2012; Piumsomboon, Dey, et al., 2019; Teo, Lee, et al., 2019). Nevertheless, other aspects of collaboration can be considered according to the evaluation scope due to the inherent flexibility provided by the CAPTURE toolkit[27] implementation, as described at the end of this sub-section.

As for the researcher(s), the toolkit provides native off-the-shelf modules to support explicit input regarding selected dimensions of collaboration, following the work proposed in chapter 4 (Figure 48 - 5):

- **Research**: area of application, research context and study type;
- **Time**: synchronicity, duration and predictability;
- **Team**: distribution, role structure, size, life-span, turnover, multidisciplinarity, technology usage, homogeneity of abilities, and knowledge of others (Figure 50 - top);
- **Task**: scope and type of task, interdependence, amount of information and movement required to fulfil the task, number of completion stages, resources necessary to achieve the goal (Figure 50 - bottom);
- **User Actuation**: capacity to passive-view, interact/explore, share/create, as well as level of symmetry;
- **Communication**: structure, mode, intent, frequency and duration;
- **Environment**: amount of noise, level of brightness, number of persons in the environment, weather conditions and resources available;
- **Notes**: interesting events, notes, comments, or difficulties, as well as if the goals were achieved and the amount of physical movement conducted by the team-members.

[27] - tinyurl.com/CAPTUREToolkit [Accessed: 21-Jul-2021]

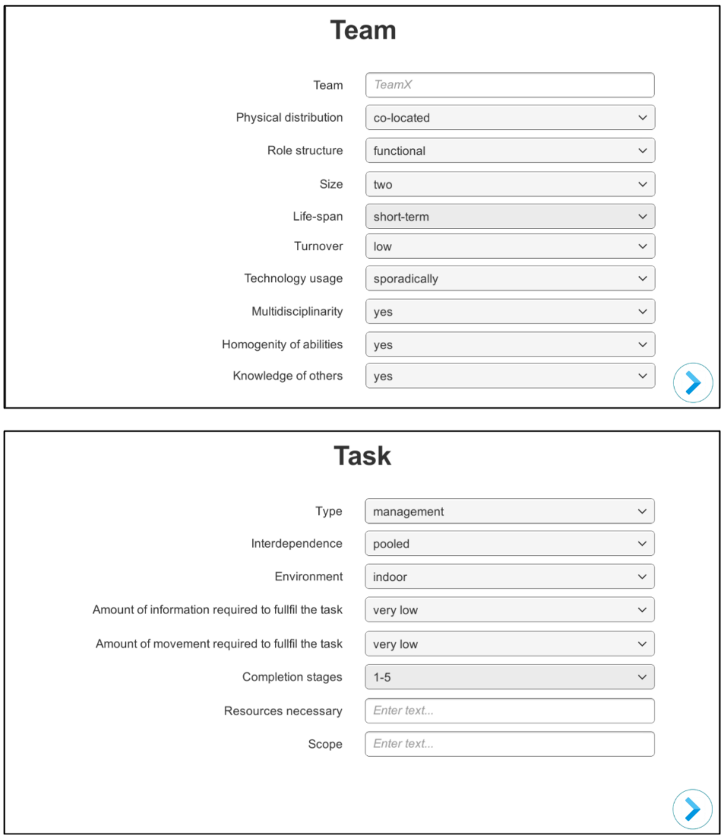

Figure 50 – CAPTURE toolkit - example of pre-defined screens associated with selected dimensions of collaboration. Top - characteristics of the Team; Bottom - characteristics of the Task.

At the system level, CAPTURE consists of a Unity Package that can easily be added to existing collaborative solutions in Unity. All data gathered from the different team-members and researcher(s) during collaboration sessions is stored in a central server for post-evaluation analysis through a visualization dashboard (Figure 51), which allows reviewing the work effort of a particular team or set of teams, as well as comparing different tools, if that is the evaluation scope.

The modules of the proposed toolkit can be integrated into existing remote tools via visual editors, i.e., with minimal need for programming skills (Figure 52). It is possible to drag and drop ready to use prefabs and editable scripts into Unity 3D projects, which can be modified according to the evaluation scope in the inspector module. Figure 52 illustrates the example of the collaborative process script, which researchers can manually edit (set the number of elements, add relevant aspects of collaboration to be assessed, etc.) according to the evaluation scope. This dynamic approach allows researchers to re-use scripts over different evaluation sessions according to the collaborative effort being considered. For development, Unity 3D was used based on C# scripts. Communication between each instance is performed over Wi-Fi through calls to a PHP server.

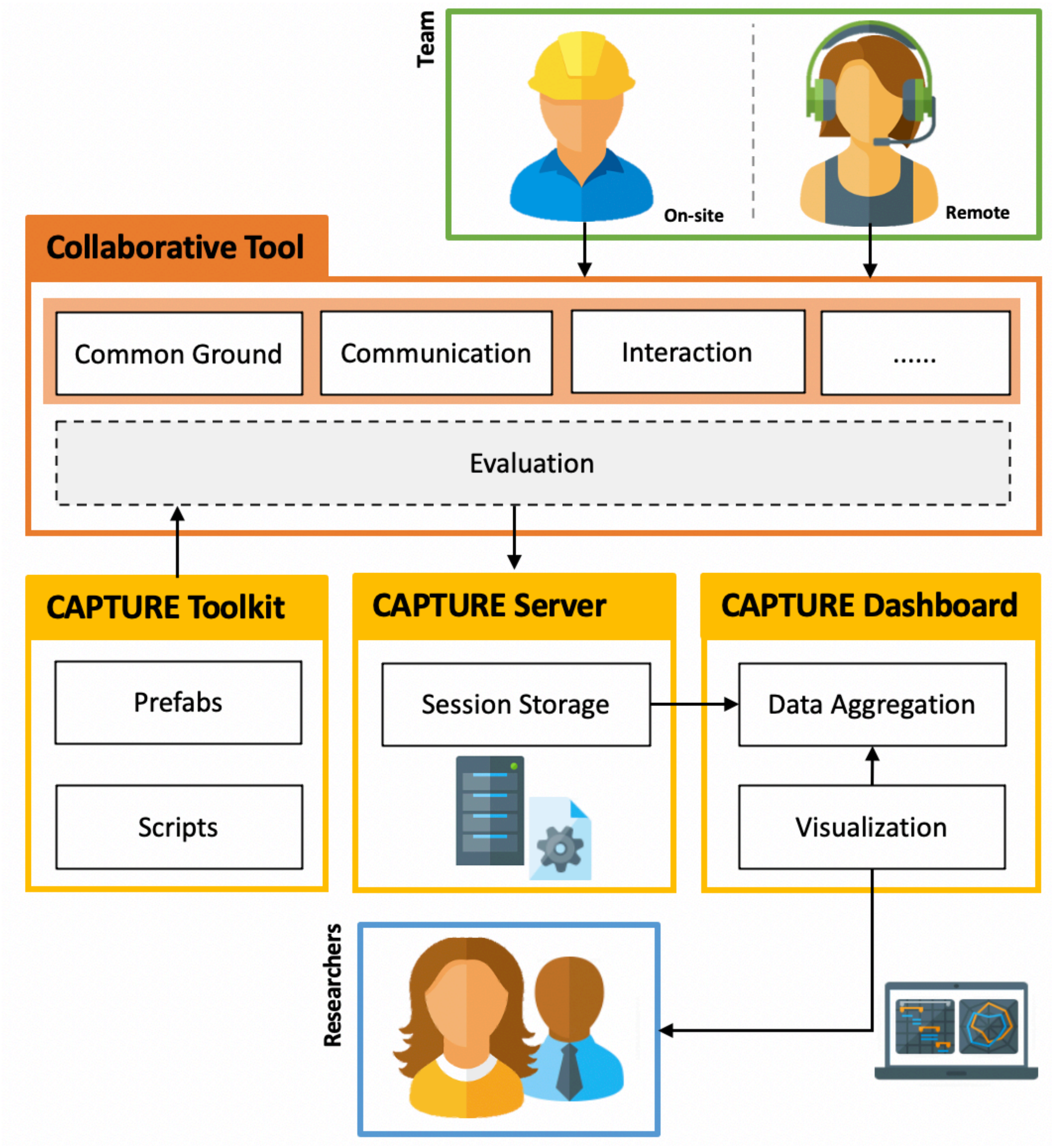

Figure 51 – CAPTURE architecture. The toolkit can be integrated into a collaborative tool via visual editor. All data collected during collaboration is stored in a central server, which can be analysed during post-task analysis through the visualization dashboard.

In short, the field needs to have more contextualized evaluation strategies, allowing to learn more regarding how technology addresses the collaborative process. All of this can support an effort towards systematized data, which may support the proposal of guidelines in the future, resulting from the experience and knowledge accumulated through the analysis from multiple research teams and different technology approaches with contextualized information. This effort will allow use these recommendations to jump-start the quality of current and novel solutions right from the very beginning of their conceptualization, which have already been proven useful in remote scenarios.

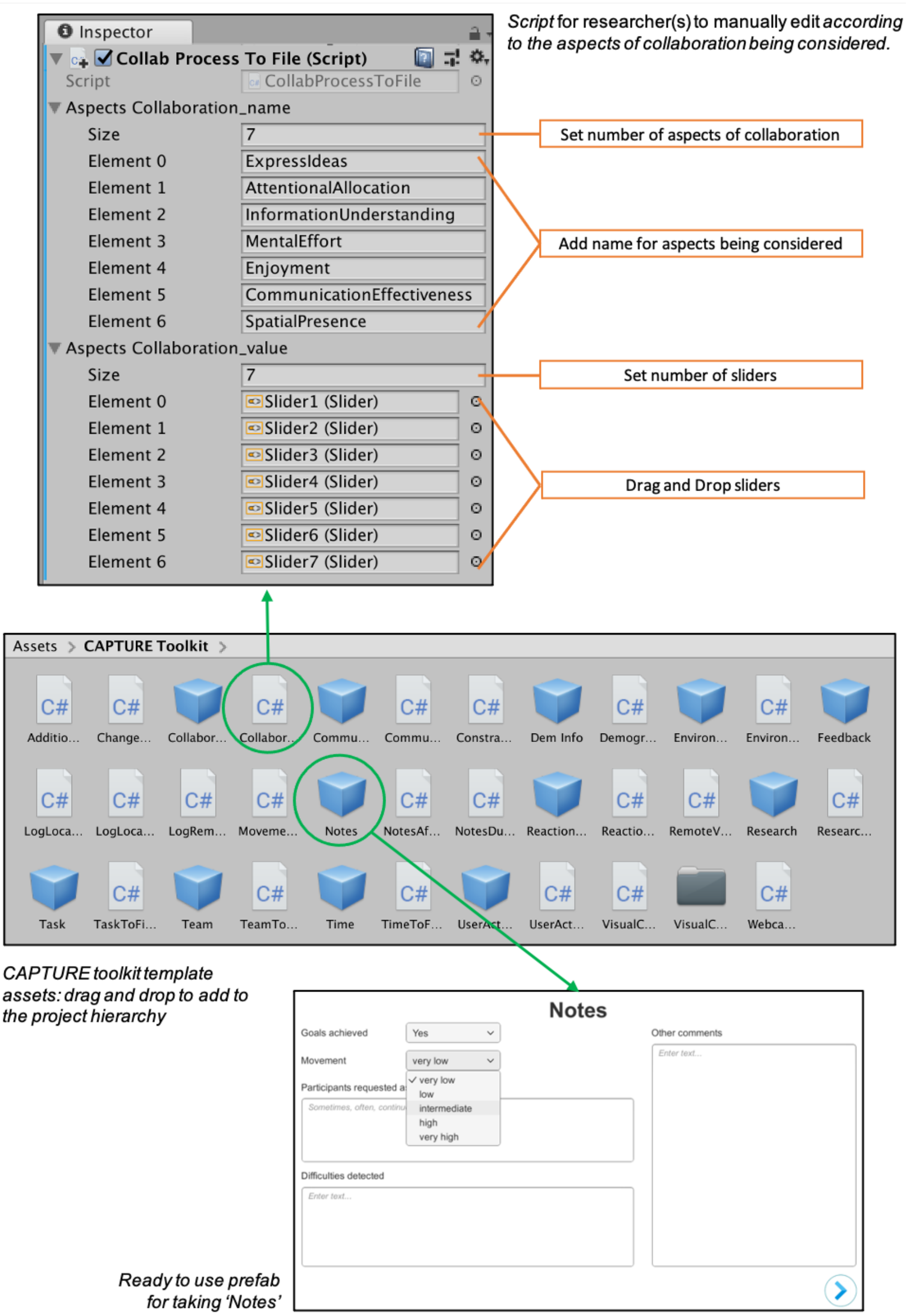

Figure 52 – Overview of the CAPTURE toolkit assets: ready to use scene prefabs and editable scripts, which researchers may modify according to the aspects of collaboration being considered for the evaluation.

5.4 User Study on a Remote Maintenance Scenario

A user study was conducted to compare the collaborative process of distributed teams using two distinct tools when instrumented with CAPTURE: Video Chat and AR-based Annotations. These were proposed following a user-centered approach with partners from the industry sector to probe how AR could provide solutions to support their collaborative needs.

5.4.1 Experimental Setup

To create a common ground between distributed team-members, two distinct methods were provided: a video chat tool and an AR-based annotation tool. Next, a brief description of the main features of each tool is provided. To clarify, the hardware used was the same for both methods, only the characteristics of the tool changed. Also, both tools were developed using the Unity 3D game engine, based on C# scripts. Communication was provided over Wi-Fi through WebRTC calls to a dedicated server. To place the augmented content in the real-world environment, we used the Vuforia library.

Video Chat Tool

The first method uses video chat features to provide support during the collaboration effort (Figure 53). On-site participants can point a handheld device to the situation context, which is shared though live video stream with the remote expert. In this context, the face of the expert is visible at all times, while the on-site participant may change between showing the task context or his face using the back and front cameras of the device. Besides, team-members can share text messages using the chat to ensure important messages are kept visible. Using these features, team-members may communicate and discuss the content being captured to express the main difficulties, identify areas of interest or the remote expert to inform where to act and what to do.

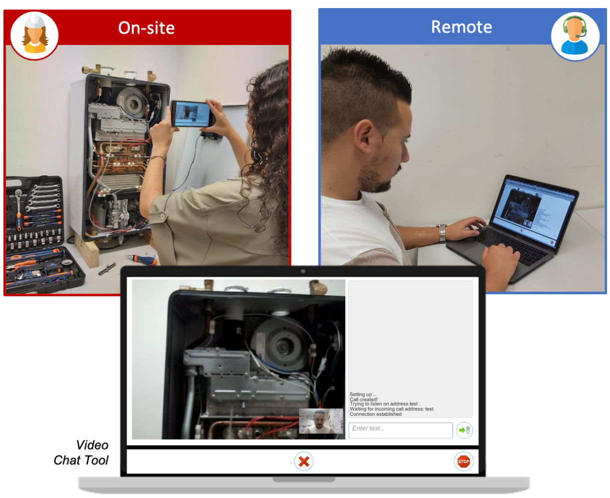

Figure 53 – Video Chat tool for remote collaboration.

AR-based Annotation Tool

The second method uses AR-based annotations as additional layers of information (Figure 54). On-site participants can point a handheld device to capture the situation context. Using audio communication and annotation features like drawing, placing pre-defined shapes or notes, as well as sorting annotations, the participant can edit the capture to illustrate difficulties, identify specific areas of interest or indicate questions. Then, the capture is sent to the remote expert, allowing provide instructions accordingly i.e., inform where to act, and what to do, using similar annotation features. Afterwards, the on-site participant receives the annotations. The handheld device can be placed on top of a surface to follow the instructions in a hands-free setting. At any time, it can be picked up to perform an augmentation of the annotations, by re-aligning with the real world. This process can be repeated iteratively until the task is successfully accomplished.

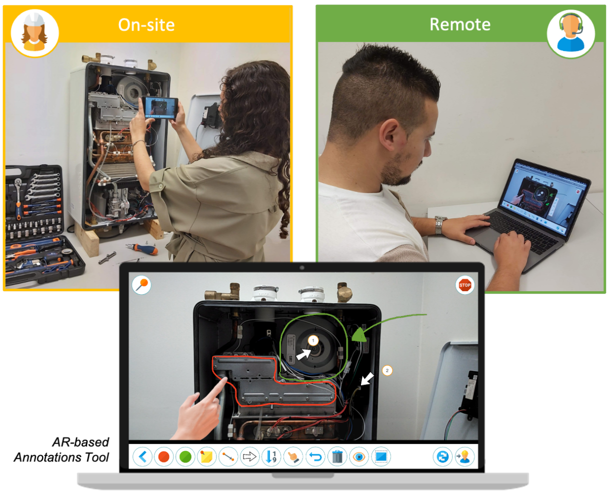

Figure 54 – AR-based Annotation tool for remote collaboration.

5.4.2 Experimental Design

A within-group experimental design was used. The null hypothesis (H0) considered was that the two experimental conditions are equally usable and acceptable to conduct the selected maintenance tasks. The independent variable was the information display method provided during the collaborative process, with two levels corresponding to the experimental conditions: **C1 - Video Chat and C2 - AR-based Annotations**. For both experimental conditions, the tools used provided a similar level of user actuation for both team-members, having identical features to view (C1 and C2), create, share and interact with augmented content (C2). Performance measures and participants' opinion were the dependent variables. Participants' demographic data, as well as previous experience with AR and collaborative tools were registered as secondary variables.

5.4.3 Tasks

We focused on a case study where an on-site participant using a handheld device had to perform a maintenance procedure on a boiler while being assisted by a remote expert using a computer. The tasks required accomplishing the following steps (Figure 55): 1- replace interconnected components, 2- plug and unplug some energy modules, 3- remove a specific sensor, as well as 4- integrate new components into the equipment. For each condition, different tasks were used to minimize bias, i.e., learning effect. Nevertheless, we defined these tasks based on feedback from our industry partners regarding their usual work activities and needs, while ensuring a similar level of difficulty and resources for each task. Each task was a defined-problem with 4 completion stages, forcing team-members to communicate in a continuous way while acting alternately (reciprocal interdependence) in an indoor environment with controlled illumination conditions and reduced noise. Besides the participants and researchers, no other individuals were present. The on-site participant needed to use different hand tools to perform the procedures, although low physical movement was required.

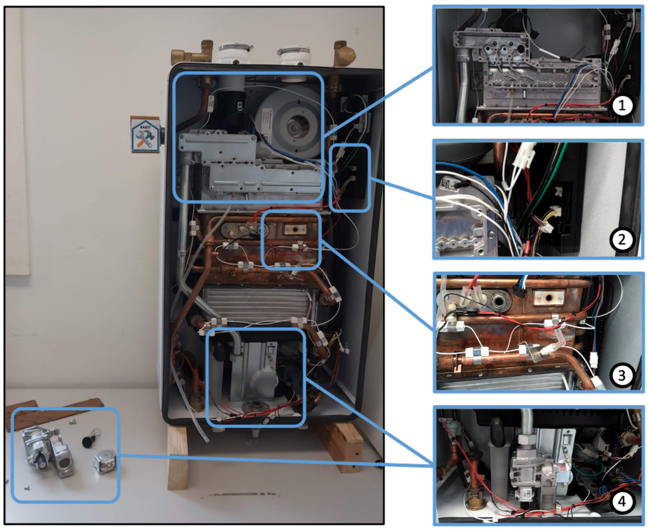

Figure 55 – Illustration of some of the completion stages associated with the maintenance tasks used in the study: 1- replace interconnected components; 2- plug and unplug some energy modules; 3 - remove a specific sensor; 4 - integrate new components into the equipment.

5.4.4 Measurements

All data was collected through the CAPTURE toolkit for all conditions, including standard measures found in literature like task performance based on the overall total time, i.e., time needed to complete the tasks, answer to questionnaires and participation in a brief interview, as well as task time, i.e., time required for successfully fulfil the task in a collaborative manner.

Besides, novel measures, taking advantage of the toolkit off-the-shelf modules, i.e., information on selected dimensions of collaboration (e.g., time, team; task; user actuation, communication, environment); the overview of the collaborative process (e.g., easy to communicate or express ideas, level of spatial presence, enjoyment, mental effort, information understanding and attention allocation) at the end of the tasks; participants emotional state, before and after the task fulfilment; participants preferences and opinion, also at the end. Hence, the toolkit was integrated into an existing video chat tool, as well as the AR-based tool using stabilized annotations, previously described in chapter 3.

5.4.5 Procedure

Participants were instructed on the experimental setup, the tasks and gave their informed consent. Then, they were introduced to both tools and a time for adaptation was provided. Participants would act as on-site technicians with condition C1 and then C2, always in this order, while a researcher was the remote counterpart to ensure the instructions were correctly transmitted. We used this approach to facilitate collaboration, as having participants also act as the remote counterpart would add an additional level of complexity, which in discussion with our industrial partners did not appear necessary. To elaborate, in most real-life cases they encounter, there are a small number of remote experts ready to assist, while the number of on-site technicians keeps growing. Therefore, following this situation, it appears logic to start by having the participants acting as on-site technicians. Nevertheless, it still allows to have a granular view of the work effort, since not all collaborative processes are created equal. Hence, the researcher also followed the same procedure during the evaluation. We argue that the data collected from this role convey variability in the way collaboration occurred and in what works or not, depending on the team-members, which demonstrates the ability of the measures used to have some granularity in the evaluation of how the collaborative process took place.

Participants started with a demographic questionnaire. In the next stage, they completed the maintenance tasks while observed by a researcher who assisted them if necessary and registered any relevant event. Immediately after completing the tasks using the conditions, participants answered a post-study questionnaire regarding the collaborative process, as well as their preferences towards the tool used. Then, a small interview was conducted to understand participants' opinion regarding their collaboration with each condition. The data collection was conducted under the guidelines of the Declaration of Helsinki. Also, all measures were followed to ensure a COVID-19 safe environment during each session of the user study.

5.4.6 Participants

We recruited 26 participants (9 female - 34.7%), whose ages ranged from 20 to 63 years old (M = 33.1, SD = 11.7). Participants had various professions, e.g., Master and PhD students, Researchers and Faculty members from different fields, as well as Software Engineers, Front-End Developers and an Assembly Line Operator. With respect to individual and team profile, 14 participants had prior experience with AR and 24 with collaborative tools. With the exception of 1 team, all collaborators had knowledge of each other prior to the study.

5.5 Results and Discussion

This section presents and discusses the main results obtained from the analysis of the data collected through CAPTURE.

5.5.1 Overall total time and task time

As for the total duration, sessions lasted 32 minutes on average (SD = 3.10) using condition C1 and 28 minutes on average (SD = 3.03) using condition C2 (Figure 56) including the aforementioned steps. Regarding task duration, it lasted 16 minutes on average (SD = 2.68) using condition C1 and 12 minutes on average (SD = 2.66) using condition C2. Therefore, participants were quicker on average to perform the tasks when using condition C2, despite having higher data variability when compared to condition C1.

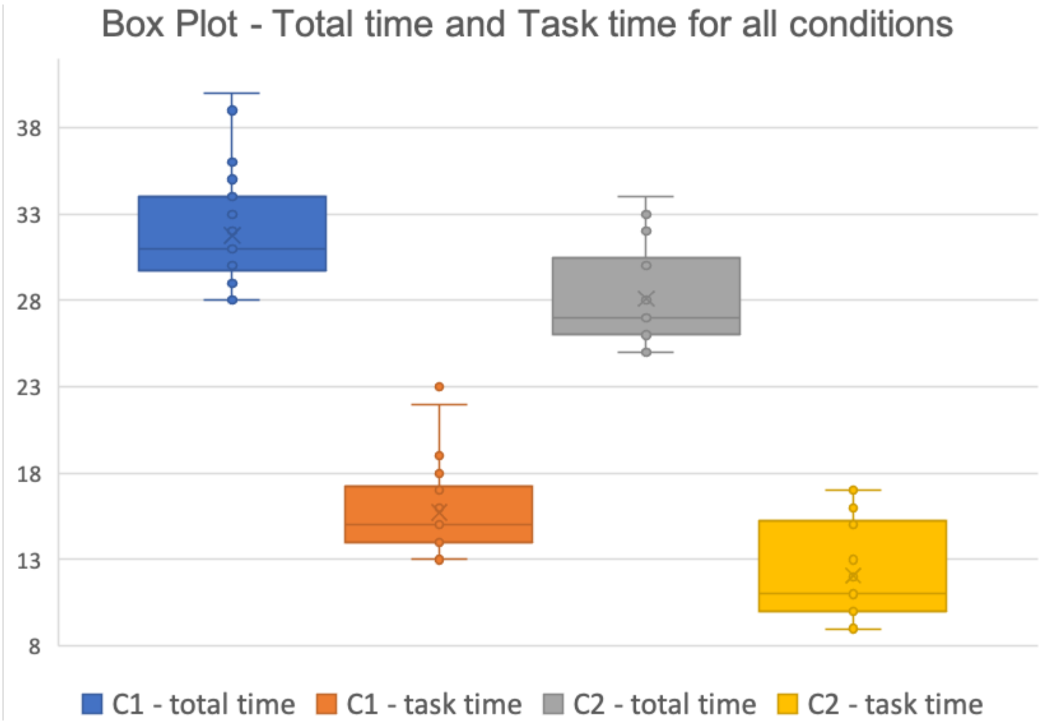

Figure 56 – Total time and task time with the two conditions (in minutes). C1: video chat tool; C2: AR-based annotation tool.

5.5.2 Overview of the collaborative process

Regarding condition C1, participants rated the collaborative process (Likert-type scale: 1- Low; 7- High) as following (Figure 57- top): express ideas (median= 4.5), attentional allocation (median= 4), information understanding (median= 5), mental effort (median= 5), enjoyment (median= 4), communication (median= 5), spatial presence (median= 5.5). As for condition C2, participants rated the collaborative process as following (Figure 57 - bottom): express ideas (median= 6), attentional allocation (median= 7), information understanding (median= 7), mental effort (median= 2), enjoyment (median= 6), communication (median= 6), spatial presence (median= 5).

Hence, it is possible to understand that for the majority of aspects of collaboration, i.e., easy to share ideas properly, level of attention allocation, level of information understanding, level of enjoyment and easy to communicate, condition C2 was rated higher by the participants. Regarding the level of mental effort, participants rated higher condition C1, possibly due to the diminished level of attentional allocation this condition had, which lead to some communication arguing in order to understand where to perform some activities. Therefore, these results suggest that the AR-based annotation tool was better in such aspects of collaboration when compared to the video alternative.

In contrast, for condition C1 the level of spatial presence was higher. This might be associated to the fact that this condition supported live video sharing between team-members, which may have an impact on participants feeling of togetherness with their collaborative counterparts, since it was possible to see the remote expert at all times during the task duration. On the other side, condition C2 provided stabilized AR-based annotations on top of captures/images of the task context. This condition did not allow see the remote expert during the task procedures, which may have affected participants reaction towards the level of spatial presence, although not with any major difference. In this context, a smaller data variability can also be observed for easy to share ideas properly, level of information understanding, level of mental effort, easy to communicate and level of spatial presence, when analysing the box plots of condition C1 and C2, as illustrated by Figure 57.

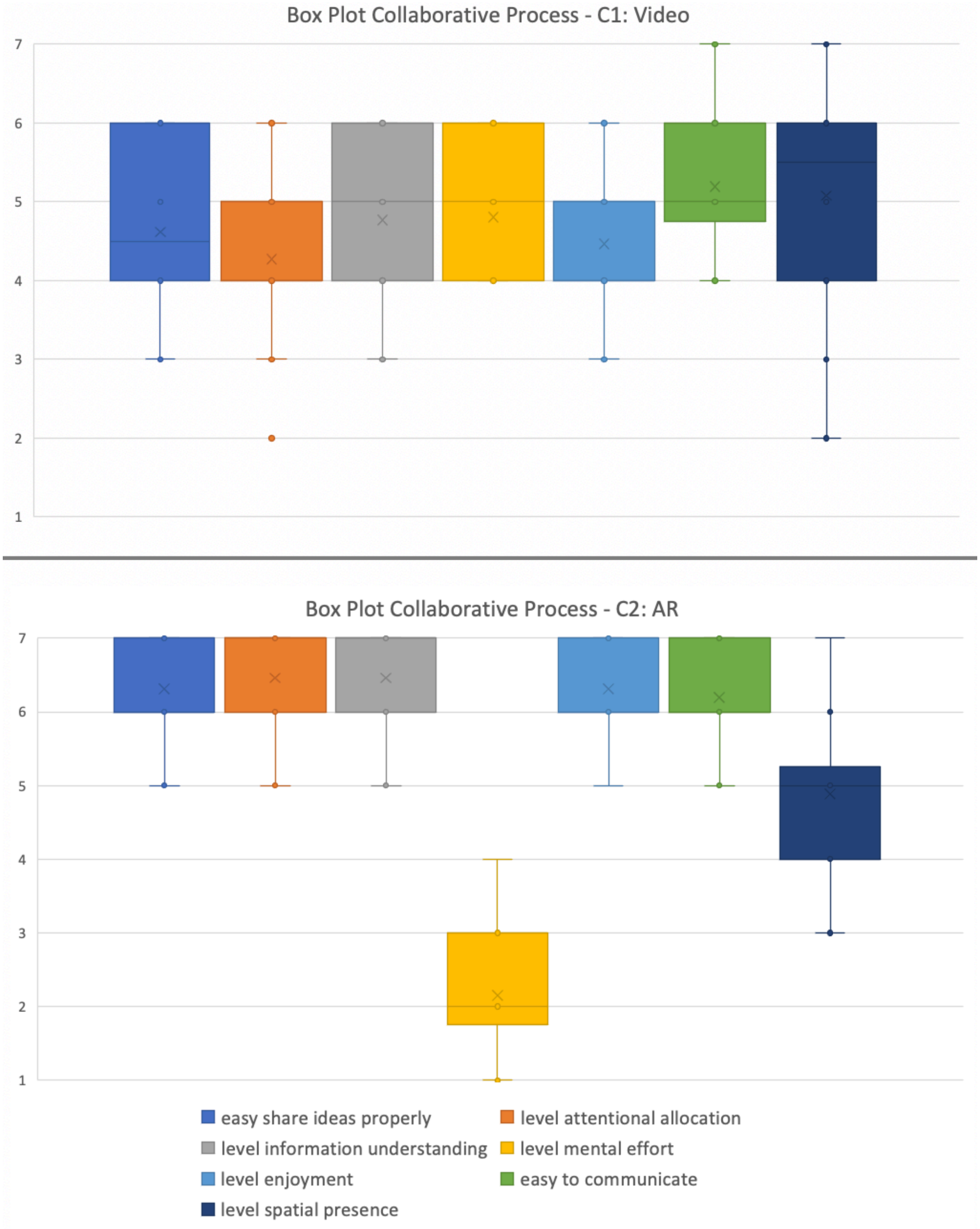

Figure 57 – Overview of the collaborative process outcomes for all teams during a scenario of remote maintenance, including all the selected measures collected: easy to share ideas properly, as well as communicate, level of attentional allocation, information understanding, mental effort, enjoyment, spatial presence. Top - C1: video chat tool; Bottom - C2: AR-based annotation tool. Data displayed using a Likert-type scale: 1- Low; 7- High.

Through the visualization dashboard of the CAPTURE toolkit, it is possible to analyze the collaborative process at the end of an evaluation session for a specific team (Figure 58), or set of different teams (Figure 59).

To elaborate, it is possible to analyze the aspects of collaboration obtained from the use of different tools for the elements of the same team, as explored in this study, which is illustrated in Figure 58 through a random selection. Naturally, following the results presented above, when using condition C2, the team had a better collaborative performance when compared to the results of condition C1. Nevertheless, by analysing the elements of each team individually, such type of visualization allows to identify aspects of collaboration that could be useful to improve over time, or that may be relevant to update in the collaborative tool being used. For example, when using condition C2, the on-site participant rated the level of spatial presence lower. This fact may suggest that to improve the feeling of togetherness, the AR-based annotation tool might benefit from including video sharing in its features.

Furthermore, it is possible to compare the collaborative process of different teams, e.g., teams that knew and don’t knew each other prior to the study. In this vein, Figure 59 illustrates that elements from the team that knew each other rated higher the dimensions: express ideas, information understanding, spatial presence, communication, and enjoyment (Figure 59 - bottom). Contrarily, the level of mental effort was lower, suggesting the collaborative effort was smoother and easier to conduct for these individuals, when compared to the values reported by a team that didn’t know each other previously (Figure 59 - top).

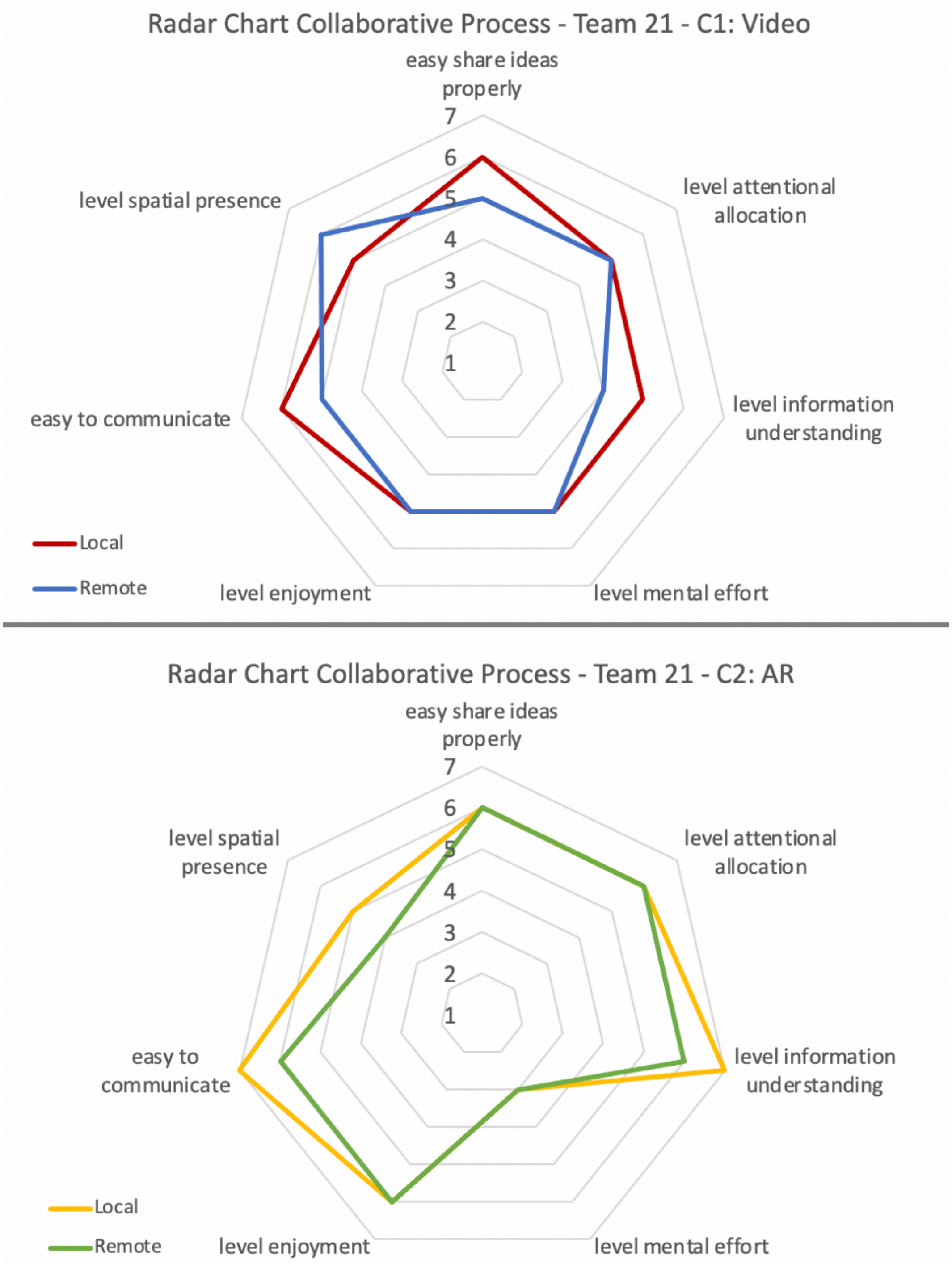

Figure 58 – Collaborative process for the same team during remote maintenance using the two tools: Top - C1: video chat tool; Bottom - C2: AR-based annotation tool. Data displayed using a Likert-type scale: 1- Low; 7- High.

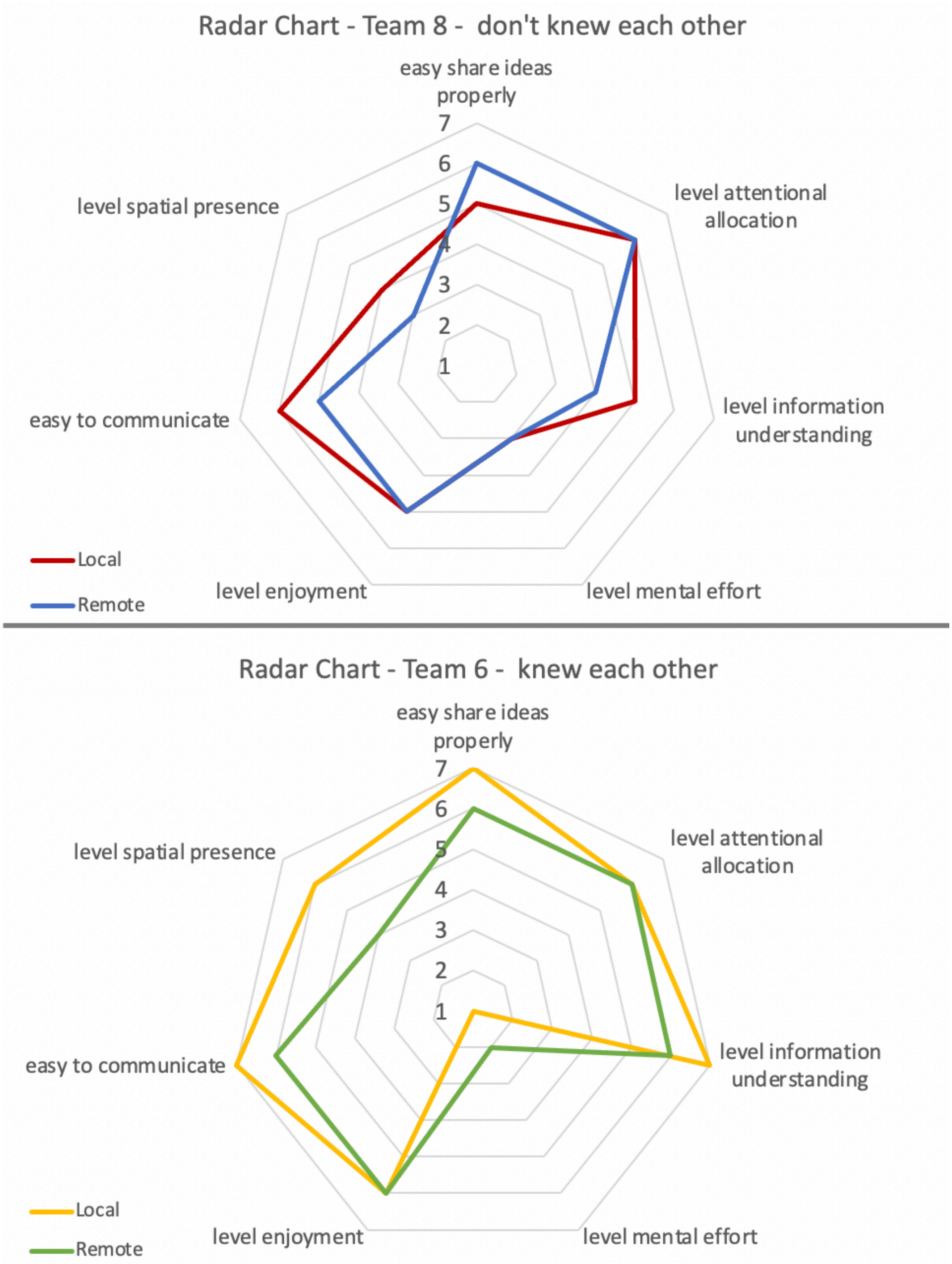

Figure 59 – Collaborative process of two different teams using the AR tool: Top - don't knew each other; Bottom - knew each other prior to the study. Data displayed using a Likert-type scale: 1- Low; 7- High.

5.5.3 Participants preferences and opinion

With respect to participants experience with the tools, 44 reaction cards were selected to characterize condition C1, including 5 neutral, 9 negative and 30 with positive meaning. Likewise, 46 were selected to characterize condition C2, including 3 neutral, 1 negative and 40 with positive meaning (Figure 60). The following top 10 reaction cards represent participants most selected expressions to characterize each condition (Figure 61): C1 - accessible, collaborative, helpful, flexible, simplistic, familiar, usable, unrefined, expected and time-consuming; C2 - helpful, empowering, collaborative, appealing, easy-to-use, engaging, flexible, novel, innovative and advanced.

Figure 60 – Participants total reaction cards regarding the collaborative tools. C1: video chat tool; C2: AR-based annotation tool}. A larger font size means that the word was selected by more participants (higher frequency). Red - negative meaning; gray - neutral meaning; green - positive meaning.

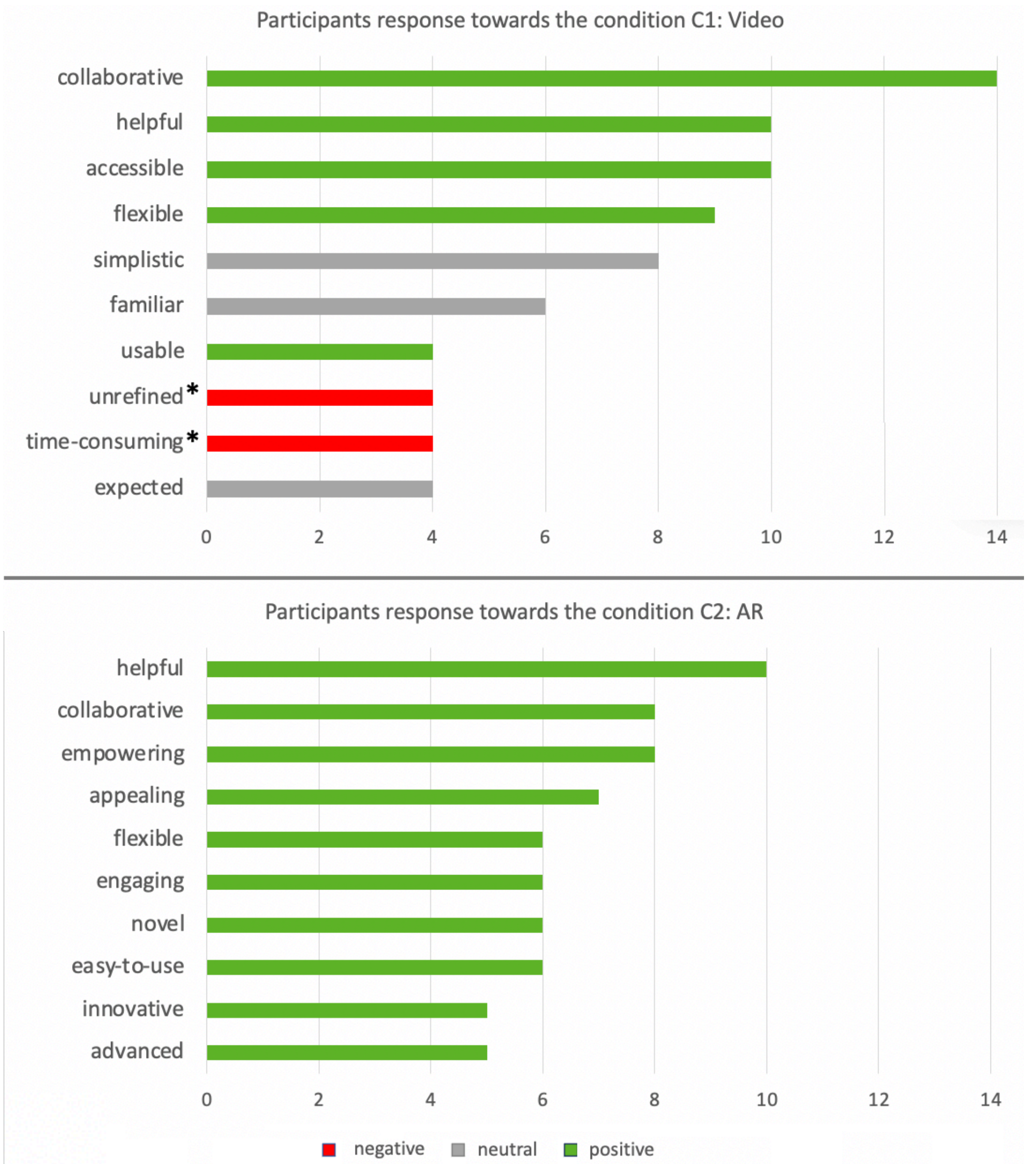

Figure 61 – Participants top 10 reaction cards towards the collaborative tools. C1: video chat tool; C2: AR-based annotation tool. Red (*) - negative meaning; gray - neutral meaning; green - positive meaning.

However, when analysing participants emotional state, collected before and after the tasks, a clearer perspective is attained. To elaborate, regarding condition C1, participants emotional state before the study varied among joy (11 out of 26), surprise (3 out of 26), excitement (8 out of 26) and contempt (4 out of 26) (Figure 62 - top). Then, after the study, it varied among joy (7 out of 26), surprise (1 out of 26), excitement (1 out of 26) and contempt (17 out of 26) (Figure 62- top). As for condition C2, participants emotional state before the study varied among joy (12 out of 26), surprise (3 out of 26), excitement (7 out of 26) and contempt (4 out of 26) (Figure 62 - bottom). Then, after the study, it varied among joy (6 out of 26), surprise (4 out of 26) and excitement (6 out of 26) (Figure 62 - bottom).

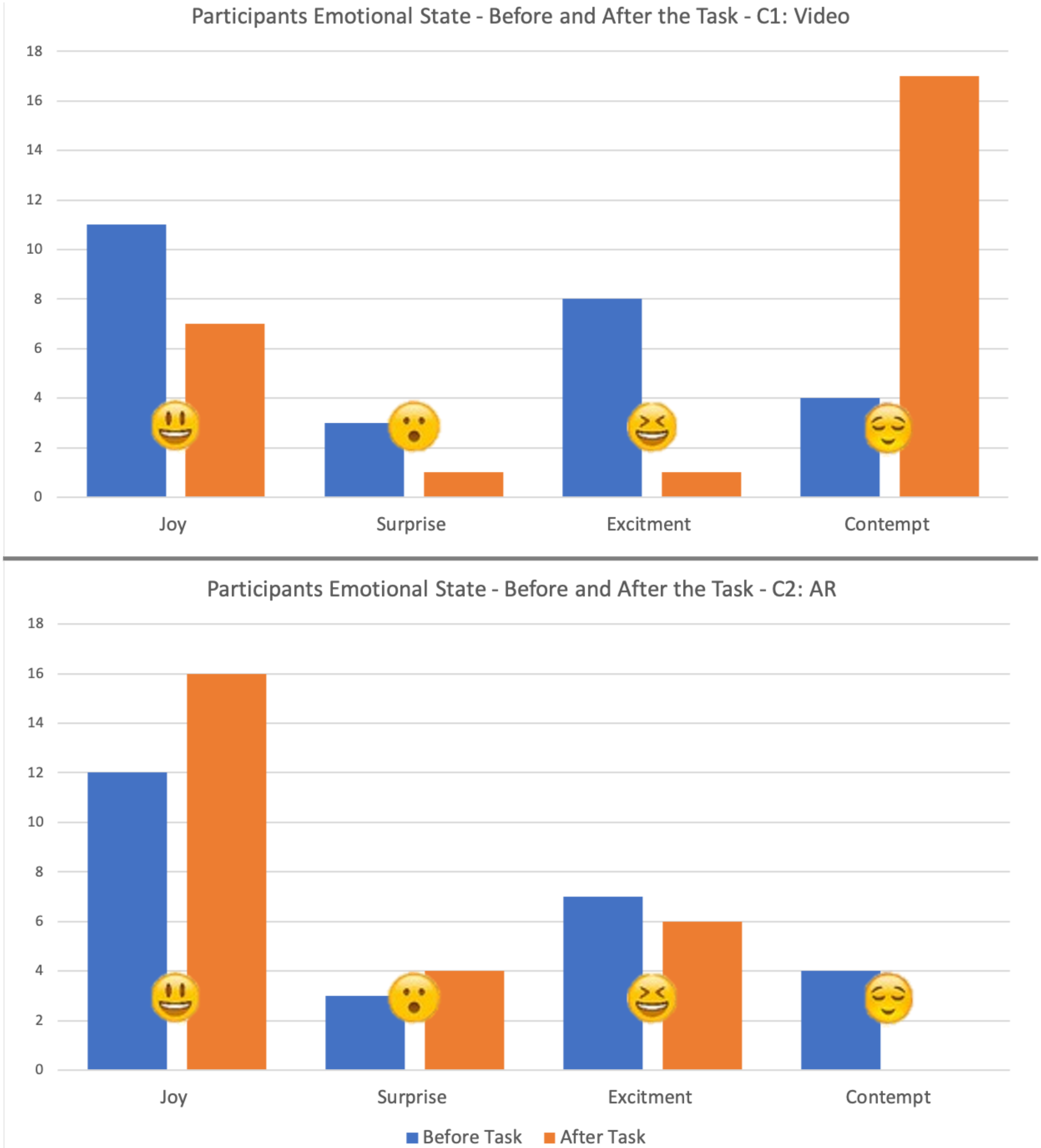

Figure 62 – Participants emotional state before (top) and after (bottom) the tasks for each condition. C1: video chat tool; C2: AR-based annotation tool.

Hence, it is possible to verify that for condition C1, there was a decrease in the number of participants feeling joy, surprise and excitement at the end of the study, which lead to a significant rise associated to the emotional state of contempt. Contrarily, regarding condition C2, there were no occurrences of contempt, while joy and surprise had higher number of participants expressing those feelings. As for excitement, although the number of participants that reported such feeling is lower, it is very close to the values reported at the beginning of the study. As such, condition C2 presents higher values for emotions correlated with positive connotation, e.g., joy, surprise and excitement when compared to condition C1, which only presents a higher value for contempt (neutral connotation). In addition, Figure 63 presents participants satisfaction regarding the tools used through a box plot representation, which illustrates clearly that condition C2 was preferred when compared to condition C1, following the participants' emotional state indications.

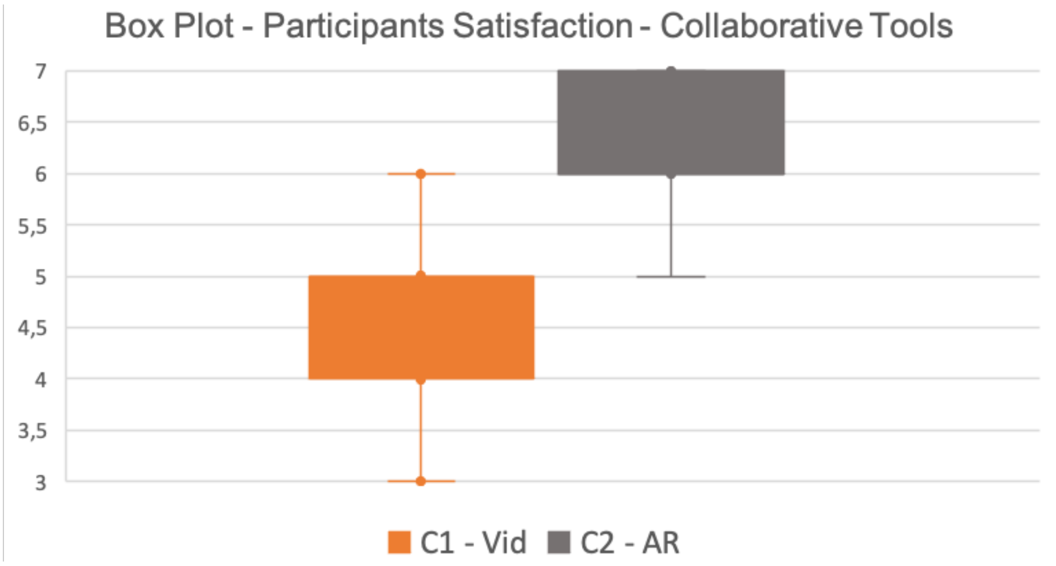

Figure 63 – Participants satisfaction towards the tools. C1: video chat tool; C2: AR-based annotation tool. Data displayed using a Likert-type scale: 1- Low; 7- High.

5.5.4 Final remarks

Motivated by existing limitations of current frameworks, which are not well suited to characterize how collaboration occurs in remote scenarios supported by AR, we presented a conceptual framework to support researchers in designing better evaluations based on retrieving contextualized data for more comprehensive analysis of the collaborative process. While typical evaluation methods disregard high amounts of available information, the proposed approach allows researchers to conduct a more user centric evaluation, where contextual aspects must be part of future evaluations, getting data from typical predefined methods, as well as real time data from the collaborative process.

To summarize the added value of our proposal, and how it compares to existing approaches, the conceptual framework instantiated through the CAPTURE toolkit allows to retrieve additional amounts of contextual data, and selected aspects of collaboration according to the evaluation scope (typically not reported in the literature, but which are very informative/valuable to understand the focus of the work). Hence, elicit more comprehensive analysis using the visualization dashboard to better define how research should progress and how the tools can evolve. In this vein, it is also possible to explore the collaborative process of a particular team, as well as compare different teams or distinct collaborative tools. Given the toolkit flexible architecture, it may be integrated into existing tools via visual editor by non-experts in programming, which expands its adoption by the community.

Another aspect that must be emphasized, is the capacity to adapt to the available data collection instruments. Although self-reporting was used to gather the emotional response, CAPTURE can adapt to support the inclusion of external sensors (e.g., biomedical devices), if necessary for different scenarios and evaluation goals. With all things considered, it is possible to better understand the phenomenon, i.e., recognize when selected aspects of collaboration affect the work effort. By having these insights, it is possible to more easily identify key issues that need to be tackled to ensure a proper shared understanding is attained by distributed team-members in future sessions of remote collaboration. By doing so, the research community can evolve from simple evaluations on how technology works, to more complex evaluations aimed to capture a better perspective on the different factors of collaboration supported by AR, which may lead to a more effective collaborative process over time. Hence, we have shown that a better characterization of the collaborative process can be successfully used to provide an additional perspective on the nuances of remote collaboration mediated by AR, which without contextual data would not be possible.

Altogether, due to the flexibility and range of the proposed conceptual model, the instrumentation through the CAPTURE toolkit establishes itself as a general-purpose evaluation approach, providing data that otherwise would be difficult to obtain and analyze. While we must be prudent with generalizing our findings, we expect our insights to be valuable for future reproduction in other domains beside maintenance context.

In addition, the continuous observation of contextual data in other tools and with other users may allow, in the future, the research community to move into a phase of producing guidelines for remote scenarios mediated by AR, which are supported by experimental data and can guide the initial development of novel collaborative solutions.

Later, we intend to pursue the creation of guidelines to facilitate researchers' actions and elicit more complete evaluations in such scenarios through a set of recommendations. Furthermore, we also intend to conduct field studies in industry context to test and demonstrate how our proposal may lead to a richer and comprehensive characterization of the collaborative process, when compared to how it is being currently reported. By conducting evaluation through a contextualized approach, we argue project managers and technicians in industrial scenarios 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. For example, by doing so, they can obtain insights of using this type of evaluation on the shop floor and asses the use of AR-based solutions for collaborative purposes, while making informed decisions on the available resources, which may in turn have an impact on productivity, time and costs. The outcomes can also lead to improvements in individual behaviour and team collaboration in specific situations and tasks over time.

6 Conclusions and Future Work

Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning.

Winston Churchill

Understand the added value of AR to the work effort in remote scenarios is paramount to improve the characterization of the collaborative process and the efficiency of distributed team-members.

In this vein, perform evaluations in such scenarios presents an essential, but difficult endeavor, due to the multitude of aspects that may influence the way collaboration occurs, particularly the need to consider each team-member involved in the work effort, the nature of the task being considered, the characteristics of the tool providing a common ground, as well as the contextual data from the surrounding environments during this activity.

Considering the current state of the art, this thesis focused on four main contributions:

- Critical analysis to identify the main limitations and situate the maturity of the field;
- Definition of a common ground for systematization of perspectives based on the proposal of dimensions of collaboration, conceptual models and a human-centered taxonomy;
- Understand the needs of target-users through a HCD methodology, leading to the development of an AR-based tool using annotations to connect distributed individuals;
- Creation of a conceptual framework to allow the research community to obtain an additional perspective into the collaborative process.

This work was accomplished using the Engineering Design Methodology to define, develop, evaluate and improve all the contributions obtained, while also profiting from a research project with partners from the industry sector, which gave us the opportunity to have some insights on the real needs and tasks of target-users through the involvement of domain experts.

This section summarizes and discusses the main contributions and limitations of this thesis, and outlines directions for future research.

6.1 Main Results and Discussion

At the beginning of this thesis, we pointed to the necessity of creating an alternative approach for evaluating team-members during scenarios of remote collaboration mediated by AR, considering the innate dynamicity of such activities and including context as an important data source for more complete characterizations of the collaborative process.

Considering the results obtained, some important contributions were made in this direction, namely to the state of the art through a **systematic study which resulted in a critical analysis that identified the main limitations and opportunities of the field**, **while situating its maturity** between the Replication and Empiricism phases of the BRETAM model. Also, the proposal of a roadmap of important research actions that need to be addressed to help the field advance into new phases.

Another important contribution has been performed by addressing the lack of frameworks to characterize how Collaborative AR occurs. The **proposal of a conceptual model, and human-centered taxonomy** helps to establish a common ground for systematization of perspectives and discussion for the field. With these, we expect to influence and improve how research in this field is reported by providing a structured list of the defining characteristics. This led to the **creation of an interactive visualization tool** for dimensional categorization, which allows to identify existing patterns in the form of gaps, trends and opportunities. Besides, we also proposed an ontology to better understand the relations between the proposed dimensions and the common elements of the evaluation process. This is an important step into the direction of being able to more thoroughly classify and discuss the contributions of AR to the collaborative work effort.

A relevant achievement was also obtained from **applying a human-centered design methodology to understand the needs of target-users in real scenarios**, while most works found in literature focus on creating AR-based solutions for rather simple toy problems, which require minimal collaboration and use of AR. The process of identifying requirements with such experienced individuals confirms the need for traceability, offering useful qualitative feedback on how to support remote collaboration using AR. Moreover, by merging these outcomes with literature methods in an AR-based prototype exploring annotations and performing its evaluation through a user study, we were able to show the prototype capacity to merge the shared perceived realities of different professionals and enriching each individual experience.

To facilitate the creation/application of evaluation tests, an additional contribution was **an evaluation framework for contextualized data gathering and analysis**, including conceptual models and an ontology, allowing to support researchers in designing and conducting distributed evaluations in a more structured manner. Hence, considering data monitoring regarding the level of collaboration, behaviour and performance of each intervening party, individually and as a team, as well as consider contextual data. While typical evaluation methods disregard high amounts of available information, the proposed approach allows to conduct more user-centric evaluations, by getting data from typical predefined methods, as well as real-time data collection during the collaborative process, which results in a richer set of data, for more comprehensive analysis.

To instantiate this vision, the **CAPTURE toolkit was developed** to provide an additional perspective based on the evaluation scope, as well as help make considerations and draw conclusions about the use of AR-based solutions for remote scenarios. We introduced extensible specifications for several elements like selected dimensions of collaboration and pre-defined measurements elements, to allow researchers define specific evaluation situations and obtain “in situ” data about them. By having a grasp on several aspects (e.g., ease to communicate and to share ideas, level of attention allocation, information understanding, etc.), which are usually ignored, it is possible to analyze the collaborative process of a particular team or compare between a set of teams, as well as different collaborative tools. The toolkit can be integrated into a collaborative tool via visual editor. All data collected during collaboration is stored in a central server, which can be analyzed during post-task analysis through a visualization dashboard. To prove the feasibility and overall value of our proposal, we created a proof-of-concept scenario.

We **illustrate the use of the toolkit through a case study on real-life remote maintenance tasks, comparing two distinct methods instrumented with the proposed toolkit**: sharing of video and AR-based annotations. Results show that remote team-members using AR-based annotations performed the tasks faster and collaborated in an easier way, when compared to the traditional video alternative. The use of AR was also preferred by the majority of team-members, reporting higher levels of joy, surprise and excitement. Based on their opinions, AR appears as a valuable option for analysis and discussion, since it supports more natural and intuitive interaction, leading to new insights, which contributes to increase empathy, interest and collaboration. Likewise, it is also possible to identify when selected aspects of collaboration that may affect the work effort. For example, some team-members rated lower the level of spatial awareness. By having these insights, it is possible to revisit the AR-based tool to tackle these limitations for future sessions of remote collaboration. Thus, our expectations for the use of the toolkit have been reached: conduct comparative analysis of distributed teams may benefit researchers in better understanding the collaborative phenomenon supported by AR, designing novel methods and improve the efficiency of the collaborative effort. Altogether, due to the flexibility and range of the CAPTURE toolkit, the solution establishes itself as a general-purpose evaluation approach, providing data that otherwise would be difficult to obtain and analyze. While we must be prudent with generalizing our findings, we expect our insights to be valuable for future reproduction in other domains besides the maintenance context.

With all things considered, we have shown that a better characterization of the collaborative process can be successfully used to provide an additional perspective on the nuances of remote collaboration mediated by AR. Hence, we hope that our insights can help the research community to design and conduct better evaluations, and in turn lead to the design of better and more efficient collaborative AR solutions, improving remote collaboration which is the ultimate goal.

6.2 Future work

In the scope of the issues addressed in this thesis, many and diverse opportunities arise for future research. Here are the ones that appear more promising.

A natural step is the **creation of guidelines** to facilitate researchers' actions and elicit more complete characterizations and evaluations in remote scenarios, which may result in improving the quality of novel AR-based solutions.

Equally important, is to **integrate individuals from multidisciplinary areas** into the refinement of the measurements used by the CAPTURE toolkit (e.g., Telerehabilitation, Bioengineering, Affective Computing, Psychology, etc.), which could help to better understand how remote collaborators behave through AR using a holistic methodology, i.e., by focusing on additional dimensions of collaboration besides the ones considered.

Equally relevant, is to **offer ready to use evaluation services and dynamic visualizations**, which may be configured according to the evaluation goals. It is paramount to **continue to develop the visualization dashboard of the CAPTURE toolkit** to support novel requirements and data gathering approaches. For example, provide features to facilitate the process of measuring the common ground, i.e., paying attention to how team-members communicate: counting the number of times specific words are used (deictic vs non-deictic), counting the number of clarifying questions asked, the amount of silence between interactions, among others. In this vein, improve the dashboard to present audio/video synchronized with other user-related events captured during the evaluation process.

Another essential point is to **perform longer, more complex field studies with domain experts** to test the *CAPTURE* toolkit in real settings. This is an opportunity to continue to profit from a rich R&D context, while also contributing to the lack of field evaluations. It would also be relevant to perform statistical analysis on the obtained results, aiming to connect the dimensions used to the collaborative process rating, thus allowing to better **understand the role of the human-in-the-loop**.

The **development of authoring tools** must also be tackled to diminish existing technical issues that limit the adoption of AR in remote scenarios. Although this work presented a possible solution to handle some existing constraints through the creation of step-by-steps instructions, it is important to improve acceptance, in particular, include target-users in the creation and evaluation processes. Also, explore the **combination of existing AR visualization methods to establish a common ground**, e.g., including 3D reconstructions as input for a shared model approach, while integrating annotations into the process, thus creating tools that combine the advantages of 2D and 3D approaches, which may be essential to handle more complex remote scenarios.

Also interesting, is to **understand how to deliver contextualized information**, i.e., how information can be shared without cluttering the collaborators field of view and without interfering with the tasks. A problem mentioned in some situations by participants when the annotations appeared in an intrusive way, thus occluding/cluttering important areas of the environment. Thus, **exploring new mechanisms to filter and display AR-based information, as well as interact with such content**, allowing team-members to select what should be presented or discarded at a given moment using different types of interaction, according to the type of device being used.

In addition, **explore novel interaction paradigms, resorting to multimodal approaches, and different interaction devices.** These may allow the use of additional modalities, the combination of multiple devices or the use of different notifications to improve collaboration in environments with specific particularities, e.g., larger dynamical environments like industrial assembly lines.

To finish, also relevant is **collaboration among multiple team-members** to address how ownership of virtual content may impact the collaborative process in such scenarios, which requires a better characterization and evaluation of the collaborative effort moving forward.

References

Akkil, D., & Isokoski, P. (2018). Comparison of gaze and mouse pointers for video-based collaborative physical task. *Interacting with Computers*, *30*(6), 524–542.

Al-Azawi, R., & Shakkah, M. S. (2018). Embedding augmented and virtual reality in educational learning method: Present and future. *International Conference on Information and Communication Systems, ICICS 2018*, 218–222.

Altug, Y., & Mahdy, A. M. (2016). A Perspective on Distributed and Collaborative Augmented Reality. *International Journal of Recent Trends in Human Computer Interaction (IJHCI)*, *7*(2), 23–41.

Alves, J., Marques, B., Carlos, F., Paulo, D., & Sousa, B. S. (2021). Comparing augmented reality visualization methods for assembly procedures. *Journal of Virtual Reality*, *2020*, 1–14.

Alves, J., Marques, B., Cruz, I., Sousa, A., Pinto, G., Guedes, D., Reis, F., Ramos, A., Ribeiro, C., Dias, P., & Santos, B. S. (2018). Remote Interaction in Virtual and Augmented Reality: a Case Study using Multiple Devices. *International Conference on Graphics and Interaction, ICGI 2018*.

Alves, J., Marques, B., Dias, P., & Santos, B. S. (2020). Using Augmented Reality and Step by Step Verification in Industrial Quality Control. *International Conference on Human Systems Engineering and Design: Future Trends and Applications, IHSED*.

Alves, J., Marques, B., Dias, P., & Sousa, B. (2021). Using augmented reality for industrial quality assurance: a shop floor user study. *The International Journal of Advanced Manufacturing Technology*, 1–13.

Alves, J., Marques, B., Neves, M., Maio, R., Santos, A., Justo, I., Rainho, R., Costa, D., Santos, B. S., & Dias, P. (2019). Augmented Reality Platform to foster Collaborative Manipulation of 3D virtual objects: A Prototype. *International Conference on Graphics and Interaction, ICGI 2019*.

Alves, J., Marques, B., Oliveira, M., Araujo, T., Dias, P., & Santos, B. S. (2019). Comparing Spatial and Mobile Augmented Reality for Guiding Assembling Procedures with Task Validation. *IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC 2019*.

Andreas, D., & Billinghurst, M. (2011). Evaluating Augmented Reality Systems. In *Handbook of Augmented Reality* (pp. 289–307). Springer Publishing Company, Incorporated.

Antunes, P., Herskovic, V., Ochoa, S. F., & Pino, J. A. (2014). Reviewing the quality of awareness support in collaborative applications. *Journal of Systems and Software*, *89*(1), 146–169.

Araujo, R. M., Santoro, F. M., & Borges, M. R. S. (2004). A conceptual framework for designing and conducting groupware evaluations. *International Journal of Computer Applications in Technology*, *19*(3–4), 139–150.

Araujo, R. M., Santoro, F. M., & Borges, M. R. S. (2003). The CSCW lab for groupware evaluation. *International Conference on Computer Supported Cooperative Work in Design*, 148–153.

Ardito, C., Desolda, G., Lanzilotti, R., Malizia, A., & Matera, M. (2020). Analysing trade-offs in frameworks for the design of smart environments. *Behaviour & Information Technology*, *39*(1), 47–71.

Arias, E., Eden, H., Fischer, G., Gorman, A., & Scharff, E. (2000). Transcending the Individual Human Mind-Creating Shared Understanding through Collaborative Design. *ACM Transactions on Computer-Human Interaction*, *7*(1), 84–113.

Aromaa, S., Aaltonen, I., Kaasinen, E., Elo, J., & Parkkinen, I. (2016). Use of Wearable and Augmented Reality Technologies in Industrial Maintenance Work. *Proceedings of 20th International Academic Mindtrek Conference*, To be published 2016.

Aschauer, A., Reisner-Kollmann, I., & Wolfartsberger, J. (2021). Creating an Open-Source Augmented Reality Remote Support Tool for Industry: Challenges and Learnings. *Procedia Computer Science*, *180*, 269–279.

Aschenbrenner, D., Rojkov, M., Leutert, F., Verlinden, J., Lukosch, S., Latoschik, M. E., & Schilling, K. (2018). Comparing Different Augmented Reality Support Applications for Cooperative Repair of an Industrial Robot. *Adjunct Proceedings - 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2018*, 69–74.

Ashtari, N., Bunt, A., McGrenere, J., Nebeling, M., & Chilana, P. K. (2020). Creating Augmented and Virtual Reality Applications: Current Practices, Challenges, and Opportunities. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13.

Augstein, M., & Neumayr, T. (2019). A Human-Centered Taxonomy of Interaction Modalities and Devices. *Interacting with Computers*, *31*(1), 27–58.

Azuma, R. T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385.

Bacca, J., Baldiris, S., Fabregat, R., & Graf, S. (2014). Augmented Reality Trends in Education: A Systematic Review of Research and Applications. *Educational Technology & Society*, *17*(4), 133–149.

Bai, H., Sasikumar, P., Yang, J., & Billinghurst, M. (2020). A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13.

Bai, Z., & Blackwell, A. F. (2012). Analytic review of usability evaluation in ISMAR. *Interacting with Computers*, *24*(6), 450–460.

Bannai, Y., Tamaki, H., Suzuki, Y., Shigeno, H., & Okada, K. (2006). A tangible user interface for remote collaboration system using mixed reality. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4282 LNCS*, 143–154.

Barakonyi, I., Prendinger, H., Schmalstieg, D., & Ishizuka, M. (2007). Cascading hand and eye movement for augmented reality videoconferencing. *IEEE Symposium on 3D User Interfaces 2007 - Proceedings, 3DUI 2007*, 71–78.

Barnum, C. M. (2010). *Usability Testing Essentials: Ready, Set...Test!* (1st ed.). Morgan Kaufmann Publishers Inc.

Barroso, J., Fonseca, L., Marques, B., Dias, P., & Sousa Santos, B. (2020). Remote Collaboration using Mixed Reality: Exploring a shared model approach through different interaction methods. *European Conference on Computer-Supported Cooperative Work, ECSCW*, 1–6.

Barsom, E., Graafland, M., & Schijven, M. (2016). Systematic review on the effectiveness of augmented reality applications in medical training. *Surgical Endoscopy and Other Interventional Techniques*, *30*(10), 4174–4183.

Belen, R., Nguyen, H., Filonik, D., Del Favero, D., & Bednarz, T. (2019). A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering*, *3*(2), 181–223.

Bellini, H., Chen, W., Sugiyama, M., Shin, M., Alam, S., & Takayama, D. (2016). Virtual & Augmented Reality: Understanding the race for the next computing platform. In *Goldman Sachs*.

Benford, S., Greenhalgh, C., & Reynard, G. (1998). Understanding and Constructing Shared Spaces with Mixed-Reality Boundaries. *ACM Transactions on Computer-Human Interaction*, *5*(3), 185–223.

Bhattacharya, B., & Winer, E. H. (2019). Augmented reality via expert demonstration authoring (AREDA). *Computers in Industry*, *105*, 61–79.

Bhorkar, G. (2017). A Survey of Augmented Reality Navigation. *Seminar on Software Systems, Technologies and Security*. http://arxiv.org/abs/1708.05006

Billinghurst, M. (2021). Grand Challenges for Augmented Reality. *Frontiers in Virtual Reality*, *2*(March), 1–4.

Billinghurst, M., Bee, S., Bowskill, J., & Kato, H. (2003). Asymmetries in collaborative wearable interfaces. *Proceedings of International Symposium on Wearable Computers (ISWC)*, 133–140.

Billinghurst, M., Clark, A., & Lee, G. (2015). A Survey of Augmented Reality. *Foundations and Trends in Human–Computer Interaction*, *8*(2–3), 73–272.

Billinghurst, M., Cordeil, M., Bezerianos, A., & Margolis, T. (2018). Collaborative Immersive Analytics. In K. Marriott, F. Schreiber, T. Dwyer, K. Klein, N. H. Riche, T. Itoh, W. Stuerzlinger, & B. H. Thomas (Eds.), *Immersive Analytics* (pp. 221–257). Springer International Publishing.

Billinghurst, M., Gupta, K., Katsutoshi, M., Lee, Y., Lee, G., Kunze, K., & Sugimoto, M. (2017). Is It in Your Eyes? Explorations in Using Gaze Cues for Remote Collaboration. In *Collaboration Meets Interactive Spaces* (pp. 1–483).

Billinghurst, M., & Kato, H. (1999). Collaborative Mixed Reality. *International Symposium on Mixed Reality, ISMR 99*, 261–284. http://www.springerlink.com/index/10.1007/978-3-642-87512-0_15

Billinghurst, M., & Kato, H. (2002). Collaborative Augmented reality. *Communications of the ACM*, *45*(7), 64–70.

Bolstad, C. A., & Endsley, M. R. (2005). Choosing Team Collaboration Tools: Lessons from Disaster Recovery Efforts. *Ergonomics in Design*, *13*(4), 7–14.

Bottani, E., & Vignali, G. (2019). Augmented reality technology in the manufacturing industry: A review of the last decade. *IISE Transactions*, *51*(3), 284–310.

Bottecchia, S., Cieutat, J. M., Merlo, C., & Jessel, J. P. (2009). A new AR interaction paradigm for collaborative teleassistance system: The POA. *International Journal on Interactive Design and Manufacturing*, *3*(1), 35–40.

Brás, S., Soares, S. C., Cruz, T., Magalhães, T., Marques, B., Dantas, C., Fernandes, N., & Fernandes, J. M. (2018). The feasibility of an augmented reality system to study the psychophysiological correlates of fear-related responses. *Brain and Behavior*, *June*, 1–10.

Braz, J. M., & Pereira, J. M.-. (2008). TARCAST: Taxonomy for Augmented Reality CASTing with Web Support. *Virtual Reality*, *7*(4), 47–56. http://www.inesc-id.pt/ficheiros/publicacoes/4625.pdf

Brockmann, T., Krueger, N., Stieglitz, S., & Bohlsen, I. (2013). A Framework for Collaborative Augmented Reality Applications. *Proceedings of the Nineteenth Americas Conference on Information Systems*, 1–10.

Bruno, F., Barbieri, L., Marino, E., Muzzupappa, M., D'Oriano, L., & Colacino, B. (2019). An augmented reality tool to detect and annotate design variations in an Industry 4.0 approach. *International Journal of Advanced Manufacturing Technology*, *105*(1–4), 875–887.

Campbell, M., Kelly, S., Jung, R., & Lang, J. (2017). *The State of Industrial Augmented Reality 2017*.

Campbell, M., Kelly, S., Jung, R., & Lang, J. (2018). *The State of Industrial Augmented Reality: A Spotlight on Industrial Innovation*.

Carmigniani, J., & Furht, B. (2011). *Handbook of Augmented Reality*. http://link.springer.com/10.1007/978-1-4614-0064-6

Carvalho, R. M., de Castro Andrade, R. M., de Oliveira, K. M., de Sousa Santos, I., & Bezerra, C. I. M. (2017). Quality characteristics and measures for human–computer interaction evaluation in ubiquitous systems. *Software Quality Journal*, *25*(3), 743–795.

Ceyssens, J., Coppers, S., & Vanacken, D. (2020). Towards Traceable Design Rationale in Augmented Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–8). Association for Computing Machinery.

Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (1999). What are ontologies, and why do we need them? *IEEE Intelligent Systems and Their Applications*, *14*(1), 20–26.

Chenechal, M. Le, Duval, T., Gouranton, V., Royan, J., & Arnaldi, B. (2016). Vishnu: Virtual immersive support for HelpiNg users an interaction paradigm for collaborative remote guiding in mixed reality. *2016 IEEE 3rd VR International Workshop on Collaborative Virtual Environments, 3DCVE 2016*, 9–12.

Chittaro, L., & Buttussi, F. (2015). Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety. *IEEE Transactions on Visualization and Computer Graphics*, *21*(4), 529–538.

Choi, S. H., Kim, M., & Lee, J. Y. (2018). Situation-dependent remote AR collaborations: Image-based collaboration using a 3D perspective map and live video-based collaboration with a synchronized VR mode. *Computers in Industry*, *101*(May), 51–66.

Cidota, M., Lukosch, S., Datcu, D., & Lukosch, H. (2016). Comparing the Effect of Audio and Visual Notifications on Workspace Awareness Using Head-Mounted Displays for Remote Collaboration in Augmented Reality. *Augmented Human Research*, *1*(1), 1–15.

Clark, H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, 127–149.

Collazos, C. A., Gutiérrez, F. L., Gallardo, J., Ortega, M., Fardoun, H. M., & Molina, A. I. (2019). Descriptive theory of awareness for groupware development. *Journal of Ambient Intelligence and Humanized Computing*, *10*(12), 4789–4818.

Congdon, B. J., Wang, T., & Steed, A. (2018). Merging environments for shared spaces in mixed reality. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*.

Datcu, D., Lukosch, S. G., & Lukosch, H. K. (2016). Handheld Augmented Reality for Distributed Collaborative Crime Scene Investigation. *International Conference on Supporting Group Work*, 267–276.

de Belen, R. A. J., Nguyen, H., Filonik, D., Favero, D. Del, & Bednarz, T. (2019). A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering*, *3*(2), 181–223.

de Souza Cardoso, L. F., Mariano, F. C. M. Q., & Zorzal, E. R. (2020). A survey of industrial augmented reality. *Computers and Industrial Engineering*, *139*, 1–12.

Dey, A., Billinghurst, M., Lindeman, R. W., & Ii, J. E. S. (2018). A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014. *Frontiers in Robotics and AI*, *5*(April).

Dey, A., Billinghurst, M., Lindeman, R. W., & Swan, J. E. (2018). A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014. *Frontiers in Robotics and AI*, *5*, 5–6.

Dieter, G., & Schmidt, L. (2013). *Engineering Design* (Fifth Edit). McGraw-Hill Education. https://books.google.pt/books/about/Engineering_Design.html?id=_Sg1AAAAQBAJ&redir_esc=y

Dourish, P., & Bellotti, V. (1992). Awareness and Coordination in Shared Workspaces. *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work*, 107–114.

Egger, J., & Masood, T. (2020). Augmented reality in support of intelligent manufacturing – A systematic literature review. *Computers and Industrial Engineering*, *140*, 1–22.

Ellis, C. A., Gibbs, S. J., & Rein, G. (1991). Groupware: Some Issues and Experiences. *Commun. ACM*, *34*(1), 39–58.

Elvezio, C., Sukan, M., Oda, O., Feiner, S., & Tversky, B. (2017). Remote collaboration in AR and VR using virtual replicas. *ACM SIGGRAPH 2017 VR Village (SIGGRAPH '17)*, 1–2.

Ens, B., Lanir, J., Tang, A., Bateman, S., Lee, G., Piumsomboon, T., & Billinghurst, M. (2019). Revisiting Collaboration through Mixed Reality: The Evolution of Groupware. *International Journal of Human-Computer Studies*, *131*, 81–98.

Fakourfar, O., Ta, K., Tang, R., Bateman, S., & Tang, A. (2016). Stabilized Annotations for Mobile Remote Assistance. *Conference on Human Factors in Computing Systems, CHI*, 1548–1560.

Fernandes, A. M., & Albuquerque, P. B. (2012). Tactual perception: a review of experimental variables and procedures. *Cognitive Processing*, *13*(4), 285–301.

Fernández del Amo, I., Erkoyuncu, J. A., Roy, R., Palmarini, R., & Onoufriou, D. (2018). A systematic review of Augmented Reality content-related techniques for knowledge transfer in maintenance applications. *Computers in Industry*, *103*, 47–71.

Fernández Del Amo, I., Erkoyuncu, J. A., Roy, R., & Wilding, S. (2018). Augmented Reality in Maintenance : An information-centred design framework. *Procedia Manufacturing*, *19*(2017), 148–155.

Ferrari, A., Spoletini, P., & Gnesi, S. (2016). Ambiguity and tacit knowledge in requirements elicitation interviews. *Requirements Engineering*, *21*(3), 333–355.

Fiorentino, M., Uva, A. E., Gattullo, M., Debernardis, S., & Monno, G. (2014). Augmented reality on large screen for interactive maintenance instructions. *Computers in Industry*, *65*(2), 270–278.

Fischer, J. E., Reeves, S., Brown, B., & Lucero, A. (2018). Beyond “Same Time, Same Place”: Introduction to the Special Issue on Collocated Interaction. *Human–Computer Interaction*, *33*(5–6), 305–310.

Flotyński, J. (2021). Creating explorable extended reality environments with semantic annotations. *Multimedia Tools and Applications*, *80*(5), 6959–6989.

Gaines, B. R. (1991). Modeling and forecasting the information sciences. *Information Sciences*, *57–58*, 3–22.

García-Pereira., I., Gimeno., J., Morillo., P., & Casanova-Salas., P. (2020). A Taxonomy of Augmented Reality Annotations. *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 1: GRAPP,* 412–419.

Gattullo, M., Scurati, G. W., Fiorentino, M., Uva, A. E., Ferrise, F., & Bordegoni, M. (2019). Towards augmented reality manuals for industry 4.0: A methodology. *Robotics and Computer-Integrated Manufacturing*, *56*(March 2018), 276–286.

Gauglitz, S., Lee, C., Turk, M., & Höllerer, T. (2012). Integrating the physical environment into mobile remote collaboration. *MobileHCI’12 - Proceedings of the 14th International Conference on Human Computer Interaction with Mobile Devices and Services*, 241–250.

Gauglitz, S., Nuernberger, B., Turk, M., & Höllerer, T. (2014a). In touch with the remote world: Remote Collaboration with Augmented Reality Drawings and Virtual Navigation. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, 197–205.

Gauglitz, S., Nuernberger, B., Turk, M., & Höllerer, T. (2014b). World-Stabilized Annotations and Virtual Scene Navigation for Remote Collaboration. *Proceedings of ACM Symposium on User Interface Software and Technology*, 449–459.

Geng, J., Song, X., Pan, Y., Tang, J., Liu, Y., Zhao, D., & Ma, Y. (2020). A systematic design method of adaptive augmented reality work instruction for complex industrial operations. *Computers in Industry*, *119*, 103229.

Gergle, D., Kraut, R. E., & Fussell, S. R. (2013). Using Visual Information for Grounding and Awareness in Collaborative Tasks. *Human–Computer Interaction*, *28*(1), 1–39.

Gervasi, R., Mastrogiacomo, L., & Franceschini, F. (2020). A conceptual framework to evaluate human-robot collaboration. *The International Journal of Advanced Manufacturing Technology*, *108*(3), 841–865.

Grubert, J., & Zollmann, S. (2017). Towards Pervasive Augmented Reality: Context- Awareness in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, *23*(6), 1706–1724.

Grubert, J., Zollmann, S., Langlotz, T., & Regenbrecht, H. (2017). Towards Pervasive Augmented Reality: Context- Awareness in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, *23*(6), 1706–1724.

Grudin, J., & Poltrock, S. (2011). Computer Supported Cooperative Work. In *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Interaction Design Foundation.

Grudin, J., & Poltrock, S. E. (1997). Computer-supported cooperative work and groupware. *Advances in Computers*, *45*, 269–320.

Günther, S., Kratz, S., Avrahami, D., & Mühlhäuser, M. (2018). Exploring Audio , Visual , and Tactile Cues for Synchronous Remote Assistance. *PETRA 2018*, 339–344.

Gupta, K., Lee, G., & Billinghurst, M. (2016). Do You See What I See ? The Effect of Gaze Tracking on Remote Collaboration using a Head Mounted System . *IEEE Transactions on Visualization and Computer Graphics*, *2626*(1), 2413–2422.

Gurevich, P., Lanir, J., & Cohen, B. (2015). Design and Implementation of TeleAdvisor: a Projection-Based Augmented Reality System for Remote Collaboration. *Computer Supported Cooperative Work, CSCW*, *24*(6), 527–562.

Gutwin, C., & Greenberg, S. (1999). The Effects of Workspace Awareness Support on the Usability of Real-Time Distributed Groupware. *ACM Trans. Comput.-Hum. Interact.*, *6*(3), 243–281.

Hadwin, A. F., Gress, C. L. Z., & Page, J. (2006). Toward Standards for Reporting Research: A Review of the Lliterature on Computer-Supported Collaborative Learning. *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, 1007–1011.

Hall, M., McMahon, C. A., Bermell-Garcia, P., Johansson, A., & Ravindranath, R. (2018). Capturing synchronous collaborative design activities: A state-of-the-art technology review. *International Design Conference, DESIGN*, 347–358.

Halskov, K., & Hansen, N. B. (2015). The diversity of participatory design research practice. *Journal of Human Computer Studies*, *74*, 81–92.

Hamadache, K., & Lancieri, L. (2009). Strategies and taxonomy, tailoring your CSCW evaluation. *Lecture Notes in Computer Science*, *5784*, 206–221.

Han, D.-I., Tom Dieck, M. C., & Jung, T. (2017). User experience model for augmented reality applications in urban heritage tourism. *Journal of Heritage Tourism*, *6631*, 1–16.

Henderson, S., Member, S., & Feiner, S. (2011). Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. *IEEE Transactions on Visualization and Computer Graphics, TVCG*, *17*(10), 1355–1368.

Hernandez-de-Menendez, M., Morales-Menendez, R., Escobar, C. A., & McGovern, M. (2020). Competencies for Industry 4.0. *International Journal on Interactive Design and Manufacturing*, *14*(4), 1511–1524.

Herskovic, V., Pino, J. A., Ochoa, S. F., & Antunes, P. (2007). Evaluation methods for groupware systems. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4715 LNCS*, 328–336.

Hofmann, E., & Rüsch, M. (2017). Industry 4.0 and the current status as well as future prospects on logistics. *Computers in Industry*, *89*, 23–34.

Hoppe, A. H., Reeb, R., van de Camp, F., & Stiefelhagen, R. (2018). Interaction of distant and local users in a collaborative virtual environment. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10909 LNCS*, 328–337.

Huang, W., Alem, L., & Tecchia, F. (2013). HandsIn3D: Supporting remote guidance with immersive virtual environments. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8117 LNCS*(PART 1), 70–77.

Huang, W., Kim, S., Billinghurst, M., & Alem, L. (2019). Sharing hand gesture and sketch cues in remote collaboration. *Journal of Visual Communication and Image Representation*, *58*, 428–438.

Hugues, O., Fuchs, P., & Nannipieri, O. (2011). New Augmented Reality Taxonomy- Technologies and Features of Augmented Environment. In *Handbook of Augmented Reality* (pp. 47–64). Springer, New York, NY.

Irlitti, A., Piumsomboon, T., Jackson, D., & Thomas, B. H. (2019). Conveying spatial awareness cues in xR collaborations. *IEEE Transactions on Visualization & Computer Graphics*, *25*(11), 3178–3189.

Irlitti, A., Smith, R. T., Itzstein, S. Von, Billinghurst, M., & Thomas, B. H. (2017). Challenges for Asynchronous Collaboration in Augmented Reality. *Adjunct Proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2016*, 31–35.

Isenberg, P., Elmqvist, N., Scholtz, J., Gernea, D., Ma, K., & Hagen, H. (2011). Collaborative Visualization: Definition, Challenges, and Research Agenda. *Information Visualization*, *10*(4), 310–326.

Ishii, H., Kobayashi, M., & Arita, K. (1994). Iterative Design of Seamless Collaboration Media. *Communications of the ACM*, *37*(8), 83–97.

Izard, C. E. (2007). Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. *Perspectives on Psychological Science*, *2*(3), 260–280.

Jacko, J. A. (2012). *Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, Third Edition*. CRC Press.

Jalo, H., Pirkkalainen, H., Torro, O., Kärkkäinen, H., Puhto, J., & Kankaanpää, T. (2018). How can collaborative augmented reality support operative work in the facility management industry? *IC3K 2018 - Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, *3*(Ic3k), 41–51.

Jerald, J. (2015). The VR book: Human-centered design for virtual reality. In *ACM Books*. ACM Books.

Jetter, J., Eimecke, J., & Rese, A. (2018). Augmented reality tools for industrial applications: What are potential key performance indicators and who benefits? *Computers in Human Behavior*, *87*(February), 18–33.

Johansen, R. (1988). *GroupWare: Computer Support for Business Teams*. The Free Press.

Johson, S., Gibson, M., & Mutlu, B. (2015). Handheld or Handsfree? Remote Collaboration via Lightweight Head-Mounted Displays and Handheld Devices. *ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1825–1836.

Kangsoo Kim, Mark Billinghurst, Gerd Bruder, Henry Been-Lirn Duh, & Gregory F. Welch. (2018). revisiting Trends in. *IEEE Transactions on Visualization and Computer Graphics*, *24*, 2947–2962.

Kim, K., Billinghurst, M., Bruder, G., Duh, H. B., & Member, S. (2018). Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008 – 2017). *IEEE Transactions on Visualization and Computer Graphics, TVCG*, *24*(11), 2947–2962.

Kim, S., Billinghurst, M., & Kim, K. (2020). Multimodal interfaces and communication cues for remote collaboration. *Journal on Multimodal User Interfaces*, *14*(4), 313–319.

Kim, S., Billinghurst, M., Lee, C., & Lee, G. (2018). Using freeze frame and visual notifications in an annotation drawing interface for remote collaboration. *KSII Transactions on Internet and Information Systems*, *12*(12), 6034–6056.

Kim, S., Billinghurst, M., & Lee, G. (2018). The Effect of Collaboration Styles and View Independence on Video-Mediated Remote Collaboration. *Computer Supported Cooperative Work: CSCW: An International*, *27*(3–6), 569–607.

Kim, S., Lee, G. A., Ha, S., Sakata, N., & Billinghurst, M. (2015). Automatically freezing live video for annotation during remote collaboration. *Conference on Human Factors in Computing Systems - Proceedings*, *18*, 1669–1674.

Kim, S., Lee, G., Billinghurst, M., & Huang, W. (2020). The combination of visual communication cues in mixed reality remote collaboration. *Journal on Multimodal User Interfaces*, *14*(4), 321–335.

Kim, S., Lee, G., Huang, W., Kim, H., Woo, W., & Billinghurst, M. (2019). Evaluating the Combination of Visual Communication Cues for HMD-based Mixed Reality Remote Collaboration. *Conference on Human Factors in Computing Systems - Proceedings*, 1–13.

Kim, S., Lee, G., Sakata, N., & Billinghurst, M. (2014). Improving Co-Presence with Augmented Visual Communication Cues for Sharing Experience through Video Conference. *IEEE International Symposium on Mixed and Augmented Reality 2014, ISMAR 2014*, 83–92.

Komiyama, R., Miyaki, T., & Rekimoto, J. (2017). JackIn space: Designing a seamless transition between first and third person view for effective telepresence collaborations. *ACM International Conference Proceeding Series*.

Kong, X. T. R., Luo, H., Huang, G. Q., & Yang, X. (2019). Industrial wearable system: the human-centric empowering technology in Industry 4.0. *Journal of Intelligent Manufacturing*, *30*(8), 2853–2869.

Kozar, O. (2010). Towards Better Group Work: Seeing the Difference between Cooperation and Collaboration. *English Teaching Forum*, *2*(4), 23.

Krauß, V., Boden, A., Oppermann, L., & Reiners, R. (2021). Current Practices, Challenges, and Design Implications for Collaborative. *CHI Conference on Human Factors in Computing Systems*.

Kreijns, K., Kirschner, P. A., & Vermeulen, M. (2013). Social Aspects of CSCL Environments: A Research Framework. *Educational Psychologist*, *48*(4), 229–242.

Kwon, B. C., Kim, H., Wall, E., Choo, J., Park, H., & Endert, A. (2017). AxiSketcher: Interactive Nonlinear Axis Mapping of Visualizations through User Drawings. *IEEE Transactions on Visualization and Computer Graphics*, *23*(1), 221–230.

Kwon, J. U., Hwang, J. I., Park, J., & Ahn, S. C. (2019). Fully asymmetric remote collaboration system. *IEEE Access*, *7*, 54155–54166.

Lalanne, D., Nigay, L., Palanque, philippe, Robinson, P., Vanderdonckt, J., & Ladry, J.-F. (2009). Fusion Engines for Multimodal Input: A Survey. *Proceedings of the 2009 International Conference on Multimodal Interfaces*, 153–160.

Lamberti, F., Manuri, F., Sanna, A., Paravati, G., Pezzolla, P., & Montuschi, P. (2014). Challenges, Opportunities, and Future Trends of Emerging Techniques for Augmented Reality-based Maintenance. *IEEE Transactions on Emerging Topics in Computing*, *2*(4), 411–421.

Lee, C. P., & Paine, D. (2015). From the matrix to a model of coordinated action (MoCA): A conceptual framework of and for CSCW. *CSCW 2015 - Proceedings of the 2015 ACM International Conference on Computer-Supported Cooperative Work and Social Computing*, 179–194.

Lee, G. A., Kim, S., Lee, Y., Dey, A., Piumsomboon, T., Norman, M., & Billinghurst, M. (2017). Mutually Shared Gaze in Augmented Video Conference. *International Symposium on Mixed and Augmented Reality, ISMAR 2017*, *7*, 79–80.

Lee, G. A., Teo, T., Kim, S., & Billinghurst, M. (2017). Mixed reality collaboration through sharing a live panorama. *SIGGRAPH Asia 2017 Mobile Graphics and Interactive Applications, SA 2017*.

Lee, G. A., Teo, T., Kim, S., & Billinghurst, M. (2019). A User Study on MR Remote Collaboration Using Live 360 Video. *Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2018*, 153–164.

Li, W., Nee, A., & Ong, S. (2017). A State-of-the-Art Review of Augmented Reality in Engineering Analysis and Simulation. *Multimodal Technologies and Interaction*, *1*(3), 17.

Liao, Y., Deschamps, F., de Freitas Rocha Loures, E., & Ramos, L. F. P. (2017). Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal. *International Journal of Production Research*, *55*(12), 3609–3629.

Limbu, B. H., Jarodzka, H., Klemke, R., & Specht, M. (2018). Using sensors and augmented reality to train apprentices using recorded expert performance: A systematic literature review. *Educational Research Review*, *25*, 1–22.

Lindeman, R. W., & Noma, H. (2007). A Classification Scheme for Multi-Sensory Augmented Reality. *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology*, 175–178.

Longo, F., Nicoletti, L., & Padovano, A. (2017). Smart operators in industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context. *Computers and Industrial Engineering*, *113*, 144–159.

Lorenz, M., Knopp, S., & Klimant, P. (2018). Industrial Augmented Reality: Requirements for an Augmented Reality Maintenance Worker Support System. *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 151–153.

Ludwig, T., Stickel, O., Tolmie, P., & Sellmer, M. (2021). shARe-IT: Ad hoc Remote Troubleshooting through Augmented Reality. *Computer Supported Cooperative Work: CSCW: An International Journal*, *Cicourel 1990*.

Lukosch, S., Billinghurst, M., Alem, L., & Kiyokawa, K. (2015). Collaboration in Augmented Reality. *Computer Supported Cooperative Work, CSCW*, *24*(6), 515–525.

Mackay, W. E. (1998). Augmented Reality: Linking real and virtual worlds A new paradigm for interacting with computers. *Proceedings of the Working Conference on Advanced Visual Interfaces - AVI '98* , 13–21. https://www.lri.fr/~mackay/pdffiles/AVI98.AugmentedReality.pdf

Madeira, T., Marques, B., Alves, J., Dias, P., & Santos, B. S. (2020). Exploring annotations and hand tracking in Augmented Reality for remote collaboration. *International Conference on Human Systems Engineering and Design: Future Trends and Applications, IHSED*, 83–89.

Mahmood, T., Fulmer, W., Mungoli, N., Huang, J., & Lu, A. (2019). Improving information sharing and collaborative analysis for remote geospatial visualization using mixed reality. *Proceedings - 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2019*, 236–247.

Mani, M., Haapala, K., Smullin, M., & Morris, K. (2016). Using industry focus groups and literature review to identify challenges in sustainable assessment theory and practice. *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 1–10. https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=920843

Marques, B., Alves, J., Esteves, R., Dias, P., Fonseca, E., & Santos, B. S. (2018). Conducting Maintenance Tasks on Boilers with Augmented Reality in Multiple Devices. *International Conference on Graphics and Interaction, ICGI 2018*.

Marques, B., Alves, J., Neves, M., Maio, R., Justo, I., Santos, A., Rainho, R., Ferreira, C., Dias, P., & Santos, B. S. (2020). Interaction with Virtual Content using Augmented Reality: a User Study in Assembly Procedures. *Proceedings of the ACM on Human-Computer Interaction*, *4*(196), 1–17.

Marques, B., Araújo, T., Silva, S., Alves, J., Dias, P., & Santos, B. S. (2021). Visually exploring a Collaborative Augmented Reality Taxonomy. *Information Visualization, IV*, 1–6.

Marques, B., Barroso, J., Dias, P., & Sousa Santos, B. (2021). Exploring MR for Remote Collaboration: Comparing a 3D shared model approach with 2D annotations. *International Conference on Graphics and Interaction, ICGI 2021*, 1–4.

Marques, B., Dias, P., Alves, J., Fonseca, E., & Santos, B. S. (2019). Configuration and use of pervasive augmented reality interfaces in a smart home context: a prototype. *International Conference on Human Systems Engineering and Design: Future Trends and Applications, IHSED 2019*.

Marques, B., Dias, P., Rocha, A., Alves, J., & Santos, B. S. (2019). An Augmented Reality Framework for Supporting Technicians during Maintenance Procedures. *International Conference on Graphics and Interaction, ICGI 2019*, 1–4.

Marques, B., Esteves, R., Alves, J., Ferreira, C., Dias, P., & Santos, B. S. (2019). Investigating different Augmented Reality Approaches in Circuit Assembly: a User Study. *Eurographics 2019*, 45–48.

Marques, B., Silva, S., Dias, P., & Santos, B. S. (2021a). Evaluating Augmented Reality based Remote Collaboration: a contextualized approach. In *Human-Automation Interaction:: Mobile Computing* (pp. 1–14). Springer International Publishing.

Marques, B., Silva, S., Dias, P., & Santos, B. S. (2021b). A Toolkit to Evaluate and Characterize the Collaborative Process in Scenarios of Remote Collaboration Supported by AR. *IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 336–337.

Marques, B., Silva, S., Dias, P., & Santos, B. S. (2021c). An Ontology for Evaluation of Remote Collaboration using Augmented Reality. *European Conference on Computer-Supported Cooperative Work, ECSCW*, 1–8.

Marques, B., Silva, S., Rocha, A., Dias, P., & Santos, B. S. (2021). Remote Asynchronous Collaboration in Maintenance scenarios using Augmented Reality and Annotations. *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE VR 2021*, 567–568.

Marques, B., Silva, S. S., Alves, J., Araujo, T., Dias, P. M., & Sousa Santos, B. (2021). A Conceptual Model and Taxonomy for Collaborative Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, 1–18.

Marques, B., Teixeira, A., Silva, S., Alves, J., Dias, P., & Santos, B. S. (2021). A critical analysis on remote collaboration mediated by Augmented Reality: making a case for improved characterization and evaluation of the collaborative process. *Computers and Graphics*, 1–15.

Marques, B., Teixeira, A., Silva, S., Alves, J., Dias, P., & Santos, B. S. (2020). A Conceptual Model for Data Collection and Analysis for AR-based Remote Collaboration Evaluation. *IEEE International Symposium on Mixed and Augmented Reality, ISMAR Posters*, 1–2.

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, *12*(4), 1160–1177.

Martinetti, A., Rajabalinejad, M., & Van Dongen, L. (2017). Shaping the Future Maintenance Operations: Reflections on the Adoptions of Augmented Reality Through Problems and Opportunities. *Procedia CIRP*, *59*(TESConf 2016), 14–17.

Masai, K., Kunze, K., sugimoto, M., & Billinghurst, M. (2016). Empathy Glasses. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 1257–1263.

Masoni, R., Ferrise, F., Bordegoni, M., Gattullo, M., Uva, A. E., Fiorentino, M., Carrabba, E., & Di Donato, M. (2017). Supporting Remote Maintenance in Industry 4.0 through Augmented Reality. *Procedia Manufacturing*, *11*, 1296–1302.

Masood, T., & Egger, J. (2019). Augmented reality in support of Industry 4.0—Implementation challenges and success factors. *Robotics and Computer-Integrated Manufacturing*, *58*(August), 181–195.

Masood, T., & Egger, J. (2020). Adopting augmented reality in the age of industrial digitalisation. *Computers in Industry*, *115*, 103112.

Matthews, B., See, Z. S., & Day, J. (2021). Crisis and extended realities: remote presence in the time of COVID-19. *Media International Australia*, *178*(1), 198–209.

McInnerney, J. M., & Roberts, T. S. (2004). Collaborative or Cooperative Learning? In *Online Collaborative Learning: Theory and Practice* (Vol. 31, Issue 2, pp. 203–214).

Mekni, M., & Lemieux, A. (2014). Augmented Reality: Applications, Challenges and Future Trends. *Applied Computational Science*, 205–214.

Merino, L., Schwarzl, M., Kraus, M., Sedlmair, M., Schmalstieg, D., & Weiskopf, D. (2020). Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009–2019). *IEEE International Symposium on Mixed and Augmented Reality, ISMAR*, 1–14.

Meyer, M., & Dykes, J. (2019). Criteria for rigor in visualization design study. *IEEE Transactions on Visualization and Computer Graphics, TVCG*, *26*(1), 87–97.

Miaskiewicz, T., & Kozar, K. A. (2011). Personas and user-centered design: How can personas benefit product design processes? *Design Studies*, *32*(5), 417–430.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, *77*(12), 1321–1329.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. *Photonics for Industrial Applications*, 282–292.

Miller, M. R., Jun, H., Herrera, F., Yu Villa, J., Welch, G., & Bailenson, J. N. (2019). Social interaction in augmented reality. *PloS One*, *14*(5), e0216290.

Miyashita, T., Meier, P., Tachikawa, T., Orlic, S., Eble, T., Scholz, V., Gapel, A., Gerl, O., Arnaudov, S., & Lieberknecht, S. (2008). An augmented reality museum guide. *IEEE International Symposium on Mixed and Augmented Reality 2008, ISMAR 2008*, 103–106.

Mourtzis, D., Zogopoulos, V., & Vlachou, E. (2017). Augmented Reality Application to Support Remote Maintenance as a Service in the Robotics Industry. *Procedia CIRP*, *63*, 46–51.

Müller, J., Rädle, R., & Reiterer, H. (2017). Remote Collaboration With Mixed Reality Displays: How Shared Virtual Landmarks Facilitate Spatial Referencing. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6481–6486.

Narumi, T., Nishizaka, S., Kajinami, T., Tanikawa, T., & Hirose, M. (2011). Meta Cookie+: An Illusion-Based Gustatory Display. In R. Shumaker (Ed.), *Virtual and Mixed Reality - New Trends* (pp. 260–269). Springer Berlin Heidelberg.

Neale, D. C., Carroll, J. M., & Rosson, M. B. (2004). Evaluating computer-supported cooperative work: Models and frameworks. *ACM Conference on Computer Supported Cooperative Work, CSCW*, 112–121.

Nebeling, M., Speicher, M., Wang, X., Rajaram, S., Hall, B. D., Xie, Z., Raistrick, A. R. E., Aebersold, M., Happ, E. G., Wang, J., Sun, Y., Zhang, L., Ramsier, L. E., & Kulkarni, R. (2020). MRAT: The Mixed Reality Analytics Toolkit. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12.

Neumann, U., & Majoros, A. (1998). Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. *IEEE Conference on Virtual Reality, IEEE VR*, 4–11.

Nguyen, T. T. H., & Duval, T. (2015). A survey of communication and awareness in collaborative virtual environments. *2014 International Workshop on Collaborative Virtual Environments, 3DCVE 2014*, 1–8.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*(3), 336–359.

Nielsen, J., & Landauer, T. K. (1993). A Mathematical Model of the Finding of Usability Problems. *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, 206–213.

Nilsson, S., Johansson, B. J. E., & Jönsson, A. (2009). A co-located collaborative augmented reality application. *International Conference on Virtual Reality Continuum and Its Applications in Industry, VRCAI 09*, 179.

Norman, M., Lee, G. A., Smith, R. T., & Billinghurst, M. (2019). The Impact of Remote User's Role in a Mixed Reality Mixed Presence System. *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*.

Normand, J.-M., Servières, M., & Moreau, G. (2012). A new typology of augmented reality applications. *Proceedings of the 3rd Augmented Human International Conference, AH '12*, 1–8.

Noy, N. F., & McGuinness, D. L. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. *Stanford Knowledge Systems Laboratory Technical Report*, *15*(2), 1–25.

Obermair, F., Althaler, J., Seiler, U., Zeilinger, P., Lechner, A., Pfaffeneder, L., Richter, M., & Wolfartsberger, J. (2020). Maintenance with Augmented Reality Remote Support in Comparison to Paper-Based Instructions: Experiment and Analysis. *2020 IEEE 7th International Conference on Industrial Engineering and Applications, ICIEA 2020*, 942–947.

Oda, O., Elvezio, C., Sukan, M., Feiner, S., & Tversky, B. (2015). Virtual Replicas for Remote Assistance in Virtual and Augmented Reality. *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, 405–415.

Olalere, T. (2012). Methodology in Accounting Research: A Critique of Taxonomy. *SSRN Electronic Journal*, 1–40.

Olsson, T., Jarusriboonchai, P., Woźniak, P., Paasovaara, S., Väänänen, K., & Lucero, A. (2020). Technologies for Enhancing Collocated Social Interaction: Review of Design Solutions and Approaches. *Computer Supported Cooperative Work (CSCW)*, *29*(1), 29–83.

Ong, S. K., Yuan, M. L., & Nee, A. Y. C. (2008). Augmented reality applications in manufacturing: A survey. *International Journal of Production Research*, *46*(10), 2707–2742.

Ostergaard, K. J., & Summers, J. D. (2009). Development of a systematic classification and taxonomy of collaborative design activities. *Journal of Engineering Design*, *20*(1), 57–81.

Pace, F. De, Manuri, F., & Sanna, A. (2018). Augmented Reality in Industry 4.0. *American Journal of Computer Science and Information Technology*, *6*(1), 1–7.

Paelke, V. (2014). Augmented Reality in the Smart Factory. *Emerging Technology and Factory Automation*, 1–4.

Palmarini, R., Erkoyuncu, J. A., Roy, R., & Torabmostaedi, H. (2018). A systematic review of augmented reality applications in maintenance. *Robotics and Computer-Integrated Manufacturing*, *49*(March 2017), 215–228.

Patel, H., Pettitt, M., & Wilson, J. R. (2012). Factors of collaborative working : A framework for a collaboration model. *Applied Ergonomics*, *43*(1), 1–26.

Pece, F., Steptoe, W., Wanner, F., Julier, S., Weyrich, T., Kautz, J., & Steed, A. (2013). PanoInserts: Mobile spatial teleconferencing. *Conference on Human Factors in Computing Systems - Proceedings*, 1319–1328.

Pereira, C., Almeida, N., Martins, A. I., Silva, S., Rosa, A. F., Oliveira e Silva, M., & Teixeira, A. (2015). Evaluation of complex distributed multimodal applications: Evaluating a TeleRehabilitation system when it really matters. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9194*, 146–157.

Pereira, C., Teixeira, A., & Silva, M. O. (2016). *Dynamic Evaluation for Reactive Scenarios, Ph. D. Dissertation. University of Aveiro (MAPi)*.

Pereira, C., Teixeira, A., & Silva, M. O. (2014). Live Evaluation within Ambient Assisted Living Scenarios. *International Conference on PErvasive Technologies Related to Assistive Environments, PETRA*, 1–6.

Petersen, N., & Stricker, D. (2015). Cognitive Augmented Reality. *Computers and Graphics (Pergamon)*, *53*, 82–91.

Piumsomboon, T., Day, A., Ens, B., Lee, Y., Lee, G., & Billinghurst, M. (2017). Exploring enhancements for remote mixed reality collaboration. *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications on - SA '17*, 1–5.

Piumsomboon, T., Dey, A., Ens, B., Lee, G., & Billinghurst, M. (2019). The effects of sharing awareness cues in collaborative mixed reality. *Frontiers Robotics AI*, *6*, 1–18.

Piumsomboon, T., Lee, G. A., Hart, J. D., Ens, B., Lindeman, R. W., Thomas, B. H., & Billinghurst, M. (2018). Mini-me: An adaptive avatar for Mixed Reality remote collaboration. *Conference on Human Factors in Computing Systems - Proceedings*, *2018-April*, 1–13.

Piumsomboon, T., Lee, G. A., Irlitti, A., Ens, B., Thomas, B. H., & Billinghurst, M. (2019). On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. *Conference on Human Factors in Computing Systems - Proceedings*, 1–17.

Plummer-D'Amato, P. (2008). Focus group methodology Part 1: Considerations for design. *International Journal of Therapy and Rehabilitation*, *15*(2), 69–73.

Pope, H. (2018). Virtual and Augmented Reality Applications. In *Library Technology Reports*.

Poppe, E., Brown, R., Johnson, D., & Recker, J. (2012). Preliminary evaluation of an augmented reality collaborative process modelling system. *Proceedings of the 2012 International Conference on Cyberworlds, Cyberworlds 2012*, 77–84.

PTC. (2020). *Vuforia Chalk Enables Secure Communication to Increase Safety and Efficiency - Toyota Case Study*. https://www.ptc.com/en/case-studies/toyota-selects-vuforia-chalk

Quandt, M., Knoke, B., Gorldt, C., Freitag, M., & Thoben, K. D. (2018). General Requirements for Industrial Augmented Reality Applications. *Procedia CIRP*, *72*, 1130–1135.

Quint, F., Loch, F., & Bertram, P. (2017). The Challenge of Introducing AR in Industry - Results of a Participative Process Involving Maintenance Engineers. *Procedia Manufacturing*, *11*(June), 1319–1323.

Radkowski, R., Herrema, J., & Oliver, J. (2015). Augmented Reality-Based Manual Assembly Support With Visual Features for Different Degrees of Difficulty. *International Journal of Human-Computer Interaction*, *31*(5), 337–349.

Ranatunga, D., Feng, D., Adcock, M., & Thomas, B. (2013). Towards object based manipulation in remote guidance. *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013*, *October*, 1–6.

Ratcliffe, J., Soave, F., Bryan-Kinns, N., Tokarchuk, L., & Farkhatdinov, I. (2021). Extended Reality (XR) Remote Research: a Survey of Drawbacks and Opportunities. *CHI Conference on Human Factors in Computing Systems*, 1–13.

Regenbrecht, H., Lum, T., Kohler, P., Ott, C., Wagner, M., Wilke, W., & Mueller, E. (2004). Using augmented virtuality for remote collaboration. *Presence: Teleoperators and Virtual Environments*, *13*(3), 338–354.

Regenbrecht, H. T., Wagner, M. T., & Baratoff, G. (2002). Magicmeeting: A collaborative tangible augmented reality system. *Virtual Reality*, *6*(3), 151–166.

Rhee, T., Thompson, S., Medeiros, D., Dos Anjos, R., & Chalmers, A. (2020). Augmented Virtual Teleportation for High-Fidelity Telecollaboration. *IEEE Transactions on Visualization and Computer Graphics*, *26*(5), 1923–1933.

Richardson, T., Gilbert, S., Holub, J., Thompson, F., MacAllister, A., Radkowski, R., Winer, E., Davies, P., & Terry, S. (2014). Fusing Self-Reported and Sensor Data from Mixed-Reality Training. *Interservice/Industry Training, Simulation, and Education Conference, I/ITSEC 2014*, *14158*, 1–12. http://www.frederickt.com/pubs/iitsec2014.pdf

Rokhsaritalemi, S., Sadeghi-Niaraki, A., & Choi, S. M. (2020). A review on mixed reality: Current trends, challenges and prospects. *Applied Sciences (Switzerland)*, *10*(2).

Röltgen, D., & Dumitrescu, R. (2020). Classification of industrial Augmented Reality use cases. *Procedia CIRP*, *91*, 93–100.

Roschelle, J., & Teasley, S. D. (1995). The Construction of Shared Knowledge in Collaborative Problem Solving. In C. O'Malley (Ed.), *Computer Supported Collaborative Learning* (pp. 69–97). Springer Berlin Heidelberg.

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). *Industry 4.0 - The Future of Productivity and Growth in Manufacturing Industries*.

Ryskeldiev, B., Cohen, M., & Herder, J. (2018). StreamSpace: Pervasive Mixed Reality Telepresence for Remote Collaboration on Mobile Devices. *Journal of Information Processing*, *26*(0), 177–185.

Sandor, C., Fuchs, M., Cassinelli, A., Li, H., Newcombe, R., Yamamoto, G., & Feiner, S. (2015). Breaking the Barriers to True Augmented Reality. *ArXiv E-Prints*, *1512.05471*(December), 1–13.

Sanna, A., Manuri, F., Lamberti, F., Member, S., Paravati, G., & Pezzolla, P. (2015). Using Handheld Devices to Support Augmented Reality-based Maintenance and Assembly Tasks. *IEEE International Conference on Consumer Electronics, ICCE 2015*, 178–179.

Sasikumar, P., Gao, L., Bai, H., & Billinghurst, M. (2019). Wearable remotefusion: A mixed reality remote collaboration system with local eye gaze and remote hand gesture sharing. *Adjunct Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2019*, 393–394.

Schmalstieg, D., & Höllerer, T. (2016). *Augmented Reality: Principles and Practice*.

Schmutz, J., Hoffmann, F., Heimberg, E., & Manser, T. (2015). Effective coordination in medical emergency teams: The moderating role of task type. *European Journal of Work and Organizational Psychology*, *24*(5), 761–776.

Schneider, M., Rambach, J., & Stricker, D. (2017). Augmented reality based on edge computing using the example of remote live support. *IEEE International Conference on Industrial Technology*, 1277–1282.

Schulz, H.-J., Hadlak, S., & Schumann, H. (2011). The Design Space of Implicit Hierarchy Visualization: A Survey. *IEEE Transactions on Visualization and Computer Graphics*, *17*(4), 393–411.

Scott, S. D., Graham, T. C. N., Wallace, J. R., Hancock, M., & Nacenta, M. (2015). "Local Remote" Collaboration: Applying Remote Group Awareness Techniques to Co-Located Settings. *Proceedings of the ACM Conference Companion on Computer Supported Cooperative Work and Social Computing*, 319–324.

Sedlmair, M., Meyer, M., & Munzner, T. (2012). Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2431–2440.

Sereno, M., Wang, X., Besancon, L., Mcguffin, M. J., & Isenberg, T. (2020). Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics*.

She, M., & Li, Z. (2017). Design and Evaluation of a Team Mutual Awareness Toolkit for Digital Interfaces of Nuclear Power Plant Context. *International Journal of Human–Computer Interaction*, *33*(9), 744–755.

Šmite, D., Wohlin, C., Galviņa, Z., & Prikladnicki, R. (2014). An empirically based terminology and taxonomy for global software engineering. *Empirical Software Engineering*, *19*(1), 105–153.

Smith, J. A. (2015). Qualitative psychology: A practical guide to research methods. In *Qualitative psychology: A practical guide to research methods.* Sage Publications, Inc.

Smith, K. A. (1995). Cooperative learning: effective teamwork for engineering classrooms. *Proceedings - Frontiers in Education Conference*, *1*, 208–213.

Speicher, M., Hall, B. D., & Nebeling, M. (2019). What is mixed reality? *Conference on Human Factors in Computing Systems - Proceedings*, 1–15.

Stasko, J., Catrambone, R., Guzdial, M., & McDonald, K. (2000). An evaluation of space-filling information visualizations for depicting hierarchical structures. *Int. J. Hum. Comput. Stud.*, *53*, 663–694.

Stokols, D., Misra, S., Moser, R. P., Hall, K. L., & Taylor, B. K. (2008). The Ecology of Team Science: Understanding Contextual Influences on Transdisciplinary Collaboration. *American Journal of Preventive Medicine*, *35*(2, Supplement), S96–S115.

Sukumar, P. T., Avellino, I., Remy, C., Devito, M. A., Dillahunt, T. R., McGrenere, J., & Wilson, M. L. (2020). Transparency in qualitative research: Increasing fairness in the CHI review process. *Conference on Human Factors in Computing Systems - Proceedings*.

Suomela, R., & Lehikoinen, J. (2004). Taxonomy for visualizing location-based information. *Virtual Reality*, *8*(2), 71–82.

Tait, M., & Billinghurst, M. (2014). View independence in remote collaboration using AR. *ISMAR 2014 - IEEE International Symposium on Mixed and Augmented Reality - Science and Technology 2014, Proceedings*, 309–310.

Tait, M., & Billinghurst, M. (2015). The Effect of View Independence in a Collaborative AR System. *Computer Supported Cooperative Work: CSCW: An International Journal*, *24*(6), 563–589.

Teixeira, A. (2014). A critical analysis of speech- based interaction in healthcare robots: Making a case for the increased use of speech in medical and assistive robots. *Speech Technology and Text Mining in Medicine and Health Care*, 3–29.

Teo, T., Hayati, A. F., Lee, G. A., Billinghurst, M., & Adcock, M. (2019). A technique for mixed reality remote collaboration using 360 panoramas in 3D reconstructed scenes. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*.

Teo, T., Lawrence, L., Lee, G. A., Billinghurst, M., & Adcock, M. (2019). Mixed reality remote collaboration combining 360 video and 3D reconstruction. *Conference on Human Factors in Computing Systems - Proceedings*, 1–14.

Teo, T., Lee, G. A., Billinghurst, M., & Adcock, M. (2019). Investigating the Use of Different Visual Cues to Improve Social Presence within a 360 Mixed Reality Remote Collaboration*. *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*.

Teo, T., Lee, G. A., Billinghurst, M., & Adcock, M. (2018). Hand Gestures and Visual Annotation in Live 360 Panorama-based Mixed Reality Remote Collaboration. *Proceedings of the 30th Australian Conference on Computer-Human Interaction, OzCHI*, 406–410.

Teruel, M. A., Navarro, E., López-Jaquero, V., Montero, F., & González, P. (2017). A Comprehensive Framework for Modeling Requirements of CSCW Systems. *J. Softw. Evol. Process*, *29*(5), n/a.

Tönnis, M., & Plecher, D. A. (2011). Principles in Augmented Reality_ - Classification and Categorization Guidelines - Version 1 . 0. *Technical University of Munchen*.

Tullis, T., & Albert, B. (2013). Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics: Second Edition. In *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics: Second Edition*.

Tzimas, E., Vosniakos, G. C., & Matsas, E. (2019). Machine tool setup instructions in the smart factory using augmented reality: a system construction perspective. *International Journal on Interactive Design and Manufacturing*, *13*(1), 121–136.

Usman, M., Britto, R., Börstler, J., & Mendes, E. (2017). Taxonomies in software engineering: A Systematic mapping study and a revised taxonomy development method. *Information and Software Technology*, *85*, 43–59.

Van Krevelen, D. W. F., & Poelman, R. (2010). A Survey of Augmented Reality Technologies, Applications and Limitations. *The International Journal of Virtual Reality*, *9*(2), 1–21.

van Lopik, K., Sinclair, M., Sharpe, R., Conway, P., & West, A. (2020). Developing augmented reality capabilities for industry 4.0 small enterprises: Lessons learnt from a content authoring case study. *Computers in Industry*, *117*, 103208.

Waldow, K., Fuhrmann, A., & Grünvogel, S. M. (2019). Investigating the Effect of Embodied Visualization in Remote Collaborative Augmented Reality. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11883 LNCS*, 246–262.

Wang, D., Gao, Q., Tan, H., Liu, Z., Zhou, L., Jia, L., & Li, Z. (2020). Coordination breakdowns in nuclear power plant control rooms: cause identification and behaviour-sequence analysis. *Ergonomics*, *63*(6), 660–681.

Wang, P., Bai, X., Billinghurst, M., Zhang, S., Han, D., Sun, M., Wang, Z., Lv, H., & Han, S. (2020). Haptic Feedback Helps Me? A VR-SAR Remote Collaborative System with Tangible Interaction. *International Journal of Human-Computer Interaction*, *36*(13), 1242–1257.

Wang, P., Bai, X., Billinghurst, M., Zhang, S., Zhang, X., Wang, S., He, W., Yan, Y., & Ji, H. (2021). AR/MR Remote Collaboration on Physical Tasks: A Review. *Robotics and Computer-Integrated Manufacturing*, *72*.

Wang, P., Zhang, S., Bai, X., Billinghurst, M., He, W., Sun, M., Chen, Y., Lv, H., & Ji, H. (2019). *2 . 5DHANDS : a gesture-based MR remote collaborative platform*.

Wang, P., Zhang, S., Bai, X., Billinghurst, M., Zhang, L., Wang, S., Han, D., Lv, H., & Yan, Y. (2019). A gesture- and head-based multimodal interaction platform for MR remote collaboration. *The International Journal of Advanced Manufacturing Technology*, *105*(7), 3031–3043.

Wang, P., Zhang, S., Billinghurst, M., Bai, X., He, W., Wang, S., Sun, M., & Zhang, X. (2019). A comprehensive survey of AR/MR-based co-design in manufacturing. *Engineering with Computers*.

Wang, S., Parsons, M., Stone-McLean, J., Rogers, P., Boyd, S., Hoover, K., Meruvia-Pastor, O., Gong, M., & Smith, A. (2017). Augmented reality as a telemedicine platform for remote procedural training. *Sensors*, *17*(10), 1–21.

Wang, T. Y., Sato, Y., Otsuki, M., Kuzuoka, H., & Suzuki, Y. (2019). Effect of full body avatar in augmented reality remote collaboration. *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*, 1221–1222.

Wang, X., & Dunston, P. S. (2006). Groupware concepts for augmented reality mediated human-to-human collaboration. *… and Decision Making in Civil and …*, 1836–1842.

Wang, X., Ong, S. K., & Nee, A. Y. C. (2016). A comprehensive survey of augmented reality assembly research. *Advances in Manufacturing*, *4*(1), 1–22.

Wildman, J. L., Thayer, A. L., Rosen, M. A., Salas, E., Mathieu, J. E., & Rayne, S. R. (2012). Task types and team-level attributes: Synthesis of team classification literature. In *Human Resource Development Review* (Vol. 11, Issue 1).

Woodburn, L., Yang, Y., & Marriott, K. (2019). Interactive Visualisation of Hierarchical Quantitative Data: An Evaluation. *2019 IEEE Visualization Conference (VIS)*, 96–100.

XMReality. (2020). *How to Successfully Transfer Knowledge & Expertise*. https://xmreality.com/whitepaper/knowledge-expertise/

Yamada, S., & Chandrasiri, N. P. (2018). Evaluation of Hand Gesture Annotation in Remote Collaboration Using Augmented Reality. *25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 - Proceedings*, 727–728.

Yoon, B., Kim, H. Il, Lee, G. A., Billinqhurst, M., & Woo, W. (2019). The effect of avatar appearance on social presence in an augmented reality remote collaboration. *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*, 547–556.

Yoshioka, T., Herman, G., Yates, J., & Orlikowski, W. (2001). Genre Taxonomy: A Knowledge Repository of Communicative Actions. *ACM Trans. Inf. Syst.*, *19*(4), 431–456.

Zhou, F., Duh, H., & Billinghurst, M. (2008). Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR. *IEEE International Symposium on Symposium on Mixed and Augmented Reality, ISMAR 2008*, 193–202.

Zhu, J., Ong, S. K., & Nee, A. Y. C. (2014). A context-aware augmented reality system to assist the maintenance operators. *International Journal on Interactive Design and Manufacturing*, *8*(4), 293–304.

Zillner, J., Mendez, E., & Wagner, D. (2018). Augmented Reality Remote Collaboration with Dense Reconstruction. *Adjunct Proceedings - 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2018*, 38–39.

Zollmann, S., Grasset, R., Langlotz, T., Lo, W. H., Mori, S., & Regenbrecht, H. (2020). Visualization Techniques in Augmented Reality: A Taxonomy, Methods and Patterns. *IEEE Transactions on Visualization and Computer Graphics*, 1–20.