A Systematic Exploration of Collaborative Immersive Systems for Sense-Making in STEM

Article · December 2024

DOI: 10.1080/0144929X.2024.2441963

2 authors, including:

Brandon Syiem
The University of Sydney

A Systematic Exploration of Collaborative Immersive Systems for Sense-Making in STEM

BRANDON V. SYIEM, School of Computer Science
The University of Sydney, Australia
SELEN TÜRKAY, School of Computer Science
Queensland University of Technology, Australia

Scientific sense-making in STEM fields is a complex, yet essential activity, that greatly benefits from collaborations. However, challenges associated with collaboration, such as the geographic separation of experts, access to specialized equipment, and meaningful data representation, often hinder this process. Solutions to collaborative challenges have been extensively explored in CSCW and HCI literature. Among such solutions, immersive systems offer novel data visualizations, interactions, and representations that can support collaborative sense-making in STEM fields. Recognizing the increasing interest from HCI researchers on the intersection of collaboration and immersive systems, we conduct a systematic review to answer pertinent questions regarding the research landscape, the design and implementation of collaborative immersive systems for STEM sense-making. We find that current research leans towards synchronous collaborations, AR technology, and sense-making for learning in science domains. We further discuss prevalent trends and considerations observed in our findings, to inform future research directions.

Additional Key Words and Phrases: Immersive Systems, Virtual Reality, Augmented Reality, Mixed Reality, STEM, Sense-making, Collaboration, Systematic Review

1 INTRODUCTION

Complex challenges in scientific domains have often motivated experts to collaborate in search of solutions [1]. Historical evidence of such collaboration can be found as early as 335 BC, when Aristotle founded the Peripatetic School in Athens. This school served as a hub for philosophers, mathematicians, and scientists to gather, store, and interpret information across fields [2]. In contemporary times, the need for collaboration in scientific fields has become more pronounced. As scientific problems become increasingly niche and demand interdisciplinary perspectives, the need for experts from diverse backgrounds to collaborate has intensified [3]. In response to such needs, a growing number of technological tools and solutions have been developed to support scientific collaboration, such as emails, videoconferencing applications, scheduling software, digital libraries, shared authoring tools, and shared remote access to instrumentation [1].

These technological advancements are often employed to support collaboration across various stages involved in the scientific inquiry process, which underpins knowledge creation. For example, extensive work has been done in the field of computer-supported cooperative work (CSCW) on the use of technology to support the collection, storage, and management of data used for scientific purposes [4]. Among the various stages in scientific inquiry, the process of *scientific sense-making* has garnered attention from various perspectives, including psychology [5],

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The Version of Record of this manuscript has been published and is available in Behaviour & Information Technology, published 18 December 2024, https://www.tandfonline.com, https://doi.org/10.1080/0144929X.2024.2441963

Authors' addresses: Brandon V. Syiem, School of Computer Science
The University of Sydney, Sydney, NSW, Australia, brandon.syiem@sydney.edu.au; Selen Türkay, School of Computer Science
Queensland University of Technology, Brisbane, QLD, Australia, selen.turkay@qut.edu.au.

philosophy [6], and human-computer interactions (HCI) [7, 8]. 'Sense-making' broadly describes the process of 'making sense' of data/information to achieve understanding [7], while the term 'scientific' alludes to the rigorous methodologies employed in the sense-making process [9]. However, the semantics of both 'scientific' [6, 10] and 'sense-making' [5] have stirred debates within the scientific community, their precise definitions often eluding consensus. For instance, multiple theories have been developed to describe the sense-making process [11–13]. Regardless of this uncertainty, technological solutions have sought to support, replicate, and/or enhance the scientific sense-making process [14–20], given its crucial role in scientific fields.

Immersive technologies/systems, such as augmented reality (AR), virtual reality (VR), mixed reality (MR), and extended reality (XR), offer great potential for supporting collaborative scientific sense-making through immersive 3D visualizations, manipulations, and data sharing. Recent work has demonstrated the use of these systems for enabling and enhancing collaboration [21, 22], sense-making [19, 20, 23, 24], and domain-specific problem-solving in science, technology, engineering, and mathematics (STEM) [25–28]. Such growing interest has led to numerous surveys on how immersive systems are designed and subsequently shape modern collaborative work [29–34], and scientific work in STEM-related disciplines [35–39]. However, despite the pivotal role of sense-making in STEM [1, 40, 41], current surveys fail to address the complexities of *designing and using* collaborative immersive systems to support sense-making practices in specialized domains. Additionally, it remains unclear how to reconcile knowledge of existing sense-making theories with the developing trends of immersive systems and real-world scientific practice. Immersive systems present a wide range of novel capabilities for viewing, processing, and interacting with data and collaborators in specialized application domains in STEM. However, realising this potential requires a deeper understanding of how to design and leverage these systems effectively. As such, we aim to complement existing research by conducting a comprehensive literature review to answer the following relevant questions: *How are collaborative immersive systems applied in supporting the process of scientific sense-making in STEM?; How have considerations related to sense-making processes in STEM influenced the design of relevant collaborative immersive systems?*

Our review resulted in 39 papers focused on the use of collaborative immersive systems for sense-making in STEM. Our analysis revealed prevalent trends and notable considerations related to the design of these systems for scientific sense-making in STEM and further highlighted features of collaborative systems, immersive systems, and sense-making that were less explored. Specifically, we find that current research leaned towards AR technology, real-time (synchronous) collaborations, and sense-making for learning primarily in the science domain. We discuss our findings in light of real-world scientific collaboration practices and the influence of sense-making processes on current research practices related to collaborative immersive systems in STEM.

This work contributes to HCI and STEM research communities by offering an in-depth analysis of collaborative immersive systems tailored for sense-making within STEM disciplines. Our findings highlight predominant trends, notably a significant focus on the *science* domain with an emphasis on the use of sense-making for *learning*. We uncover the adaptability of the HCI community in response to real-world dynamics, such as the rise in *remote* collaborative systems post-2019. We further spotlight the influence of different sense-making characteristics on collaborative immersive system and experimental design. We further discuss notable methods of analysis, observed in our sample, that are used to provide valuable insights into the sense-making process. Additionally, we identify research gaps, such as the underrepresentation of *asynchronous* collaborations and the integration of physical elements in AR setups, to help guide future innovations. In essence, this research bridges the gap between STEM and HCI, illuminating prevailing research trajectories and uncharted avenues and highlighting the potential of collaborative immersive systems for supporting sense-making in STEM.

2 BACKGROUND

We split our discussion of relevant literature into three parts; 1) On Scientific Sense-making, 2) Collaboration in Immersive Systems, and 3) Immersive Systems for STEM.

2.1 On Scientific Sense-making

To discuss the concept of 'scientific sense-making', we first introduce the term 'scientific discovery'. 'Scientific discovery' refers to new findings in relation to objects, properties, events, processes, theories, hypothesis or methods, that is a direct result of the scientific inquiry process [42]. This processes can be loosely summarised by the following steps; 1) generate hypotheses, 2) collect relevant data, and 3) make sense of the data to confirm or reject the generated hypotheses [43, 44]. The sense-making process, listed in step 3, has been a topic of ongoing interest for academics in fields such as psychology, philosophy, history, anthropology, and sociology [6]. Klahr and Simon [6] states five reasons for this interest in the scientific sense-making process; for its human and humane value, its mythology, to understand complex human thinking, to understand the developmental course of the process, and to design systems that can replicate the process of scientific sense-making. Additionally, prior work has demonstrated another reason for studying the mechanisms of scientific sense-making – and sense-making in general – to innovate on systems and methods to support learners and practitioners of (scientific) sense-making [17, 18, 45–47].

To this end, previous work has studied different methods to support and replicate both scientific and general sense-making. For example, research as early as the 1980s have attempted to replicate the process of scientific sense-making through computer programs, such as the DENDRAL [15] or KAKEDA [14]. More recently, the conversation around the use of technologies in relation to scientific sense-making has broadened from strictly *replicating* sense-making processes to *supporting* learners [16, 17] and practitioners [18] of sense-making. Significant efforts have also been directed in supporting 'general' sense-making in the field of HCI [46–49] and artificial intelligence [50, 51]. But what differentiates 'scientific' sense-making from 'general' sense-making?

Prior work disagrees on the precise distinction between scientific and everyday sense-making [52]. Some define 'scientific' as the use of scientific methods (precise definitions, critical questioning, meticulous evidence seeking, objective reasoning) for sense-making [9], while others argue that scientific sense-making is named so, as a consequence of the discovery being scientific and not on account of the process itself [6]. Others still, consider the relationship between scientific sense-making and everyday sense-making as fundamentally continuous [53], i.e., the absence of a dichotomy between knowledge and language used in everyday experiences (such as those expressed by children) and those used in science (such as those expressed in academic writing). This lack of consensus on the precise boundaries between "scientific" and "everyday" sense-making, leads us to adopt a broader perspective on the sense-making process for this review (including both "scientific" and "everyday" sense-making).

Sense-making, itself, is a complex topic that has been likened to phenomena such as creativity, curiosity, comprehension, mental-modelling and situational awareness, with none capturing the complete essence of the term [5]. Five major theories have surfaced in the literature that describe the phenomena of sense-making from different perspectives [11–13]. We adopt Kolko's [12] terminology and refer to these approaches as Dervin's theory, Russel's theory, Hoffman, Klein & Moon's theory, Snowden's, and Weick's theory of sense-making.

Dervin's theory positions sense-making as an individual and continuous process that substantiates *learning* of complicated ideas through actions, as opposed to studying them abstractly [12, 54]. They describe sense-making as a process of seeking, using, creating and rejecting knowledge and information to 'bridge' 'gaps' that individuals encounter when moving through time and space [54]. Russell et al. [8] view sense-making as a process for *information processing and mental-modelling* of data specific for a set task or problem [8, 12, 55]. They describe sense-making as 'the process of searching for a representation and encoding data in that representation to answer task-specific questions'. Russel's work on sense-making lies primarily within the context of information systems and reducing costs of information processing. A broader perspective on sense-making was presented by Klein et al. [5]. They describe sense-making as a process to model connections for the purpose of *problem-solving/decision-making* for a specific and contained problem [12]. Specifically, Klein et al. [5] refers to sense-making as a motivated and continuous effort to understand connections between people, events and places to anticipate future trajectories and act effectively. Similar to Klein, Snowden's [56] theory of sense-making centres on *problem-solving/decision-making*, but from an organizational perspective. They proposed a 'sense-making framework', called Cynefin, that supports collective sense-making for tackling unspecified problems and considering intractable problems in new ways. Lastly, Weick et al. [57] pivots their discussion of sense-making around organizational behaviour. They view organizations, not as an static individual entity, but as a continuously changing collective of interacting individuals that 'orgnizes' information for orgnizational growth and planning [12].

These theories largely agree that sense-making is a *process*, but disagree on the approaches/methods involved and the purpose of the process (i.e., what is sense-making a process of?) [11]. Within these five theories of sense-making, we discern four broad categories of *sense-making as a process of*; *learning*, *mental-modelling/information processing*, *problem-solving/decision-making*, and *organizational growth and planning*. We also acknowledge that the particularities of each approach/method within the different sense-making theories is beyond the scope of this paper, and is extensively discussed in prior work (see Turner et al. [11], Harteveld [55] and Kolko [12]).

These categories of sense-making processes highlight key differences that may influence the design and evaluation of supporting tools and technologies. For instance, technologies designed to support sense-making as a process of *learning* may wish to incorporate user actions that facilitates understanding of broader abstract concepts. In contrast, sense-making as a process for *problem-solving/decision-making* may require technological solutions that provide specialized tools and interactions for a specific and contained problem. Such differences also influence the methods used in evaluating the success of these solutions, with the former assessing the long-term and broader effects of the sense-making process, and the latter focusing on performance related to the specific problem task. As such, these categories of sense-making will serve as one of several lenses through which we scrutinize existing research to better understand how current collaborative immersive systems support different sense-making processes.

2.2 Collaboration in Immersive Systems

The use of technology for collaboration has been extensively pursued in the context of work and education, leading to the formation of their respective sub-fields; Computer-Supported Cooperative Work (CSCW) and Computer-Supported Collaborative Learning (CSCL). Such attention on collaboration can be attributed to its importance in addressing the complex challenges and opportunities that arise in science and society [3]. With the advent of accessible immersive technology hardware (such as the Meta Quest 2), an increasing amount of research has been directed towards understanding the use of immersive technology in supporting collaboration. This can be attributed to the unique affordances provided by immersive technology that enables the design of collaborative tools unavailable through other technology. For instance, immersive technology can enable deictic and non-verbal communication between remote and/or co-located collaborators present in the same virtual space [21, 58], a feature not easily replicable with other technology for remote users.

Despite the growing interest in collaborative immersive systems, recent reviews have mainly limited their discussion to perspectives of CSCW theories and frameworks, and/or individual human factors (including user experience and interactions). For example, Ens et al. [31] strongly drew on concepts related to CSCW and classified the papers included in their review along five dimensions; Time and Space [59] (a well known matrix in the CSCW literature), Symmetry (based on the collaborators roles), Artificiality (based on the reality-virtuality continuum [60], i.e., the degree to which the collaboration space was virtual or real), Focus (based on the primary target of collaborative activity), and Scenario (based on the application use case of the system).

Prior reviews have also discussed relevant works along isolated concepts related to CSCW. For instance, Ouverson and Gilbert [32] expanded the discussion presented by Ens et al. [31] along the axis of 'Symmetry', and focused their review on *asymmetric* collaborations in immersive systems. Similarly, Schäfer et al. [29] discusses real-time remote collaboration (Time and Space concepts in CSCW literature) in MR systems, but does so primarily through the lens of human factors. Schäfer et al. [29] creates a taxonomy to classify papers along contributions related to the virtual *environment*, the virtual *avatars*, and the provided *interactions*. They discuss the significance of these concepts in relation to human factors such as telepresence, social presence, embodiment, interaction preference, and awareness, among others.

Human factors, in relation to the individuals within the collaborative scenario, has also taken centre stage in recent reviews on collaborative immersive systems. van den Oever et al. [33] and Laskay et al. [34] presents discussion related to user performance, cognitive challenges, situational awareness, perception, and mental workload, in addition to methodological and design approaches employed in collaborative immersive systems for maritime operation and preoperative planning respectively. Other reviews have presented significant discussion through both perspectives; discussing CSCW concepts in relation to collaborative immersive systems, and the human factors associated with the design of such systems. For instance, de Belen et al. [30] discuss both the user experience and interaction aspects of collaborative immersive systems, along with the CSCW 'Space' dimension (which they term 'collaborative setups'). Additionally, they analyse the reviewed papers' contributions based on the application domains and hardware used for immersive applications.

These reviews highlight an increasing interest in the use of immersive systems for collaborations. They also spotlight pertinent decision that need to be considered when designing collaborative immersive systems. The dimensions utilized by Ens et al. [31] provides an insightful overview of the design considerations and variations between collaborative immersive systems explored in current literature. However, these dimensions alone fail to capture the nuances and complexities arising from the use of immersive systems for collaboration in specific domains, such as seen in those presented by van den Oever et al. [33] and Laskay et al. [34]. Such specialized application domains, as seen in STEM fields, can benefit greatly from expanding the current literature to better understand the design, effects, and evaluation of collaborative immersive systems used to support their core needs and activities. As such, in this review, we aim to complement the existing body of work on collaborative immersive systems by focusing on the crucial process of sense-making in STEM application domains.

2.3 Immersive systems for STEM

Compared to collaboration in immersive systems, the topic of immersive system use in STEM has benefited from a wider set of perspectives in recent reviews, albeit in a limited context, i.e., education. Consequently, reviews have mainly discussed the topic from the perspectives of learning and instructional theories. For example, a review on AR use for STEM by Ibáñez and Delgado-Kloos [35] presents their discussion through the lens of instructional processes (strategies and techniques). They define instructional *strategy* according to Akdeniz [61, p.57-105] as the approaches followed by instructors to achieve the fundamental aims of instruction, and identify three distinct categories in the papers they reviewed: presentation, discovery and cooperative learning. Ibáñez and Delgado-Kloos [35] refer to instructional *techniques* as the rules, procedures, tool and skills to deploy instruction techniques [62], and identify five distinct categories in the papers they reviewed: observation, inquiry, game, role-play, and concept maps. In addition, Ibáñez and Delgado-Kloos [35] sheds light on the measures used to evaluate the affective and cognitive outcomes, along with the design considerations for AR applications in STEM, specifically in the context of AR features employed (Location-based, Marker-based, Camera-based registration, etc.). The prominence of instructional processes related to *discovery* and *collaboration* in prior work, as seen in the review, highlight the importance of these activities within STEM education and potentially in STEM professions.

The focus on instructional strategies and techniques in relation to immersive system use in STEM was also adopted by Theodoropoulos and Lepouras [36] and later by Mystakidis et al. [37]. The former narrows the scope to programming education, but broadens the discussion to include technology, design, and user interaction [36]. They extensively discuss and highlight both the pedagogical outcomes (cognitive and affective outcomes, learning gains, motivation, etc.) and processes (cognitive behaviour, collaboration, affective processes, etc.), as well as the user experience and interaction aspects (satisfaction, usability, engagement, etc.) of the reviewed papers. The latter review by Mystakidis et al. [37] directs their attention to the use of AR for higher education in STEM, discussing the instructional strategies and techniques employed, along with design considerations for the immersive interventions through the lens of cognitive multimedia learning theory [63]. In addition to the three categories of instructional strategies identified by Ibáñez and Delgado-Kloos [35], Mystakidis et al. [37] discuss two additional categories of *activity-based* and *experiential* strategies. They also categorize their instructional techniques as; observation, simulation, project, problem-solving, and question-answer. These reviews highlight the core themes that appear in the discussion around instructional use of immersive system in STEM education, and clearly indicate the prominence of both collaboration and discovery (including instructional techniques related to problem-based learning) in the literature.

Prior work has also extensively discussed the learning outcomes and contributions to learners of immersive systems use in STEM education [38, 39]. Categorization into themes of the different learning outcomes and learner contributions was carried out in both the reviews presented by Sırakaya and Alsancak Sırakaya [39] and Ajit [38]. Themes related to learning outcomes largely overlapped between the two reviews, such as visualization, fun learning, collaborative learning, concretize abstract concepts, and student-centered learning. Conversely, themes related to learner contributions had fewer overlaps, with Sırakaya and Alsancak Sırakaya [39] focusing on more specialized themes, such as outcomes related to spatial or cognitive abilities, and Ajit [38] discussing broader outcomes, such as knowledge construction. Both reviews also commonly discuss the benefits of immersive systems in increasing interaction amongst students and between students and teachers.

Other works have explored the prevalence of immersive systems research in various STEM domains, and the research methodologies employed in evaluating such systems [64, 65]. For instance, Ciupe et al. [64] investigated the most prominent fields in engineering education where immersive systems were employed and found that Production Engineering and Computer Engineering comprised the majority of the application domain. They also found that the most common type of study involved a solution proposal, and the most common intervention involved simulations of 3D environments and tools. A more detailed view of the research methodologies employed in evaluating virtual reality systems in computer science education was presented by Agbo et al. [65]. Their findings provide insights on the employed research methods (qualitative, quantitative, mixed method, design and development, etc.), data collection tools (observations, surveys, system generated data, questionnaires, etc.) and analysis methods (descriptive analysis, learning analytics, reflective analysis, etc.). Additionally, Agbo et al. [65] present an analysis of the author keyword co-occurrence pattern of the reviewed articles, which highlights a strong relationship between ‘virtual reality’ and keywords of ‘visualization’, ‘virtual environment’, ‘immersive learning’, ‘games’, ‘gamification’ and ‘collaborative learning’.

The reviews discussed in this section shows a trend in employing immersive systems to facilitate collaboration and sense-making (among others) within STEM education. This may be attributed to the fact that collaboration and sense-making (including discovery, problem-solving and decision-making) are core activities within a majority of STEM disciplines (both within and beyond the educational context) [1, 6]. Despite the critical importance of these activities, current literature limits their focus on perspectives of instructional learning and outcomes, only addressing collaboration and sense-making concepts peripherally. In this work, we call attention to these crucial elements of scientific enquiry and conduct a systematic review to better understand the current considerations and approaches used in realizing collaborative immersive systems for sense-making in STEM.

3 METHODS

In this systematic review, we aim to understand the current state of collaborative immersive systems used for scientific sense-making in the fields of science, technology, engineering and mathematics (STEM). Specifically, this review is centered around the following research questions: *how are collaborative immersive systems applied in supporting the process of scientific sense-making in STEM?; how have considerations related to sense-making processes in STEM influenced the design of relevant collaborative immersive systems?* By answering these questions, we provide insights on the relationship between properties of collaborative immersive systems and the sense-making processes that they support, and how such relationships have evolved over time. We draw on these insights to reflect on the current state of collaborative immersive systems for scientific sense-making in STEM, and propose future directions.

We reviewed articles from academic databases that include substantial Human-Computer Interaction related papers, namely the ACM Digital Library, Web of Science, and Scopus. We adopt the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [66] guidelines to conduct our review. The inclusion of a paper in our review was determined by the following criteria:

- **Use of Immersive System:** The paper needs to include an immersive [67] augmented, virtual, mixed, or extended reality system. We include mobile-based or screen-based displays if they present digital information that *matches* the users’ body movements or interactions with the physical world [67].
- **Designed for and evaluated with collaborative tasks:** The paper must describe system design or development for the purposes of collaboration between 2 or more users.
- **Includes User Study:** The paper must include an empirical user study that evaluates a system, validates a methodology, or tests a theoretical concept/framework. We rely on the description of the experimental design, procedures and/or findings to understand the specific sense-making processes involved during the use of collaborative immersive systems.
- **Is related to a STEM field:** The application context of the collaborative immersive system described in the paper must be related to a STEM discipline. As the categorization of a field as a STEM field is not often consistent [68], we decided to include papers that fall under the high-frequency STEM education or occupation fields as presented by Koonce et al. [69].

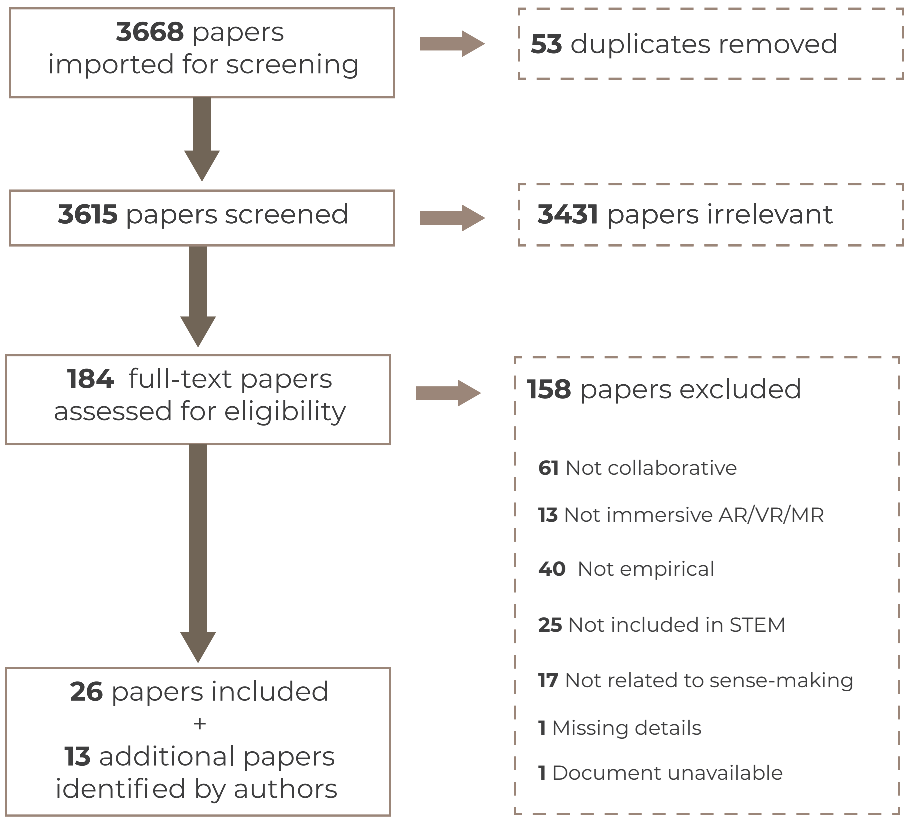

Fig. 1. Overview of our literature review detailing each stage of sampling process as per the PRISMA guidelines.

- **Involves Sense-making:** The paper must discuss sense-making processes, including tasks related to organizational growth and planning, learning, problem-solving/decision-making, or mental-modelling/information processing [11, 12, 55] (see section 2.1 for details on sense-making processes).

3.1 Sampling

This review spans papers published between 2011 and 2023. This review was initially started in 2021 and aimed to review papers published in the past 10 years. However, as circumstances did not allow us to complete the analysis in 2021, this review was redone in 2023 — adding papers published in 2022 and 2023. The search was conducted using the following search query:

```
[Abstract: "virtual reality" OR "augmented reality" OR
"mixed reality" OR "extended reality" OR "immersive reality"]
AND
[All: collab* OR teamwork OR co-operat*]
AND
[All: sense-making OR discovery]
```

```
AND
[All: science OR technology OR engineering OR mathematics OR "STEM"]
```

We included terms relevant to immersive systems in the abstract as this was a core requirement for our review. Additionally, we searched for terms related to collaboration and sense-making in the whole document as we wanted to review papers that included both these terms but did not necessarily focus on either application, i.e., terms related to both, either or neither may appear in the abstract or title, but were necessary to appear in the whole document.

The initial dataset included 3668 articles. A team of five researchers, specializing in HCI and immersive systems, proceeded to remove all duplicate papers from the initial search result, and screened all papers for relevance based on their title and abstract. We followed with a full review of the remaining papers, including papers based on our inclusion criteria. We conducted frequent meetings to ensure that all researchers were consistent with the inclusion and exclusion of papers.

A total of 26 papers fit our criteria and were included in our review. An additional 13 papers were identified using manual search techniques that have been used in prior work [70–73] and are compatible with the PRIMA guidelines [66]. We used a combination of backward and forward citation snowballing [74] and leveraged our contact with study authors of relevant papers [66]. We conducted backward snowballing by examining the reference lists found in the papers included through our database search and the relevant reviews discussed in section 2. Forward snowballing was achieved by examining sources that cite the papers included through our database search and the relevant reviews discussed in section 2 [1]. Our screening using snowballing was restricted to the title only, and we only considered articles that were related to immersive systems and at least two of the three other keywords used in our search string, i.e., collaboration (including teaching/instructing), sense-making, and STEM. If we found literature surveys on immersive systems (that could potentially include relevant sources based on their title) or new articles that met our eligibility criteria, we further screened them using the snowball approach. We limited our search to a depth of 2. As per the PRISMA guidelines [66], we have highlighted the sources included through our manual search in table 5 of our appendix section to increase transparency. Figure 1 shows a summary of our review process as per the PRISMA procedure.

3.2 Analysis

We extracted meta-data and information related to the problem statement, system description, experiment design, user study, and data analysis from the included papers. Following data extraction, one researcher led the analysis of the extracted data and engaged in open coding to identify elements related to collaboration, immersive system use and sense-making. We followed with an axial coding exercise to organize our codes into themes [75].

We used the developed themes to categorize each paper along dimensions related to the discussed system's 1) collaborative properties, 2) immersive qualities, and 3) sense-making use cases. The first six dimensions, related to the CSCW concepts of *time* and *space*, the *symmetry* of collaborator capabilities, the extent of *artificiality* of the immersive system, the *focus* of collaboration, and the collaborative *scenario*, are directly taken from the work presented by Ens et al. [31]. We adopt these dimension as they encompass important distinction pertaining to the collaborative and immersive properties of the described, developed, and/or evaluated system. The seventh dimension is derived from Lee and Paine's [76] matrix which describes the scale of collaboration enabled by the immersive system i.e, the number of collaborators that the system allows. Finally, the eight dimension categorizes papers based on the primary purpose of the sense-making process involved in the experimental task presented in the paper. This dimension is based on the

[1]We used Google Scholar for forward snowballing

different theories of sense-making discussed in our background (section 2.1). We summarize each dimensions and their possible values as follows:

(1) **Time:** Includes the value *synchronous* for systems that enable collaborators to use the system at the same time, and *asynchronous*, for systems that enable collaborators to work at different times while maintaining state. Similar to Ens et al. [31], we also include the category *both* for systems that enable both synchronous and asynchronous workflows.
(2) **Space:** Includes *co-located* category for systems that require users to be in the same physical space to work, and *remote* category for systems that enable users to work while physically separated. The category *both* is also included to classify systems that enabled both co-located/remote collaborations.
(3) **Symmetry:** Categorizes papers based on the similarity/dissimilarity of interactions and visualizations afforded to different collaborators. Includes the value *symmetric* when all collaborators have the same interactions and content access, and the value *asymmetric* for when collaborators are not afforded the same capabilities. We also included the category *both* for papers that discussed systems that involved both symmetric and asymmetric collaborations.
(4) **Focus:** Derived from the dimension with the same name presented in Ens et al. [31], this dimension categorizes papers based on the focus of collaboration. Includes *workspace* referring to collaborations primarily focused on the workspace, *object* where the focus on collaboration is a physical object or a digital recreation of a physical object that is transmitted to a collaboration, *person* when the focus of collaboration is the 3D, augmented or physical representation of the collaborator(s), and *environment* where collaboration is centered on the surroundings of the participating collaborators.
(5) **Scenario:** This category includes *shared workspace* for papers focusing on a combined physical or artificial workspace, *remote expert* for collaborative systems that involve remote guidance or interaction between a local novice and a remote expert, and *Telepresence* for systems designed around communication between collaborators, and *co-annotation* for papers focusing on enabling collaborators the ability to author and register content into the immersive environment. These categories adopted from the review presented by Ens et al. [31], but exclude the category *shared experience* which categorizes systems that focus on the experiences of each individual and not the task. This was excluded as our review specifically focuses on the sense-making task that the collaborative immersive system enables.
(6) **Artificiality:** Refers to the degree of digital and physical content incorporated by the system [31, 77]. Includes the category *mostly physical* for systems that require users to primarily attend to or interact with the physical environment, the category *mostly digital* for systems that places more utility on digital visualization and interactions, and *hybrid* for systems that emphasizes both physical and digital content and interactions.
(7) **Scale:** Indicates the number of collaborators the system is designed for. While Scale was treated as a continuous value ranging from 2 to N in the matrix presented by Lee and Paine [76], we adopt a simpler categorization including *Dyad* (for specifically 2 collaborators) and *Multi-user* (For systems enabling more than 2 collaborators).
(8) **Sense-making purpose:** The primary purpose of the sense-making process in the experimental task used to evaluate the collaborative immersive system. This can take the value *learning*, *organizational growth and planning*, *problem-solving/decision-making* , and *mental-modelling/information processing*. These categories are derived from established sense-making theories (discussed in section 2.1 and summarised in prior work [11, 12, 55]). We focus on the *primary purpose* of sense-making due to its multifaceted nature within tasks in STEM domains.

> Consider the tasks related to word problems (commonly observed in education, such as mathematics) and qualitative data analysis. Both tasks involves processes related to parsing of data (e.g., text describing a "real-world" scenario for word problems or participant related transcripts for qualitative research) extracting key information (*mental-modelling/information processing*), deciding on and applying appropriate methods to arrive at a solution (*problem-solving/decision-making*), and making connections between data and abstract concepts (*learning*). However, the primary purpose of sense-making for these task differ. For instance, the sense-making process in most word problems is used to support *learning*, with success being evaluated through broader learning outcomes of abstract concepts. In contrast, sense-making in qualitative analysis serves varied purposes depending on the task, such as *problem-solving/decision-making* to address specific research problems, or *mental-modelling/information processing* to realize connections between concepts distilled from the data. We posit that these distinctions are important as they provide insights for the design and evaluation of technological solutions catered towards the different sense-making purposes. For instance, *learning* focused tasks may require a refined feedback loop to strengthen learning goals, whereas, tasks focused on *mental-modelling/information processing* may require tools for effectively searching for and filtering through information.

Examining the papers using the identified dimensions enables us to gain a better understanding of the landscape surrounding the use of collaborative immersive application for sense-making processes in STEM. Specifically, our analysis provides insights on the relevant collaborative properties adopted for tasks involving sense-making, the sense-making processes that are supported by immersive systems, and the relationship between collaborative factors, sense-making, and STEM domains.

4 FINDINGS

We report the findings from our analysis of the 39 included papers in three separate subsections: 1) the distribution across dimensions related to collaboration, 2) the immersive technology used and the degree of artificiality, and 3) the sense-making purposes used in STEM areas supported through collaborative immersive systems.

4.1 Distribution across Collaborative Dimensions

We present Table 1 which displays the various dimensions and their corresponding categories associated with the collaborative nature of the immersive systems. This table also indicates the number and percentage of papers corresponding to each category. Additionally, Figure 2 illustrates the distribution of papers across these collaborative dimensions (represented by the vertical axes) over time (grouped in 3 years, except for the last 2 years). Detailed plots for each collaborative dimension, distributed over the sampled years, are provided in the appendix section.

Time. We found that all 39 papers (100%) addressed *synchronous* collaboration under the Time dimension. This means that all collaboration using the immersive systems described took placed in real-time. These papers focused on scenarios and tasks where one collaborator's action or prompt was often followed by another collaborator's reaction or response within a reasonable span of time. For example, Maria et al. [78] describes a system where a remote radiologist, using a desktop, assists a surgeon in an operating room to make sense of and identify structures, such as tumors or nodes, on a shared image of a patient's medical scan. Such a procedure necessitates synchronous collaboration, given that the expert's actions are contingent upon the local user's queries or actions. Similar task scenarios are depicted in other surveyed works, like those by Neroni et al. [79], Nikolic et al. [80] and Sarkar et al. [81], where collaboration required real-time discussion and interactions amongst collaborators.

Table 1. Collaborative dimensions and their respective categories as exhibited by systems described in the included papers. Note that categories for dimensions that were not recorded during our analysis do not appear in the table.

Collaborative Dimension	Number of Papers	Percentage
Time		
Synchronous	39	100%
Space		
Co-locates	32	82%
Remote	6	15%
Both	1	3%
Symmetry		
Symmetric	27	69%
Asymmetric	11	28%
Both	1	3%
Scale		
Dyad	16	41%
Multi-user	23	59%
Focus		
Workspace	38	97%
Object	1	3%
Scenario		
Shared Workspace	36	92%
Remote Expert	3	8%

Space. Compared to the Time dimension,the Space dimension showed greater variation. The majority of papers (32 papers, 82%) fell under the *co-located* category. These papers primarily leveraged the affordances of physical co-location, such as face-to-face communication, awareness of collaborators, non-verbal communication, and deictic expressions to enable collaboration within an immersive environment. For instance, Kang et al. [82] developed a system that enabled groups of students to discuss, inquire and solve simple mathematical problems through experimentation with everyday physical objects. These items were collectively decided to be displayed in front of a tablet camera equipped with a custom built AR application. Similarly, the immersive application described by Benk et al. [83] enabled dyads in the same room to explore, discuss, and interact with visualizations of an machine learning (ML) model predictions. Following this, they could collaboratively design their own ML model.

Six papers (15%) were found to describe *remote* immersive systems within the Space dimension. These systems were primarily designed to address challenges stemming from collaborators' inability to share the same physical space. For instance, Maria et al. [78], Aschenbrenner et al. [84] and Gasques et al. [85] enabled users located in-situ to collaborate with remote experts using immersive systems for tasks that require expert assistance. The remaining papers under the *remote* category facilitated peers to collaboratively work (interact, discuss, and observe) around a shared visualization of relevant workspace. For example, Sedlák et al. [86], Šašinka et al. [87] enabled students, remotely located, to interact and observe 2D contour plots of geographical locations along with their 3D counterparts, to form better association between the different representations. Finally, 1 paper (3%) describes an immersive system that enables both co-located and remote users. O'Connor et al. [88] presented and evaluated a framework that allows collaborators to explore physics simulations in a shared virtual environment.

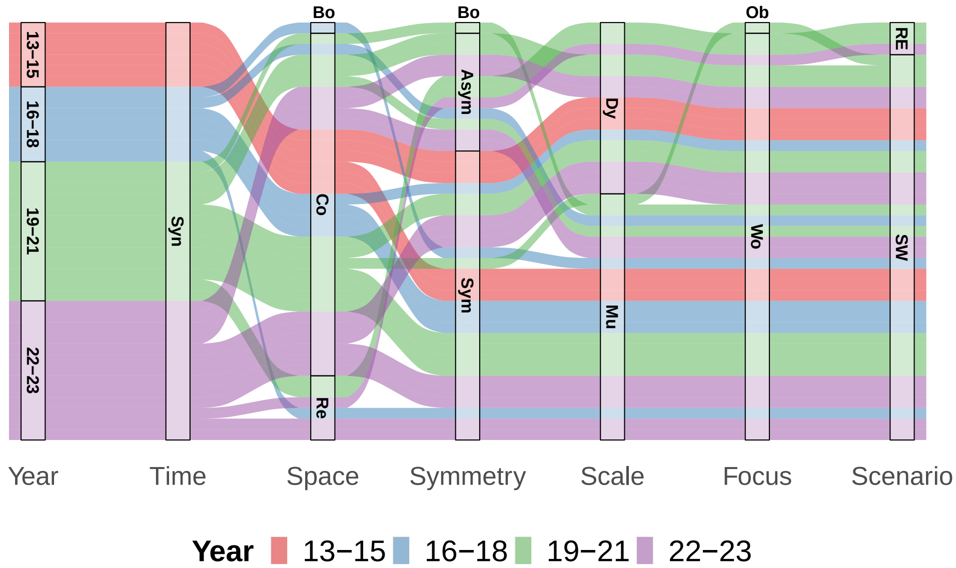

Fig. 2. Distribution of papers over time across the different collaborative dimensions. Each vertical axes represents a collaborative dimension; excluding the first which represents the years (2013-2023) divided in buckets. The division of each axes represent the different categories and the size of each division represents the number of papers that fall under that category for that dimension. The width of the path along the different categories over each axis represent the number of papers that fall under that category for that axis. The abbreviations under each axis are: Time (Syn = Synchronous), Space (Re = Remote, Co = Co-located, Bo = Both), Symmetry (Sym = Symmetric, Asym = Asymmetric, Bo = Both), Scale (Dy = Dyad, Mu = Multi-User), Focus (Wo = Workspace, Ob = Object), Scenario (SW = Shared Workspace, RE = Remote Expert).

Symmetry. We identified 27 papers (69%) that fall under the *symmetric* category of the Symmetry dimension. This category consists of the majority of our samples and includes papers where every collaborator has the same affordances with respect to the interactions and visualizations offered by the immersive system. For instance, Williams-Bhatti et al. [89] investigates the effects of a multi-user VR collaborative application on knowledge acquisition, and performance for doctors and nurses. While the roles that the doctors and nurses take within the system described by Williams-Bhatti et al. [89] are different, the basic affordances offered by the systems to every user is the same, hence the categorization as a *symmetric* collaborative system.

In contrast, *Asymmetric* collaborative systems was the focus of 11 papers (28%) in our sample, and comprise of papers that involve immersive systems offering different interactions and visualization to users based on their role. In the application described by Webb et al. [90], one user (designated as the pilot) immersed themselves in a virtual environment using a head-mounted display while a collaborator (termed the co-pilot) viewed the pilot's interactions and provided instructions available on physical materials. Finally, only one paper (3%) was designated as *both.* [91] investigated student's attitude towards learning science and acceptability of MR applications through multiple MR applications that enabled either *symmetric* or *asymmetric* collaboration. This included an application where a single student using a VR headset interacted with a virtual object while other students observed a projection of the interaction on a large screen. Conversely, another application presented the same visualizations and interactions to all student through an AR application on tablets or smartphones.

Scale. Of our sample, 16 papers (41%) detailed immersive systems that facilitated *Dyadic* collaborations. Papers involving dyadic collaborations made up the entirety of expert-novice applications, as can be seen by tracing the lines passing through the *Asym* category in the Symmetry dimension to the *RE* category in the Scenario dimension in Figure 2. Dyadic collaboration were also notably prevalent in paired learning applications, and applications that required collaborators to perform different tasks. For example, [92] explored how student pairs could learn about a science related topic through an AR systems that incorporated multiple physical images augmented with educational visual content. The physical images triggered additional information once the correct combination of images were placed together in front of the AR device. Other dyadic collaborative systems were designed for users assigned primarily to different tasks, such as seen in the works by Gasques et al. [85], Thompson et al. [93], Uz-Bilgin et al. [94] and Wang et al. [95].

Multi-user collaborative systems was featured in 23 papers (59%) in our sample. These papers explored a wider range of applications when compared to systems focused on *dyadic* collaborations. Immersive systems designed for *multi-user* collaboration included applications in design [96], review of 3D models and recreations [80, 97], decision-making and problem-solving [79, 88], collaborative learning [82, 87], and collaboration for specialized tasks, such as surgery [89]. We also found that papers categorized as *multi-user* were published consistently in our yearly groupings, with each grouping spanning three years, except the final one, which covered only two years, as illustrated in Figure 2.

Focus. The vast majority of the papers were categorized as *workspace* (38 papers, 97%) under the focus dimension, indicating that the focus of collaboration was primarily on the physical and/or virtual object(s) relevant to the collaborative task. For instance, Barrett et al. [98] developed an immersive system where collaborators focus on both physical and virtual elements of their shared workspace. In their implementation students explored concepts related to chemical engineering through experimentation with tangible objects to receive corresponding virtual feedback on a MR tabletop display. Other studies primarily directed collaborators' attention towards a virtual workspace. Such approaches are exemplified in the works of Williams-Bhatti et al. [89], Tan et al. [99], de Back et al. [100] and Neroni et al. [79]. A comprehensive list of all incorporated papers, along with their categorization across the identified dimensions, can be found in Table 5 in the appendix section.

Only a single paper (3%) fell under the *object* category in the Focus dimension. All other categories, as described in section 3.2, were notable absent in our sample. The system presented by Kang et al. [82] focuses collaborating students' attention on everyday physical objects. These objects can be scanned using an AR device, enabling students to solve rudimentary mathematical problems. For instance, they could sum up the count of similar physical objects detected, or divide a numerator (represented by a set of physical objects) by a denominator (represented by a different set of physical objects).

Scenario. In our sample, 36 papers (92%) were categorized as *Shared Workspace*, while 3 papers (8%) were identified as *Remote Expert* with respect to the Scenario dimension. Papers categorized as *Shared Workspace* focused on using immersive systems in enabling collaborators to engage with tasks and topics of interest with novel interactions and visualizations, or provide collaborators with tools and/or visualizations to better communicate around a shared workspace. For instance, Kang et al. [101] incorporated multiple physiological sensors to enable groups of children to interact with a shared display for collaborative science learning. Similarly, Barrett et al. [98] and Chen and Wang [102] (among others) enabled users to use tangible objects to interact with a display for collaborative experimentation and inquiry. Other works in our sample focused on the use of immersive systems to better communication between collaborators within virtual environments; either by enabling co-located users to explore, interact with, and discuss

Table 2. Categories of immersive technology used. AR + VR category includes papers that utilize both AR and VR systems. Single categories, such as AR, VR, CAVE VR, and MR only discuss/use systems falling under that category. Note that MR systems include systems that includes AR and augmented virtuality [60], i.e., focusing on the physical objects/interactions within the immersive environment.

Immersive Technology	Number of Papers	Percentage
Augmented Reality	14	36%
Virtual Reality	11	28%
Cave Automatic Virtual Reality	4	10%
Mixed Reality	8	21%
Augmented and Virtual Reality	2	5%

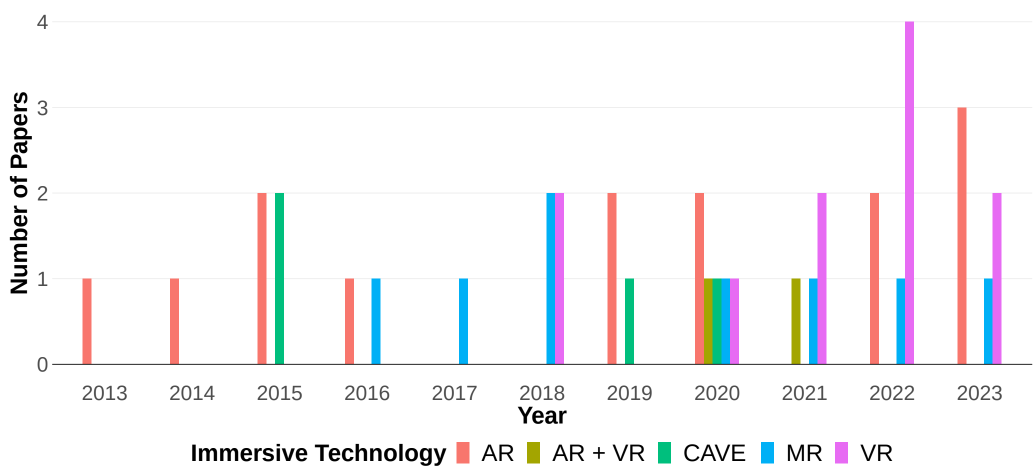

Fig. 3. Distribution of papers over the years based on the immersive technology used.

a shared visualization together [80, 100, 103], or by enabling remote (and/or a mix of remote and co-located [88]) collaborators to inhabit and interact with a shared virtual workspace [87, 97].

Papers classified under the *Remote Expert* category emphasized the introduction of novel interactions and visualization. These systems were designed to bolster the ability of remote experts to better guide and communicate with a local novice, along with means for the local novice to better gain support from the remote expert. For instance, the immersive system in Gasques et al. [85] enabled remote experts to observe the local novice user's workspace through 3D recreations, point of view of the novice user, and external cameras in the novice user's workspace. Their system also enabled the novice user to observe annotation created by the remote expert, along with the remote user's virtual avatar. Remaining papers classed as *Remote Expert* include papers by Aschenbrenner et al. [84] and Maria et al. [78]. They describe similar immersive systems to the one presented by Gasques et al. [85], which aim to enhance communication and guidance between a remote expert and local user through shared annotations and mutually accessible perspective views.

4.2 Immersive Technology use and Artificiality

Immersive Technology. Table 2 shows the number and percentage of papers that used various categories of immersive systems. These categories have been streamlined for simplicity, omitting granular details such as the specific type of

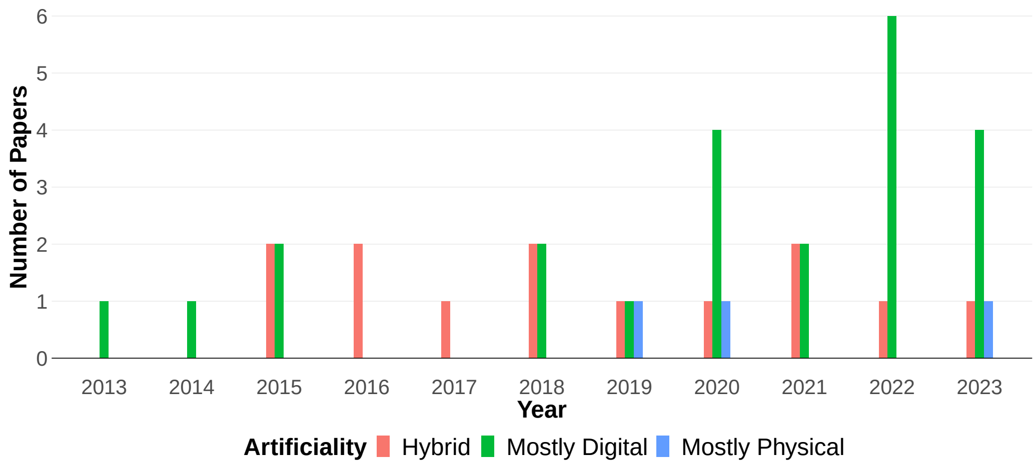

Fig. 4. Distribution of papers over the years based on the ARTIFICIALITY dimension described in section 3.2.

AR used (e.g., spatial, head-mounted, mobile). A more detailed breakdown is available in Table 5 in the appendix. We found that 14 papers (36%) reported focusing on augmented reality for their study, 11 papers (28%) used virtual reality, 4 papers (10%) used a cave automatic virtual reality system, 8 papers (21%) focused on mixed reality technology, and 2 papers (5%) reported using both augmented and virtual reality for their work.

Figure 3 shows the distribution of papers over time, and the immersive technology utilized. AR papers appear in every year from 2013-2023 except 2017, 2018 and 2021, whereas VR papers for collaborative immersive system for sense-making in STEM initially appeared in 2018, and reappeared starting 2021 all the way through to 2023, with the most number of papers appearing in 2022. The other categories (MR, CAVE, and AR+VR), with relatively smaller number of papers, were published between the years 2015 and 2023.

Table 3. The number and percentage of papers based on the ARTIFICIALITY dimension.

ARTIFICIALITY	NUMBER OF PAPERS	PERCENTAGE
Mostly Physical	3	8%
Mostly Digital	23	59%
Hybrid	13	33%

Artificiality. We found three papers (8%) in our sample that employed immersive technology where the focus was primarily on real physical objects or the real-world, i.e., *Mostly Physical* category. These papers primarily used virtual content to supplement information about a task that users are performing in their physical context. For example, some integrated virtual content authored by a remote user to assist a local user with a task related to their immediate physical surrounding [78, 84], while others deployed algorithms to generate relevant virtual content for the user(s) [82].

Papers categorized as *Mostly Digital* accounted for the majority of our sample, with 23 papers (59%). These papers focused on primarily two areas. The first area of focus relates to the effects of shared 3D visualizations and interactions in immersive systems on collaborative tasks (learning, decision-making, etc.) typically performed in-situ, in-person, and/or through other non-immersive technology (non-immersive VR, collaboration through skype, etc) [83, 91, 96, 99, 100, 104]. For instance, Tan et al. [99] explored the effects of 3D recreations of steel reinforcement bars in an immersive AR environment on student learning and collaboration when compared to 2D drawings and non-immersive 3D models on a 2D desktop. The second area of focus related to the creation of novel visualizations and interactions for workspace/objects with no physical counterparts, or physical counterparts that are impractical to work with under normal circumstances [93–95, 105]. For instance, Matovu et al. [105] developed a system allowing students to experiment with hydrogen bond simulations through embodied interactions in VR.

Thirteen papers (33%) were categorized as *Hybrid* under the Artificiality dimension. The majority of these papers (12 out of 13 papers) focused on the use of physical objects and sensors as a means of interaction with virtual objects or workspaces in an immersive environment. For example, Ens et al. [106], Radu and Schneider [107], and Matcha and Rambli [103], among others, use physical object to interact with an immersive system, either tabletop display with AR [106] or augmented reality displays [103, 107]. Additionally, one paper [85] in the *hybrid* category enabled individual collaborators to either view a fully virtual environment or a physical environment augmented with virtual content. Table 3 and Figure 4 provide details on the number and distribution of papers over time based on the Artificiality dimension.

4.3 STEM and Sense-making Purpose

Table 4. The number and percentage of papers based on the sense-making purpose dimension within the observed STEM areas. sense-making purpose categories and STEM areas that were not observed in our sample do not appear in the table.

STEM and Sense-making purpose	Number of Papers	Percentage
Science		
Learning	18	46%
Mental-Modelling/Information Processing	2	5%
Problem-solving/Decision-making	10	26%
Engineering		
Learning	2	5%
Problem-solving/Decision-making	5	13%
Mathematics		
Learning	2	5%

Table 4 shows the distributions of papers focusing on the different purposes of sense-making grouped by STEM fields. We found that most papers focused on *science* applications (30 papers, 77%), with a majority centered on the use of sense-making for the purpose of *learning* (18 papers, 46%). Remaining papers falling under the *science* umbrella centered on sense-making for *mental-modelling/information processing* (2 papers, 5%) and *problem-solving/decision-making* (10 papers, 26%). Papers under *engineering* accounted for 18% of our sample with 7 papers, and centered on *learning* (2 papers, 5%), and *problem-solving/decision-making* (5 papers, 13%). Papers focusing on the STEM area of *mathematics* included 2 papers (5%) both of which emphasized sense-making for *learning*.

The papers focusing on the use of sense-making for *learning* primarily harnessed immersive systems to realize diverse learning paradigms, including embodied learning [101, 107], computer-supported collaborative learning [87, 103], scientific discovery learning [86], cognitive affective model of immersive learning [86], experiential learning [107], and game-based learning [87]. For instance, Sedlák et al. [86] developed a collaborative immersive VR environment based

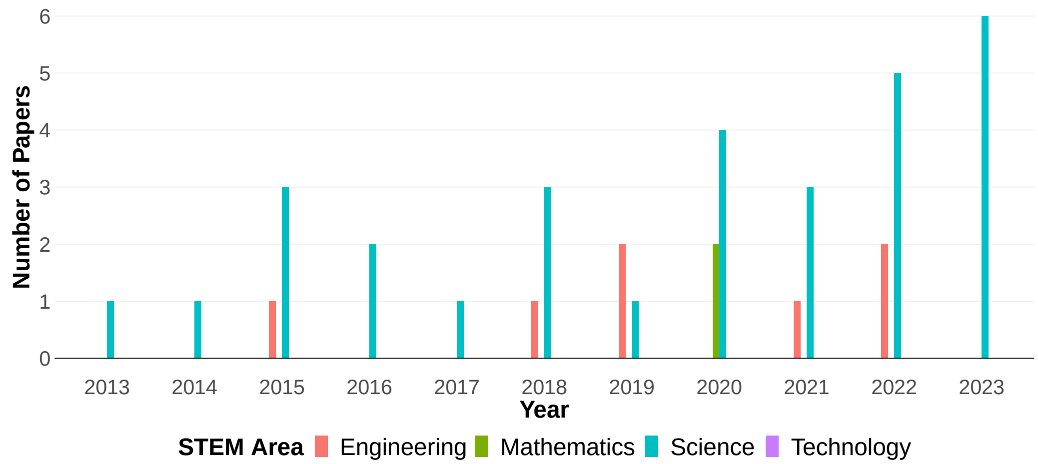

Fig. 5. Distributions of papers over the years based on the *STEM Area* of application.

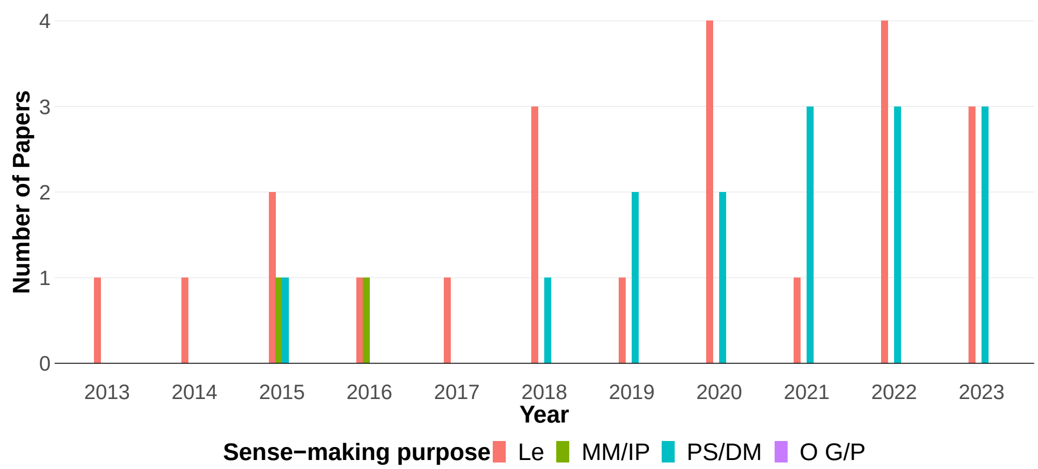

Fig. 6. Distribution of papers over time based on the Sense-making purpose dimension as described in section 3.2. Here, the abbreviated legends refer to: Le - Learning, MM/IP - Mental-modelling/Information processing, PS/DM - Problem-solving/Decision-making, and O G/P - Organizational growth and planning.

on *learning through problem-solving* [108] and *scientific discovery learning* [109], enabling students to learn geographical concepts related to hypsography.

Mental-modelling/information processing centered papers primarily focused on the use of sense-making in understanding connections, and cause-effect relationships, for a specific topic or context. For instance, Philips et al. [110] adapts and assesses MacEachren and Kraak's [111] 3D immersive geovisualization cube for supporting research students in hypothesis generation, assessment, and modelling of flood risks in a given region. Finally, papers highlighting sense-making for *problem-solving/decision making* primarily tackled tasks with well-defined problems or narrow set of end goals. For example, Neroni et al. [79] proposed a new method for investigating design activities, by observing how

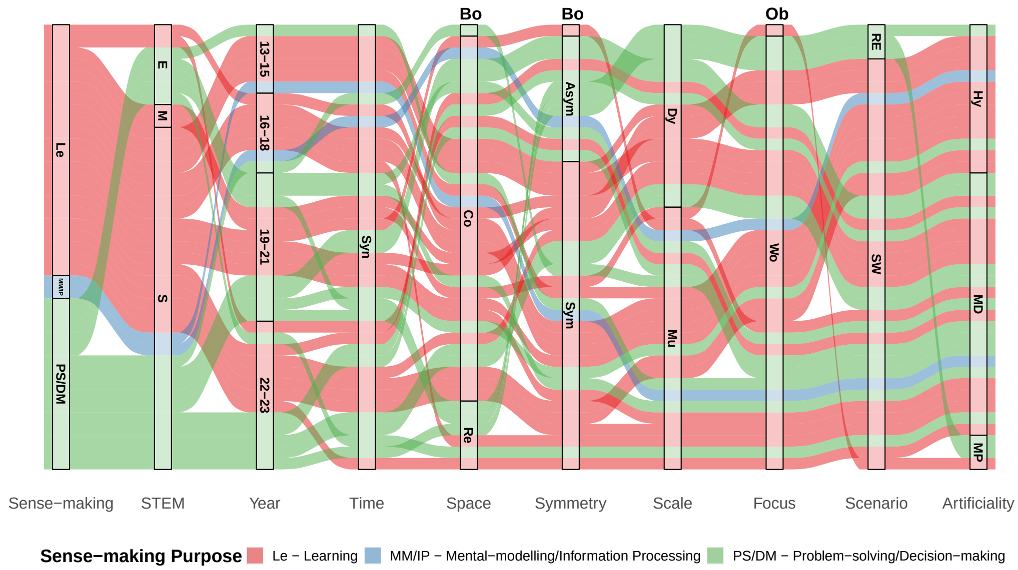

Fig. 7. Distribution of papers based on the SENSE-MAKING PURPOSE across all dimensions. The abbreviations under each axis are: Time (Syn = Synchronous), Space (Re = Remote, Co = Co-located, Bo = Both), Symmetry (Sym = Symmetric, Asym = Asymmetric, Bo = Both), Scale (Dy = Dyad, Mu = Multi-User), Focus (Wo = Workspace, Ob = Object), Scenario (SW = Shared Workspace, RE = Remote Expert), and Artificiality (Hy = Hybrid, MD = Mostly Digital, MP = Mostly Physical).

users build and test constructs to solve a specific problem in immersive virtual reality game that adheres to physics constraints.

Figure 5 and 6 show the temporal distribution of our sample, based on the *STEM Area* and the SENSE-MAKING PURPOSE dimensions. We found that papers focusing on collaborative immersive systems for sense-making in *science* have been published in every year between 2013 and 2023. Conversely, papers focusing on *mathematics* have seen the least number of publication and were only published in the year 2020 (apart from *technology* which saw no publications during the span of our review). Papers with an *engineering* focus were primarily published between the years 2015 and 2022. The use of sense-making in our sample exhibits similar trends, with *learning* focused papers appearing in every year. *Mental-modelling/information processing* related papers had the least number of publications (excluding the absent *organizational growth and planning* sense-making focus), with one publication in each of 2015 and 2016. Papers using sense-making for *problem-solving/decision-making* were primarily published between 2018 and 2023.

Lastly, Figure 7 shows the intersection between the diverse categories of SENSE-MAKING PURPOSE with the different categories along the other identified dimension described in section 3.2. It can be discerned, for example, that papers focused on *mathematics* all centered on the use of sense-making for *learning*, were all published between 2019-2021, used *co-located* immersive systems, were *symmetric* in nature, differed in SCALE (with presence in both *dyadic* and *multi-user* applications), differed in FOCUS (with presence in *workspace* and *object*), and fell under the *Shared workspace* SCENARIO.

5 DISCUSSION & FUTURE DIRECTIONS

Our analysis reveals recurring patterns and prevalent properties of collaborative immersive systems for scientific sense-making in STEM fields. Notably, our findings show that research in recent years has demonstrated particular interests in the *science* domain, with a focus of sense-making as a process for *learning* (82%). We found a noticeable preference towards certain collaborative dimensions, specifically *synchronous* (100%) and *co-located*(82%) systems. This might reflect the natural dynamics of many STEM collaborations but it also raises questions about the underrepresentation of other forms of collaboration. This discussion is based on the observed trends and gaps in our findings and is divided into two subsections. The first subsection discusses the progression of collaborative immersive systems observed in our findings in relation to real-world scientific practices. Specifically, this subsection focuses on the temporal shifts in the *collaborative and immersive properties* (see section 3.2) of systems described in the included papers.The second subsection discusses our findings in relation to the sense-making theories introduced in section 2. This subsection focuses on relevant extracted information (see section 3.2) used to categorize papers along the *sense-making purpose* dimension.

5.1 Parallels and Dissimilarities to Real-world Scientific Collaboration

Our findings indicate that the evolution of collaborative immersive systems for sense-making in STEM largely reflect the challenges and opportunities that arise from available technology and real-world scientific needs. A notable trend is observed when examining figure 2 indicating that a majority of the papers focusing on *remote* collaborative systems appear during or after 2019. This surge can be tracked back to the uncertainties introduced by social distancing norms and policies, which reshaped the collaboration modalities available to researchers and and scientists. While only three papers [78, 86, 97] explicitly attribute their focus to the influence of these social distancing policies, we hypothesize that these restrictions catalyzed a renewed interest in exploring remote collaborations. This hypothesis finds support in works like Šašinka et al. [87] and Gasques et al. [85], published after 2019. These papers not only delve into remote collaboration but also resonate with the inherent need of scientific communities that often span vast geographical divides, as highlighted by Sonnenwald [1].

The trajectory of immersive system development and its expanding capabilities provide a lens into how technology has shaped collaborative immersive systems over time. In earlier research endeavors, the prevalent use of marker-based, mobile, and/or screen-based AR devices was evident, as seen in Figure 3. These devices, largely due to their accessibility and capabilities, catered primarily to co-located collaborations. However, a pivotal shift towards VR systems becomes evident from 2018 onward. This transition aligns with significant advancements in VR technology during that period, exemplified by the release of the first untethered VR headset, Oculus Go in 2018 [2]).

Our findings also reflect the interests and need to support educational endeavours using immersive systems in STEM expressed in recent prior work (as discussed in section 2.3), with more than half of the sampled papers (22 out of 39 papers) focused on the use of collaborative immersive systems to support sense-making for *learning* (see table 4). This emphasis on learning might be rooted in the potential of immersive systems to bring to life a wide spectrum of learning theories and paradigms. Such potential and its implications are both evident in our findings (see section 4.3), and echoed in prior research (discussed in our background 2.3). However, some important questions arose that are beyond the scope of this paper. These questions relate to the influence of collaborative immersive systems on current

[2]See Oculus Go

learning paradigms. Future research is needed to understand if and how collaborative immersive systems disrupt existing learning taxonomies and how these taxonomies should be adapted and used in response to this technology.

Research on collaborative immersive system for sense-making in STEM have evolved rapidly, adapting to address real-world challenges and harnessing emerging opportunities and needs. However, there seems to be a glaring oversight: the absence on certain collaborative dimensions that, while perhaps not of immediate urgency, remain crucial for holistic STEM collaboration. A case in point is the overwhelming focus on *synchronous* collaborations in the sampled papers. This is surprising given that the real-world scientific landscape thrives on a blend of both synchronous and asynchronous collaborations, as evidenced by prior work [1, 112]. This sole focus on *synchronous* collaborative immersive systems for sense-making in STEM is puzzling, especially when previous research has already showcased its feasibility and benefits of asynchronous collaboration in immersive systems. For example, early work presented by Irlitti et al. [113] demonstrates how mobile AR devices can be used to author content in real-time and placed in-situ of a workspace for collaborators to access as and when necessary, without the need for collaborators to be present at the same time. While the application presented by Irlitti et al. [113] was not designed for scientific collaboration, such asynchronous collaboration are commonplace, and an important part of long-term collaboration in scientific domains [114]. More recent works have also explored new paradigms in asynchronous collaborations [115], opening the door for transformative tools for sense-making in STEM. This observed trend (or lack thereof) underscores a larger implication: Are we, in our quest to harness the latest and most immediate opportunities, inadvertently sidelining other impactful avenues? As the field progresses, it is crucial to ensure a balanced exploration, one that embraces both immediate innovations and foundational principles of collaboration.

Our findings also indicate that the use of immersive systems for long-term collaborations in STEM sense-making has been largely neglected. While many of today's non-immersive collaborative tools (like emails, shared worksheets, video/audio recordings) inherently supported prolonged interactions, it is concerning to note a dearth of research effort in exploring immersive systems for persistent collaborations. This gap becomes even more pronounced given our observation that most studies pivot around the theme of a collaborative *workspace* (Focus:*Workspace* and Scenario:*Shared Workspace*). As such, there is a clear avenue for future research to delve into the additional CSCW related dimensions, especially those that cater to persistent collaborative practices. For example, the paradigms introduced by Lee and Paine [76] which encompasses interdisciplinarity, the freshness of collaboration (nascence), dynamics of collaborator turnover, and the nature of collaboration longevity (short- vs long-term collaboration), could be pivotal in shaping the next wave of immersive collaborative systems.

A separate, yet equally intriguing trend from our findings, is the pronounced tilt towards *mostly digital* content in immersive systems, especially when many of these systems are grounded in AR technology and emphasize *co-located* collaborations. On the surface, this bias could be attributed to the natural inclination of users to lean on familiar co-located collaboration modalities, like physical gestures and deictic expressions, supplemented by the rich digital visualizations AR offers [83, 104]. However, sidelining the integration of physical objects, interactions, and spaces might be counterproductive. The exclusion of these physical elements can inhibit the sense-making processes [116], reduce user presence [117, 118], and have negative consequences on awareness within the immersive environment [119, 120]. We did find papers that have capitalized on the benefits of tangible interactions (example, Ens et al. [106]) and incorporating digital content in physical spaces (example, Maria et al. [78]). However, this group of papers represented a minority, and focus on highly specific use cases. As we progress in this domain, it's vital to ensure that immersive systems seamlessly integrate the digital with the tangible and the immediate with the long-term, to fully harness their potential for collaboration and sense-making in STEM.

While significant advances have been made in realizing and understanding collaborative immersive system use for supporting sense-making processes in STEM, our findings highlight new opportunities and shortcomings in current research and use of these systems. Growing interests in collaborative immersive systems for *learning* present fresh opportunities to explore novel learning paradigms and expand existing learning taxonomies enabled through such systems. However, the excessive focus on sense-making as a process of learning, synchronous collaborations, mostly digital content, and transient work paradigms, fail to reflect the range of real-world collaborative practices observed in scientific domains. As such, future work is needed to address these shortcomings and expand our knowledge on the implication of collaborative immersive systems for the different sense-making processes in STEM, and beyond.

5.2 Influence of Sense-making in STEM on Collaborative Immersive System Design

As detailed in section 3, our analysis of the experimental design, procedures and findings of the experimental tasks used in our sample papers highlight the multifaceted nature of sense-making tasks in STEM domains. Many of the sense-making tasks observed involved multiple stages, making them resistant to simple categorization based solely on the methods/approaches/processes used, as outlined in section 2.1). However, the *primary purpose* of sense-making often stands out distinctly. For instance, the task used by Sedlák et al. [86] involved problem-solving, where student participants were required to draw conclusions about a terrain based on given contour plots. While this evokes elements of problem-solving sense-making characteristics, the end goal of the system was to support students in learning *broader concepts* related to hypsography, which were assessed through pre- and post-test scores; thereby placing this under the SENSE-MAKING PURPOSE: *Learning* category.

This categorization based on sense-making purposes provides a lens through which we scrutinize our sample for similarities, differences and patterns in the design of collaborative immersive system and related experiments. For instance, papers related to *learning* using collaborative immersive systems in STEM often incorporated elements of inquiry and exploration in their system and task design, such as discovering the effects of an embodied action on the designed MR environment, and the subsequent use of the action within the learning application [105, 121–125]. Such design considerations reflect Dervin's view of sense-making as a process of discovering and bridging 'gaps' through actions [54], thus substantiating learning [12].

Similarly, papers focused on *problem-solving/decision-making* largely designed their tasks around established problems with known decisions/solutions/actions that allow for assessments in correctness and user performance. For example, Aschenbrenner et al. [84] and Sacks et al. [96] both employed tasks where user solutions/decisions can be timed, and marked for correctness. This is not always the case in our sample, as few papers [80, 85] opted for tasks with uncertain outcomes and employed purely qualitative measures to assess the interactions, collaborations, and the problem-solving processes involved. While these tasks differ in prior availability of known solutions/decisions, they share similarities in their focus on a *specific and contained* task, which is reminiscent of sense-making for problem-solving/decision-making described by Klein et al. [5]. The sampled papers categorized as sense-making for problem-solving/decision-making were less inclined towards Snowden's [56] view, which includes idea generation for *unspecified* problems. This may indicate that unspecified problems are overlooked in current research related to collaborative immersive systems, or suggests a lack of research tools and methodologies to evaluate systems designed for uncertain and/or unspecified tasks.

Our categorization also enabled us to discern trends in the current literature. We find a predominant emphasis on sense-making for *learning*, followed by *problem-solving/decision-making*, and finally on *mental-modelling/information processing* in our sample. Across the categories found in our sample, we observed a lack of discussion around the sense-making *process*. Papers primarily discussed sense-making in relation to the experimental task used, and focused on the related outcome measures hypothesized to be affected by the collaborative immersive system (e.g., learning gains, or accuracy in problem-solving tasks). While the processes involved in sense-making were not *explicitly* discussed, we found that concepts related to the sense-making process permeated the discussion around system and experimental design, and influenced the experimental task, evaluation and analysis in the sampled papers.

Additionally, we observed differences in evaluation methods between papers focusing on the various sense-making categories. Papers focused on *problem-solving/decision-making* primarily employed methods concerning the participants' performance, such as accuracy and completion times, within the contained experimental task (for example, see works by Aschenbrenner et al. [84], Sacks et al. [96], etc.), while papers focused on *learning* focused on evaluations that assessed the participants' gain in understanding of *broader learning objectives* beyond the immediate experimental task (for example, see papers by Lin et al. [92], Barrett et al. [98], Matovu et al. [105], etc.). Lastly, while papers categorized as *mental-modelling/information processing* were sparse in our sample, we noticed a trend towards using tasks that focused on a specific context or topic, and evaluating participants' abilities in gathering and connecting information within that context, as opposed to a specific problem. For instance, [101] designed and evaluated a system where users, visualized as different organs on a projected screen, explored the relationship between different human organs, their respective biological systems, and the effects of physical activities, through individual and collective embodied action. Such trends and influences observed between sense-making categories and experimental task/evaluation may help inform future studies focused on collaborative immersive systems for sense-making in STEM fields.

However, while our focus on the *primary purpose* of sense-making offers the mentioned insights into the design and experimental patterns and trends within the current literature, it neglects the complexity and interconnections between the different sub-processes involved in sense-making tasks. These sub-processes may belong to sense-making categories other than the primary purpose of sense-making (see section 3.2), and future work is needed to better understand the interconnections between the primary purpose of sense-making, the sense-making sub-processes involved, and their influence on collaborative immersive system design and evaluation.

Irrespective of the different categories of sense-making, we observed that the sampled papers employed various means of analysing both qualitative and quantitative measures to evaluate the effects of collaborative immersive system on sense-making tasks in STEM. For both qualitative and quantitative measures, we found an emphasis on analysing the outcome measures related to the sense-making task, such as learning gains, performance measures, and change in user understanding (through qualitative coding and interpretation of data). This provides valuable insights on the benefits, or detriments, of the proposed system for sense-making in STEM. However, less attention was paid towards the influence of the collaborative immersive system on the *process* of sense-making. We deem this as an important oversight, as our categorization of sense-making suggest patterns in the choice of task, and in turn, the processes involved in approaching those tasks.

Despite this, few papers in our sample offered rich insights into the sense-making process, albeit without specifically focusing on the sense-making process. These papers, in addition to analysing measures of sense-making *outcomes*, employ techniques that were effective in distilling deeper insights into the sense-making *process*. Methods such as lag sequential analysis [126] were used in the works by Lin et al. [92] and Sarkar et al. [81], and served as way to further analyse qualitative codes to find sequential patterns in the sense-making process. As an example, Lin et al. [92] used lag sequential analysis to elaborate on the steps that users undertook in constructing the problem space, the conceptual space, and their interaction, for a physics activity involving sense-making for learning. Another method of analysis, called 'Epistemic Network Analysis' [127, 128], was used by Uz-Bilgin et al. [94] and Thompson et al. [93] to find and present the relationship between coded data that appeared in conversation between users of a collaborative immersive system. This form of analysis represents each code as a node in a network — each node representing and idea or topic of conversation observed in the use of a collaborative immersive system — and the connections between nodes represent their presence in the same 'segmentation of time' [129]. The presentation of such findings, along with the effects on sense-making outcomes, enables richer dialogue on the effects and use of collaborative immersive systems for sense-making in STEM, and possibly for other application domains.

Our findings highlight the complexity of sense-making tasks and demonstrate the influence of underlying sense-making processes on current research practices. Notably, papers placed in different categories of Sense-making purpose exhibited observable patterns in system design, choice of experimental task, and/or evaluation methods. Patterns related to system design were minimal and observed primarily within the Sense-making Purpose: *Learning* category (e.g., enabling interactions for inquiry and exploration). However, more prominent patterns were observed in the choice of experimental task, evaluation, and analysis methods between the different categories. Understanding these patterns, along with knowledge of the sense-making processes involved for a task, can aid researchers in better designing and evaluating collaborative immersive systems to support the task. Further, our analysis did not reveal trends in the results and takeaways of the empirical studies presented in our sample. This is expected, given the vast differences in the technology used, experimental tasks, and STEM domains explored in our sample. Future work could focus on narrower aspects to investigate takeaways and results related to specific sets of technology, tasks, or STEM domains for sense-making in collaborative immersive systems.

Our findings also underline the focus of current literature on sense-making outcomes, and highlight the need to employ a broader range of evaluation methods, such as 'Lag Sequential Analysis' [126] and 'Epistemic Network Analysis' [127, 128], to elicit richer insights on how collaborative immersive systems support sense-making processes, and the desired outcomes, in STEM application fields.

6 CONCLUSION

In this paper, we present a systematic review of collaborative immersive systems for sense-making in STEM. Our review focused on the nature of collaboration, enabled through immersive systems, to support sense-making in STEM, and the influence of sense-making characteristics prevalent in STEM fields on the design of collaborative immersive systems and related research. Our findings show that recent work has focused on collaborative properties that leverage technological advancements, and provide solutions to address the challenges and needs of real-world scientific collaborations. A pronounced focus on the *science* domain, particularly within the context of *learning*, was observed as a dominant trend. The preference for *synchronous* and *co-located* systems mirrors prevalent STEM collaboration practices, but also points to overlooked opportunities in the wider research landscape. The field's adaptability is evident in the post-2019 surge in *remote* collaborative systems, likely spurred by global social distancing measures. Additionally, the progression from AR to VR systems underscores the symbiotic relationship between research directions and available technologies. Despite these advancements, there is a marked emphasis on sense-making *outcomes* in current research, with little attention paid to the *processes* underlying sense-making in collaborative settings within STEM domains. This focus, while valuable, may overshadow the nuanced dynamics and complex interactions between the processes involved in different sense-making tasks and the collaborative immersive systems designed to support them. It is essential to strike a balance in our research to better understand both sense-making outcomes and processes to achieve a comprehensive understanding of STEM collaborative sense-making. The underrepresentation of engineering, mathematics, and technology domains indicates unexplored potentials. Furthermore, gaps like neglected asynchronous collaboration research and sparse integration of physical elements in AR systems pinpoint untapped innovative avenues. As the field continues to mature, ensuring a balanced exploration that blends immediate innovations with foundational principles will be pivotal.

ACKNOWLEDGEMENTS

This project was supported by an Australian Research Council Discovery Project Grant (ARC DP210103889) funded by the Australian Government. We would like to thank Jessica Walker, Cael Gallagher, and Alexandra Douglass-Bonner for their contributions during the initial search and screening of papers for this systematic review.

REFERENCES

[1] Diane H Sonnenwald. Scientific collaboration. *Annu. Rev. Inf. Sci. Technol.*, 41(1):643–681, 2007.
[2] Robert W Sharples. The peripatetic school. In *Routledge History of Philosophy Volume II*, pages 168–208. Routledge, 2003.
[3] Kara L Hall, Amanda L Vogel, Grace C Huang, Katrina J Serrano, Elise L Rice, Sophia P Tsakraklides, and Stephen M Fiore. The science of team science: A review of the empirical evidence and research gaps on collaboration in science. *American psychologist*, 73(4):532, 2018.
[4] Melissa H Cragin and Kalpana Shankar. Scientific data collections and distributed collective practice. *Computer Supported Cooperative Work (CSCW)*, 15:185–204, 2006.
[5] Gary Klein, Brian Moon, and Robert R Hoffman. Making sense of sensemaking 1: Alternative perspectives. *IEEE intelligent systems*, 21(4):70–73, 2006.
[6] David Klahr and Herbert A Simon. Studies of scientific discovery: Complementary approaches and convergent findings. *Psychological Bulletin*, 125 (5):524, 1999.
[7] Peter Pirolli and Daniel M Russell. Introduction to this special issue on sensemaking, 2011.
[8] Daniel M Russell, Mark J Stefik, Peter Pirolli, and Stuart K Card. The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 269–276, 1993.
[9] Overson Shumba and Alex Simpande. Social media use and scientific sense-making among mathematics and science student teachers at a public university. *European Journal of Education Studies*, 8(2), 2021.
[10] Karl Popper. *Realism and the aim of science: From the postscript to the logic of scientific discovery*. Routledge, 2013.
[11] John R Turner, Jeff Allen, Suliman Hawamdeh, and Gujjula Mastanamma. The multifaceted sensemaking theory: A systematic literature review and content analysis on sensemaking. *Systems*, 11(3):145, 2023.
[12] Jon Kolko. Sensemaking and framing: A theoretical reflection on perspective in design synthesis. 2010.
[13] Christine Urquhart, Louisa Mei Chun Lam, Bonnie Cheuk, and Brenda Dervin. Sense-making/sensemaking. In *Sense-Making/Sensemaking*. Oxford University Press, 2020.
[14] Deepak Kulkarni and Herbert A Simon. The processes of scientific discovery: The strategy of experimentation. *Cognitive science*, 12(2):139–175, 1988.
[15] Bruce G Buchanan and Edward A Feigenbaum. Dendral and meta-dendral: Their applications dimension. In *Readings in artificial intelligence*, pages 313–322. Elsevier, 1981.
[16] Heather Toomey Zimmerman, Susan M Land, Chrystal Maggiore, and Chris Millet. Supporting children's outdoor science learning with mobile computers: Integrating learning on-the-move strategies with context-sensitive computing. *Learning, Media and Technology*, 44(4):457–472, 2019.
[17] Heather Toomey Zimmerman and Susan M Land. Supporting children's place-based observations and explanations using collaboration scripts while learning-on-the-move outdoors. *International Journal of Computer-Supported Collaborative Learning*, 17(1):107–134, 2022.
[18] David Leslie. Does the sun rise for chatgpt? scientific discovery in the age of generative ai. *AI and Ethics*, pages 1–6, 2023.
[19] Ying Yang, Tim Dwyer, Michael Wybrow, Benjamin Lee, Maxime Cordeil, Mark Billinghurst, and Bruce H Thomas. Towards immersive collaborative sensemaking. *Proceedings of the ACM on Human-Computer Interaction*, 6(ISS):722–746, 2022.
[20] Brandon Victor Syiem, Ryan M. Kelly, Tilman Dingler, Jorge Goncalves, and Eduardo Velloso. Addressing attentional issues in augmented reality with adaptive agents: Possibilities and challenges. *International Journal of Human-Computer Studies*, page 103324, 2024. ISSN 1071-5819. doi: https://doi.org/10.1016/j.ijhcs.2024.103324. URL https://www.sciencedirect.com/science/article/pii/S1071581924001083.
[21] Andrew Irlitti, Mesut Latifoglu, Thuong Hoang, Brandon Victor Syiem, and Frank Vetere. Volumetric hybrid workspaces: Interactions with objects in remote and co-located telepresence. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2024.
[22] Jens Emil Sloth Grønbæk, Ken Pfeuffer, Eduardo Velloso, Morten Astrup, Melanie Isabel Sønderkær Pedersen, Martin Kjær, Germán Leiva, and Hans Gellersen. Partially blended realities: Aligning dissimilar spaces for distributed mr meetings. In *Proc. CHI*, 2023. ISBN 9781450394215. doi: 10.1145/3544548.3581515. URL https://doi.org/10.1145/3544548.3581515.
[23] Lee Lisle, Kylie Davidson, Edward JK Gitre, Chris North, and Doug A Bowman. Sensemaking strategies with immersive space to think. In *2021 IEEE virtual reality and 3D user interfaces (VR)*, pages 529–537. IEEE, 2021.
[24] Kylie Davidson, Lee Lisle, Kirsten Whitley, Doug A Bowman, and Chris North. Exploring the evolution of sensemaking strategies in immersive space to think. *IEEE transactions on visualization and computer graphics*, 29(12):5294–5307, 2022.

[25] Erick Martins Ratamero, Dom Bellini, Christopher G Dowson, and Rudolf A Römer. Touching proteins with virtual bare hands: Visualizing protein–drug complexes and their dynamics in self-made virtual reality using gaming hardware. *Journal of computer-aided molecular design*, 32: 703–709, 2018.

[26] Stefanie Zollmann, Christof Hoppe, Stefan Kluckner, Christian Poglitsch, Horst Bischof, and Gerhard Reitmayr. Augmented reality for construction site monitoring and documentation. *Proceedings of the IEEE*, 102(2):137–154, 2014.

[27] Alexander Zable, Lloyd Hollenberg, Eduardo Velloso, and Jorge Goncalves. Investigating immersive virtual reality as an educational tool for quantum computing. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, pages 1–11, 2020.

[28] Hannes Kaufmann, Dieter Schmalstieg, and Michael Wagner. Construct3d: a virtual reality application for mathematics and geometry education. *Education and information technologies*, 5:263–276, 2000.

[29] Alexander Schäfer, Gerd Reis, and Didier Stricker. A survey on synchronous augmented, virtual, andmixed reality remote collaboration systems. *ACM Computing Surveys*, 55(6):1–27, 2022.

[30] Ryan Anthony J de Belen, Huyen Nguyen, Daniel Filonik, Dennis Del Favero, and Tomasz Bednarz. A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering*, 3(2):181–223, 2019.

[31] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun A Lee, Thammathip Piumsomboon, and Mark Billinghurst. Revisiting collaboration through mixed reality: The evolution of groupware-supplemental visualization. 2019.

[32] Kaitlyn M Ouverson and Stephen B Gilbert. A composite framework of co-located asymmetric virtual reality. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–20, 2021.

[33] Floris van den Oever, Morten Fjeld, and Bjørn Sætrevik. A systematic literature review of augmented reality for maritime collaboration. *International Journal of Human–Computer Interaction*, pages 1–16, 2023.

[34] Nicholas MB Laskay, Jordan A George, Laquanda Knowlin, Todd P Chang, James M Johnston, and Jakub Godzik. Optimizing surgical performance using preoperative virtual reality planning: A systematic review. *World Journal of Surgery*, pages 1–11, 2023.

[35] María-Blanca Ibáñez and Carlos Delgado-Kloos. Augmented reality for stem learning: A systematic review. *Computers & Education*, 123:109–123, 2018.

[36] Anastasios Theodoropoulos and George Lepouras. Augmented Reality and programming education: A systematic review. *International Journal of Child-Computer Interaction*, 30:100335, December 2021. ISSN 2212-8689. doi: 10.1016/j.ijcci.2021.100335. URL https://www.sciencedirect.com/science/article/pii/S2212868921000544.

[37] Stylianos Mystakidis, Athanasios Christopoulos, and Nikolaos Pellas. A systematic mapping review of augmented reality applications to support STEM learning in higher education. *Education and Information Technologies*, 27(2):1883–1927, March 2022. ISSN 1573-7608. doi: 10.1007/s10639-021-10682-1. URL https://doi.org/10.1007/s10639-021-10682-1.

[38] Gloria Ajit. A Systematic Review of Augmented Reality in STEM Education. *Studies of Applied Economics*, 39(1), October 2021. ISSN 1697-5731. doi: 10.25115/eea.v39i1.4280. URL https://ojs.ual.es/ojs/index.php/eea/article/view/4280. Number: 1.

[39] Mustafa Sırakaya and Didem Alsancak Sırakaya. Augmented reality in STEM education: a systematic review. *Interactive Learning Environments*, 0 (0):1–14, February 2020. ISSN 1049-4820. doi: 10.1080/10494820.2020.1722713. URL https://doi.org/10.1080/10494820.2020.1722713. Publisher: Routledge _eprint: https://doi.org/10.1080/10494820.2020.1722713.

[40] Cynthia Passmore, Julia Svoboda Gouvea, and Ronald Giere. Models in science and in learning science: Focusing scientific practice on sense-making. In *International handbook of research in history, philosophy and science teaching*, pages 1171–1202. Springer, 2013.

[41] Noriko Hara, Paul Solomon, Seung-Lye Kim, and Diane H Sonnenwald. An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information science and Technology*, 54(10):952–965, 2003.

[42] Jutta Schickore. Scientific discovery. 2014.

[43] Anuj Karpatne, Gowtham Atluri, James H Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on knowledge and data engineering*, 29(10):2318–2331, 2017.

[44] Judea Pearl and Dana Mackenzie. *The book of why: the new science of cause and effect.* Basic books, 2018.

[45] Alessandro Antonietti and Manuela Cantoia. To see a painting versus to walk in a painting: an experiment on sense-making through virtual reality. *Computers & Education*, 34(3-4):213–223, 2000.

[46] Nitesh Goyal, Gilly Leshed, Dan Cosley, and Susan R Fussell. Effects of implicit sharing in collaborative analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 129–138, 2014.

[47] Nitesh Goyal and Susan R. Fussell. Intelligent interruption management using electro dermal activity based physiological sensor for collaborative sensemaking. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3), sep 2017. doi: 10.1145/3130917. URL https://doi.org/10.1145/3130917.

[48] Johannes Zagermann, Ulrike Pfeil, Philipp von Bauer, Daniel Fink, and Harald Reiterer. *"It's in My Other Hand!" – Studying the Interplay of Interaction Techniques and Multi-Tablet Activities*, page 1–13. Association for Computing Machinery, New York, NY, USA, 2020. ISBN 9781450367080. URL https://doi.org/10.1145/3313831.3376540.

[49] Peter Pirolli and Stuart Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, volume 5, pages 2–4. McLean, VA, USA, 2005.

[50] Floris Bex, Susan Van den Braak, Herre Van Oostendorp, Henry Prakken, Bart Verheij, and Gerard Vreeswijk. Sense-making software for crime investigation: how to combine stories and arguments? *Law, Probability & Risk*, 6(1-4):145–168, 2007.

[51] Ajit Sharma. Artificial intelligence for sense making in survival supply chains. *International Journal of Production Research*, pages 1–24, 2023.
[52] Francis Crick. *What mad pursuit: A personal view of scientific discovery.* Basic Books, 1988.
[53] Beth Warren, Cynthia Ballenger, Mark Ogonowski, Ann S Rosebery, and Josiane Hudicourt-Barnes. Rethinking diversity in learning science: The logic of everyday sense-making. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 38(5):529–552, 2001.
[54] Brenda Dervin. Sense-making theory and practice: An overview of user interests in knowledge seeking and use. *Journal of knowledge management*, 2(2):36–46, 1998.
[55] Casper Harteveld. Making sense of studying games: using sensemaking as a perspective for game research. In *Learn to game, game to learn: Proceedings of the 40th ISAGA Conference. Singapore: National University of Singapore*, 2009.
[56] Cynthia F Kurtz and David J Snowden. The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM systems journal*, 42 (3):462–483, 2003.
[57] Karl E Weick, Kathleen M Sutcliffe, and David Obstfeld. Organizing and the process of sensemaking. *Organization science*, 16(4):409–421, 2005.
[58] Andrew Irlitti, Mesut Latifoglu, Qiushi Zhou, Martin N Reinoso, Thuong Hoang, Eduardo Velloso, and Frank Vetere. Volumetric mixed reality telepresence for real-time cross modality collaboration. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. doi: 10.1145/3544548.3581277. URL https://doi.org/10.1145/3544548.3581277.
[59] Christine V Bullen and Robert Johansen. Groupware, a key to managing business teams? 1988.
[60] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12): 1321–1329, 1994.
[61] Celal Akdeniz. Instructional process and concepts in theory and practice. *Instructional Process and Concepts in Theory and Practice*, 2016.
[62] G Feryal Gündüz. Instructional techniques. *Instructional process and concepts in theory and practice: improving the teaching process*, pages 147–232, 2016.
[63] Richard E Mayer. Multimedia learning. In *Psychology of learning and motivation*, volume 41, pages 85–139. Elsevier, 2002.
[64] Aurelia Ciupe, Serban Meza, and Bogdan Orza. Systematic assessment of interactive instructional technologies in higher engineering education. In *International Conference on Interactive Collaborative Learning*, pages 797–804. Springer, 2020.
[65] Friday Joseph Agbo, Ismaila Temitayo Sanusi, Solomon Sunday Oyelere, and Jarkko Suhonen. Application of Virtual Reality in Computer Science Education: A Systemic Review Based on Bibliometric and Content Analysis Methods. *Education Sciences*, 11(3):142, March 2021. ISSN 2227-7102. doi: 10.3390/educsci11030142. URL https://www.mdpi.com/2227-7102/11/3/142. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
[66] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, and PRISMA Group*. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Annals of internal medicine*, 151(4):264–269, 2009.
[67] Mel Slater and Sylvia Wilbur. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6):603–616, 1997.
[68] Valarie L Akerson, Angela Burgess, Alex Gerber, Meize Guo, Taukir Ahmed Khan, and Steven Newman. Disentangling the meaning of stem: Implications for science education and science teacher education. *Journal of Science Teacher Education*, 29(1):1–8, 2018.
[69] David A Koonce, Jie Zhou, Cynthia D Anderson, Dyah A Hening, and Valerie Martin Conley. What is stem? In *2011 ASEE Annual Conference & Exposition*, pages 22–1684, 2011.
[70] Joni Salminen, Kathleen Guan, Soon-gyo Jung, Shammur A Chowdhury, and Bernard J Jansen. A literature review of quantitative persona creation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2020.
[71] Jesse W Grootjen, Henrike Weingärtner, and Sven Mayer. Uncovering and addressing blink-related challenges in using eye tracking for interactive systems. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–23, 2024.
[72] Laura D Cosio, Oğuz'Oz' Buruk, Daniel Fernández Galeote, Isak De Villiers Bosman, and Juho Hamari. Virtual and augmented reality for environmental sustainability: A systematic review. In *Proceedings of the 2023 chi conference on human factors in computing systems*, pages 1–23, 2023.
[73] Max Pascher, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. How to communicate robot motion intent: A scoping review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2023.
[74] Claes Wohlin. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pages 1–10, 2014.
[75] Juliet Corbin and Anselm Strauss. *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Sage publications, 2014.
[76] Charlotte P Lee and Drew Paine. From the matrix to a model of coordinated action (moca) a conceptual framework of and for cscw. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, pages 179–194, 2015.
[77] Steve Benford, Chris Greenhalgh, Gail Reynard, Chris Brown, and Boriana Koleva. Understanding and constructing shared spaces with mixed-reality boundaries. *ACM Transactions on computer-human interaction (TOCHI)*, 5(3):185–223, 1998.
[78] Sophie Maria, Helena M Mentis, Geoffroy Canlorbe, and Ignacio Avellino. Supporting collaborative discussions in surgical teleconsulting through augmented reality head mounted displays. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2023.
[79] Maria Adriana Neroni, Alfred Oti, and Nathan Crilly. Virtual reality design-build-test games with physics simulation: opportunities for researching design cognition. *International Journal of Design Creativity and Innovation*, 9(3):139–173, 2021.

[80] Dragana Nikolic, Laura Maftei, and Jennifer Whyte. Becoming familiar: How infrastructure engineers begin to use collaborative virtual reality in their interdisciplinary practice. *Journal of Information Technology in Construction*, 24:489–508, 2019.

[81] Pratiti Sarkar, Kapil Kadam, and Jayesh S Pillai. Learners' approaches, motivation and patterns of problem-solving on lines and angles in geometry using augmented reality. *Smart Learning Environments*, 7(1):1–23, 2020.

[82] Seokbin Kang, Ekta Shokeen, Virginia L Byrne, Leyla Norooz, Elizabeth Bonsignore, Caro Williams-Pierce, and Jon E Froehlich. Armath: augmenting everyday life with math learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2020.

[83] Michaela Benk, Raphael P Weibel, Stefan Feuerriegel, and Andrea Ferrario. " is it my turn?" assessing teamwork and taskwork in collaborative immersive analytics. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–23, 2022.

[84] Doris Aschenbrenner, Florian Leutert, Argun Çençen, Jouke Verlinden, Klaus Schilling, Marc Latoschik, and Stephan Lukosch. Comparing human factors for augmented reality supported single-user and collaborative repair operations of industrial robots. *Frontiers in Robotics and AI*, page 37, 2019.

[85] Danilo Gasques, Janet G Johnson, Tommy Sharkey, Yuanyuan Feng, Ru Wang, Zhuoqun Robin Xu, Enrique Zavala, Yifei Zhang, Wanze Xie, Xinming Zhang, et al. Artemis: A collaborative mixed-reality system for immersive surgical telementoring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[86] Michal Sedlák, Čeněk Šašinka, Zdeněk Stachoň, Jiří Chmelík, and Milan Doležal. Collaborative and individual learning of geography in immersive virtual reality: An effectiveness study. *Plos one*, 17(10):e0276267, 2022.

[87] Čeněk Šašinka, Zdeněk Stachoň, Michal Sedlák, Jiří Chmelík, Lukáš Herman, Petr Kubíček, Alžběta Šašinková, Milan Doležal, Hynek Tejkl, Tomáš Urbánek, et al. Collaborative immersive virtual environments for education in geography. *ISPRS International Journal of Geo-Information*, 8(1):3, 2018.

[88] Michael O'Connor, Helen M Deeks, Edward Dawn, Oussama Metatla, Anne Roudaut, Matthew Sutton, Lisa May Thomas, Becca Rose Glowacki, Rebecca Sage, Philip Tew, et al. Sampling molecular conformations and dynamics in a multiuser virtual reality framework. *Science advances*, 4(6): eaat2731, 2018.

[89] Ayesha Williams-Bhatti, David Carruthers, and Andrew Sean Wilson. Sepsis collab: A virtual reality training simulation for sepsis treatment. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 132–135. IEEE, 2023.

[90] Mary Webb, Megan Tracey, William Harwin, Ozan Tokatli, Faustina Hwang, Ros Johnson, Natasha Barrett, and Chris Jones. Haptic-enabled collaborative learning in virtual reality for schools. *Education and Information Technologies*, pages 1–24, 2022.

[91] Dogus Beyoglu, Cigdem Hursen, and Arda Nasiboglu. Use of mixed reality applications in teaching of science. *Education and Information Technologies*, 25:4271–4286, 2020.

[92] Tzung-Jin Lin, Henry Been-Lirn Duh, Nai Li, Hung-Yuan Wang, and Chin-Chung Tsai. An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system. *Computers & Education*, 68:314–321, 2013.

[93] Meredith Thompson, Cigdem Uz Bilgin, Rik Eberhart, Lucy Cho, Melat Anteneh, and Eric Klopfer. Visualizing the collaborative problem solving process in an immersive cross platform game. In *2021 7th international conference of the immersive learning research network (iLRN)*, pages 1–6. IEEE, 2021.

[94] Cigdem Uz-Bilgin, Meredith Thompson, and Melat Anteneh. Exploring how role and background influence through analysis of spatial dialogue in collaborative problem-solving games. *Journal of Science Education and Technology*, 29:813–826, 2020.

[95] Annie Wang, Meredith Thompson, Dan Roy, Katharine Pan, Judy Perry, Philip Tan, Rik Eberhart, and Eric Klopfer. Iterative user and expert feedback in the design of an educational virtual reality biology game. *Interactive Learning Environments*, 30(4):677–694, 2022.

[96] Rafael Sacks, Jennifer Whyte, Dana Swissa, Gabriel Raviv, Wei Zhou, and Aviad Shapira. Safety by design: dialogues between designers and builders using virtual reality. *Construction Management and Economics*, 33(1):55–72, 2015.

[97] Sunseng Tea, Kriengsak Panuwatwanich, Rathavoot Ruthankoon, and Manop Kaewmoracharoen. Multiuser immersive virtual reality application for real-time remote collaboration to enhance design review process in the social distancing era. *Journal of Engineering, Design and Technology*, 20 (1):281–298, 2022.

[98] Rainier Barrett, Heta A Gandhi, Anusha Naganathan, Danielle Daniels, Yang Zhang, Chibueze Onwunaka, April Luehmann, and Andrew D White. Social and tactile mixed reality increases student engagement in undergraduate lab activities. *Journal of Chemical Education*, 95(10):1755–1762, 2018.

[99] Yi Tan, Wenyu Xu, Keyu Chen, Chunyan Deng, and Peng Wang. An interactive and collaborative augmented reality environment for civil engineering education: steel reinforcement bars teaching as an example. *Engineering, Construction and Architectural Management*, 2022.

[100] Tycho T de Back, Angelica M Tinga, Phong Nguyen, and Max M Louwerse. Benefits of immersive collaborative learning in cave-based virtual reality. *International Journal of Educational Technology in Higher Education*, 17:1–18, 2020.

[101] Seokbin Kang, Leyla Norooz, Vanessa Oguamanam, Angelisa C Plane, Tamara L Clegg, and Jon E Froehlich. Sharedphys: Live physiological sensing, whole-body interaction, and large-screen visualizations to support shared inquiry experiences. In *Proceedings of the the 15th international conference on interaction design and children*, pages 275–287, 2016.

[102] Cheng-ping Chen and Chang-Hwa Wang. Employing augmented-reality-embedded instruction to disperse the imparities of individual differences in earth science learning. *Journal of Science Education and Technology*, 24:835–847, 2015.

[103] Wannisa Matcha and Dayang Rohaya Awang Rambli. Time on task for collaborative augmented reality in science experiment. *environment*, 15:16, 2016.

[104] Torvald F Ask, Kaur Kullman, Stefan Sütterlin, Benjamin J Knox, Don Engel, and Ricardo G Lugo. A 3d mixed reality visualization of network topology and activity results in better dyadic cyber team communication and cyber situational awareness. *Frontiers in big Data*, 6:1042783, 2023.

[105] Henry Matovu, Mihye Won, David Franklin Treagust, Dewi Ayu Kencana Ungu, Mauro Mocerino, Chin-Chung Tsai, and Roy Tasker. Change in students' explanation of the shape of snowflakes after collaborative immersive virtual reality. *Chemistry Education Research and Practice*, 24(2): 509–525, 2023.

[106] Barrett Ens, Sarah Goodwin, Arnaud Prouzeau, Fraser Anderson, Florence Y Wang, Samuel Gratzl, Zac Lucarelli, Brendan Moyle, Jim Smiley, and Tim Dwyer. Uplift: A tangible and immersive tabletop system for casual collaborative visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1193–1203, 2020.

[107] Iulian Radu and Bertrand Schneider. What can we learn from augmented reality (ar)? benefits and drawbacks of ar for inquiry-based learning of physics. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–12, 2019.

[108] Cindy E Hmelo-Silver, Manu Kapur, and Miki Hamstra. Learning through problem solving. In *International handbook of the learning sciences*, pages 210–220. Routledge, 2018.

[109] Qi Chen and Jianwei Zhang. Collaborative discovery learning based on computer simulation. *Collaborative learning, reasoning, and technology*, pages 127–146, 2013.

[110] Andrea Philips, Ariane Walz, Andreas Bergner, Thomas Graeff, Maik Heistermann, Sarah Kienzler, Oliver Korup, Torsten Lipp, Wolfgang Schwanghart, and Gerold Zeilinger. Immersive 3d geovisualization in higher education. *Journal of Geography in Higher Education*, 39(3):437–449, 2015.

[111] Alan M MacEachren and Menno-Jan Kraak. Exploratory cartographic visualization: advancing the agenda, 1997.

[112] Nelson R Manohar and Atul Prakash. The session capture and replay paradigm for asynchronous collaboration. In *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work ECSCW'95: 10–14 September, 1995, Stockholm, Sweden*, pages 149–164. Springer, 1995.

[113] Andrew Irlitti, Stewart Von Itzstein, Leila Alem, and Bruce Thomas. Tangible interaction techniques to support asynchronous collaboration. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 1–6. IEEE, 2013.

[114] Francis T Marchese and Natasha Brajkovska. Fostering asynchronous collaborative visualization. In *2007 11th International Conference Information Visualization (IV'07)*, pages 185–190. IEEE, 2007.

[115] Andreas Rene Fender and Christian Holz. Causality-preserving asynchronous reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2022.

[116] Tahir Mahmood, Erik Butler, Nicholas Davis, Jian Huang, and Aidong Lu. Building multiple coordinated spaces for effective immersive analytics through distributed cognition. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, pages 1–11. IEEE, 2018.

[117] Cédric Fleury, Alain Chauffaut, Thierry Duval, Valérie Gouranton, and Bruno Arnaldi. A generic model for embedding users' physical workspaces into multi-scale collaborative virtual environments. In *ICAT 2010 (20th International Conference on Artificial Reality and Telexistence)*, 2010.

[118] Adalberto L Simeone, Eduardo Velloso, and Hans Gellersen. Substitutional reality: Using the physical environment to design virtual reality experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3307–3316, 2015.

[119] Brandon Victor Syiem, Ryan M Kelly, Eduardo Velloso, Jorge Goncalves, and Tilman Dingler. Enhancing visitor experience or hindering docent roles: attentional issues in augmented reality supported installations. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 279–288. IEEE, 2020.

[120] Brandon Victor Syiem, Ryan M Kelly, Jorge Goncalves, Eduardo Velloso, and Tilman Dingler. Impact of task on attentional tunneling in handheld augmented reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–14, 2021.

[121] Xintian Tu, Joshua Danish, Megan Humburg, Mengxi Zhou, Nitasha Mathayas, Noel Enyedy, and Tessaly Jen. Understanding young children's science learning through embodied communication within an mr environment. *International Journal of Computer-Supported Collaborative Learning*, 18(2):231–258, 2023. doi: 10.1007/s11412-023-09395-z.

[122] Selena Steinberg and Joshua Danish. Balancing agency and accountability to support learning in playful, embodied science activities. In *Proceedings of the 16th International Conference of the Learning Sciences-ICLS 2022, pp. 1149-1152.* International Society of the Learning Sciences, 2022.

[123] Maggie Dahn, Noel Enyedy, and Joshua Danish. How teachers use instructional improvisation to organize science discourse and learning in a mixed reality environment. International Society of the Learning Sciences, Inc.[ISLS]., 2018.

[124] Bria Davis, Xintian Tu, Joshua A Danish, and Noel Enyedy. The impact of play, gesture, and teacher prompts on student explanations about the particulate nature of matter. *Making a Difference: Prioritizing Equity and Access in CSCL.*

[125] Danielle Keifert, Chi Xiao, Noel Enyedy, and Joshua Danish. Learners as phenomena: Expansive inquiry as students embody water particles. *Learning, Culture and Social Interaction*, 31:100572, 2021. doi: 10.1016/j.lcsi.2021.100572.

[126] Roger Bakeman and John M Gottman. *Observing interaction: An introduction to sequential analysis.* Cambridge university press, 1997.

[127] David Williamson Shaffer, Wesley Collier, and Andrew R Ruis. A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, 3(3):9–45, 2016.

[128] David Williamson Shaffer. *Quantitative ethnography.* Lulu. com, 2017.

[129] Kathleen Quardokus Fisher, Laura Hirshfield, Amanda Siebert-Evenstone, Golnaz Arastoopour Irgens, and Milo Koretsky. Network analysis of interactions between students and an instructor during design meetings. In *2016 ASEE Annual Conference & Exposition*, 2016.

[130] Hao-Chiang Koong Lin, Sheng-Hsiung Su, Sheng-Tien Wang, and Shang-Chin Tsai. Influence of cognitive style and cooperative learning on application of augmented reality to natural science learning. *International Journal of Technology and Human Interaction (IJTHI)*, 11(4):41–66, 2015.

[131] Hung-Yuan Wang, Henry Been-Lirn Duh, Nai Li, Tzung-Jin Lin, and Chin-Chung Tsai. An investigation of university students' collaborative inquiry learning behaviors in an augmented reality simulation and a traditional simulation. *Journal of Science Education and Technology*, 23: 682–691, 2014.

[132] James Planey, Robin Jephthah Rajarathinam, Emma Mercier, and Robb Lindgren. Gesture-mediated collaboration with augmented reality headsets in a problem-based astronomy task. *International Journal of Computer-Supported Collaborative Learning*, 18(2):259–289, 2023. doi: 10.1007/s11412-023-09398-w.

A APPENDIX: ADDITIONAL GRAPHS AND TABLES

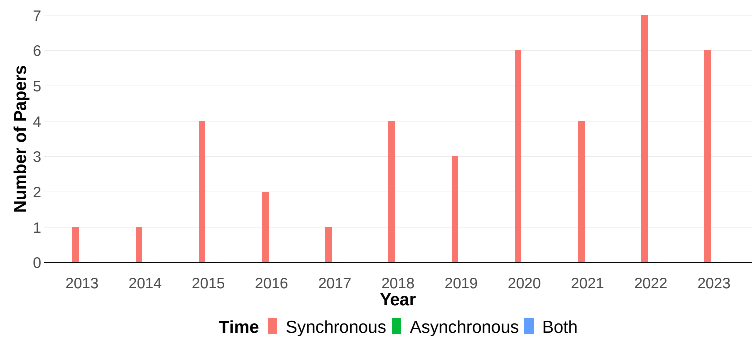

Fig. 8. Distribution of papers over the years based on the Time dimension described in section 3.2.

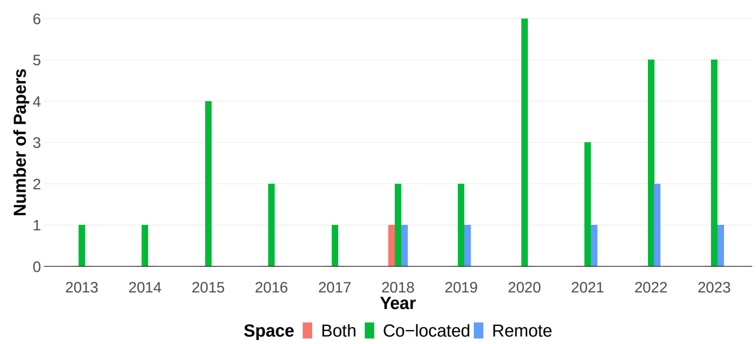

Fig. 9. Distribution of papers over the years based on the Space dimension described in section 3.2.

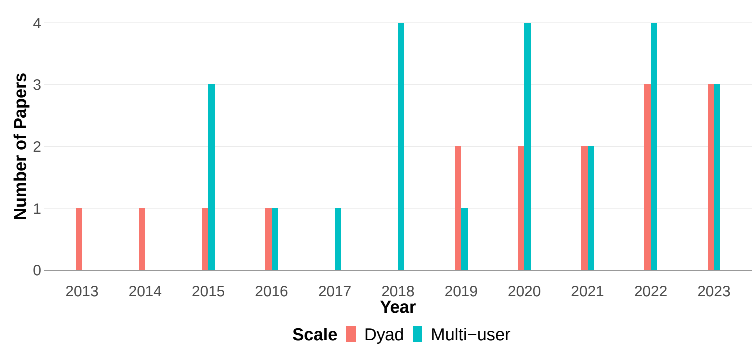

Fig. 10. Distribution of papers over the years based on the Scale dimension described in section 3.2.

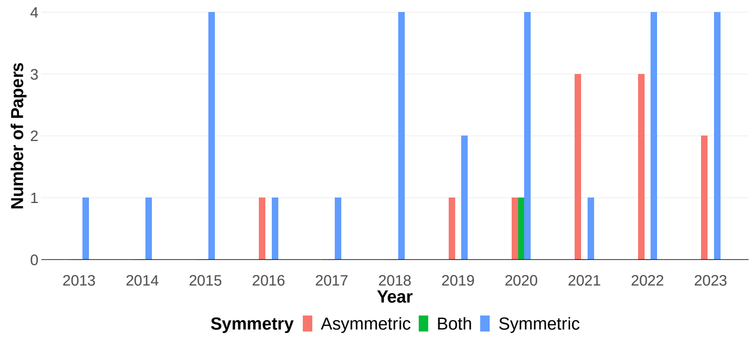

Fig. 11. Distribution of papers over the years based on the Symmetry dimension described in section 3.2.

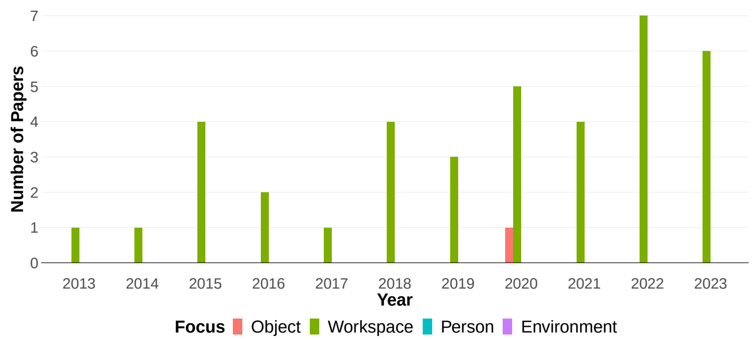

Fig. 12. Distribution of papers over the years based on the Focus dimension described in section 3.2.

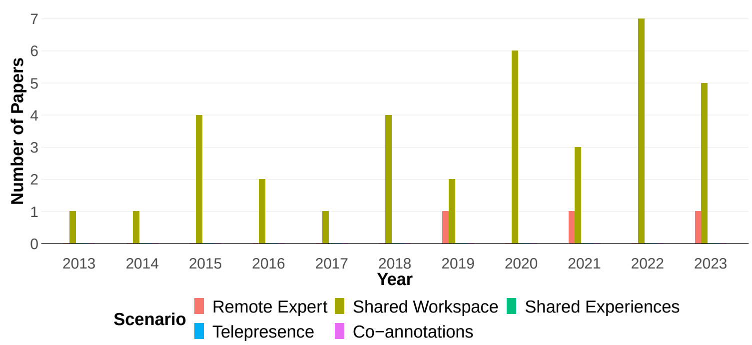

Fig. 13. Distribution of papers over the years based on the Scenario dimension described in section 3.2.

Table 5. Summary table of all the papers included in our review, along with their respective categories along the collaborative dimensions, artificiality dimension, and the sense-making purpose dimension as described in section 3.2. The table also presents the STEM field of focus, and immersive technology used/discussed in each paper. Rows with a gray background in the first column highlight papers identified by authors but did not appear in our database search.

Ref	STEM Field	Collaborative Dimensions						Immersive	Artificiality	Sense-making Purpose
		Time	Space	Symmetry	Scale	Focus	Scenario	Technology		
[84]	Engineering - Robotics	Synchronous	Remote	Asymmetric	Dyad	Workspace	Remote Expert	HMD AR Handheld AR Spatial AR	Mostly Physical	Problem-solving Decision-making
[91]	Science - Health - Planetary	Synchronous	Co-located	Both	Multi-user	Workspace	Shared Workspace	Handheld AR HMD VR	Mostly Digital	Learning
[98]	Engineering - Chemical	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	MR -Tabletop Display -Tangibles	Hybrid	Learning
[102]	Science - Earth Science	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	Screen-based AR	Hybrid	Learning
[100]	Science - Neuroanatomy	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	CAVE	Mostly Digital	Learning
[85]	Science - Medicine	Synchronous	Remote	Asymmetric	Dyad	Workspace	Remote Expert	HMD AR HMD VR	Hybrid	Problem-solving Decision-making
[101]	Science - Health - Biology	Synchronous	Co-located	Asymmetric	Multi-user	Workspace	Shared Workspace	MR - Projected Display - Physiological Sensors -Motion Capture	Hybrid	Mental-modelling Information Processing
[82]	Mathematics - Arithmetic - Geometry	Synchronous	Co-located	Symmetric	Multi-user	Object	Shared Workspace	Handheld AR	Mostly Physical	Learning
[130]	Science - Chemistry - Geology	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	Screen-based AR	Hybrid	Learning
[92]	Science - Physics	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	Handheld AR	Mostly Digital	Learning
[103]	Science - Physics	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	Screen-based AR	Hybrid	Learning
[79]	Engineering - Industrial Engineering Physics	Synchronous	Co-located	Asymmetric	Multi-user	Workspace	Shared Workspace	HMD VR Projected Display	Mostly Digital	Problem-solving Decision-making
[80]	Engineering - Construction	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	CAVE	Mostly Digital	Problem-solving Decision-making
[88]	Science - Biology	Synchronous	Both	Symmetric	Multi-user	Workspace	Shared Workspace	HMD VR	Mostly Digital	Problem-solving Decision-making
[110]	Science - Geology	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	CAVE	Mostly Digital	Mental-modelling Information Processing

Continued on next page

Table 5 – continued from previous page

Ref	STEM Field	Collaborative Dimensions						Immersive Technology	Artificiality	Sense-making Purpose
		Time	Space	Symmetry	Scale	Focus	Scenario			
[107]	Science - Physics	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	HMD AR	Hybrid	Learning
[96]	Engineering - Construction	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	CAVE	Mostly Digital	Problem-solving Decision-making
[81]	Mathematics - Geometry	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	Handheld AR	Mostly Digital	Learning
[87]	Science - Geology	Synchronous	Remote	Symmetric	Multi-user	Workspace	Shared Workspace	HMD VR	Mostly Digital	Learning
[97]	Engineering - Construction	Synchronous	Remote	Symmetric	Multi-user	Workspace	Shared Workspace	HMD VR	Mostly Digital	Problem-solving Decision-making
[93]	Science - Biology	Synchronous	Co-located	Asymmetric	Dyad	Workspace	Shared Workspace	HMD VR Handheld VR	Mostly Digital	Problem-solving Decision-making
[94]	Science - Biology	Synchronous	Co-located	Asymmetric	Dyad	Workspace	Shared Workspace	HMD VR Handheld VR	Mostly Digital	Problem-solving Decision-making
[95]	Science - Biology	Synchronous	Co-located	Asymmetric	Dyad	Workspace	Shared Workspace	HMD VR Handheld VR	Mostly Digital	Problem-solving Decision-making
[131]	Science - Physics	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	Handheld AR	Mostly Digital	Learning
[90]	Science - Biology	Synchronous	Co-located	Asymmetric	Dyad	Workspace	Shared Workspace	HMD VR	Mostly Digital	Learning
[104]	Science - Computer Science	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	HMD AR	Mostly Digital	Problem-solving Decision-making
[83]	Science - Computer Science	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	HMD AR	Mostly Digital	Problem-solving Decision-making
[106]	Science - Computer Science	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	HMD AR MR: -Tabletop Display -Tangibles	Hybrid	Problem-solving Decision-making
[78]	Science - Medicine	Synchronous	Remote	Asymmetric	Dyad	Workspace	Remote Expert	HMD AR	Mostly Physical	Problem-solving Decision-making
[105]	Science - Chemistry	Synchronous	Co-located	Symmetric	Dyad	Workspace	Shared Workspace	HMD VR	Mostly Digital	Learning
[86]	Science - Geology	Synchronous	Remote	Symmetric	Multi-user	Workspace	Shared Workspace	HMD VR	Mostly Digital	Learning
[99]	Engineering - Civil	Synchronous	Co-located	Asymmetric	Multi-user	Workspace	Shared Workspace	HMD AR	Mostly Digital	Learning
[89]	Science - Medicine	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	HMD VR	Mostly Digital	Learning
[121]	Science - Physics	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	MR -Projected Display -Motion Capture	Hybrid	Learning

Continued on next page

Table 5 – continued from previous page

Ref	STEM Field	Collaborative Dimensions						Immersive Technology	Artificiality	Sense-making Purpose
		Time	Space	Symmetry	Scale	Focus	Scenario			
[132]	Science - Astronomy	Synchronous	Co-located	Asymmetric	Multi-user	Workspace	Shared Workspace	HMD AR	Mostly Digital	Problem-solving Decision-making
[122]	Science - Ecology	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	MR -Projected Display -Motion Capture	Hybrid	Learning
[125]	Science - Physics	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	MR -Projected Display -Motion Capture	Hybrid	Learning
[124]	Science - Physics	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	MR -Projected Display -Motion Capture	Hybrid	Learning
[123]	Science - Ecology	Synchronous	Co-located	Symmetric	Multi-user	Workspace	Shared Workspace	MR -Projected Display -Motion Capture	Hybrid	Learning