

Exploring the Relationship Between Quality of Experience and Quality of Service in Collaborative Mixed Reality

John O Murray*

Department of Electrical and
Computer Engineering
University of Massachusetts, Amherst
Amherst, Massachusetts, USA
jomurray@umass.edu

Yasra Chandio*

Department of Electrical and
Computer Engineering
University of Massachusetts, Amherst
Amherst, Massachusetts, USA
ychandio@umass.edu

Michael Zink

Department of Electrical and
Computer Engineering
University of Massachusetts, Amherst
Amherst, Massachusetts, USA
zink@ecs.umass.edu

Abstract

Collaborative Mixed Reality (CollabMR) enables multiple users to interact with both the physical and virtual worlds in real time. One of the major challenges CollabMR faces is the timely synchronization of shared virtual content, yet how network Quality of Service (QoS) maps to user Quality of Experience (QoE) remains poorly understood. To address this gap, we present an exploratory, multi-layer framework linking QoS inputs (latency, bandwidth) to system responsiveness and perceptual, cognitive, and behavioral QoE. We evaluate this framework through a controlled within-subject human study with 60 participants (30 pairs) using HoloLens 2 headsets and a collaborative 3D puzzle task under four network conditions. To validate, we analyze network-level measurements, interaction logs, task performance metrics, and post-task subjective questionnaires to examine how variations in system responsiveness manifest across our conceptual model.

CCS Concepts

• **Computing methodologies** → **Mixed / augmented reality**; • **Human-centered computing** → **User studies**.

Keywords

Mixed Reality, Networking, Collaboration

ACM Reference Format:

John O Murray, Yasra Chandio, and Michael Zink. 2026. Exploring the Relationship Between Quality of Experience and Quality of Service in Collaborative Mixed Reality. In *ACM International Conference on Interactive Media Experiences (IMX '26)*, June 09–11, 2026, Athlone, Ireland. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3788851.3805019>

1 Introduction

Mixed Reality (MR) has transitioned from a niche visualization tool to a foundational technology for high-stakes remote collaboration in fields such as healthcare [38], engineering [4], education [25], and industrial training [32]. Unlike traditional video conferencing, Collaborative Mixed Reality (CollabMR) allows participants to

jointly inhabit a shared spatial environment, interacting with virtual 3D objects as if they were physical entities [17]. For instance, in a remote classroom setting, an instructor can share a virtual 3D representation of an object, showing it to students from different angles and positions, an aircraft maintenance technician can practice complex engine maintenance on a virtual jet engine [54], or surgeons can rehearse procedures on digital twins of patient organs [75]. The ultimate goal of these systems is to provide a “real-world” experience in which object updates and partner actions are reflected across all displays instantaneously, regardless of their physical locations.

Despite the potential of CollabMR, delivering a high-fidelity experience remains a significant technical challenge. The system must maintain a sustained synchronization of state, actions, and feedback over networked environments [55]. In a collaborative setting, an action performed by one user, such as removing a fan blade from a jet engine, must be realistically and consistently reflected across all participants’ displays. However, achieving this real-world immediacy is difficult. Factors such as local headset processing and network constraints, including bandwidth limitations and latency, often delay object updates. To make CollabMR truly immersive, prior studies have reported strict performance thresholds. They suggest that high-fidelity experiences require throughput between 400 to 600 Mbps with latencies as low as 5 to 20 ms [56, 71]. Other established benchmarks suggest an end-to-end 15 ms Round Trip Time (RTT) [21] as the limit for maintaining synchronization. Ultra-low latency and high bandwidth are often assumed to be prerequisites for presence [59], a key indicator of user experience in MR¹.

Even under this assumption, it remains unclear how network Quality of Service (QoS) maps to user Quality of Experience (QoE). QoS is defined by objective network metrics such as bandwidth, packet loss, delay, and jitter, which can be directly measured in networked systems [7]. In contrast, QoE is inherently subjective and is typically quantified through user assessments. While prior research has concentrated on QoE for video streaming [3, 20, 31, 46, 48, 62, 68, 79] or general immersive realities [33, 39, 45, 57, 72], no QoE models have been established to characterize the relationship between QoE and QoS specifically in the context of CollabMR. To explore these complexities, we propose a multi-layered conceptual framework that explores the gap between these objective network metrics and subjective user experiences, requiring a structured approach to understanding how technical limitations manifest as experiential disruptions.

*Both authors contributed equally to this research.



This work is licensed under a Creative Commons Attribution 4.0 International License. *IMX '26, Athlone, Ireland*

© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2448-0/26/06
<https://doi.org/10.1145/3788851.3805019>

¹Presence reflects the subjective feeling of being in a virtual environment, either individually or with others [59].

Our framework begins at *QoS Inputs*, where latency and bandwidth are varied independently to define the boundaries of content transmission. These inputs directly shape *System Responsiveness*, which covers the system’s observable real-time behavior, such as synchronization lag and update consistency. The user experience then further divides across three layers: *Perceptual QoE* that captures presence and co-presence; *Cognitive QoE* measures the mental effort required to compensate for system inconsistencies; and *Behavioral QoE* validates how these factors translate into functional outcomes like completion time and coordination errors. Tracing this pathway identifies how system constraints accumulate to impair the collaborative experience.

To evaluate our exploratory model, we conducted a user study ($N = 60$) involving 30 pairs of participants using Microsoft HoloLens 2 headsets [43] connected via a controlled network. We designed a collaborative 3D puzzle application that requires participants to solve problems using shared virtual objects. Each group performed the task repeatedly under four different network conditions while the system logged objective performance metrics and network measurements. Following each task, participants completed subjective questionnaires to provide feedback on their experience. This design allows us to analyze the transitions from technical network constraints to system responsiveness, and finally to the resulting user experience. In our study, we ask the following research questions:

RQ1: How do network parameters impact user experience in Collaborative Mixed Reality?

RQ2: To what extent do degraded QoS parameters increase the cognitive effort and mental demand required for users to maintain synchronization and monitor shared object states?

RQ3: To what extent do degraded QoS parameters lead to measurable impairments in collaborative task performance?

2 Background and Related Work

2.1 Collaborative Mixed Reality Systems

CollabMR systems enable multiple users to simultaneously view, manipulate, and interact with shared virtual objects that are spatially registered within a common physical reference frame [6]. Unlike single-user XR, which primarily maintains consistency between a user’s tracked pose and a locally rendered scene [63, 65], CollabMR must accommodate spatial and temporal consistency of shared objects across multiple devices.

CollabMR also differs from XR conferencing systems, which are typically optimized for stable audiovisual, or avatar-based communication [28, 77]. In contrast, CollabMR prioritizes low-latency synchronization of user-manipulable virtual objects within a unified spatial frame. This distinction introduces stricter requirements on update propagation, spatial alignment, and consistency, particularly for co-located users or remote collaborators interacting with the same objects. As a result, CollabMR systems pose unique system challenges spanning shared state management, spatial alignment, and real-time networking. Typical architectures consist of head-mounted displays running custom client applications, a networking layer for propagating object state updates, and a server responsible for coordinating users and maintaining session-level object state.

2.2 Networked State Synchronization in MR

CollabMR systems require maintaining a consistent shared virtual state under strict network delay. From a distributed systems perspective, they resemble the continuously updated replicated state systems, where object creation, movement, and deletion events must be propagated across clients within a bounded delay [26]. Unlike traditional distributed applications, MR systems operate on continuous spatial data, requiring frequent transmission of object transforms.

To support this, our CollabMR system utilizes an asynchronous networking mechanism optimized for low-latency state propagation. Perfect synchrony is infeasible due to latency and packet variability; instead, systems aim for sufficient temporal consistency to preserve coherent interaction [18, 19]. These constraints are especially pronounced on head-mounted devices such as the Microsoft HoloLens 2, where rendering, interaction handling, and network transmission must occur within a tight per-frame budget. As a result, our CollabMR operates under a soft real-time synchronization model, tolerating transient inconsistencies while prioritizing timely update delivery [5, 67]. These synchronization requirements motivate the QoS evaluated in this work.

2.3 Quality of Service in CollabMR

QoS in CollabMR refers to objective, system-level properties of the networking and synchronization infrastructure. Common QoS metrics include end-to-end latency, available bandwidth, packet jitter, and packet loss [47]. Among these, latency and bandwidth are most frequently controlled and reported in CollabMR systems, as they directly constrain the timeliness and fidelity of shared object updates. While prior work has reported acceptable QoS bounds for collaborative virtual environments [27, 50], these thresholds are highly application- and hardware-dependent. Rather than assuming universal limits, our work treats QoS metrics as controllable variables whose effects must be evaluated relative to the demands of continuous spatial interaction in modern MR systems.

2.4 Quality of Experience in CollabMR

QoE captures users’ subjective perception of system performance, interaction quality, and overall satisfaction [8, 73]. QoE has been widely studied in multimedia systems, immersive media, and single-user VR, where factors such as latency, frame rate, and visual fidelity influence immersion and comfort [9, 73]. However, existing QoE models primarily focus on individual or passive experiences and do not explicitly address shared object interaction or collaborative dynamics. In CollabMR, users must coordinate actions and adapt to the behavior of others, introducing additional cognitive and behavioral dimensions [2, 27].

2.4.1 Limitations of Existing QoS–QoE Mappings for CollabMR.

Existing QoS–QoE mappings typically assume single-user feedback loops and passive content consumption [8, 73]. These assumptions do not hold in CollabMR, where interaction is interdependent and delays compound across local and networked processing stages. Delayed or inconsistent object updates can affect coordination strategies and task execution even when subjective experience

remains stable [35, 44, 50]. Understanding CollabMR, therefore, requires examining how system-level constraints propagate through responsiveness, perception, cognitive effort, and collaborative performance, motivating the structured framework and exploratory study presented in our work.

3 Design

CollabMR allows distributed users to jointly manipulate virtual content anchored in shared physical environments. However, delivering this experience requires synchronizing state, actions, and feedback across networked systems. This section presents the theoretical motivation and conceptual model that guide our investigation into how system-level constraints shape collaboration in MR. Instead of assuming a fixed or causal mapping between system metrics and experience, our design is intended to systematically understand how network constraints are experienced by users [52] in CollabMR. We begin by identifying the key constructs and justifying their inclusion using relevant theories. We then describe a layered conceptual model that maps potential relationships among these constructs. This framework allows us to generate testable hypotheses and select appropriate metrics across system, perceptual, and task layers.

3.1 Theoretical Foundations

We focus on three foundational constructs: (1) *presence*, (2) *cognitive load*, and (3) *task performance* because they collectively represent the progression from system behavior to user experience and observable outcome in CollabMR [14]. These constructs are not arbitrary: each captures a key breakdown point in the collaboration pipeline. Together, they operationalize QoE as it emerges from underlying system-level constraints, allowing us to examine how changes in QoS manifest at multiple stages of the collaborative experience.

3.1.1 Presence: Capturing Perceptual Quality. The subjective sense of “being there” is fundamental to MR [59]. In collaborative settings, this extends to co-presence: the sense that another user is spatially and temporally available. These states are sustained by timely, consistent feedback in response to user actions. When the system fails to maintain responsive behavior, the continuity of interaction is disrupted, weakening users’ sense of immersion and shared spatial awareness [52]. This additional mental effort is a known cost that predicts drop-offs in user engagement and coordination quality [16]. Prior work shows that inconsistencies (mismatched object states [10], delayed reactions [12]) reduce immersion, break presence [13], and degrade user experience [76]. Our framework includes presence as a necessary perceptual component impacted by responsiveness. Spatial and temporal lags disrupt user engagement and impair social perception of others in shared space [64]. Without the perception of being immersed in a shared space, users struggle to coordinate actions [49], maintain spatial awareness [78], and interpret partner behavior [61]. Thus, presence serves as a proxy for QoE and an early, sensitive indicator of QoS degradation. However, presence alone cannot explain how users compensate for or adapt to degraded responsiveness, motivating the inclusion of cognitive and behavioral components in our framework.

3.1.2 Cognitive Load: Capturing User Effort. When the system is less responsive due to network-induced delays or inconsistent update delivery (delayed object updates or intermittent feedback), users must invest additional mental effort to interpret outcomes or repeat actions [23]. In the context of MR, when the system is inconsistent or unresponsive, users must actively compensate by mentally tracking object states, confirming collaborators’ actions, or waiting for delayed feedback. This might involve rechecking the object state, waiting for confirmation, or adjusting actions to account for delay. This redirection of effort for compensatory behaviors increases cognitive load, which can degrade attention, reduce engagement, and lead to poor coordination [24]. Including cognitive load allows us to test whether perceptual and performance effects of degraded QoS are direct or effort-driven. Cognitive Load Theory [69] posits that human working memory and attention are limited resources. When these are redirected to manage system inconsistencies, users have fewer resources available for core task execution or collaborative interaction. Prior work in interactive systems confirms that system instability increases mental demand, even when overall task complexity remains unchanged [36].

3.1.3 Task Performance: Capturing Functional Outcomes. Regardless of subjective experience, collaborative systems must enable users to complete tasks efficiently and accurately. Task performance represents the concrete output of the system under different conditions. It provides final-layer validation that presence and load effects translate into actual impairment. Prior work suggests that task efficiency and accuracy decline as cognitive demand increases [51]. The inability to maintain a shared state or anticipate partner actions in collaborative contexts introduces delays and coordination failures [14]. Ultimately, the system fails in its functional objective if users cannot complete a collaborative task efficiently and accurately. Errors, delays, and misalignments in object state reflect concrete consequences of degraded system experience. Including task performance allows us to test whether upstream perceptual and cognitive effects yield observable behavioral outcomes. In this sense, task performance can serve as a validation layer, revealing whether experiential degradation caused by system unresponsiveness results in measurable collaborative breakdowns.

3.1.4 Bringing QoE Together. In combination, presence, cognitive load, and task performance form a layered and observable expression of QoE in CollabMR. Unlike traditional applications, where QoE may be captured by a single satisfaction rating or engagement score, we present a more structured view of what the CollabMR experience demands. Here, QoE emerges through:

- **Perceptual fidelity**, reflected in presence and co-presence;
- **Cognitive cost**, reflected in the mental effort users must expend to coordinate in degraded conditions; and
- **Functional outcomes**, reflected in how effectively users complete collaborative tasks.

These components are theoretically distinct but operationally linked. For instance, degraded responsiveness may directly reduce presence or increase cognitive load, which in turn reduces performance and user satisfaction. We include all three constructs to capture (1) *whether QoE degrades under constrained conditions and* (2) *how and where that degradation occurs*. This multi-dimensional

view of QoE allows us to evaluate different aspects of user experience and test multiple explanatory pathways. It also aligns with recent calls in immersive systems research to move beyond isolated user experience metrics and consider the full experiential and behavioral arc [11, 49].

3.2 Conceptual Model

In the previous section, we established our theoretical foundation for understanding user experience in CollabMR.

Here, we connect these experiential constructs to system-level QoS constraints. These constraints shape how quickly and consistently shared object states and interactions are propagated across collaborative users [53], thereby influencing the temporal and spatial coherence of collaboration. As upstream system inputs, QoS parameters can introduce timing mismatches, reduce state fidelity, and limit responsiveness, creating conditions under which coordination and shared perception become unstable [21]. While many network metrics and resource functions may affect collaborative systems (jitter, packet loss, delay, and bandwidth) [74], we focus in this initial study on *latency* and *bandwidth* as primary system-level constraints. This focus is motivated by their central role in shaping interaction dynamics: delayed action feedback can disrupt the co-perception of object manipulations [34] and impair temporal coordination [30], while insufficient bandwidth can reduce update frequency or state resolution, harming spatial alignment [40] and mutual awareness [58]. Together, these parameters define the operational conditions under which shared interaction unfolds and bound the range of system behaviors experienced by users. They also represent a fundamental tradeoff space in real-world MR deployments, where system designers must balance update rate versus precision, responsiveness versus visual fidelity, and consistency versus frame loss under practical network constraints.

In this section, we connect these experiences to QoS constraints. These system-level inputs affect how quickly and consistently shared object states and interactions are reflected across users [53]. Thus, they are upstream levers that can introduce timing mismatches, lower visual or state fidelity, and limit responsiveness.

3.2.1 Why Latency and Bandwidth? Latency and bandwidth are the two QoS parameters [1, 37] we manipulate because they are both theoretically grounded and practically impactful in networked MR systems. Latency governs how long it takes for one user’s action to be reflected in the partner’s view, directly shaping action-to-feedback delay and temporal coupling. These parameters are (1) *independent*, unlike jitter or packet loss, which are harder to control and may introduce stochastic randomness or noise [15]. Latency and bandwidth can be systematically varied and modeled under controlled conditions. (2) *Sufficient to affect distinct dimensions of responsiveness*, since latency primarily affects temporal immediacy, while bandwidth affects update continuity and representational fidelity. (3) *Representative of real-world deployment constraints*, as most MR systems must balance frame rate, quality, and interactivity within limited bandwidth or in high-latency mobile, wireless, or cloud-edge deployments [70].

3.2.2 Overview of the model. To structure this relationship between system constraints and user experience, our model consists of five

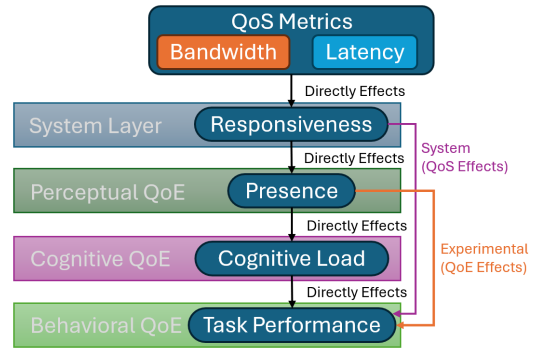


Figure 1: Conceptual model of QoS-QoE relationships in CollabMR. Arrows indicate hypothesized directional relationships.

interconnected layers, each representing a distinct but related class of variables:

Layer 1: QoS Inputs. The network layer, where latency and bandwidth are varied independently. These controllable system-level parameters bound *how quickly and how completely shared content is transmitted*.

Layer 2: System Responsiveness. This layer reflects the system’s observable real-time behavior, including *action-to-feedback delay, object synchronization lag, and update rate consistency*. These metrics are shaped by QoS inputs.

Layer 3: Perceptual QoE. This layer captures users’ subjective experience of the system’s responsiveness, including *presence, and co-presence*, which shape immersion and mutual awareness.

Layer 4: Cognitive QoE. This layer represents the cognitive effort required to maintain coordination under degraded responsiveness, operationalized through *perceived mental demand and monitoring overhead*.

Layer 5: Behavioral QoE (Performance). The final layer reflects how system and experiential conditions manifest in collaborative task outcomes, including *completion time, coordination errors, and execution smoothness*.

3.2.3 System-to-Experience Pathways. Figure 1 provides an overview of the model’s structure and flow. At the first level, network QoS inputs (latency and bandwidth) shape the system’s responsiveness by determining how quickly and reliably user actions are reflected across devices. Higher latency increases action-to-feedback delay and shared object update lag. Limited bandwidth constrains the frequency and fidelity of transmitted updates, particularly during periods of interaction bursts or sustained state change. Together, these factors limit the system’s ability to maintain timely and continuous synchronization of shared virtual content.

System responsiveness, in turn, influences multiple experiential constructs. It affects users’ sense of presence by shaping the continuity and reliability of interaction feedback: as responses become delayed or inconsistent, users are less likely to perceive themselves as immersed in a coherent shared environment or as co-located with their collaborator. Responsiveness also introduces cognitive demands. When the system fails to behave as expected, users must monitor, adapt, and compensate for uncertainty in system state, increasing mental effort and coordination overhead.

Finally, both perceptual disruption (presence) and increased cognitive load feed into observable task performance. Reduced presence may lead to disengagement or confusion, while increased cognitive effort can slow task execution and increase coordination errors. Responsiveness may also affect performance directly by delaying the propagation of critical actions between users, leading to inefficiencies and misaligned operations. As such, *task performance functions as an integrative outcome that captures the compounded effects of degraded responsiveness, impaired perception, and elevated cognitive effort.*

3.2.4 Mapping Hypotheses to the Model. Our hypotheses are derived from the theoretical pathways embedded in our conceptual model. Each hypothesis set targets a specific dimension of user experience or behavior and traces its origin back to the manipulated QoS conditions. To this end, we test the following hypotheses. For clarity, each hypothesis is labeled as *Hypothesis Number (Conceptual Pathway): Hypothesis Statement (Corresponding Research Question)*, where the conceptual pathway indicates the direction of influence examined in our model.

- **H1:** (*QoS Inputs* → *Responsiveness* → *Perceptual QoE*): Variations in latency and bandwidth impact perceptual experience (presence) in CollabMR (**RQ1**).
- **H2:** (*QoS Inputs* → *Responsiveness* → *Cognitive QoE*): Increased latency and reduced bandwidth lead to higher cognitive load (**RQ2**).
- **H3:** (*Responsiveness* → *Cognitive QoE* → *Behavioral QoE*): Increased cognitive effort negatively impacts participants' collaborative task performance, reflected in longer task completion times and increased coordination errors (**RQ3**).

In this formulation, task outcomes are a downstream result of both system constraints and the user's perceptual and cognitive responses. Together, these hypotheses allow us to test whether performance or experience metrics degrade under poorer QoS.

4 User Study

In this section, we present the details of our controlled human subject study to understand how different network conditions impact user experience in CollabMR. Our study aims to observe collaborative performance under varying network constraints, focusing on how network latency and bandwidth limitations influence user perception (presence and cognitive load) and task performance reflected in interaction patterns.

To this end, we evaluate four network conditions: **low latency (LL)**, where no artificial delay or bandwidth constraints are imposed; **high latency (HL)**, where end-to-end network delay is increased to simulate delayed action feedback; **high bandwidth (HB)**, which allows near-unrestricted data transmission; and **low bandwidth (LB)**, where available bandwidth is constrained to limit update frequency and fidelity. These conditions allow us to isolate and systematically study how latency and bandwidth affect collaboration in CollabMR.

4.1 Participants

We conducted an a priori power analysis for a within-subject study with four conditions (**LL**, **HL**, **LB**, **HB**). Assuming a medium effect size ($f = 0.25$), four repeated measures, $\alpha = 0.05$, desired power

of 0.95, and a nonsphericity correction of $\epsilon = 1$, the analysis recommended a minimum sample size of 48 participants (24 pairs). We recruited 60 participants, forming 30 collaborating pairs, where each pair interacted together in a shared MR environment during the collaborative task. Participants had a mean age of 26.37 years (SD = 10.22) and included 43 male-identifying, 16 female-identifying, and 1 non-binary-identifying individuals. All participants had normal or corrected-to-normal vision and provided informed consent. Demographic information and prior familiarity with immersive technologies are summarized in Table 1. The study protocol was reviewed and approved by our Institutional Review Board.

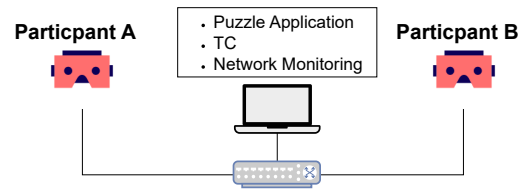


Figure 2: CollabMR Experimental Setup and Network Configuration.

4.2 Apparatus and Network Setup

We used two HoloLens 2 [43] head-mounted displays and a local server to coordinate shared state across participants, as shown in Figure 2. The HoloLens 2 provides an untethered MR experience with inside-out tracking and gesture recognition. We developed a custom collaborative puzzle application in Unreal Engine (using the Microsoft Mixed Reality Toolkit) to render 3D virtual puzzle pieces and enable multi-user interactions. Each headset rendered the shared virtual content and supported direct manipulation of virtual objects using native hand-tracking interactions. Participants were invited to a shared lab room with a designated 10ft x5ft space. They were informed that they could move freely within this area during the task.

The server hosted the collaboration logic, including object state synchronization, task progression, and event logging. We used a dedicated local server (a high-performance PC) to handle multi-user state synchronization via a standard Ethernet switch, allowing us to avoid variability introduced by wireless interference and background traffic. All HoloLens devices were connected to this server over a controlled Gigabit Ethernet network, ensuring a stable baseline with minimal uncontrolled latency or packet loss. To emulate network constraints, we introduced artificial latency and bandwidth limits using Linux Traffic Control (tc) [41]. The server applied `tc qdisc netem` rules [42] to each HoloLens client's traffic to impose precise one-way network delays and throughput caps. This setup allowed us to systematically control networking conditions in software.

For each experimental trial, the desired latency and bandwidth setting were configured on the server before the task began. We implemented four network profiles across our four experimental conditions as listed in Table 2. **LL** (no limitation), **HL** (50-70 ms RTT delay), crossed with **HB** (unthrottled, up to 867 Mbps) and **LB** (throttled, 10 Mbps). These values were chosen to represent a

Table 1: Demographic and media usage results across all participants. Key for frequency, lowest to highest: never/almost never; rarely (< 2 times); occasionally (a few times); frequently in the past; frequently (> 2 times/month), with used to representing those who have had experience before, but not lately.

Value	Results
Gender	1 non-binary; 16 female; 43 male
Age	mean = 26.37; standard deviation = 10.22
Prior familiarity with participant in pair	yes=33;no=27
Familiarity with Augmented Reality	never=23; rarely=15;occasionally=20; frequently=1;used to=1
Familiarity with Virtual Reality	never=8; rarely=21;occasionally=28; frequently=2;used to=1
Familiarity with Video Games	never=6; rarely=10;occasionally=12; frequently=22;used to=10

best-case local network scenario versus a degraded network scenario akin to congested or long-distance connections [22]. The HoloLens application exchanged only state-update messages (object transformations and event triggers) so that the chosen bandwidth limits, including the LB condition, remained above the minimum required for basic functionality. This design choice allowed us to preserve experimental control by aligning conversational timing with system-mediated interaction timing. This setup allowed us to systematically vary latency and bandwidth while holding all other system factors constant.

Table 2: Bandwidth and Latency Conditions

Condition	Bandwidth	Latency
Low Latency (LL)	No Limit	No Limit
High Bandwidth (HB)	867 Mbps	No Limit
Low Bandwidth (LB)	10 Mbps	No Limit
High Latency (HL)	No Limit	50–70 ms RTT

4.3 Experimental Task

We designed a *shared 3D puzzle assembly task* that requires tight coordination and continuous awareness of a partner’s actions. We chose this task to reflect a common CollabMR scenario in which users must manipulate shared virtual objects anchored in physical space while maintaining a consistent understanding of object state and partner intent.

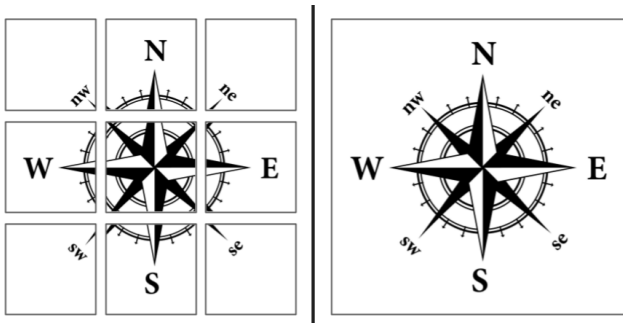


Figure 3: The nine individual pieces of the puzzle (separated), shown on the left. The completed puzzle, shown on the right.

4.3.1 Primary Task. The primary task involved collaboratively assembling a 3×3 virtual puzzle composed of nine distinct pieces into a fixed configuration (as shown in Figure 3). Both participants were in the same room when performing the task and, as such, were in each other’s direct lines of sight. At the beginning of each trial, a single anchor piece was placed at a known reference location (as shown in Figure 4a). Subsequent pieces were spawned one at a time and had to be jointly positioned and oriented correctly by the two participants (as shown in Figure 4b). Only one movable piece was active at any given time, requiring participants to coordinate turns, communicate intent, and monitor object state updates. A piece snapped into place automatically when it was moved within a predefined spatial and angular tolerance of its target position, providing immediate visual confirmation. Once placed, the next piece was spawned, and the process continued until the puzzle was complete (as shown in Figure 4c). Each puzzle assembly constituted one trial.

4.3.2 Virtual Scene and Cues. The virtual puzzle was rendered in a shared MR scene anchored to the physical environment using a printed QR code as a spatial reference. This reference ensured consistent alignment of virtual content across both headsets and allowed participants to reason about object positions relative to a common frame of reference. The virtual environment itself was intentionally minimal to avoid visual distractions and to focus attention on the shared task. Visual cues were limited to the puzzle pieces themselves and their snapping behavior upon correct placement. No additional instructional prompts were shown during the task, allowing collaboration strategies to emerge naturally.

4.3.3 Interaction and Feedback. Participants interacted with puzzle pieces using the HoloLens 2 hand-tracking system, supporting both near-field direct grabs and far-field ray-based interactions. These interaction modalities reflect common MR manipulation techniques and allowed participants flexibility in how they approached the task. Feedback was primarily visual and action-based. When a piece was grabbed, moved, or released, the corresponding state update was propagated to the partner’s view. Correct placement triggered an immediate snap-to-position animation, serving as confirmation for both users.

4.4 Measures

We collected a combination of objective system-level metrics, interaction-level performance measures, and subjective user-reported assessments. Interaction logs captured timestamps for object grabs, movements, releases, spawn events, and snap completions, along with the



Figure 4: Visual stages of the collaborative puzzle assembly task. (a) initial placement of the anchor puzzle piece at the shared reference location, (b) active manipulation and alignment of subsequent pieces, and (c) a nearly completed puzzle assembly.

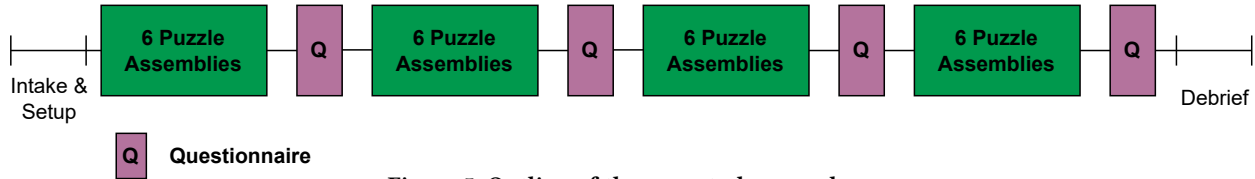


Figure 5: Outline of the user study procedure.

interaction type used. These logs were used to compute task completion times, coordination delays, and error-related behaviors. Network-level measurements were collected at the server, including packet timing and throughput, allowing us to associate interaction events with underlying network behavior. This enabled analysis of how imposed latency and bandwidth constraints manifested at the system level.

Subjective experience was assessed using standardized questionnaires administered after each block of trials under a given network condition. Presence with the IPQ [60] and PQ [66] questionnaires, cognitive load through NASA-Task Load Index (NASA-TLX) [29]. The PQ evaluates factors such as the possibility to act and examine, realism, self-evaluation, and interface quality. The IPQ measures spatial and general presence, realism, and involvement. The presence scores are derived from 33 items (14 IPQ and 19 PQ), the cognitive load score is derived from 5 items from NASA-TLX, and group behavior from 5 items from our custom-designed survey on a 7-point scale. We also deployed a custom questionnaire [14] to measure participants' perspectives on how their group interacted. The group behavior characterization questionnaire assesses key aspects, including contribution awareness, shared attention, proximity impact, conversational support, and overall group collaboration. We use this custom group characterization questionnaire because of our high trial repetition count to reduce response burden.

4.5 Pilot Study

Prior to the main study, we conducted a pilot study to validate task parameters and ensure feasibility within the allotted session time. We evaluated different spatial and angular thresholds for puzzle piece snapping and determined that a tolerance of 5 cm in position and 15 degrees in orientation provided clear feedback without trivializing the task. We also explored different numbers of puzzle repetitions per condition. Configurations with more than six repetitions led to excessive session duration, increased headset battery drain, and participant fatigue. Based on these observations, we selected six puzzle assemblies per condition to balance data richness and participant comfort.

4.6 Study Procedure

The study procedure is shown in Figure 5. Each study session began with informed consent, a demographic questionnaire, and a brief introduction to the HoloLens 2 interaction techniques. Participants were informed that they would collaboratively assemble multiple puzzles under varying conditions, but were not told the specific focus on network behavior to reduce expectancy effects. For each experimental condition, participants completed six consecutive puzzle assemblies. The order of conditions was counterbalanced across pairs using a Latin square design to mitigate order effects. Before each condition, the server applied the corresponding network constraints and initialized logging.

During each trial, participants collaboratively assembled the puzzle until completion, after which the next trial began automatically. Upon completing all six trials for a condition, participants removed their headsets and independently filled out a post-task questionnaire on cognitive load, presence, and perceived collaboration. This cycle repeated for all four conditions. The time each group took to complete the task varied; however, the total session duration, including consent, briefing, training, calibration, the task, and surveys, lasted under an hour.

5 Experiment Analysis

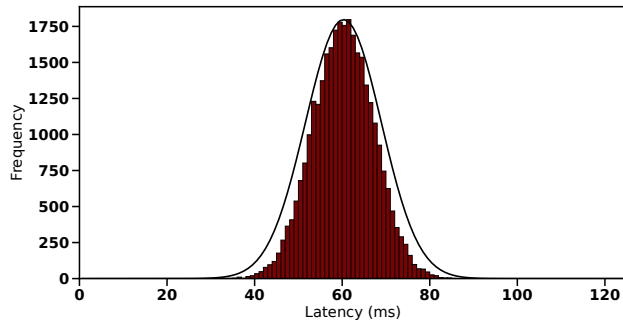
We report results by progressively moving through the layers of our conceptual model (§3.2), from network-level behavior to system responsiveness, user experience, and collaborative performance. Unless otherwise noted, all results are reported across all 30 participant pairs and four experimental conditions (LL, HL, LB, HB).

5.1 Network Conditions and Responsiveness Validation

We first verified that the imposed network conditions produced the intended QoS differences, corresponding to **Layer 1 (QoS Inputs)** and their direct manifestation in **Layer 2 (Responsiveness)** in our conceptual model (§3.2.2). LL, LB, and HB conditions exhibited negligible packet latency (all < 1 ms on average), while the HL

Table 3: Network performance metrics across experimental conditions.

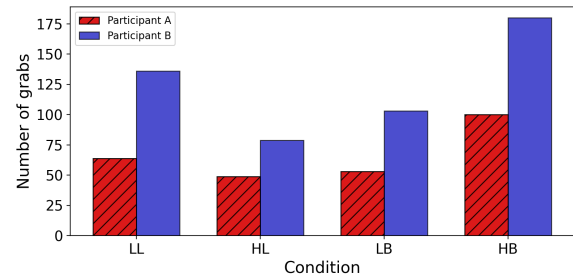
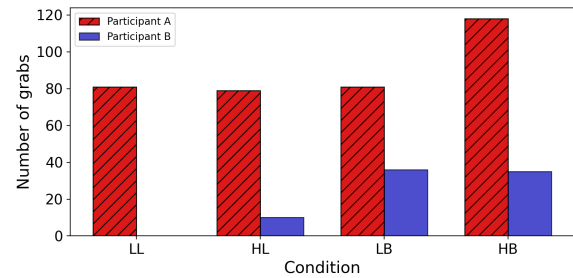
Measure	LL	HL	LB	HB
Average packet latency (ms)	0.2321	60.2542	0.2817	0.2866
Average bandwidth utilization (Mbps)	0.0591	0.0626	0.0741	0.0556
Packet loss (%)	0.1414	1.2133	0.0321	0.1531
Average jitter (ms)	0.0670	6.9654	0.0812	0.0796
Effective object jitter (ms)	16.9360	88.3110	14.8340	16.1966

**Figure 6: Histogram of packet latency measurements collected under the HL condition across all experimental traces.**

condition showed a substantially increased mean latency of approximately 60 ms, consistent with the configured delay range. Figure 6 further confirms that latency in the **HL** condition followed a stable latency distribution consistent with the imposed `tc netem` configuration, centered around the intended delay range ($\sim 50 - 70$ ms RTT), with a mean near 60 ms.

Bandwidth utilization remained low across all conditions, consistent with the application’s communication profile, which primarily involved lightweight state-update messages rather than continuous high-rate data streams. As a result, the reduced throughput in the **LB** condition did not induce packet loss or congestion. In this setting, bandwidth throttling constrained available capacity without disrupting the transmission of interaction-critical updates.

Average network jitter increased in the **HL** condition, as expected due to delayed packet delivery. We manipulated latency through an injected delay that resulted in packets having RTTs within a 50-70 ms range following a normal distribution to simulate typical real-world latency, which generates jitter from the varied inter-packet latencies. However, raw packet-level jitter alone does not directly reflect users’ perceptual experience of object motion. To better capture perceptually relevant instability, we computed an application-level metric of *effective object jitter* based on the number of skipped movement updates. State-update packets were indexed, and packets arriving out of order (with indices lower than the most recently accepted update) were discarded to prevent visible oscillations from back-and-forth state corrections, and we mark these as lost packets. While this mechanism preserves visual stability under mild delay, sustained packet reordering can cause multiple consecutive updates to be skipped, leading to abrupt state corrections in object position, which we refer to as object snapping. This effect is captured indirectly through elevated effective object jitter values (Table 3), which exceed multiple display frames in the **HL** condition. All conditions except **HL** exhibited effective object jitter

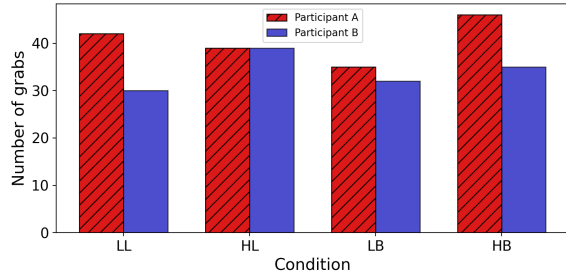
**(a) Participant B contributes more than participant A.****(b) Participant A contributing more than participant B.****Figure 7: Example groups with highly asymmetric interaction effort during puzzle assembly, measured by grab activity.**

on the order of a single display frame, consistent with expected rendering behavior. In contrast, the **HL** condition produced effective object jitter exceeding five times the display update interval of the HoloLens 2, indicating substantially reduced temporal consistency in object motion despite stable packet delivery.

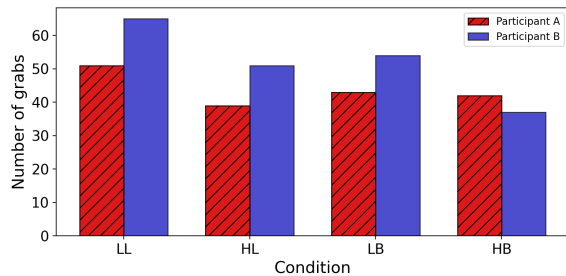
5.2 Interaction Patterns and Coordination Strategies

Before aggregating results across conditions, we examined interaction behavior at the group level to understand how participants coordinated during the collaborative task and to **contextualize later condition-level analyses**. Analysis of headset logs revealed substantial variability in how effort was distributed between partners. Figure 7 and Figure 8 show representative examples of groups with highly asymmetric and highly balanced effort distributions, respectively. Specifically, Figure 7a and Figure 7b illustrate two groups (Group 4 and Group 14) exhibiting large asymmetries in grab activity, in which one participant contributed substantially more interaction effort than their partner. In contrast, Figure 8a

and Figure 8b show two groups (Group 16 and Group 2) with near-balanced effort distributions. Importantly, this asymmetry was not systematically associated with participant role (A vs. B) or with any specific network condition. Across the full trace that we collected in our study, effort dominance was nearly evenly split, with comparable numbers of groups showing greater contribution from either participant based on total piece-held time and grab counts.



(a) Near-equal activity with participant A dominating.



(b) Near-equal activity with participant B dominating.

Figure 8: Example groups showing balanced interaction effort between collaborators during puzzle assembly.

Beyond these quantitative differences, we observed recurring qualitative collaboration strategies across groups. Some pairs adopted a shared approach, where one participant transported puzzle pieces while the other performed final placement. Other pairs used a split strategy, either alternating responsibility across puzzle iterations or allowing one participant to complete most of a given puzzle while the partner observed before switching roles. A smaller subset of groups exhibited competitive behavior, in which both participants attempted to grab newly spawned pieces opportunistically or override the partner’s actions. The groups shown in Figure 7 and Figure 8 are presented as illustrative examples of these broader interaction patterns, rather than as representative averages. These interaction patterns varied widely between groups and emerged independently of network conditions, providing important context for interpreting subsequent task performance and collaboration perception measures. We also like to clarify that we do not interpret condition-level differences in interaction balance or grab duration here, as these metrics primarily reflect group-specific coordination styles rather than systematic effects of network constraints.

5.3 Subjective Experience and Cognitive Effort

We next examined subjective measures corresponding to **Layer 3 (Perceptual QoE)** and **Layer 4 (Cognitive QoE)** in our conceptual

model (§3.2.2) to assess how participants experienced collaboration under varying experimental conditions. Table 4 summarizes mean questionnaire scores aggregated across participants and conditions.

Across all four conditions, participants reported consistently high levels of presence, collaboration quality, and social engagement. Measures of perceptual experience, including IPQ, PQ, and the combined presence score, remained stable across **LL**, **HL**, **LB**, and **HB**. Mean presence ratings differed by less than the standard deviation across conditions, suggesting that participants did not perceive a noticeable degradation in immersion even when system responsiveness was reduced. Similarly, subjective collaboration constructs, such as cohesion, attention, conversation quality, and overall collaboration, showed minimal variation across conditions. Scores for these constructs were uniformly high, indicating that participants maintained a strong sense of shared activity and mutual engagement regardless of latency or bandwidth constraints. Together, these results suggest that collaborative experience was robust to the tested network impairments, even when objective performance metrics showed modest degradation.

Table 4: Subjective experience measures (mean ± standard deviation) across experimental conditions.

Construct	LL	HL	LB	HB
Presence Questionnaires				
IPQ	3.91±1.20	3.92±1.24	3.88±1.14	3.87±1.05
PQ	5.40±1.05	5.29±1.02	5.34±0.96	5.42±0.91
Combined Presence	4.65±1.00	4.61±1.00	4.61±0.92	4.64±0.88
Cognitive Effort Questionnaire				
NASA-TLX	2.38±1.12	2.43±1.30	2.47±1.22	2.20±0.95
Custom Group Behavior Questionnaire				
Cohesion	5.78±1.52	5.87±1.61	6.12±1.29	6.27±1.06
Attention	5.57±1.39	5.62±1.34	5.83±1.25	5.68±1.27
Proximity	4.52±2.09	4.43±2.11	4.55±2.07	4.47±1.96
Conversation	5.32±1.85	5.45±1.73	5.53±1.76	5.75±1.62
Collaboration	6.13±1.49	6.17±1.34	6.25±1.22	6.28±1.09

Cognitive effort followed a similar pattern. NASA-TLX scores remained low and tightly clustered across conditions, with no clear increase under higher latency or reduced bandwidth. This indicates that participants did not perceive degraded network conditions as substantially more mentally demanding, despite measurable increases in interaction effort and task duration under **HL**. In other words, users appeared able to compensate behaviorally for reduced responsiveness without experiencing a corresponding increase in perceived workload.

Table 5 provides a finer-grained view of perceptual experience by reporting PQ and IPQ subscales scores, separated by network condition and participant role. For PQ, the subscales capture dimensions of action control (ACT), interface quality (IFQUAL), perceived realism (REAL), examination and exploration (EXAM), and overall experienced quality (EVAL) of the virtual environment. Across all conditions, these dimensions remained largely stable, indicating that users’ sense of control, realism, and engagement with the shared scene were not systematically disrupted by increased latency or reduced bandwidth. Similarly, IPQ subscales assessing general presence (GP), spatial presence (SP), involvement (INV), and experienced realism (REAL) showed only minor, unsystematic fluctuations across conditions. No consistent trend emerged that

Table 5: PQ and IPQ subscale Mean \pm Standard deviation per participant role, computed from aggregated scores per condition.

Participants per Condition	PQ					IPQ			
	ACT	IFQUAL	REAL	EXAM	EVAL	GP	SP	INV	REAL
LL									
A+B	5.69 \pm 1.27	5.20 \pm 1.57	5.09 \pm 1.45	5.54 \pm 1.35	5.98 \pm 1.36	4.88 \pm 1.60	4.59 \pm 1.64	3.22 \pm 1.79	3.51 \pm 1.71
A	5.81 \pm 1.35	5.17 \pm 1.65	5.09 \pm 1.53	5.44 \pm 1.44	5.82 \pm 1.47	4.63 \pm 1.87	4.42 \pm 1.67	3.17 \pm 1.74	3.29 \pm 1.61
B	5.57 \pm 1.17	5.23 \pm 1.49	5.09 \pm 1.36	5.63 \pm 1.24	6.15 \pm 1.22	5.13 \pm 1.23	4.75 \pm 1.60	3.26 \pm 1.83	3.73 \pm 1.76
HL									
A+B	5.47 \pm 1.38	5.02 \pm 1.55	5.00 \pm 1.55	5.52 \pm 1.31	6.05 \pm 1.37	4.80 \pm 1.58	4.44 \pm 1.62	3.35 \pm 1.85	3.61 \pm 1.82
A	5.58 \pm 1.36	5.17 \pm 1.49	4.96 \pm 1.61	5.47 \pm 1.37	6.02 \pm 1.52	4.77 \pm 1.71	4.43 \pm 1.77	3.45 \pm 1.94	3.67 \pm 1.83
B	5.37 \pm 1.38	4.88 \pm 1.58	5.04 \pm 1.48	5.57 \pm 1.24	6.08 \pm 1.20	4.83 \pm 1.44	4.45 \pm 1.47	3.25 \pm 1.74	3.54 \pm 1.81
LB									
A+B	5.48 \pm 1.28	5.09 \pm 1.46	5.08 \pm 1.43	5.53 \pm 1.23	6.03 \pm 1.35	4.77 \pm 1.57	4.47 \pm 1.61	3.34 \pm 1.79	3.45 \pm 1.72
A	5.52 \pm 1.40	5.09 \pm 1.50	5.01 \pm 1.54	5.32 \pm 1.26	5.87 \pm 1.56	5.00 \pm 1.48	4.27 \pm 1.74	3.59 \pm 1.78	3.45 \pm 1.72
B	5.44 \pm 1.16	5.09 \pm 1.42	5.15 \pm 1.32	5.73 \pm 1.15	6.18 \pm 1.07	4.53 \pm 1.63	4.66 \pm 1.45	3.08 \pm 1.76	3.45 \pm 1.72
HB									
A+B	5.70 \pm 1.17	5.19 \pm 1.49	5.16 \pm 1.39	5.49 \pm 1.28	6.03 \pm 1.17	4.83 \pm 1.50	4.53 \pm 1.60	3.12 \pm 1.63	3.55 \pm 1.64
A	5.70 \pm 1.19	5.21 \pm 1.49	5.02 \pm 1.43	5.43 \pm 1.33	5.83 \pm 1.34	4.80 \pm 1.58	4.47 \pm 1.70	3.26 \pm 1.66	3.58 \pm 1.56
B	5.71 \pm 1.16	5.17 \pm 1.49	5.30 \pm 1.33	5.56 \pm 1.23	6.22 \pm 0.93	4.87 \pm 1.41	4.59 \pm 1.50	2.98 \pm 1.59	3.52 \pm 1.73

would suggest degraded perceptual experience under higher latency or lower bandwidth. In particular, involvement and spatial presence scores remained comparable across all network settings, suggesting that participants maintained a coherent sense of co-presence with their collaborator despite changes in responsiveness.

Differences between participant roles (A vs. B) were small and inconsistent across both PQ and IPQ subscales. This indicates that perceptual experience was not strongly shaped by interaction dominance or role asymmetry, even in groups exhibiting uneven distributions of interaction effort. Taken together, the subscale-level analysis reinforces the aggregate results in Table 4, showing that perceptual QoE remained resilient within the range of network changes examined in our study.

5.4 Task Performance and Behavioral Outcomes

We next examined behavioral performance metrics corresponding to **Layer 5 (Behavioral QoE)** in our conceptual model (§3.2.2). Table 6 summarizes task completion time, interaction effort, and balance metrics across conditions. On average, participants took longer to complete the task under the **HL** condition than under **LL** and **LB**, reflected both in overall task completion time and in longer average time per puzzle. The **HL** condition also showed increased average time active per puzzle piece and longer puzzle durations, suggesting that delayed responsiveness modestly slowed task execution and interaction pacing.

Across conditions, interaction balance ratios remained almost the same, indicating that effort was generally distributed between collaborators rather than dominated by a single participant. Average grab durations were comparable across conditions for both participants, with only small numerical differences and no consistent monotonic trend in latency or bandwidth. While **HB** showed the longest average task and puzzle completion times, these differences did not follow a systematic pattern tied to network constraints. We therefore treat these interaction-level metrics as descriptive indicators of collaboration behavior rather than as evidence of condition-driven performance effects.

Figure 9a and Figure 9b further show how interaction effort varied across groups and conditions. While absolute effort levels differed substantially between groups, a consistent pattern emerged

under increased latency. In 19 out of 30 groups, total held time was higher in **HL** than in **LL**, and a similar pattern was observed for grab counts. This suggests that increased latency, which delayed system responsiveness, led participants to spend more time actively manipulating objects, even when overall task strategies differed between groups. Importantly, this effect appeared as a shift in interaction intensity rather than a change in collaboration style, reinforcing that latency primarily influenced execution efficiency rather than coordination roles.

At a finer granularity, piece-level analysis revealed that interaction effort was not uniformly distributed across the task. Figure 10a and Figure 10b show that the central puzzle piece consistently required more interaction across all conditions, as reflected by both longer held times and higher grab counts. This effect was stable across network settings and likely driven by task-specific properties, such as the near-symmetric geometry of the central piece, which increased alignment ambiguity. Although not directly attributable to network conditions, this finding highlights how intrinsic task difficulty can amplify the impact of reduced system responsiveness, particularly under higher latency.

Notably, bandwidth constraints alone did not affect task performance. The **LB** condition exhibited task completion times, held durations, and grab counts comparable to **LL**, indicating that the application’s communication demands remained well below the imposed bandwidth threshold. *Together, these results suggest that latency, unlike bandwidth, was the dominant network factor shaping interaction effort and execution efficiency in our CollabMR task.*

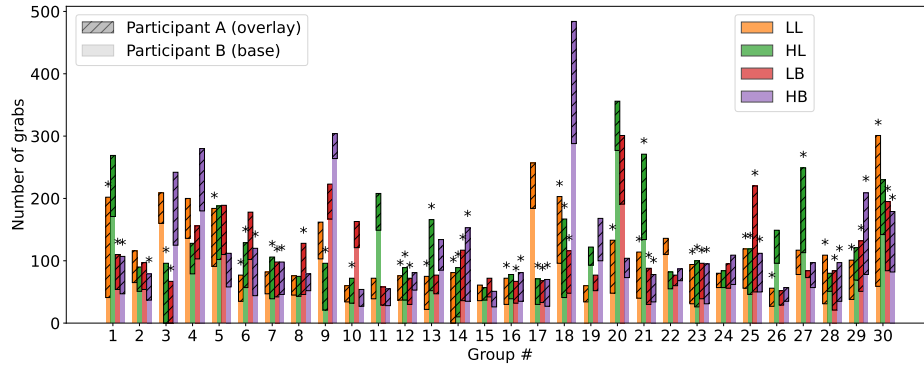
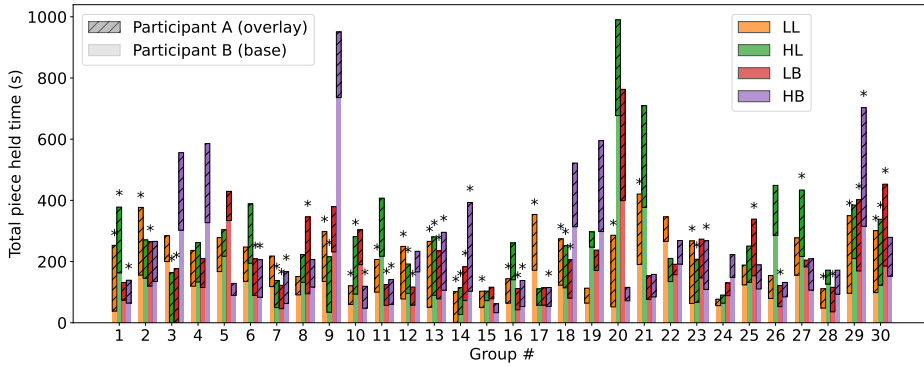
6 Discussion

6.1 Interpreting results

Our study is intended as an **exploratory investigation rather than a confirmatory test** of fixed causal relationships between network QoS and user experience in CollabMR. The imposed network conditions produced clear and measurable differences in system responsiveness, as validated by packet-level and application-level metrics (Table 3, Figure 6). Our goal was to observe how these differences manifest across perceptual, cognitive, and behavioral layers under realistic collaborative behavior, rather than to establish

Table 6: Summary of task performance and interaction metrics across experimental conditions.

Measure	LL	HL	LB	HB
Average task completion time (s)	597.395	659.525	575.659	685.154
Average time active per piece (avg \pm std)	10.539 \pm 16.294	11.854 \pm 21.290	10.133 \pm 15.441	12.119 \pm 24.886
Interaction balance ratio A/B	1.036	0.925	0.942	0.839
Average grab duration (s)	2.083	2.133	2.045	2.160
Average grab duration A (s)	2.083	2.133	2.045	2.160
Average grab duration B (s)	1.741	2.158	1.984	2.137
Average time per puzzle (s)	99.566	109.921	96.479	114.192

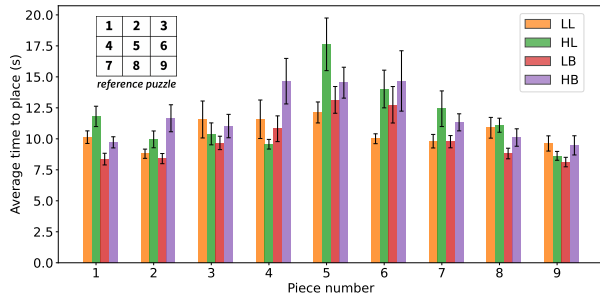
**(a) Grab count per group across conditions.****(b) Object held time per group across conditions.****Figure 9: Group-level interaction balance variability across experimental conditions.**

strict causal mediation. These differences propagated to interaction behavior and task execution, but did not consistently surface in users' reported experience. This divergence between objective responsiveness and subjective perception is how we interpret our findings across the layers of our model.

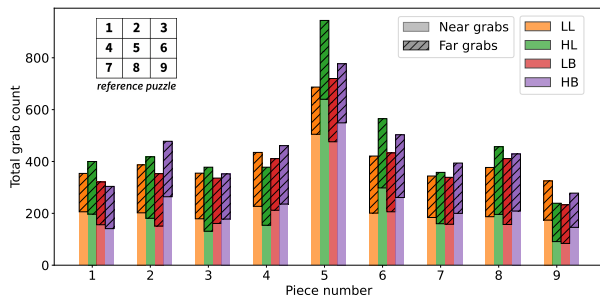
Perceptual experience remained stable across all conditions. Presence and co-presence measures showed minimal variation across latency and bandwidth manipulations (Table 4, Table 5), with condition-level differences well within the reported variance. In the context of our exploratory study, these results suggest that within the tested range, the network perturbations were insufficient to reliably disrupt perceptual QoE given the observed task strategies. Behavioral observations help contextualize this result. Many groups were not visually attending to objects during continuous motion, and several adopted collaboration strategies in which a single participant completed entire puzzle iterations independently. In these cases, object motion was either not directly observed or

was locally rendered, reducing exposure to network-mediated delays. Thus, we do not find evidence to support **H1** within the scope of our study.

A similar pattern emerged for cognitive effort. Despite increased interaction cost under higher latency, reflected in longer task durations and greater object manipulation time (Table 6, Figure 9), NASA-TLX scores remained low and consistent across all conditions (Table 4). Rather than indicating an absence of cognitive impact, this suggests that participants were able to adapt behaviorally without consciously reporting increased mental demand. System-level design choices further help explain this outcome: state-update indexing and out-of-order packet rejection preserved visual stability, preventing oscillatory corrections that might otherwise increase monitoring effort. As an exploratory finding, this indicates that cognitive compensation may occur without being reflected in self-report measures. Accordingly, **H2** is not supported within the sensitivity of our measures.



(a) Average held time per puzzle piece across all groups.



(b) Average grab count per puzzle piece across all groups.

Figure 10: Piece-level interaction effort during collaborative puzzle assembly.

Behavioral performance exhibited more consistent condition-linked effects. Under higher latency, participants took longer to complete tasks, spent more time actively manipulating puzzle pieces, and showed increased interaction effort across the majority of groups (Table 6, Figure 9a, Figure 9b). These effects were robust across groups despite wide variation in collaboration style. Importantly, these performance differences emerged without corresponding changes in perceived workload or presence, suggesting a dissociation between experienced quality and execution efficiency. Because we do not observe mediation through cognitive QoE or a direct linkage between subjective measures and performance, **H3** is supported descriptively but not through a confirmed experiential pathway; so it is partially accepted.

6.2 Implications

As we established in the earlier section, *our findings are not intended to define hard thresholds or causal mappings between network QoS and QoE*. Instead, they **highlight the conditions under which such mappings may or may not surface in any CollabMR tasks**. The results suggest that perceptual and cognitive robustness in CollabMR can mask underlying performance degradation. Relying solely on subjective QoE measures risks overlooking latent execution costs introduced by reduced responsiveness. This reinforces the need for further analyses that combine system metrics, behavioral traces, and self-report data when studying CollabMR. Task structure plays a central role in whether network effects surface at the experiential level. Tasks that permit independent completion

or minimize shared object motion can significantly dampen perceptual sensitivity to latency. Future confirmatory studies should deliberately constrain collaboration strategies (enforced handoffs or interdependent actions) to increase exposure to network-mediated effects.

Within the scope of our task, latency emerged as the dominant QoS factor influencing collaboration, while bandwidth constraints had a negligible impact. This observation is task- and workload-dependent, reflecting the lightweight communication profile of state-based MR interactions. Applications involving remote rendering or continuous media streams are likely to shift this balance. Finally, these findings motivate adaptive CollabMR systems that reason jointly about network conditions, task structure, and observed interaction patterns. Rather than treating QoS effects as universally perceptual, systems should account for how users adapt behaviorally, often invisibly, to maintain collaboration under degraded responsiveness.

7 Limitations and Future Work

7.1 System-Level Limitations

A primary limitation of our study lies in the range and nature of the imposed network constraints. While latency and bandwidth were systematically controlled, the selected values were intentionally conservative to preserve task feasibility and user comfort. As a result, the imposed constraints did not substantially disrupt basic system functionality, particularly under the bandwidth-limited condition, where application-level traffic demands remained well below the throttling threshold. This limits our ability to draw strong conclusions about bandwidth-driven degradation in collaborative MR tasks that involve richer data streams, such as remote rendering or continuous media transmission.

Future studies should therefore explore tighter and more aggressive network constraints, as well as additional QoS dimensions beyond mean latency and throughput. In particular, network jitter and bursty delay patterns are likely to play a critical role in collaborative interaction, as they can induce perceptually salient object instability even when average latency remains moderate. Similarly, examining higher-latency regimes would allow for a more precise characterization of when latency transitions from being behaviorally compensable to perceptually disruptive, especially in relation to object snapping and temporal incoherence.

Another system-level limitation concerns the interaction between network delay and object synchronization mechanisms. Our protocol employed packet indexing and update rejection to suppress oscillatory corrections, which effectively masked some visible effects of packet reordering. While this design choice ensured visual stability, it also reduced the perceptual salience of delay under moderate latency. Refining the networking and synchronization protocol to explicitly expose or parameterize object jitter will be an important direction for future work, enabling a more direct study of how low-level network behavior manifests in user experience. Finally, the collaborative task itself did not enforce tight interdependence between participants. In some groups, participants adopted coordination strategies that localized interaction to a single user at a time, thereby reducing exposure to network-induced delay. Future

system designs should incorporate task mechanics that require explicit handoffs, shared manipulation, or simultaneous object control to more directly couple network responsiveness with collaborative performance.

7.2 Study Scope and Design Limitations

Our study was designed as an exploratory investigation; as such, it has several limitations related to scope and generalizability. While we studied 30 collaborating pairs, each participant only engaged with a single partner for all assigned tasks. This limits our ability to account for individual differences in collaboration style across partners or individual-specific effects from group dynamics. Future studies could involve participants collaborating across multiple group configurations or rotating roles across sessions to better isolate individual versus group-level effects.

Second, the study focused on short, task-bounded interactions within a single session. Longer-term or longitudinal collaboration may reveal adaptation effects, strategy shifts, or fatigue-related changes in how users cope with degraded responsiveness. Additionally, our within-subject design does not allow for causal claims regarding mediation between perceptual experience, cognitive effort, and performance outcomes; observed relationships should be interpreted as descriptive and associative rather than causal. Next, participants were physically co-located, even though collaboration occurred within a shared MR environment. This physical proximity may have provided auxiliary cues or coordination affordances that are absent in fully remote collaboration. As a result, the robustness of perceptual experience observed in our study may not generalize to settings where collaborators are physically separated.

7.3 Future Work

Future work should extend this investigation to distributed collaboration settings where participants are physically separated, initially in adjacent rooms without visual contact, and later across geographically distant locations. Such settings would more closely reflect real-world collaborative MR deployments and would likely amplify the role of network latency, jitter, and packet loss on both coordination and experience. Additional task designs that require tightly coupled, simultaneous interaction, such as shared object manipulation, enforced turn-taking, or interdependent subtasks, would further clarify how network constraints affect collaboration when behavioral compensation is limited.

Finally, future studies could leverage refined sensing and logging to better connect low-level system behavior with higher-level experiential outcomes, supporting stronger modeling of how responsiveness disruptions propagate across perceptual, cognitive, and behavioral layers. Together, these directions would build on the exploratory findings presented here and move toward a more comprehensive understanding of network-aware design for collaborative MR systems.

8 Conclusion

In this work, we explored how network QoS constraints influence user experience in CollabMR over our conceptual model. Our results show that increased latency reliably degraded task efficiency and interaction effort, while subjective presence and workload remained

robust under the tested conditions. Bandwidth limitations had little effect given the task's lightweight state updates. Together, these findings highlight a dissociation between perceived experience and behavioral cost in collaborative MR and underscore the need for future studies that enforce interdependence, expand QoS stressors, and validate causal pathways between QoS and QoE.

Acknowledgments

We thank the anonymous reviewers of IMX 2026 for their insightful feedback and constructive suggestions, which greatly improved the clarity and quality of this paper. This work was supported by the National Science Foundation under Grant numbers 1901137 and 2106463.

References

- [1] Guy T. Almes, Jamshid Mahdavi, Matt Mathis, and Vern Paxson. 1998. Framework for IP Performance Metrics. RFC 2330. doi:10.17487/RFC2330
- [2] Jeremy N Bailenson, Andrew C Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. 2004. Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators & Virtual Environments* 13, 4 (2004), 428–441.
- [3] Nabajeet Barman and Maria G Martini. 2019. QoE modeling for HTTP adaptive video streaming—a survey and open challenges. *IEEE Access* 7 (2019), 30831–30859.
- [4] Aasish Bhanu, Harnish Sharma, Kalyan Piratla, and Kapil Chalil Madathil. 2022. Application of augmented reality for remote collaborative work in architecture, engineering, and construction—a systematic review. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 66, 1829–1833.
- [5] Mark Billinghurst and Hirokazu Kato. 1999. Collaborative Mixed Reality. In *ISMIR '99: First International Symposium on Mixed Reality*. ACM / IEEE, Yokohama, Japan, 6 pages.
- [6] Mark Billinghurst and Hirokazu Kato. 2002. Collaborative augmented reality. *Commun. ACM* 45, 7 (July 2002), 64–70. doi:10.1145/514236.514265
- [7] Robert T. Braden, David D. Clark, and Scott Shenker. 1994. Integrated Services in the Internet Architecture: an Overview. RFC 1633. doi:10.17487/RFC1633
- [8] Kjell Brunnström, Katrien De Moor, Ann Dooms, Sebastian Egger-Lampl, Marie-Neige Garcia, Tobias Hossfeld, Satu Jumisko-Pyykkö, Christian Keimel, Chaker Larabi, Bob Lawlor, Patrick Le Callet, Sebastian Möller, Fernando Pereira, Manuela Pereira, Andrew Perks, Antonio Pinheiro, Ulrich Reiter, Peter Reichl, Raimund Schatz, and Andrej Zgank. 2013. *Qualinet White Paper on Definitions of Quality of Experience*. Qualinet, Dagstuhl, Germany. – pages.
- [9] Kjell Brunnström, Elijs Dima, Tahir Qureshi, Mathias Johanson, Mattias Andersson, and Mårten Sjöström. 2020. Latency impact on Quality of Experience in a virtual reality simulator for remote control of machines. *Signal Processing: Image Communication* 89 (2020), 116005. doi:10.1016/j.image.2020.116005
- [10] Yasra Chandio, Noman Bashir, Victoria Interrante, and Fatima M. Anwar. 2024. Investigating the Correlation Between Presence and Reaction Time in Mixed Reality. *IEEE Transactions on Visualization and Computer Graphics* 30, 9 (2024), 5976–5992. doi:10.1109/TVCG.2023.3319563
- [11] Yasra Chandio, Victoria Interrante, and Fatima M Anwar. 2024. Balancing Presence and Safety using Reaction Time in Mixed Reality. In *2024 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, Seattle, WA, USA, 1–5.
- [12] Yasra Chandio, Victoria Interrante, and Fatima M. Anwar. 2024. Human Factors at Play: Understanding the Impact of Conditioning on Presence and Reaction Time in Mixed Reality. *IEEE Transactions on Visualization and Computer Graphics* 30, 5 (2024), 2400–2410. doi:10.1109/TVCG.2024.3372120
- [13] Yasra Chandio, Victoria Interrante, and Fatima M. Anwar. 2025. Reaction Time as a Proxy for Presence in Mixed Reality with Distraction. *IEEE Transactions on Visualization and Computer Graphics* 31, 5 (2025), 3140–3150. doi:10.1109/TVCG.2025.3549575
- [14] Yasra Chandio, Diana Romero, Salma Elmalaki, and Fatima M. Anwar. 2026. What Sensors See, What People Feel: An Exploratory Study of Subjective Collaboration Perception in Mixed Reality. In *2026 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Daegu, Republic of Korea, 162–172. doi:10.1109/VR67842.2026.00039
- [15] Mark Claypool and Jonathan Tanner. 1999. The effects of jitter on the perceptual quality of video. In *Proceedings of the Seventh ACM International Conference on Multimedia (Part 2) (Orlando, Florida, USA) (MULTIMEDIA '99)*. Association for Computing Machinery, New York, NY, USA, 115–118. doi:10.1145/319878.319909

- [16] Jian Cui, Arjan Kuijper, Dieter W. Fellner, and Alexei Sourin. 2016. Understanding People's Mental Models of Mid-Air Interaction for Virtual Assembly and Shape Modeling. In *Proceedings of the 29th International Conference on Computer Animation and Social Agents* (Geneva, Switzerland) (CASA '16). Association for Computing Machinery, New York, NY, USA, 139–146.
- [17] Ryan Anthony J de Belen, Huyen Nguyen, Daniel Filonik, Dennis Del Favero, and Tomasz Bednarz. 2019. A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering* 3, 2 (2019), 181–223.
- [18] D. Delaney, T. Ward, and S. McLoone. 2006. On Consistency and Network Latency in Distributed Interactive Applications: A Survey – Part I. *Presence: Teleoperators and Virtual Environments* 15, 2 (2006), 218–234.
- [19] D. Delaney, T. Ward, and S. McLoone. 2006. On Consistency and Network Latency in Distributed Interactive Applications: A Survey – Part II. *Presence: Teleoperators and Virtual Environments* 15, 4 (2006), 465–482.
- [20] Florin Dobrian, Vyas Sekar, Asad Awan, Ion Stoica, Dilip Joseph, Aditya Ganjam, Jibin Zhan, and Hui Zhang. 2011. Understanding the impact of video quality on user engagement. *SIGCOMM Comput. Commun. Rev.* 41, 4 (Aug. 2011), 362–373. doi:10.1145/2043164.2018478
- [21] Carmine Elvezio, Frank Ling, Jen-Shuo Liu, and Steven Feiner. 2018. Collaborative virtual reality for low-latency interaction. In *Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. ACM, Berlin, Germany, 179–181.
- [22] Federal Communications Commission. 2014. Connect America Fund Order (FCC 14-190). https://apps.fcc.gov/edocs_public/attachmatch/FCC-14-190A1.pdf Defines minimum broadband service at 10Mbps down / 1Mbps up prior to 2015 redefinition.
- [23] Arturo Garcia-Santillan, Elena Moreno-Garcia, Juan Carlos-Castro, Jorge H Zamudio-Abdala, and Julieta Garduno-Trejo. 2012. Cognitive, affective and behavioral components that explain attitude toward statistics. *Journal of mathematics research* 4, 5 (2012), 8.
- [24] Ceenu George, Manuel Demmler, and Heinrich Hussmann. 2018. Intelligent Interruptions for IVR: Investigating the Interplay between Presence, Workload and Attention. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI EA '18). Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3170427.3188686
- [25] Jaris Gerup, Camilla B. Soerensen, and Peter Dieckmann. 2020. Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review. *International Journal of Medical Education* 11 (Jan. 2020), 1–18. doi:10.5116/ijme.5e01.eb1a
- [26] Chris Greenhalgh and Steve Benford. 1999. Supporting Rich And Dynamic Communication In Large-Scale Collaborative Virtual Environments. *Presence: Teleoper. Virtual Environ.* 8, 1 (1999), 14–35.
- [27] Chris Greenhalgh, Steve Benford, and Gail Reynard. 1999. A QoS architecture for collaborative virtual environments. In *Proceedings of the Seventh ACM International Conference on Multimedia (Part 1)* (Orlando, Florida, USA) (MULTIMEDIA '99). Association for Computing Machinery, New York, NY, USA, 121–130. doi:10.1145/319463.319479
- [28] Simon N. B. Gunkel, Rick Hindriks, Karim M. El Assal, Hans M. Stokking, Sylvie Dijkstra-Soudarissanane, Frank ter Haar, and Omar Niamut. 2021. VRComm: an end-to-end web system for real-time photorealistic social VR communication. In *Proceedings of the 12th ACM Multimedia Systems Conference* (Istanbul, Turkey) (MMSys '21). Association for Computing Machinery, New York, NY, USA, 65–79. doi:10.1145/3458305.3459595
- [29] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908.
- [30] Lukáš Hejtmánek, Michael Starrett, Emilio Ferrer, and Arne Ekstrom. 2020. How Much of What We Learn in Virtual Reality Transfers to Real-World Navigation? *Multisensory Research* 33 (2020), 1–25.
- [31] Tobias Hoßfeld, Raimund Schatz, Ernst Biersack, and Louis Plissonneau. 2013. *Internet Video Delivery in YouTube: From Traffic Measurements to Quality of Experience*. Springer Berlin Heidelberg, Berlin, Heidelberg, 264–301. doi:10.1007/978-3-642-36784-7_11
- [32] C.E. Hughes, C.B. Stapleton, D.E. Hughes, and E.M. Smith. 2005. Mixed reality in education, entertainment, and training. *IEEE Computer Graphics and Applications* 25, 6 (Nov. 2005), 24–30. doi:10.1109/MCG.2005.139 Conference Name: IEEE Computer Graphics and Applications.
- [33] Eoghan Hynes, Ronan Flynn, Brian Lee, and Niall Murray. 2022. A QoE evaluation of procedural and example instruction formats for procedure training in augmented reality. In *Proceedings of the 13th ACM Multimedia Systems Conference* (Athlone, Ireland) (MMSys '22). Association for Computing Machinery, New York, NY, USA, 287–292. doi:10.1145/3524273.3532899
- [34] Allison Jing, Kieran May, Brandon Matthews, Gun Lee, and Mark Billinghurst. 2022. The Impact of Sharing Gaze Behaviours in Collaborative Mixed Reality. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–27.
- [35] Shah Khalid, Aftab Alam, Muhammad Fayaz, Fakhruddin, Sehat Ullah, and Shabir Ahmad. 2023. Investigating the effect of network latency on users' performance in Collaborative Virtual Environments using navigation aids. *Future Generation Computer Systems* 145 (2023), 68–76. doi:10.1016/j.future.2023.02.025
- [36] Martin Kocur, Philipp Schaubhuber, Valentin Schwind, Christian Wolff, and Niels Henze. 2020. The Effects of Self- and External Perception of Avatars on Cognitive Task Performance in Virtual Reality. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) (VRST '20). Association for Computing Machinery, New York, NY, USA, Article 27, 11 pages. doi:10.1145/3385956.3418969
- [37] James F. Kurose and Keith W. Ross. 2017. *Computer Networking: A Top-Down Approach* (seventh ed.). Pearson, Hoboken, NJ, USA. <https://www.pearson.com/us/higher-education/program/PGM1101673.html>
- [38] Matteo Laspro, Leya Groysman, Alexandra N. Verzella, Laura L. Kimberly, and Roberto L. Flores. 2023. The Use of Virtual Reality in Surgical Training: Implications for Education, Patient Safety, and Global Health Equity. *Surgeries* 4, 4 (2023), 635–646. <https://www.mdpi.com/2673-4095/4/4/61>
- [39] Kuan-Yu Lee, Jia-Wei Fang, Yuan-Chun Sun, and Cheng-Hsin Hsu. 2023. Modeling Gamer Quality-of-Experience Using a Real Cloud VR Gaming Testbed. In *Proceedings of the 15th International Workshop on Immersive Mixed and Virtual Environment Systems* (Vancouver, BC, Canada) (MMVE '23). Association for Computing Machinery, New York, NY, USA, 12–17. doi:10.1145/3592834.3592877
- [40] Mingyang Li and Anastasios I. Mourikis. 2013. High-precision, consistent EKF-based visual-inertial odometry. *The International Journal of Robotics Research* 32 (2013), 690–711. <https://api.semanticscholar.org/CorpusID:3778469>
- [41] Linux. 2001. Linux Traffic Control. <https://man7.org/linux/man-pages/man8/tc.8.html>.
- [42] Linux. 2011. Linux tc NetEm. <https://man7.org/linux/man-pages/man8/tc-netem.8.html>.
- [43] Microsoft. 2024. Hololens 2. <https://learn.microsoft.com/en-us/hololens/>
- [44] Nicholas Milef, Adam Ryason, Di Qi, Samuel O. Alfred, Cullen D. Jackson, and Suvranu De. 2021. Disruptions to shared mental models from poor quality of service in collaborative virtual environments. *Scientific Reports* 11, 1 (Dec. 2021), 23556. doi:10.1038/s41598-021-02567-7
- [45] Bhagyabati Moharana, Conor Keighrey, David Scott, and Niall Murray. 2022. Subjective evaluation of group user QoE in collaborative virtual environment (CVE). In *Proceedings of the 14th International Workshop on Immersive Mixed and Virtual Environment Systems* (Athlone, Ireland) (MMVE '22). Association for Computing Machinery, New York, NY, USA, 23–29. doi:10.1145/3534086.3534333
- [46] Toni Mäki, Martin Varela, and Doreid Ammar. 2015. A Layered Model for Quality Estimation of HTTP Video from QoS Measurements. In *2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, Bangkok, Thailand, 591–598. doi:10.1109/SITIS.2015.41
- [47] K. Nahrstedt and R. Steinmetz. 1995. Resource management in networked multimedia systems. In *Computer*. *Computer* 28, 5, 52–63. doi:10.1109/2.384118
- [48] Hyunwoo Nam, Kyung-Hwa Kim, and Henning Schulzrinne. 2016. QoE matters more than QoS: Why people stop watching cat videos. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*. IEEE, San Francisco, CA, USA, 1–9. doi:10.1109/INFOCOM.2016.7524426
- [49] Nels Numan and Anthony Steed. 2022. Exploring user behaviour in asymmetric collaborative mixed reality. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*. Association for Computing Machinery, Tsukuba, Japan, 1–11.
- [50] Kyoung S. Park and Robert V. Kenyon. 1999. Effects of Network Characteristics on Human Performance in a Collaborative Virtual Environment. In *Proceedings of IEEE Virtual Reality (VR)*. IEEE, Houston, TX, USA, 104–111. Shows effect of latency and jitter on collaboration task performance in networked VR.
- [51] Harold Pashler. 1994. Dual-task interference in simple tasks: data and theory. *Psychological bulletin* 116, 2 (1994), 220.
- [52] Daniel Pimentel and Charlotte Vinkers. 2021. Copresence With Virtual Humans in Mixed Reality: The Impact of Contextual Responsiveness on Social Perceptions. *Frontiers in Robotics and AI* 8 (04 2021). doi:10.3389/frobt.2021.634520
- [53] Katja Rogers, Jana Funke, Julian Frommel, Sven Stamm, and Michael Weber. 2019. Exploring Interaction Fidelity in Virtual Reality: Object Manipulation and Whole-Body Movements. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3290605.3300644
- [54] Rolls-Royce. 2019. Rolls-Royce and Qatar Airways use Virtual Reality to train engineers. <https://www.rolls-royce.com/media/press-releases/2019/15-04-2019-ar-and-qatar-airways-use-virtual-reality-to-train-engineers.aspx>.
- [55] Damien Constantine Rompapas, Christian Sandor, Alexander Plopsi, Daniel Saakes, Joongi Shin, Takafumi Taketomi, and Hirokazu Kato. 2019. Towards large scale high fidelity collaborative augmented reality. *Computers & Graphics* 84 (2019), 24–41.
- [56] Jinjia Ruan and Dongliang Xie. 2021. Networked vr: State of the art, solutions, and challenges. *Electronics* 10, 2 (2021), 166.
- [57] Jinjia Ruan and Dongliang Xie. 2021. A survey on QoE-oriented VR video streaming: Some research issues and challenges. *Electronics* 10, 17 (2021), 2155.
- [58] Kimberly Ruth, Tadayoshi Kohno, and Franziska Roesner. 2019. Secure multi-user content sharing for augmented reality applications. In *Proceedings of the*

- 28th USENIX Conference on Security Symposium (Santa Clara, CA, USA) (SEC'19). USENIX Association, USA, 141–158.
- [59] Maria Sanchez-Vives and Mel Slater. 2005. From presence to consciousness through virtual reality. *Nature Reviews Neuroscience* 6 (2005), 332–339.
- [60] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (2001), 266–281.
- [61] Mickael Sereno, Xiyao Wang, Lonni Besançon, Michael J. McGuffin, and Tobias Isenberg. 2022. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* 28, 6 (2022), 2530–2549. doi:10.1109/TVCG.2020.3032761
- [62] Michael Seufert, Sebastian Egger, Martin Slanina, Thomas Zinner, Tobias Hofffeld, and Phuoc Tran-Gia. 2015. A Survey on Quality of Experience of HTTP Adaptive Streaming. *IEEE Communications Surveys & Tutorials* 17, 1 (2015), 469–492. doi:10.1109/COMST.2014.2360940
- [63] Mel Slater. 2009. Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philosophical Transactions: Biological Sciences* 364, 1535 (2009), 3549–3557. <http://www.jstor.org/stable/40538148>
- [64] Mel Slater, Amela Sadagic, Martin Usoh, and Ralph Schroeder. 2000. Small-group behavior in a virtual and real environment: A comparative study. *Presence* 9, 1 (2000), 37–51.
- [65] Mel Slater and Sylvia Wilbur. 1997. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators & Virtual Environments* 6 (1997), 603–616. <https://api.semanticscholar.org/CorpusID:9437981>
- [66] Vinicius Souza, Anderson Maciel, Luciana Nedel, and Regis Kopper. 2021. Measuring Presence in Virtual Environments: A Survey. *ACM Comput. Surv.* 54, 8, Article 163 (Oct. 2021), 37 pages. doi:10.1145/3466817
- [67] A. Steed, M. Slater, A. Sadagic, A. Bullock, and J. Tromp. 1999. Leadership and collaboration in shared virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. IEEE, Houston, TX, USA, 112–115. doi:10.1109/VR.1999.756941
- [68] Denny Stohr, Alexander Frömmgen, Jan Fornoff, Michael Zink, Alejandro Buchmann, and Wolfgang Effelsberg. 2016. QoE Analysis of DASH Cross-Layer Dependencies by Extensive Network Emulation. In *Proceedings of the 2016 Workshop on QoE-Based Analysis and Management of Data Communication Networks (Florianopolis, Brazil) (Internet-QoE '16)*. Association for Computing Machinery, New York, NY, USA, 25–30. doi:10.1145/2940136.2940141
- [69] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science* 12, 2 (1988), 257–285.
- [70] Klervie Toczcé, Johan Lindqvist, and Simin Nadjm-Tehrani. 2019. Performance Study of Mixed Reality for Edge Computing. In *Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing (Auckland, New Zealand) (UCC'19)*. Association for Computing Machinery, New York, NY, USA, 285–294. doi:10.1145/3344341.3368816
- [71] Sam Van Damme, Javad Sameri, Susanna Schwarzmann, Qing Wei, Riccardo Trivisonno, Filip De Turck, and Maria Torres Vega. 2024. Impact of latency on QoE, performance, and collaboration in interactive multi-user virtual reality. *Applied Sciences* 14, 6 (2024), 2290.
- [72] Shanthi Vellingiri and Prabhakaran Balakrishnan. 2017. Modeling User Quality of Experience (QoE) through Position Discrepancy in Multi-Sensorial, Immersive, Collaborative Environments. In *Proceedings of the 8th ACM on Multimedia Systems Conference (Taipei, Taiwan) (MMSys'17)*. Association for Computing Machinery, New York, NY, USA, 296–307. doi:10.1145/3083187.3084018
- [73] Sara Vlahovic, Mirko Suznjevic, and Lea Skorin-Kapov. 2022. A survey of challenges and methods for Quality of Experience assessment of interactive VR applications. *Journal on Multimodal User Interfaces* 16, 3 (Sept. 2022), 257–291. doi:10.1007/s12193-022-00388-0
- [74] A. Vogel, B. Kerherve, G. von Bochmann, and J. Gecsei. 1995. Distributed multimedia and QoS: a survey. *IEEE MultiMedia* 2, 2 (1995), 10–19. doi:10.1109/93.388195
- [75] Petr Vávra, Jan Roman, P. Zonča, Peter Ihnát, M. Němec, Kumar Jayant, Nagy Habib, and Ahmed El-Gendi. 2017. Recent Development of Augmented Reality in Surgery: A Review. *Journal of Healthcare Engineering* 2017 (08 2017), 1–9. doi:10.1155/2017/4574172
- [76] Ina Wagner, Wolfgang Broll, Giulio Jacucci, Kari Kuutii, Rod McCall, Ann Morrison, Dieter Schmalstieg, and Jean-Jacques Terrin. 2009. On the role of presence in mixed reality. *Presence: Teleoper. Virtual Environ.* 18, 4 (Aug. 2009), 249–276. doi:10.1162/pres.18.4.249
- [77] Xiangyu Wang, Peter E.D. Love, Mi Jeong Kim, and Wei Wang. 2014. Mutual awareness in collaborative design: An Augmented Reality integrated telepresence system. *Computers in Industry* 65, 2 (2014), 314–324. doi:10.1016/j.compind.2013.11.012
- [78] Ying Yang, Tim Dwyer, Michael Wybrow, Benjamin Lee, Maxime Cordeil, Mark Billinghurst, and Bruce H. Thomas. 2022. Towards Immersive Collaborative Sensemaking. *Proc. ACM Hum.-Comput. Interact.* 6, ISS, Article 588 (Nov. 2022), 25 pages. doi:10.1145/3567741
- [79] Tiesong Zhao, Qian Liu, and Chang Wen Chen. 2016. QoE in video transmission: A user experience-driven strategy. *IEEE Communications Surveys & Tutorials* 19,