GroundLink: Exploring How Contextual Meeting Snippets Can Close Common Ground Gaps in Editing 3D Scenes for Virtual Production

Gun Woo (Warren) Park
Autodesk Research
Toronto, Ontario, Canada
Department of Computer Science
University of Toronto
Toronto, Ontario, Canada
warren@dgp.toronto.edu

Frederik Brudy
Autodesk Research
Toronto, Ontario, Canada
frederik.brudy@autodesk.com

George Fitzmaurice
Autodesk Research
Toronto, Ontario, Canada
george.fitzmaurice@autodesk.com

Fraser Anderson
Autodesk Research
Toronto, Ontario, Canada
fraser.anderson@autodesk.com

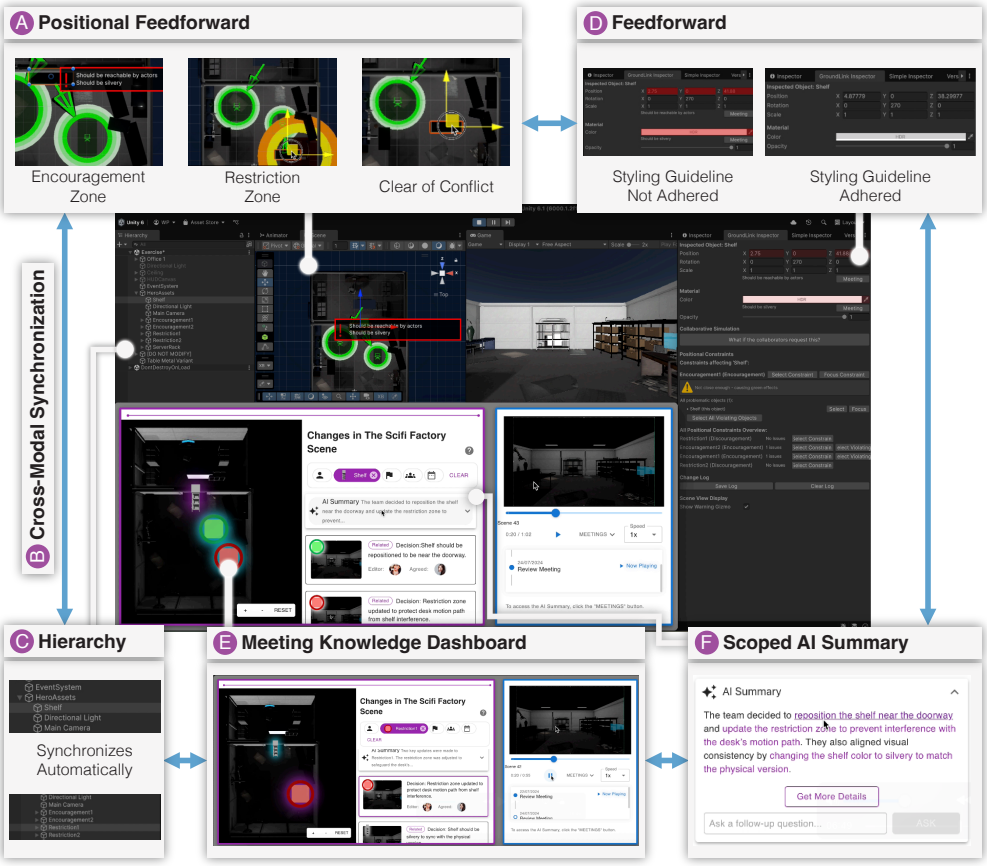


Figure 1: *GroundLink* surfaces decisions and rationale behind edits made in past meetings in virtual production projects, right within the Unity editor. (A) Positional feedforward: When the placement of an element does not follow a previous decision, it shows a green circle with an arrow to encourage placement in an encouragement zone, or a yellow/orange/red circle to encourage moving away from the restriction zone. If the placement follows the decision, all zones remain transparent. (B) Importantly, all the components in *GroundLink* synchronize, whether it is verbal (meeting recordings, summaries) or non-verbal content (3D editor). (C) The Hierarchy view in Unity also syncs, as does the Inspector (D). Just like positional feedforward, the Inspector blinks the field in red if the properties do not follow the styling guideline or decisions. (E) The meeting knowledge dashboard displays all the meeting decisions and summaries in an easily navigable format. (F) It also features an AI summary, which users can ask more scoped questions about objects and collaborators among all decisions and comments.

Abstract

Virtual Production (VP) professionals often face challenges accessing tacit knowledge and creative intent, which are important in forming common ground with collaborators and in contributing more effectively and efficiently to the team. From our formative study (N=23) with a follow-up interview (N=6), we identified the significance and prevalence of this challenge. To help professionals access knowledge, we present *GroundLink*, a Unity add-on that surfaces meeting-derived knowledge directly in the editor to support establishing common ground. It features a meeting knowledge dashboard for capturing and reviewing decisions and comments, constraint-aware feedforward that proactively informs the editor environment, and cross-modal synchronization that provides referential links between the dashboard and the editor. A comparative study (N=12) suggested that *GroundLink* help users build common ground with their team while improving perceived confidence and ease of editing the 3D scene. An expert evaluation with VP professionals (N=5) indicated strong potential for *GroundLink* in real-world workflows.

CCS Concepts

• **Human-centered computing → Interactive systems and tools**; *Empirical studies in interaction design*; *Collaborative and social computing systems and tools.*

Keywords

video meetings, remote meetings, virtual production, meeting recordings, communication

ACM Reference Format:

Gun Woo (Warren) Park, Frederik Brudy, George Fitzmaurice, and Fraser Anderson. 2026. GroundLink: Exploring How Contextual Meeting Snippets Can Close Common Ground Gaps in Editing 3D Scenes for Virtual Production. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 27 pages. https://doi.org/10.1145/3772318.3790793

1 Introduction

Virtual Production (VP), a rapidly growing branch of filmmaking, requires continuous, iterative coordination between artists, technicians, and supervisors throughout the entire filmmaking process, promising creative freedom and cost savings [57, 124]. This coordination relies heavily on frequent meetings where teams make important decisions about creative intent, technical constraints, asset specifications, and workflow priorities. These meetings serve as the primary mechanism for establishing shared understanding and aligning diverse expertise across departments, from lighting designers and camera operators to VFX artists and technical directors.

However, the decisions made in these meetings often become disconnected from the actual production work. For example, during intensive creation phases, team members could focus on immediate technical tasks and may forget or misinterpret earlier agreements. The increasing trend toward remote and distributed VP workflows [5, 108, 109] can amplify this challenge. While remote meetings enable global collaboration with top talent, they often result in less effective knowledge sharing compared to in-person interactions [86].

The challenge intensifies when a *newcomer*[1] joins an ongoing project. Newcomers inherit work shaped by dozens of prior meetings but lack efficient access to the rationale behind existing decisions. This creates a significant knowledge gap: important decisions and constraints discussed in meetings remain buried in recordings while team members work in 3D environments where those decisions should be applied.

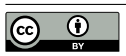
Literature suggests that knowledge workers, in general (which the VP professionals are), rarely start a project with sufficient access to materials, expertise, or experience [2, 37, 87], and that successful teams overcome these imbalances by establishing *common ground*: the shared understandings, tacit know-how, preferences, and constraints that allow members to communicate and coordinate smoothly [6, 23, 91]. However, it remains unclear whether VP professionals actually experience common ground issues in practice, and if so, what specific types of meeting-derived knowledge are most important for establishing effective collaboration.

We address these questions by investigating how to surface meeting-derived knowledge directly within the 3D content creation environment. Specifically, we ask:

- **RQ1:** What kinds of meeting-derived knowledge do virtual production teams need when working on 3D scenes?
- **RQ2:** How does surfacing meeting-derived knowledge support a team member in virtual production projects form common ground with existing team members?
- **RQ3:** How does acquiring knowledge from past meeting recordings affect team members' perceived confidence and task difficulty?

To answer **RQ1**, we first conducted a formative study, with experienced VP professionals (N=23) and follow-up interview (N=6).

[1] Can be an experienced professional.



CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3790793

This study confirmed that VP professionals do indeed experience hindrance in forming common grounds, particularly when newcomers join projects, and revealed the specific types of tacit knowledge and collaborative intent that are typically shared through meetings but become difficult to access during production work.

Guided by these empirical findings, we designed and developed *GroundLink*, a Unity add-on that augments 3D scenes with contextual meeting snippets. *GroundLink* works with existing, indexed video meeting recordings and surfaces relevant meeting snippets as users navigate and interact with 3D scene elements, creating bidirectional connections between the scene and meeting content. As users edit the scene, *GroundLink* provides contextual cues during editing, surfacing potential conflicts with earlier decisions while preserving user agency for creative judgment. It also allows users to generate summaries of past changes to help them understand constraints and collaboration needs.

To answer **RQ2** and **RQ3**, we evaluated *GroundLink* through two complementary studies. First, we conducted an exploratory, comparative within-subjects study (N=12) with participants who have 3D editing experience but no VP domain expertise. Participants edited two unfamiliar Unity scenes under time pressure, using both a powerful baseline (Microsoft Copilot) and *GroundLink*. This design allowed us to isolate the effects of our approach from confounds related to VP skill level or project familiarity. Results show that participants reported more extensive perceived common ground with the prior team and greater confidence in the quality of their results, when using *GroundLink*. To confirm domain-specific validity, we also conducted an expert evaluation (N=5) with VP professionals, who confirmed that *GroundLink* has potential to address real challenges in their workflows and would be valuable for their collaborative practices.

This paper contributes:

- *GroundLink*, a Unity add-on that helps distributed collaborators build common ground by surfacing meeting-derived knowledge within 3D editing workflows.
- Empirical findings from a formative study revealing specific knowledge gaps in virtual production collaboration and confirming that issues in forming common ground identified in literature do manifest in practice.
- A comparative user study demonstrating the perceived effectiveness of in-situ knowledge surfacing for newcomer onboarding in collaborative 3D environments, as well as an expert evaluation confirming the potential of the approach in real-world workflows.
- Design implications for integrating meeting-derived knowledge into production tools, applicable to virtual production, and other collaborative 3D workflows.

2 Background

This work addresses the challenges in virtual production. It builds on prior research in collaboration, with particular emphasis on remote work and relevant designs (feedforward and in-situ presentation of information).

2.1 Virtual Production and Collaboration

Virtual production (VP) is a fast-growing filmmaking technique that blends live-action footage with computer-generated (CG) content in real time [57, 124]. Using technologies like LED volume stages and motion capture systems, filmmakers can create high-fidelity virtual environments and props that would be difficult or expensive to build physically. These assets can be viewed and modified interactively on set, enabling rapid iteration and alignment with creative goals [108, 109]. VP enables continuous, iterative coordination among artists, technicians, and supervisors throughout the entire process [57, 109]. With VP, professionals can more easily hire top talent globally by using remote work arrangements. The industry is seeing a growing trend of geographically distributed teams, where collaborators work remotely across countries and time zones [59]. As a result, despite its relatively recent emergence, VP is increasingly being adopted [71].

However, the literature suggests that VP professionals may experience issues with establishing common ground. The film industry, in general, relies heavily on project-based hiring [10, 29, 43, 53], i.e., new project brings together new collaborators who may have never worked together before. Establishing common ground in this setting involves tacit knowledge about naming conventions, artistic preferences, technical constraints, and shared expectations. When common ground is not well established, it is often difficult for collaborators to communicate effectively [6, 90, 91]. This problem can be amplified in VP, which is at its infancy.

Remote collaboration adds another layer of complexity. VP allows for hiring top talent from around the world, but this necessitates remote collaboration, which is a growing trend [5, 108, 109]. Although real-time tools like video conferencing have become standard, they are not always sufficient for communicating nuanced design decisions or creative intent [28].

As prior work notes, remote meetings can often hinder the establishment of strong common ground [90], especially in recurring meeting settings [86]. Moreover, remote meetings are often viewed as disruptive or time-consuming [31, 78, 80], potentially leading to fewer meetings and greater divergence in mental models than needed[2]. However, remote collaboration is a growing standard in VP [5, 108, 109], and we therefore consider mechanisms to extract maximum insight from this scarcer source of data.

Given these challenges, we question whether existing collaboration tools are sufficient to support richer collaboration in the growing, remote virtual production field. Our formative study aims to determine whether common ground formation problems manifest in practice for virtual production professionals, closing a research gap where past work has only speculated about such issues.

2.2 Common Ground in (Remote) Work

2.2.1 Components of Common Ground in Virtual Production. Common ground is important for effective collaboration, especially in distributed or high-stakes work environments [6, 22, 24, 90, 91]. While cognition is often conceptualized as an individual process, it frequently involves knowledge distributed across a team [101]. It encompasses shared understandings of terminology, goals, constraints, and tacit knowledge that enable teams to communicate and coordinate smoothly. From a cognitive perspective, two complementary theoretical models help explain how common ground forms and

[2]It is easy to think that there are too many meetings in remote work [4], but the term 'meetings' here includes spontaneous interactions, such as water cooler discussions. When considering this broader definition, we can recognize a reduction in the frequency, quality, and duration of interactions with collaborators [86].

functions: Shared Mental Models (SMM) and Transactive Memory Systems (TMS).

SMM emphasizes that collaborators benefit from overlapping knowledge of tasks, processes, and team dynamics. Effective teamwork stems from shared understanding, i.e., everyone being "on the same page." Practices such as planning sessions, reflexive check-ins, and team interaction training have been shown to improve coordination and performance in ad hoc and virtual teams [16, 97].

In contrast, TMS describes common ground in terms of distributed cognition: instead of everyone knowing the same things, team members maintain an awareness of who knows what [121]. This model is especially valuable when collaborators have deep specializations, and when sharing all relevant knowledge across the team would be inefficient. TMS assumes that effective coordination relies on knowing whom to consult, rather than on uniform expertise.

Virtual production demands a balance of both paradigms, as different collaboration problems emphasize each one differently. This balance informs the direction of support that our work aims to provide. Professionals bring deep, domain-specific skills (e.g., lighting, rigging, camera tracking), highlighting the importance of TMS. Simultaneously, virtual production involves frequent, close, and cross-functional coordination to realize a shared creative intent [57], underscoring the importance of SMM. These theories suggest that a support system for common ground formation in virtual production collaboration should both provide a shared repository of knowledge and offer role-based references to help users identify experts for specific aspects of production.

2.2.2 Previous Work on Supporting Common Ground Formation. Prior systems have focused on capturing and presenting knowledge across different media and contexts to support common ground formation. Common ground can be captured passively or proactively. Passive approaches mine knowledge from corporate documentation, collaborative tools, or archived communication logs [12, 26, 79, 119]. Proactive methods, such as think-aloud computing, encourage users to verbalize their thought processes as they work, making that knowledge more accessible to others or to intelligent systems [63]. Given that meetings are abundant in virtual production [57] and that much of the constituent knowledge for common ground is potentially verbalized there, we focus our contribution on helping professionals consume this rich source of knowledge effectively.

Past work has presented this knowledge from meeting recordings using enhanced representations. One approach represents meetings in more digestible formats. Beyond traditional text-based summaries [72, 74, 84, 100, 118], MeetMap uses a large language model to generate a dynamic visual map of conversation topics based on issue-based information systems [64], for in-meeting awareness and post-meeting review [17]. Another approach provides more efficient replay mechanisms for video [49, 56], audio [114], or team-shared annotations [85]. These replays can also be spatialized using immersive mixed reality, showing what happened in a space while users were engaged in another task [21, 35].

Instead of consuming knowledge on a separate occasion, it can be embedded in conversations to help users consume meeting knowledge during the meeting itself. In cases where users miss live meetings, proxies such as AI avatars [69] or pre-recorded videos [110] simulate participation. Knowledge presentation can also be proactive: chatbots can surface relevant information [9, 75, 105], systems can trigger alerts when misunderstandings arise [62], and AR overlays can provide contextual guidance [123].

In existing systems, the consumption of constituent knowledge for common ground formation remains separate from actual work. A missed opportunity lies in embedding common ground formation into the work itself, where the act of working could simultaneously help users acquire this knowledge. This explores a direction beyond embedding knowledge during meetings, focusing instead on asynchronous work. This opportunity is particularly significant for remote virtual production professionals, where relying on communication alone requires additional effort to proactively engage with shared resources [86], and interpreting knowledge in practice is difficult given emerging terminologies and workflows [57, 124].

2.3 Surfacing Hidden Knowledge

Much of the knowledge necessary for forming common ground is tacit or hidden, residing in prior decisions, implicit intent, or shared understanding that is often not explicitly documented [91]. In co-located teams, such knowledge is shared informally through spontaneous conversations or observations. In remote settings, these opportunities are reduced [86]. Our goal is to help virtual production professionals obtain this constituent knowledge in situ, building on two relevant concepts: feedforward and the in-situ presentation of information.

2.3.1 Feedforward. Feedforward refers to interface mechanisms that preview the outcome of an interaction before the user acts, helping them anticipate consequences [30, 88, 116]. Originally used in gesture guidance [7, 36, 61, 96], VR interaction guidance [3, 83], and GUI interactions [25, 112], it has recently expanded to AI explainability [40, 81, 94, 95] and autonomous vehicles [102].

Unlike feedback, feedforward supports discovery by revealing what is possible in advance [116]. This makes it well-suited for surfacing hidden knowledge. However, poorly designed feedforward can mislead users [65], so systems should promote exploration rather than passive acceptance. While traditional feedforward communicates user interface interaction instructions (e.g., "Slide to unlock"), we explore the methods to communicate constituent knowledge of common ground in virtual production instead of the interaction instructions, such as creative decision (e.g., "The shelf should have silvery color").

2.3.2 In-situ Presentation of Information. A related strategy for surfacing hidden knowledge is the in-situ presentation of information, where relevant information is embedded directly in the user's workspace for just-in-time access [13]. This helps users relate disparate pieces of information. In programming, inline visualizations help users understand running code [8, 46]; in text-rich contexts, embedded visualizations keep users focused on the text while providing information that helps interpret it [41, 107, 120]. Comments and annotations relevant to parts of 2D canvases [33, 45, 58] or video comments [18, 52] can be placed on the relevant portions of the canvas or video, helping viewers interpret the material better and fostering collaboration. Commercial tools like Miro[3], Autodesk RV[4], and

[3] https://miro.com
[4] https://www.autodesk.com/products/flow-production-tracking/rv

SyncSketch[5] embody these principles. We also explore annotations to encourage collaboration and knowledge acquisition, but we aim to explore communicating this knowledge in a 3D editing environment.

In 3D and physical spaces, in-situ information can be projected to reveal context. Data associated with physical elements can be presented on physical spaces or objects, such as weather data or the star ratings of restaurants [15, 67, 122]. Information relevant to conversations can be presented on the communicators' faces [50, 98] to obtain information without disrupting attention to the communicator. Tools relevant for thinking about objects, such as calculators [99], or annotations in physical spaces can be provided in-situ [89]. Users can use a virtual hand in a remote physical space to refer to a location, essentially acting as a real-time in-situ annotation [66, 106], or a physical hand in a virtual space to provide real-time, in-situ annotations on their presentations [73, 103]. In virtual 3D canvases (similar to the domain of our work), such as architectural renderings, comments and annotations can be provided in-situ so that viewers can better associate the location in the 3D space with the annotations or comments [54, 55]. In this setting, not just text, but multimedia comments can be included, such as audio or video, with tree-based representations [42].

Many of these systems treat comments or annotations as static, non-reactive marks attached to assets. For collaboration systems for film production, these comments are often also decoupled from the tools where changes are actually made. By contrast, our goal is to treat meeting-derived comments and decisions as a dynamic, in-situ feedforward layer inside the primary 3D editing tool. Rather than asking users to open a distinct review environment (e.g., annotations on the video clips, instead of 3D assets), infer the meaning, and be wary of not violating some constraints suggested, we proactively surface constraints and rationales directly as they manipulate objects, with each cue linked back to its originating meeting recording clip and decision entry. We aim to help users form common ground this way.

Particularly, we recognize that in-situ annotation can guide users, for instance, for movement [1, 19]. In this work, we combine feedforward with in-situ presentation in a 3D editor environment to guide users through the meeting context while equipping them with constituent knowledge for common ground formation. Through this, we aim to make collaborators more informed, aligned, and seamlessly connected with other collaborators.

3 Formative Study

Although prior works have examined the workflows, advantages, operationalizations, and drawbacks of virtual production (Section 2.1), few have investigated how virtual production staff collaborate, which leaves open questions (**RQ1**). We conducted a formative study with virtual production professionals to address this.

3.1 Method

We sent email invitations to a company internal research community consisting primarily of customers who had indicated their primary industry to be film or whose primary software expertise included 3D design or animation software. We recruited 23 participants[6]. Participants were invited to complete a Qualtrics survey (Appendix B), providing insights into their collaboration and onboarding experiences in virtual production. Some participants expressed interests in the follow-up interview (See Appendix C for the list of questions), which the aim was to help the researchers understand some of the unique practices mentioned in the survey, in more detail. We randomly selected 6 from the 18 that were interested (demographics: Table 1 (appendix)). Participants received gift cards for participation (survey: 15 minutes, US$10; interview: 45 minutes, US$60)[7].

We performed thematic analysis [11] on all survey and interview responses[8]. Our analysis followed an inductive approach, focusing on participants' explicit descriptions of their experiences rather than applying a predefined coding frame. The lead author conducted the analysis. First, the lead author reviewed transcripts in full and noted initial impressions. Next, the author performed open, line-by-line coding in Dovetail (without using AI features), generating codes that captured meaningful units of text. The lead author then iteratively reviewed and refined the initial codes, merging, splitting, and discarding codes as needed. Related codes were grouped into candidate themes. Groups were color coded for easier access when extracting insights. Throughout this process, the lead author discussed the results of coding with co-authors on a daily basis, who provided feedback on code and theme definitions, but no second formal coding pass or inter-rater reliability was calculated, in line with reflexive thematic analysis. Finally, the lead author named and defined each theme and selected representative quotes to illustrate them in the findings. In this paper, **F***n* denotes a unique formative survey participant. Some participants (**F2,5,10,15,17,22** - see Table 1) also participated in formative study interviews, denoted as **FI***n*.

3.2 Collaboration in Virtual Production

Collaboration in virtual production is extensive and cross-functional. The majority of participants work frequently with colleagues from other roles (19/23; Figure 2A-top), notably more often than with those in their same role (12/23; Figure 2A-bottom). However, this collaboration practice brings challenges, as **F15** illustrates: *"Issues can arise weekly or even daily during fast-paced phases like scene assembly or on-set shooting, especially when multiple departments are involved."* Most often, these issues relate to communication (Figure 2B).

3.2.1 The Root Cause: Lack of Common Ground. Our analysis revealed that communication issues primarily stem from insufficient formation of common ground. The short history combined with complexity of collaboration in virtual production results in a lack of standards, as **F15** summarized: *"Because VP is still evolving, not everyone shares the same level of experience, which can lead to misunderstandings and uneven expectations."*

Specifically, 12 participants experienced issues indicating the factors, attributable to the **insufficiency in a formed common ground**

[5] https://syncsketch.com

[6] **Gender identity:** 21 M, 2 unknown; **Age:** 40.8±14.23; **Country of Residence:** 4 US, 4 India, 8 Canada, 1 Germany, 1 Honduras, 1 Spain, 1 South Korea, 1 Palestine, 1 UAE, 1 Zambia (Raw, typed inputs from participants.); **Years of experience in virtual production:** 6.2±4.18; **Number of past virtual production projects:** 28.2±44.98; **Roles:** 6 VP Artists, 4 VP Supervisors, 2 Software Developers/Engineers, 2 VP Coordinators, 2 Technicians, 1 VP Lead/Researcher, 1 VP Researcher, 1 Technical Director, 1 Animator, 1 Compositor, 1 VFX Supervisor, 1 Student

[7] All studies were approved by our institutional ethics board.

[8] Transcribed via Dovetail and verified by a researcher

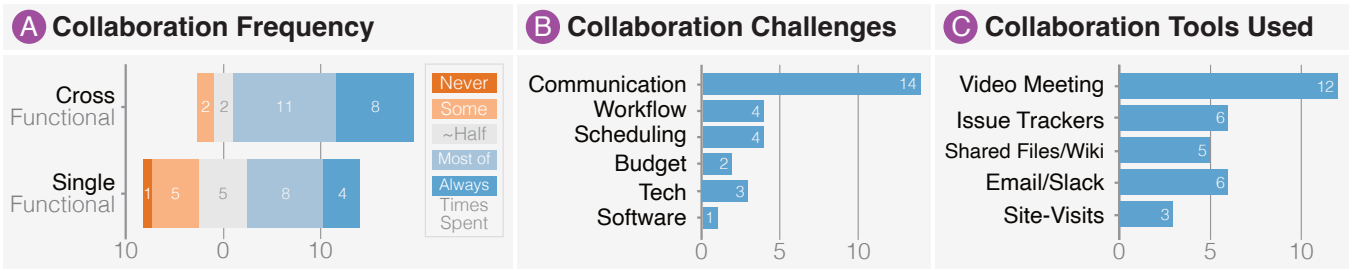


Figure 2: Formative study results on collaboration practices for virtual production. (A) shows frequency of collaboration for cross- vs. single-functional collaboration. (B) shows the challenges in current collaboration practices (multiple responses possible). (C) shows the tools that professionals currently use for collaboration (multiple responses possible).

as the main cause for communication problems. These issues encompassed inconsistencies in nomenclature, versioning style, intended workflow, tacit technical/artistic limitations, styles, and managing expectations. This lack of shared understanding manifests in several ways:

Unclear Intent Behind Actions. FI1, **FI2**, and **FI5** experienced difficulties understanding the **intention behind collaborators' actions**. When intent is unclear, **FI2** improvises: *"We know the bare minimum that we need to know [about the film/script]. But we are also given quite a lot of freedom to improvise for whatever we need to film or perform."*

These communication gaps result in production frictions with real consequences, as **FI1** explains: *"An artist will be working on [a 3D scene] and then they'll inadvertently grab something and move it without realizing it. It's happened to me many times, and then all of a sudden, [when we load] everything is out of calibration. Nothing is lining up and we have to troubleshoot."* Such misalignments are particularly problematic when dealing with "Hero Assets", which are assets important for the narrative or immersion in the scene (**FI4**). Poor coordination can even pose physical hazards, as **FI2** explains: *"The camera unfortunately came too close to [the actors] and he was spinning this whole [camera setup]. He knocked off the lens."*

Information Propagation Failures. For 4 participants, **information propagation** was the main communication issue. They do not obtain information about the most recent decisions or changes. For 3 participants, **unclear explanations of responsibilities and expectations** were the source of problems. These failures result in added production costs, as **FI1** illustrates: *"This last-minute change is going to affect everything else that we planned for and we built for. And it would have been great if you had told me that this was an option so that we could have been prepared to make those adjustments. So sometimes, the departments operate in a vacuum."*

Inadequate Knowledge Resources. Despite the availability of shared materials (Figure 2C), some professionals experience difficulties in establishing the common ground (e.g., Figure 3C). **FI6** described using internal Wiki pages and style bibles containing guidelines for each scene and project. However, **FI6** observed that collaborators often do not engage with these resources: *"I sometimes doubt that people actually look at all the information wiki pages because a lot of times we do get questions or we do get tickets for it or for another department. [Collaborators not considering the style guidelines] is more with junior artists and that comes up in dailies or during a review time or something. [Some people are like] people who don't read the instruction manuals and just hope that they can wing it."*

From these findings, we derive a design goal:

> **Design Goal 1 (DG1):** Gathering and making constituent knowledge accessible which is needed for common ground such as intention behind actions, constraints, responsibilities, nomenclature, changes, and version history.

3.2.2 Meeting-Driven Resolution with Limitations. When collaboration issues arise, participants typically rely on (mostly remote) meetings (12/23) to resolve them (Figure 2C). Issue trackers like Jira or ShotGrid, as well as shared documents like Notion, Google Docs, or Confluence are also commonly used. Convenience drives this preference, as **F10** notes: *"We use Zoom meeting since the team is all over the world"* and **FI1** states: *" It's just easier to get everybody in one room [in Zoom]."* **FI3**, **FI5**, and **FI6** used AI solutions to generate summaries of meeting recordings. **FI5** explained the method: *"We'll take the transcript and use ChatGPT to generate different versions of summaries. Strategy there is pretty basic."*

Some participants take proactive approaches to surface potential issues. **FI1** conducts daily reviews to catch issues early: *"We always try to have a daily, just to make sure if there were changes, how they're looking. I always want to do an A/B comparison. I always want to compare what it was and what we did, what changes we made. So, if there's things that got changed, I'd rather want to know about it before we load it into the volume stage and start rolling cameras."*

> **Design Goal 2 (DG2):** Proactively deliver relevant knowledge (instead of just acting as a reference place for the knowledge) from meeting histories at the moment of need during professionals' work.

3.3 Onboarding Collaborators to a 3D Scene

Onboarding as a Significant Collaboration Challenge. We define "onboarding" as not merely an activity of joining a whole new film project, but rather a more routine activity of bringing a new collaborator up to speed on a scene in a project, particularly regarding the 3D scene used for virtual production. Due to the project-driven employment nature of film production, we predicted that such onboarding would be frequent. Our findings support this (Figure 3A).

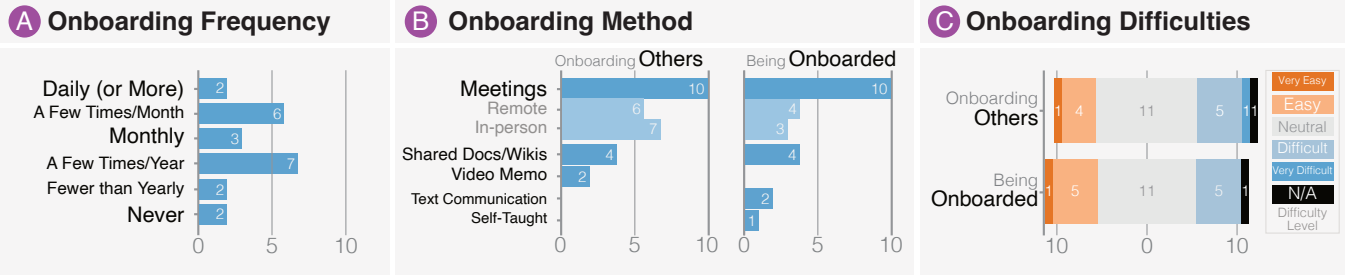


Figure 3: Formative study results on onboarding practices in virtual production. (A) shows the frequency of the onboarding activities, (B) shows the methods that collaborators use (multiple responses possible - each is out of 15/23 who reported their onboarding experiences.), and (C) shows the self-reported difficulty levels of onboarding, from the perspectives of onboarding others and being onboarded to a scene.

Since staff being onboarded start without common ground, effective onboarding is critical for project success.

Inconsistent Onboarding Experiences Reveal System Gaps. Figure 3B shows that remote video meetings are commonly used for onboarding, regardless of whether participants are onboarding others (6/15) or being onboarded themselves (4/15). However, experiences vary dramatically. While **F6** prepared personalized materials (e.g., video memos) which may help the worker being onboarded, they themselves reported that onboarding remains very difficult. Some participants perceived both being onboarded and onboarding others as challenging, with bi-polar responses overall (Figure 3C).

The quality of onboarding depends on individual organizational skills and project management, as **FI1** describes: *"We'll rewatch the lensing sessions together. We'll look at certain things and then we'll make notes. Typically, I do it because I like taking notes and have to make sure that my stuff is on point."* However, even experienced professionals struggle with poorly organized projects. **FI1** explains: *"I've worked on projects that have been so, so disorganized. It's frustrating. You're wasting so much time, you're wasting so much money, and nothing ever gets done. [...] On disorganized projects, we have to start the QA process weeks ahead of time."*

FI3 highlights how inadequate onboarding affects production quality: *"In medium and small productions, untalented professionals would make a huge setback for the whole set. Because the whole time you are trying to fix things and figure things out in a production time. It has to be a preparation time before."*

Most concerning, onboarding is often minimal or nonexistent. **F23** noted: *"Honestly, onboarding is a luxury—usually a project is delivered and it's up to the team to ingest and learn the project well enough to run it during production."* **FI5** describes a common approach: *"[For onboarding,] I usually just give them the file. Dig around, and figure it out. Let me know if you have questions."*

Design Goal 3 (DG3): Automate common ground establishment to ensure consistent, effective onboarding experiences across all projects, reducing dependency on individual organizational skills and manual knowledge transfer processes.

4 *GroundLink*

We designed *GroundLink*, a Unity add-on,[9] to operationalize our design goals (**DG1–DG3**) and to study their impact on virtual production professionals' common ground formation (**RQ2**) and confidence/ease when editing 3D scenes (**RQ3**).

GroundLink detects user interactions in the Unity editor and identifies when edits diverge from prior decisions captured in meeting recordings (Figure 1). It then surfaces in-scene feedforward cues in the editor to reveal potential consequences of the current action, and simultaneously highlights the linked decision/comment on the dashboard timeline while seeking the meeting video to the exact segment. The dashboard, editor, and feedforward cues are bidirectionally synchronized (example scenarios: Figure 4).

4.1 Meeting Knowledge Dashboard for Capturing and Reviewing Decisions and Comments (DG1)

GroundLink provides a temporal dashboard that consolidates meeting-derived, verbal knowledge, and meeting recordings in one place (Figure 1E). A zoomable minimap displays changes across meetings as a cohesive spatial overview; a filterable, chronologically ordered list captures decisions and comments with identities of editors and agreeing collaborators; and an integrated video player anchors each entry to its source segment (Figure 5). In doing so, the dashboard directly operationalizes **DG1** by gathering constituent knowledge needed by VP professionals.

Unlike prior production dashboards[10], entries in *GroundLink* synchronize with the rest of the application. Each entry has a mapping to an object, part of a 3D scene, and a video meeting snippet, binding every insight to concrete objects, moments, and collaborators. Filtering is likewise object, scene, and meeting-based, and the AI summary operates over exactly the currently scoped subset, yielding focused, on-point synthesis.

The minimap (Figure 5A) visualizes an object's location history as a dashed path and emphasizes selected elements. Selecting an object on the minimap filters the summary view to that object and highlights

[9]We implement *GroundLink* as a Unity add-on. We chose Unity because it is one of the game engines used for virtual production and is comparatively easy to learn [124].
[10]e.g., Autodesk Flow (https://www.autodesk.com/company/autodesk-platform/me) or Perforce (https://www.perforce.com/products/helix-core)

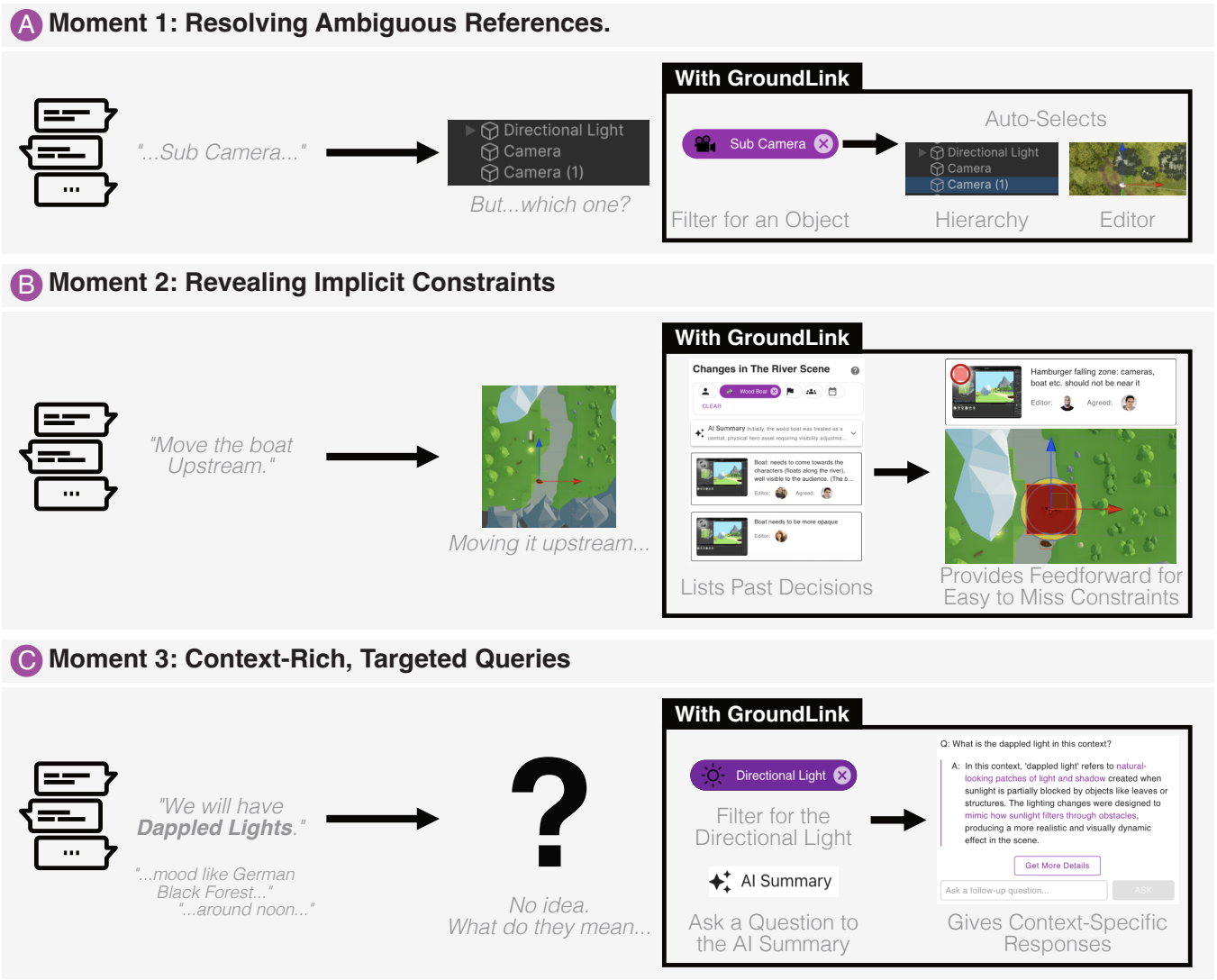


Figure 4: To illustrate how *GroundLink* supports common ground formation, consider Bob, a VFX artist newly assigned to a river scene. Traditionally, Alice (the production supervisor) would onboard him through meetings and explanations of past decisions. With *GroundLink*, Bob instead uncovers the scene's collaborative history while working.
(A) Moment 1: Resolving Ambiguous References. Bob is tasked with moving the sub camera. With *GroundLink*, clicking the object in the dashboard highlights the specific camera in the 3D editor, eliminating the need for second guessing.
(B) Moment 2: Revealing Implicit Constraints. Bob's task is to move the boat upstream, but he does not know a no-placement zone was established two meetings ago. With *GroundLink*, past decisions and a feedforward cue appear as subtle red zones in the editor, expanding and darkening as the boat nears. This helps Bob avoid placing the boat in the restricted area.
(C) Moment 3: Context-Rich, Targeted Queries. Bob is tasked with adjusting the light to create a dappled light effect, a term he does not know or understand in this scene. With *GroundLink*, he can ask the AI summary a targeted question about the scene and object of interest, avoiding unnecessary second guessing.

the corresponding entries (Figure 5B–C). All elements in the minimap and list are filterable by object, collaborator, production phase, meeting instance, and date range; when filters change, the minimap emphasizes relevant traces by increasing opacity and putting purple glow behind relevant objects while de-emphasizing others.

The summary view (Figure 5C) presents significant production decisions and comments in chronological order, with a thumbnail of the linked video snippet and a badge for the related scene element. Each entry shows a concise summary plus the editor and agreeing collaborators (Figure 5D). Hovering over a badge reveals role, name, and, for agreements, any qualifying remarks. These entries can be filtered (Figure 5B).

On the right, the video player (Figure 5E) supports rapid verification: clicking any summary entry seeks to the exact moment in the

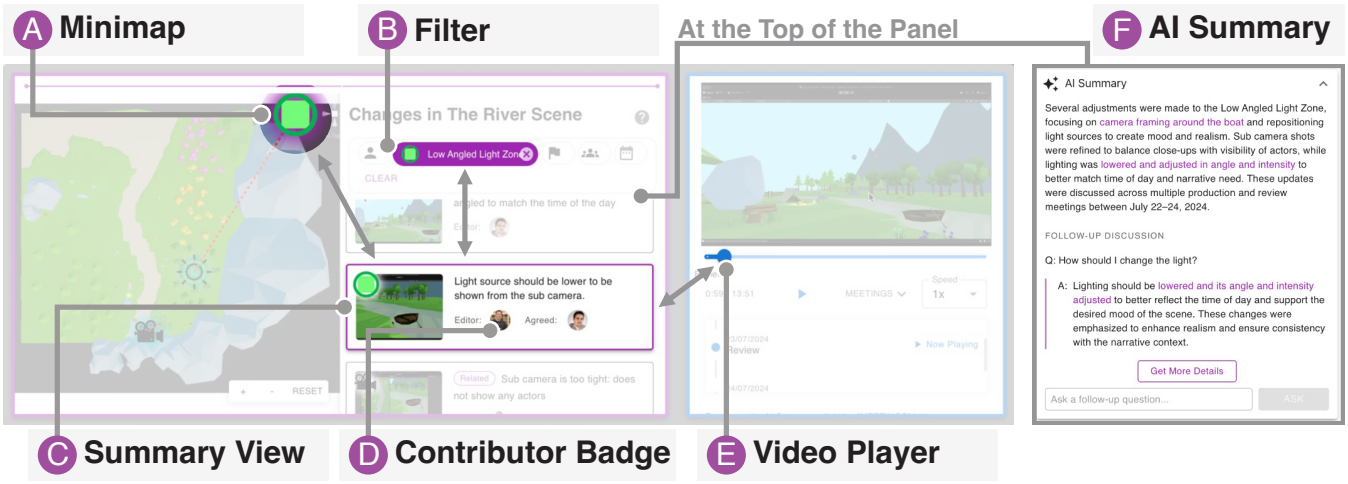


Figure 5: *GroundLink* summarizes verbal information in a sequential representation on a dashboard. (A) The minimap displays all element changes on a zoomable map. (B) Users can filter decision/comment entries by collaborator, object, meeting phase, meeting, and date range. (C) The summary view lists all changes/comments. (D) For each entry, contributor badges and agreements are attached. (E) The video player with timeline (synchronized with other panels) and meeting list allows users to watch video while reviewing decisions/comments. (F) The AI summary and chat enables quick glances and Q&A.

meeting. There is a one-to-one correspondence between list entries and meeting segments.

A collapsible AI summary (Figure 5F) sits above the list and respects the current filters (Figure 1F). Users can obtain scoped syntheses (e.g., only the "shelf" object or a collaborator "Alice"), ask follow-up questions, and jump via highlighted references to the corresponding list entries. This serves both as citations and as navigational shortcuts.

The dashboard is designed to go beyond making individual decisions more trackable, by revealing higher-level patterns about how the team works. The minimap and filters reveal how constraints accumulate over meetings, in what area in the scene, the changes happen, which movie departments tend to make particular changes/suggestions, and how often decisions are revisited. Contributor badges expose who usually edits or agrees to a given element.

We draw design rationale from Dual Coding Theory [93]. Dual Coding Theory posits partially independent verbal and non-verbal systems for interpreting two types of information. The dashboard preserves the optimal sequential representation for verbal artifacts (time-coded transcripts, decisions, comments), while other views maintain in-situ, visual representations of scene state.

4.2 Constraint-Aware Feedforward for Informing Creative Decisions and Comments (DG2)

Operationalizing **DG2**, during direct manipulation in the Unity editor, *GroundLink* overlays constraint zones (e.g., areas to avoid, preferred placements) as in-situ, visual feedforward so that relevant prior decisions appear during work rather than after-the-fact (Figure 1A). In the inspector, when a property diverges from a previously agreed guideline (e.g., a color choice), the corresponding control draws subtle attention (Figure 1D), presents a brief rationale, and offers a one-click jump to the linked decision on the dashboard.

Unlike existing tools that deliver post-hoc or cross-modal feedback (e.g., textual warnings about a 3D placement) [75, 76], *GroundLink* provides in-situ, same-modality feedforward: cues are rendered directly in the 3D editor at the locus of manipulation and update continuously as the edit unfolds. Classic feedforward communicates a definitive action (e.g., "slide to unlock") [116], whereas decisions encoded in meeting artifacts are inherently uncertain and context-dependent. If the system could infer the exact edit, it would simply apply the change and request a simple "accept"; in practice, creative judgment remains essential. *GroundLink* therefore treats ambiguity as a design resource [38]: it modulates opacity, animation, and band thickness of circles to convey proximity, and links each cue to its provenance in the meeting record for immediate verification on the dashboard. This predictive, same-modality guidance with explicit uncertainty and provenance aims to reduce context switching while avoiding false feedforward [65].

Positional constraints are communicated with circular zones that appear only in the editor view[11] (Figure 1A). Encouragement zones (Figure 6A) use a green ring and a subtle arrow from the relevant object toward the target area; as the object approaches, the ring's opacity and animation intensity diminishes and banding simplifies, becoming transparent when the object is properly within the zone. Restriction zones (Figure 6 B) convey the opposite gradient: salience increases as the object nears the restricted area and fades with distance. Each zone includes an icon at its center indicating the movie department that proposed the constraint (Figure 6 C), supporting transactive memory about the source of guidance. Interacting with a zone reveals the corresponding entry on the dashboard's summary list and seeks the meeting video to the linked segment. The summary view also shows history of decisions and comments made for the zone.

[11] I.e., not on the camera view

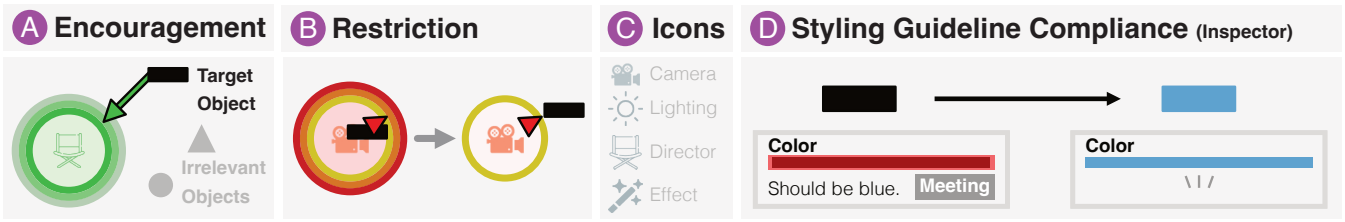


Figure 6: *GroundLink* surfaces in-situ feedforward derived from meeting recordings: (A) encouragement zone, (B) restriction zone, (C) department icons, (D) inspector feedforward for styling compliance.

Styling constraints are surfaced in the inspector (Figure 1 D). When a user selects a value that conflicts with a prior decision (e.g., having a black "shelf" when the agreed color is blue), the relevant control (e.g., color picker or opacity slider) gently blinks, accompanied by a short explanation ("Should be blue") and a meeting button that navigates to the originating discussion on the dashboard (Figure 6 D). In all cases, *GroundLink* provides uncertainty-aware feedforward tied to verifiable sources, surfacing knowledge at the moment of need while preserving creative agency.

4.3 Cross-Modal Synchronization for Linking Meetings and Editor Scene Context (DG3)

GroundLink keeps the dashboard, the Unity editor's hierarchy, the scene view, and the video player synchronized (Figure 1B). Unlike conventional production dashboards that isolate comments, meeting recordings, and the 3D editor, *GroundLink* creates fine-grained, cross-modal links between timestamped meeting snippets, change/comment summaries, and concrete 3D objects. These links transform ambiguous references (e.g., "sub camera," "this zone here") into unambiguous selections in the scene.

Users no longer have to guess which scene element corresponds to which meeting artifact, or vice versa. This synchronization happens proactively, operationalizing **DG3** by automating common ground establishment during work. This means that if a user clicks or interacts with any part of the dashboard, the rest of the dashboard (minimap, summary list, filter, and the video player) and the editor (object selection in the hierarchy and zooming into the object in the editor) are synchronized, and vice versa.

Dual Coding Theory posits partially independent verbal and non-verbal systems, with people making referential connections between two systems [93]. Production workflows often already separate the channels e.g., verbal artifacts in notes/transcripts and non-verbal artifacts in the scene, but leave the referential links implicit. This forces users to perform mental indexing across channels, especially when terminology is inconsistent or spatial configurations have changed. *GroundLink* preserves each channel's natural representation (sequential, time-coded verbal material; spatial, in-situ visual state) while externalizing the referential links between them. By turning phrases like "the hero light" into concrete, navigable selections and timepoints, the system reduces the burden of cross-modal mapping and promotes faster, more reliable interpretation without collapsing one channel into the other.

To further support **DG3**, the same mapping works in tandem with the feedforward that surface relevant prior discussion at the moment of need. For example, if a property mentioned in meetings is out of compliance, the corresponding inspector field draws attention and provides a one-click jump to the associated insight in the dashboard. This resembles the concept of "information scent" that encourages follow-up rather than interruption [20].

4.4 Implementation

GroundLink operates in two phases. The main phase consists of realtime interaction, which is the focus of this paper and is used for the user study. For this to work, another phase is required, which is a onetime preprocessing phase. While essential for the system, this preprocessing phase is not the main focus of this paper and some part is implemented using a Wizard-of-Oz method. We explain each phase in detail.

4.4.1 Realtime Interaction. Figure 7 illustrates how the two components, which comprise the add-on are configured: (1) an Electron[12] application built with React[13] and Material UI[14] that hosts the temporal dashboard and video player, and (2) Unity[15] custom Editor windows[16] that render in-editor/inspector feedforward, and log interactions. The components communicate via WebSockets, with a Node.js[17] server embedded in the Electron app. It publishes selection, filter, and playback events and editor state changes. Concretely, the Unity custom Editor window runs in the editor update loop and monitors the current editor state (e.g., selected game objects and inspector properties). On each update, it compares this state to a cached snapshot and, when it detects a change, emits an event describing the interaction. These events are streamed over the WebSocket connection to the server, which forwards them to the dashboard application and, when appropriate, requests summaries and explanations from the AI service (Real-time AI features through OpenAI's GPT-5-chat API[18]). In parallel, the Unity Editor window checks the updated state against a JSON file containing the constraint information and triggers feedforward effects in Unity by creating or removing game objects that visualize these effects. When the user clicks elements on the dashboard, it emits events back through the server to the Unity Editor window, which updates the Unity state (e.g., changing the selected object or zoom level), keeping the scene view and dashboard in sync.

[12] v36.5.0; https://www.electronjs.org
[13] v19.1.0; https://react.dev
[14] v.7.1.1; https://mui.com
[15] v6.1 (6000.1.2f1)
[16] https://docs.unity3d.com/6000.2/Documentation/Manual/UIE-HowTo-CreateEditorWindow.html
[17] v22.16.0; https://nodejs.org/en
[18] https://platform.openai.com/docs/models/gpt-5-chat-latest

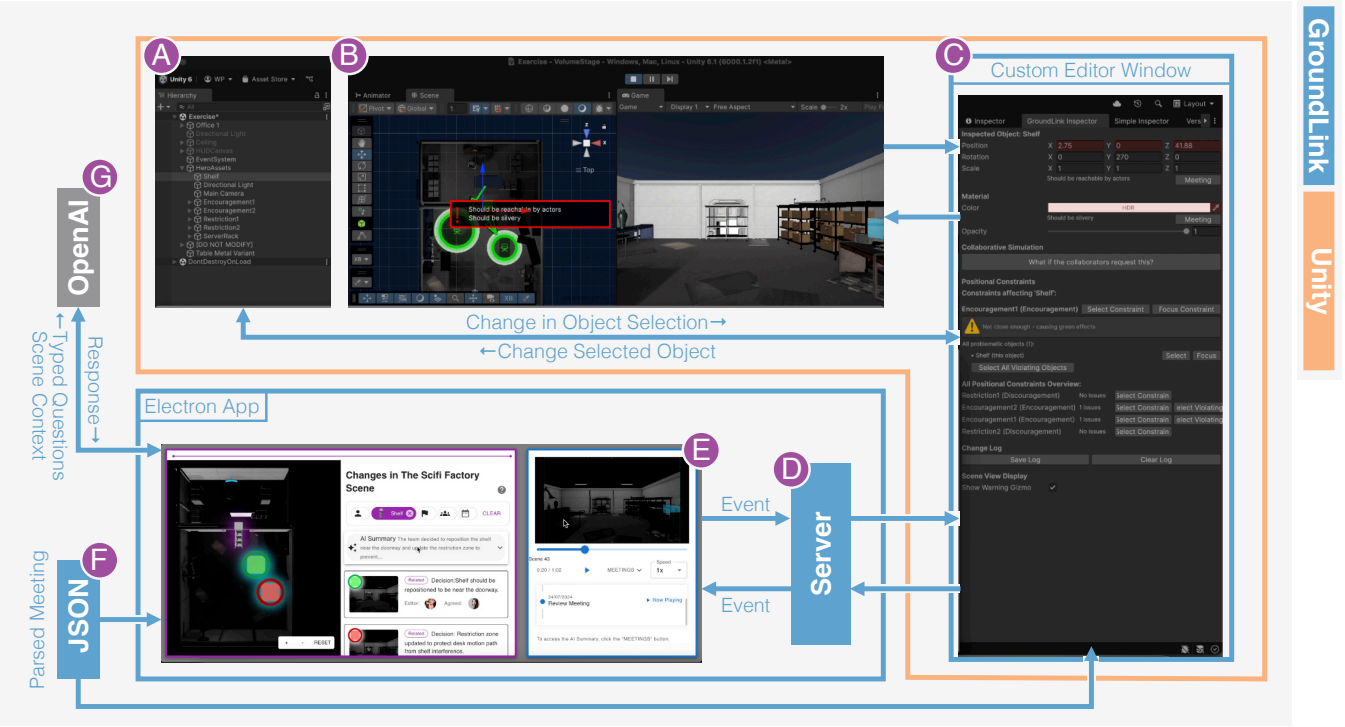


Figure 7: In the realtime interaction phase, a (junior) worker works in Unity (Figure 1). In Unity, there is a hierarchy view (A), scene/game view (B), and *GroundLink* is integrated on the (C) as a Unity custom Editor window. This window collects interactions from other Unity views (e.g., selecting a game object, changing the position/property of an object) and can also modify them (e.g., select another game object, put feedforward effects). To do this, the window communicates with a server (D) implemented in an Electron app via WebSocket. The same server also communicates with a React-based application (E), which interprets Unity changes and fetches updates to the Unity custom Editor window. In doing so, it uses the JSON (F) generated in the preprocessing phase. The React-based application uses OpenAI's API (G) to provide the real-time AI features. Note: The tools/processes we developed are marked with blue color, whereas the existing services/tools are colored in orange.

We designed *GroundLink* with a two-component architecture to help users continue working with the content-creation tools they are already familiar with, while also providing advanced features such as integrating video APIs, AI APIs, data import, and animation effects, not supported by Unity or difficult to implement with Unity.

4.4.2 Onetime preprocessing. For the user study, we used pre-indexed meeting recordings with timestamped transcripts, and scoped our exploration to knowledge consumption rather than parsing them realtime. We therefore manually annotated transcript-to-scene mappings to isolate interaction design effects rather than confound results with automatic NLP accuracy. Figure 8 illustrates this onetime, preprocessing. This is done through a parsing process we developed, which the researchers match data from different sources[19] by manually clicking the instances of change or snippets of transcripts[20]. We discuss how the parsing process[21] could be implemented in future work, later in the paper (subsection 7.4). This process of manually matching information from different sources can be characterized as a Wizard-of-Oz (WOz) method. This strategy allows researchers to build systems when current technology falls short of meeting usability requirements, so that evaluation of the rest of the system[22] can be conducted [60, 77]. WOz is a popular, validated method in the human-AI interaction community [32, 39, 47, 69, 111].

5 Exploratory Comparative User Study and Expert Evaluation

We conducted two complementary studies to evaluate *GroundLink*'s effectiveness in supporting virtual production workflows, with emphasis on newcomer experience but beyond as well. First, we performed a comparative user study (N=12) to assess how meeting-derived knowledge affects users' ability to form common ground with remote teams and complete VP tasks. Second, we conducted an expert evaluation (N=5) with VP professionals to gather domain-specific insights. Together, these studies address our research questions about common ground formation (**RQ2**) and task confidence (**RQ3**). Both studies received review and approval from our institution's internal

[19]Transcripts of the recordings, Zoom annotations including whiteboard and mouse movements, changes in the Unity Editor
[20]We developed a basic web application tool for this.
[21]We understand that the parsing process, which is part of a broader authoring process, is an important part of the system, despite not being the focus of this paper.
[22]Which is the scope of this paper.

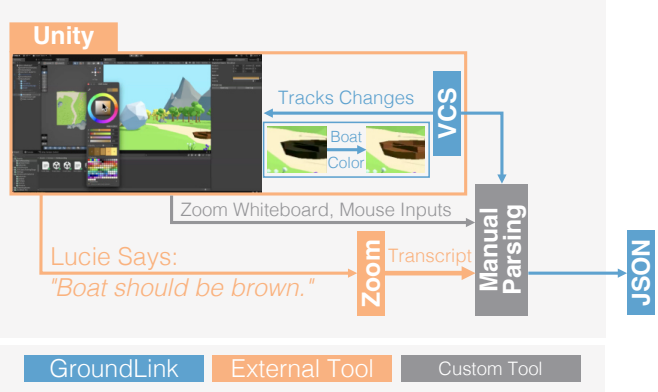


Figure 8: During meetings, the Version Control System (VCS), a component of *GroundLink* implemented as a Unity custom Editor window, tracks changes in the scene, such as changes to transforms (position, scale, rotation) and properties (color, opacity, field of view, etc.). Zoom Whiteboard/mouse inputs are also collected (manually decoded), and speech from the Zoom meeting is transcribed through Zoom and collected. All collected data is manually matched by the lead researcher using an Authoring Tool, a custom-made web app that presents the changes and helps match changes and mentions. This process generates JSON files. The JSON generated is then used in the next phase. Note: Grey colored part is also developed by us, but is not the focus of this paper.

ethics review process. Participants received a US$100 gift card for their participation, for each study.

5.1 Exploratory Comparative User Study

The goal of the comparative user study was to evaluate how meeting-derived knowledge presented through *GroundLink* can help or hinder users' virtual production workflows, and whether it helps users perceive that they form common ground with existing remote teams.

5.1.1 Participants. We recruited 12 participants[23] (completely new) through an internal mailing list. Unlike the formative study, we recruited participants with experience using 3D manipulation tools (Unity, Unreal Engine, Maya, or similar)[24]. We broadened eligibility to reach participants with more diverse backgrounds as virtual production is a growing area with people from diverse areas of expertise joining the industry [57, 124]. Therefore, this study sought to evaluate the capabilities of the system in isolation, while offloading VP-specific validation to the expert evaluation.

5.1.2 Materials and Methods. To implement the design concepts in a prototype, we created three virtual production scenes (Figure 9). These scenes were co-created by two professionals with filmmaking experience working with two of the authors; these professionals did not take part in any other survey, interview, or study. We involved filmmakers to provide ecological validity in the recordings. This co-creation process resulted in three Zoom[25] video meeting recordings per scene[26], each approximately 15 minutes long. The recordings were semi-scripted, enabling comparable scenes with similar difficulty and level of detail for the edit process. For the river and forest scenes, the recordings established prohibited zones[27], styling guidelines (e.g., opacity and color), and target zones for camera or other elements (e.g., tree, mushroom). The recordings used mouse cursor circling[28], Zoom whiteboard annotations, and placement of placeholder cubes in the scene to indicate approximate size and position of each zone, alongside verbal descriptions of styling guidelines. Researchers ensured that each of the aforementioned methods was present in both sets of recordings.

After the recordings, one of the researchers manually parsed the contents, adding the mentioned zones in Unity and encoding styling guidelines in a JSON file. These scene "constraints" had a 1:1 mapping with discrete comment/change items from the recordings; each item was identified and encoded in a JSON file.

We installed *GroundLink* with the scenes on an Intel Mac running macOS[29]. Participants took part in an in-person study using this computer. The study lasted approximately 90 minutes.

5.1.3 Task. We selected Microsoft Copilot[30] on ClipChamp[31] as the baseline (with GPT-5 mode enabled), as it represents current best practices in AI-assisted meeting review. The baseline supported raw video meeting recordings (with standard video playback controls including playback speed), AI summary and chat about the video (with links to specific video segments from the summary), transcript, and search features. As reported in our formative study (subsection 3.3), scene onboarding in VP commonly occurs through meetings, shared documents, and video memos. While the typical baseline would be simple video players with transcripts/notes, we chose Copilot to provide a stronger comparison to evaluate *GroundLink*'s in-situ feedforward and spatial mapping features against a robust text-and-video baseline with AI capabilities, rather than creating an artificial comparison with minimal tool support. This allowed us to specifically evaluate the added value of *GroundLink* features. We ran Copilot on a separate Windows laptop to avoid window switching and provide a smoother study experience. The independent variable for this study was the tool difference (treatment: use of *GroundLink*). Participants tried both tools throughout the user study in this within-subjects design.

Figure 10 shows the study sequence. After obtaining informed consent and providing a brief introduction via presentation slides, we began each session with a task. In each task, participants were assigned a scene (river or forest) and a condition (baseline or *GroundLink*). We then provided a 10-minute hands-on, activity-based walkthrough of

[23] **Age:** 33±10.6 (mean; one did not report); **Gender:** 7 M, 4 W, 1 Unknown; **Occupation:** 1 post-doctoral researcher, 2 research scientists, 1 software developer, 1 associate research & design engineer, 1 software development manager, 1 senior software developer, 2 software engineers, 3 user experience designers.

[24] 5 participants had used the tools for architecture, 7 for media and entertainment, and 6 for manufacturing. Participants had an average of 5.1±3.92 years of experience with the tools.

[25] https://www.zoom.com

[26] River and forest scenes only. For the exercise scene, we scripted the meeting and simulated the audio using ElevenLabs Text-to-Speech (https://start.elevenlabs.io). We scripted the exercise scene meeting to provide a clear, controlled illustration of the system's functionality within a short period of time. These scripted meetings helped us efficiently cover all the features.

[27] Actor zone where cameras should not go; hamburger-falling zone for the river scene; and a monster-appearance zone for the forest scene where no elements should enter.

[28] Moving the mouse cursor in a circular motion around the point of interest to emphasize.

[29] macOS 15.6; CPU: 2.4 GHz 8-Core Intel Core i9; RAM: 32 GB; GPU: AMD Radeon Pro 5500M 8 GB

[30] https://copilot.microsoft.com

[31] https://clipchamp.com/en/

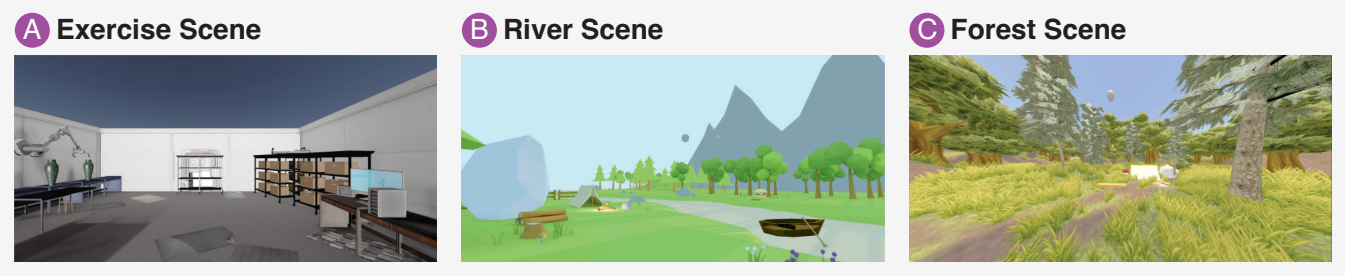


Figure 9: Three pre-assembled scenes used in the study: (A) exercise scene (a sci-fi office) that was used for demonstrating functionality to participants, (B) river scene in which the actors plan to use a boat to escape from food falling from the sky, only to realize the boat is an illusion, and (C) forest scene where actors eat fruit and talk around a campsite, followed by a monster appearing from the forest.

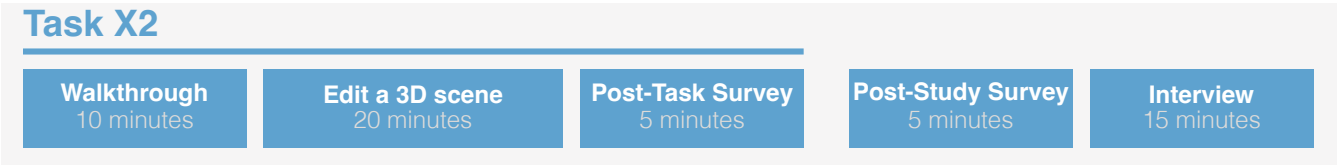


Figure 10: Task sequence for the comparative user study. It has within-subjects design with counterbalanced conditions.

the assigned condition using the exercise scene. Next, participants had 20 minutes to complete eight tasks (Table 4, Table 5). We chose the number of tasks and time budget based on similar experimental setups [89], adjusted for the unique difficulty of our VP-specific tasks. After 20 minutes, participants completed a survey consisting of the UES-SF [92] to measure user engagement, NASA-TLX [44] to measure workload, and a post-task questionnaire (Table 3) measuring shared mental models (inspired by van Rensburg and colleagues [115]), transactive memory systems (one question per dimension, modified for this study [70]), and perceived conflicts (inspired by Jehn [51]). Because our focus was exploratory, and because the experiment happened under a tight timeline[32], we did not utilize the full, validated (but long to administer) SMM and TMS instruments. Instead, we designed a small subset of customized questionnaire items, adapted from prior work, as a means of quantifying subjective responses related to shared understanding and division of expertise. These items are intended as coarse, indicative measures that complement our qualitative analysis rather than as definitive psychometric evidence. For more conclusive claims about SMM and TMS, future work should employ the complete standardized instruments and larger samples. We measured engagement and workload to determine whether *GroundLink* impacts task performance. Through the post-task questionnaire, we measured the quality of the common ground that was formed (**RQ2**), and through the confidence ratings, we measured confidence in task completion (**RQ3**).

We repeated this task once more with the scene and condition not yet tried, counterbalancing both scene and condition across participants to control for order effects. After both tasks, participants completed a post-study survey (including preference questions) and a 15-minute semi-structured interview (See Appendix D for the list of questions). The total study took approximately 90 minutes. Prior to the main study, we conducted a pilot session to refine the protocol, which resulted in adjustments to task timing and instruction clarity.

5.1.4 Measures and Analysis. We collected survey responses using Likert scales: UES-SF (5-point Likert scale), NASA-TLX (7-point Likert scale), post-task questionnaire (5-point Likert scale, Table 3) and task completion and confidence (5-point Likert scale plus "Did not attempt"). We also collected system usage data (element movements, feature usage), but given the exploratory nature of the study and the creative nature of VP (where there is no strictly correct response), we focused our analysis on survey responses.

Considering the ordinal nature of these responses, we performed Wilcoxon Signed-Rank tests using scipy 1.7.3 [117]. We report means for each condition to better illustrate the distribution, particularly given our sample size. The choice of our test makes Type I errors less likely when the distribution is possibly skewed or unknown [14], potentially enhancing the rigor of our analysis. To provide a complete picture of results, we supplement significance testing with 95% bootstrapped confidence intervals (CI) on means, calculated using arch 5.3.1 [104] in Python. We report all test statistics and CIs in Table 6. Given the exploratory nature and multiple outcomes, we center interpretation on directions and means rather than significance thresholds. We also performed preference testing, with results reported descriptively.

For qualitative analysis, we conducted 15-minute semi-structured interviews at the end of each study session. All interviews were recorded, transcribed, and analyzed using Dovetail[33] following thematic analysis procedures [11], using the same method used in our formative study (subsection 3.1).

[32]This was because participants were high-profile.

[33]https://dovetail.com

5.2 Expert Evaluation

The goal of the expert evaluation was to evaluate the specific benefits or limitations that *GroundLink* could provide from the perspective of VP professionals.

We recruited 5 experts, including participants from our formative study who had not participated in the comparative study, and additional participants through snowball sampling. Recruited expert participants are denoted as **E***n*, all with extensive[34] virtual production experience (demographics: Table 1).

The expert evaluation followed a technology probe approach [48, 95] and was conducted remotely via Zoom. Participants gained remote control of the researcher's computer to experience *GroundLink*. Before each session, we manually verified that video streaming latency and control responsiveness were adequate for the evaluation[35].

For each participant, the researcher provided the same study introduction presentation as used in the comparative study. The researcher then provided a tutorial using the exercise scene (identical to the comparative study), gathered first impressions on the features, and asked participants to experience either the river or forest scene (randomly assigned, using the same scenes from the comparative study). Participants completed 2-3 tasks with close researcher guidance[36].

Each session concluded with a 30-minute in-depth, semi-structured interview (See Appendix E for the list of questions). All interviews were recorded and transcribed using Dovetail[37], and we conducted thematic analysis on these recordings [11] using the same method used in our formative study (subsection 3.1). The study lasted approximately 60 minutes.

6 Findings

We found that the comparative study exerted appropriate task load and provided an engaging experience. While both conditions were generally comparable on engagement and task load, we observed some differences. For the UES-SF aggregate (mean; Figure 11A), *GroundLink* had a higher mean ($W = 17.5, p = 0.092$). For NASA-TLX, we did not detect a statistically significant difference ($W = 23.0, p = 0.23$), though the mean favored *GroundLink* (lower workload). At the item/subscale level, TLX-Performance (reverse-scored) had significantly lower score for *GroundLink* (Figure 11B; $p = 0.042$). For UES-SF, Focused Attention item (FA-S2 *"The time I spent using* GroundLink *just slipped away"*) had significantly higher score ($p = 0.024$) for *GroundLink*. FA-S3 (*"I was absorbed in this experience."*) also had higher score for *GroundLink*, but not significant ($p = 0.077$). The Reward item RW-S2 (*"My experience was rewarding."*) had significantly higher score for *GroundLink* ($p = 0.011$), with the largest mean difference. These were consistent with CIs (Figure 11C).

From this understanding of engagement and task load, it shows that our comparative study has provided appropriate ground for comparing both solutions, showing potentially more focused attention and a higher sense of reward and perceived performance for *GroundLink*.

When post-task questionnaire responses were aggregated, it showed a statistically significant difference between the two conditions ($W = 8.0, p = 0.012$). In the rest of this section, we report how experts and comparative study participants found the experience different, to explore the potential impacts of *GroundLink* in VP scene assembly (editing) task.

6.1 *GroundLink* Helps Close the Gaps in Common Ground (RQ2)

Our **RQ2** asked: How does surfacing meeting-derived knowledge support a team member in virtual production projects form common ground with existing team members? From the literature research, we found that supporting the formation of Shared Mental Model (SMM) and Transactive Memory System (TMS) would be important in helping users form common ground with the existing team members. Considering the complimentary nature of two theories, we analyzed the results for each separately, and we found that with *GroundLink*, both experts and comparative study participants began perceiving to form SMM and TMS with the existing team, while foreseeing fewer conflicts with them.

6.1.1 *Participants Perceive Better SMM with* GroundLink.

SMM is an important part of common ground. We found indications to suggest that with *GroundLink*, both experts and comparative user study participants perceive themselves to be more on the same page with the collaborators in meetings.

EXPERT

All experts suggested the benefit of *GroundLink* in this regard, in particular, **E4**: *"It would be amazing that we are, as a team, on one page. We are landing and standing on one page [with* GroundLink*]."* **E5** has seen the potential of *GroundLink* in enhancing existing practice of note management in VP: *"Our problem was that after the meeting, someone took notes [...]But again, the context, what people were seeing inside of the volume was missing."* Experts (**E1-4**) generally found AI summary helpful in this regard, with particular emphasis on the interaction it provides. As **E3** summarized: *"You can click these highlighted texts and it just takes you right to that part of the meeting."*
Experts also found that *GroundLink* could help surfacing important information from past meetings. **E2,3,5** mentioned the nature of meetings in creative settings, and how *GroundLink* can help collaborators obtain important information without loss, as **E3** described: *"A lot of things can get lost in translation and things can be misinterpreted or maybe once time goes by, you start to second guess certain things. But when* GroundLink *integrates the meetings, it's really useful for [preventing this]."'* **E5** emphasized the length of meetings in VP: *"If you just think about a two hour recording and you have multiple people talking. It's very cumbersome to find the right spot."*

COMPARATIVE STUDY

Similarly, comparative study participants also has shown indications to suggest which they perceived that they had a better shared mental model with the participants in the meeting recordings with *GroundLink*. For all survey questions relating to SMM, the means were higher for *GroundLink* (Figure 11D), with significantly higher score for *"I understood why earlier collaborators made the design choices I saw on this screen."* ($p = 0.026$) and *"I was aware of the constraints or conventions that applied when I edited the scene."* ($p = 0.036$).
Comparative study participants perceived that *GroundLink* helped to obtain the information needed for understanding the team. For **P10** and **P11**, the summary view contributed to general understanding of

[34] 5.6±2.51 years of experience with 14.0±9.06 virtual production projects; Roles included: Executive Producer, VP Supervisor, VP Artist, Compositor, and VP Researcher. All men.
[35] For **E4**, intermittent latency issues were reported during the session.
[36] For **E5**, researcher demonstration was done instead due to technical difficulties.
[37] https://dovetail.com

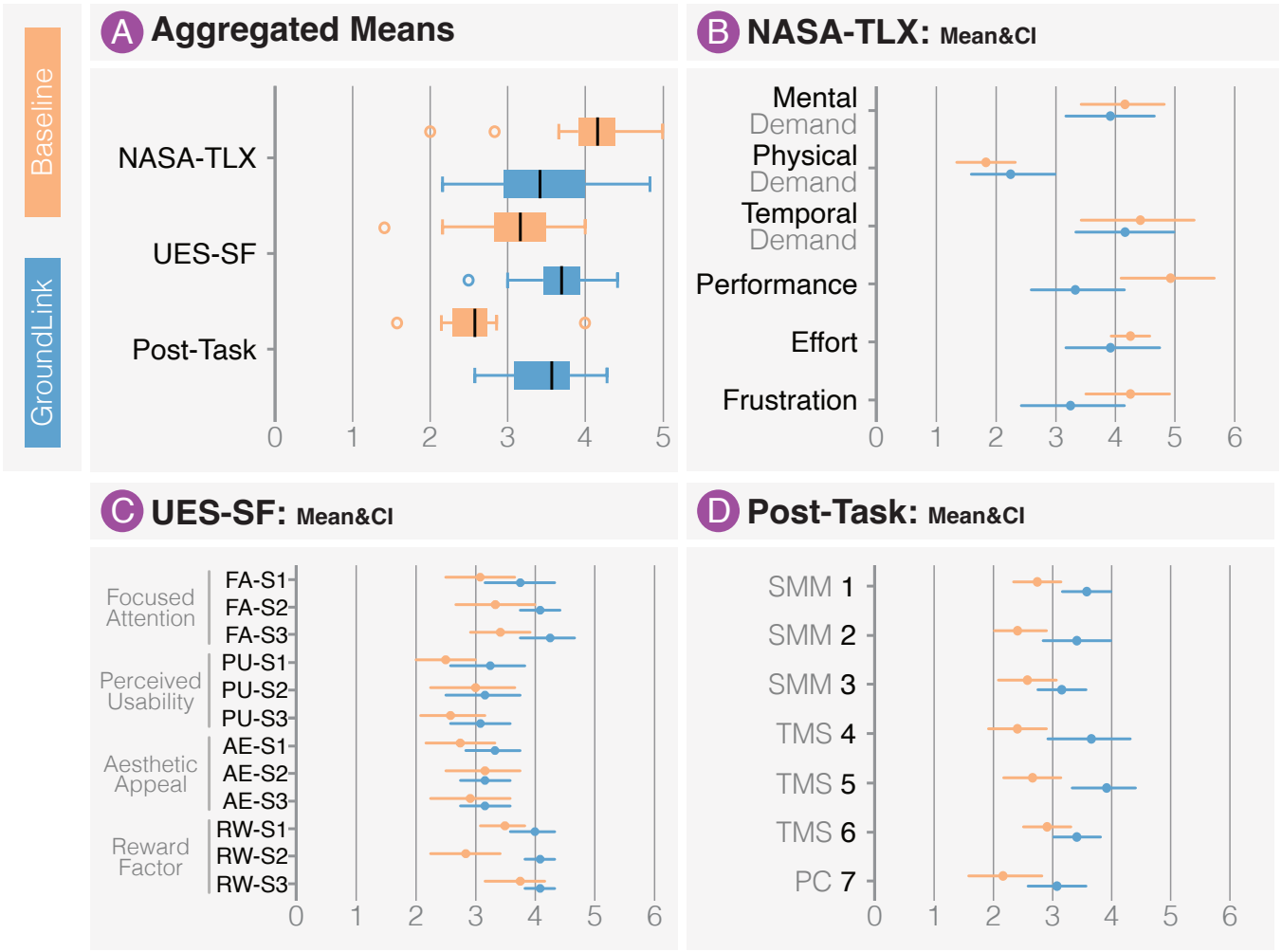


Figure 11: (A) Box plot of aggregated scores for NASA-TLX, UES-SF and Post-Task survey. (B) Mean and 95% CI for NASA-TLX (lower indicates lower workload). (C) Mean and 95% CI for UES-SF. (D) Mean and 95% CI for the post-task questionnaire (Table 3), which aggregates responses corresponding to SMM, TMS, and perceived conflict-related questions. Higher values indicate better outcomes for all measures except NASA-TLX, where lower values indicate lower workload (better outcomes). Reverse scoring is used to provide this consistency. SMM: Shared Mental Model, TMS: Transactive Memory System, PC: Perceived Conflict.

COMPARATIVE STUDY

the edits. Similarly, in understanding the mental model of existing team members, **P4,7-9,11,12** found the way that *GroundLink* aggregated all the changes and comments, and displayed them in their work application helpful. **P8** suggested: *"I didn't have to go back and forth. I could stay within my working context."* **P1,2,8,11,12** found that *GroundLink* helps in understanding more precise implications of past meeting discussions. **P1** mentioned : *"I noticed how much leeway and space there might be [in editing a scene with* GroundLink*]. [...] And I felt like having a visual representation of these ranges made it much easier for me than [having] nothing here, or trying to extract it from text and then resolving these constraints in my head."* Like experts, comparative study participants also found AI summary useful in general. **P11** used it for obtaining the overall context, and **P3,11,12** found the insights from it is precise. Similar to expert, **P2** also appreciated the idea of surfacing important decisions/comments in work application.

OVERALL

Our results indicate that proactive provision of applicable information in-situ, at the right time could have an important role in helping users of the creative software in VP to form a Shared Mental Model, extending how the model could be concretely instantiated with the support of more advanced user interface.

6.1.2 *Participants Perceive More Effective TMS with* GroundLink.

EXPERT

E1,2 particularly appreciated that the interface shows who edited/agreed the particular decisions/comments which can be useful for further discussion on the decision. To support such communication more efficiently, **E1** suggested that *GroundLink* can be equipped with more advanced share features for each comment/change item: *"I think it would be easier just to have [comments/changes] numbered as well. And then have a share button that would allow you to take those notes, drop them into an email with the context and maybe even a little*

EXPERT *image that you're sending out to an artist or somebody and say 'I need you to take a look at this right now.'"*

COMPARATIVE STUDY The means for all survey questions relating to TMS were higher for *GroundLink*, with significantly higher scores for *"I could easily tell who I would need to consult about a specific asset or decision."* ($p = 0.032$) and *"I trusted that the information I saw about prior decisions was accurate and credible."* ($p = 0.015$) (Figure 11D). This finding is consistent with the findings from interviews, with **P3,7,8,10,11** emphasizing the value of showing the editor/agreer on the interface, as **P7** described: *"Each item has the person assigned to it, and probably this person is the person who made the statement, and that's why this is the person we can reach out to [if we disagree/have a question]."*

OVERALL This result suggests that the important feature to support TMS can be to provide the right relevant profiles of professionals in-situ, on the atomic chunk of information that collaborators consume.

6.1.3 *Participants Foresee Fewer Conflicts with the Existing Team Using* GroundLink.

EXPERT **E1** foresaw reduced conflict with the existing team if *GroundLink* is used in VP: *"I do see a lot of efficiency and a lot more things getting done correctly the first time with [*GroundLink*] because it's right there in front of you. [...] It basically walks you through [the task] and holds your hand."*

COMPARATIVE STUDY We also observed a score indicating reduced perceived conflict (PC) with *GroundLink* compared to the baseline ("I worried that my changes might conflict with earlier decisions."[38]; $p = 0.062$; Figure 11D), but not significant.
Half of the participants (**P5,7-10,12**) worried less about potentially undoing others' work or not following previous decisions with *GroundLink*, as **P12** suggested: *"On Copilot it's strictly me who has to figure out what to do and then I might be messing something up. But with* GroundLink*, I think I offloaded that part of my brain that would worry about that."* Also, most participants (**P1-4,8-12**) thought the scene edited with *GroundLink* would be accepted by the existing team that created the meeting recordings. In contrast, only **P5,6** thought the scene edited with Copilot would be accepted. **P9** compared their confidence: *"With* GroundLink*, 75% versus with Copilot, it's 40%. Because there is more context with* GroundLink *and then the bounding boxes [exist]. Versus [on Copilot,] you have to piece together."*

OVERALL These results suggest that forming a common ground could result in better outcomes in terms of perceived performance and perceived reduction of conflicts.

6.1.4 *Challenge: Can Make [Junior] Professionals Too Reliant.*

GroundLink aims to provide powerful features to support forming common ground with the team, which is a realm of communication skill. While the aim is to support, some participants raised concerns about it potentially replacing the skill.

EXPERT **E1** thought *GroundLink* limits junior professionals from learning how to form common ground: *"If a junior is working on a project with [*GroundLink*] and then they go to another place and they don't have this tool, it's going to be like, 'oh, well, I don't know what to do.'"*, but regardless suggested that *GroundLink* could be more agentic, by autonomously generating edits and provide accept/reject options to users. Similarly, **E2,3** were concerned that *GroundLink*'s feedforward might result in professionals blindly following the system, as **E2** described: *"I wonder if the person will become too reliant. And then, you'll notice the first thing I did was start ignoring information. And I think a young person would absolutely do that same thing."*

COMPARATIVE STUDY Consistent with experts, **P8,9** also suggested that *GroundLink* could perform more agentic edits, which might be inconsistent with the capability of *GroundLink*. **P9** mentioned: *"Maybe instead of us dragging [elements in scene],* GroundLink *could ask for possible options and then [users] accept it and then see how it looks rather than [users] moving it."*

OVERALL These perceptions may be due to perceiving *GroundLink* as too powerful. The challenge here seems to be communicating the capabilities of *GroundLink* which remains an open question.

6.1.5 *Challenge: Authoring Process.*

In the current implementation of *GroundLink*, researchers manually authored the comments/changes data, but participants were not explicitly told about the authoring process.

EXPERT **E3** thought the authoring process could be a potential barrier: *"If it works exactly how it does here [in the study session], that's great. But I'm just suspicious of what it takes for it to get to this."* Similarly, **E1,5** suggested that the nature of feedback in video meetings demand a mediator before interpreting the suggestion. **E5** said: *"In the film industry, when you receive feedback by filmmakers, you might not want to use it exactly word by word [as they might not understand low-level details]. But also sometimes the feedback can be very harsh. So our job is to translate it in a way that makes sense to the artists, without hurting anyone's feelings."*

OVERALL Our research focused on exploring the the knowledge consumption process, and as such, authoring process was not part of our scope of exploration. However, future research seems necessary for the authoring process.

6.2 *GroundLink* Enabled More Confident, Easier VP Scene Editing (RQ3)

Our RQ3 asked: How does acquiring knowledge from past meeting recordings affect team members' perceived confidence and task difficulty? From both experts and comparative study participants, we found that *GroundLink* might contribute to the increased perceived confidence and reduced task difficulty.

EXPERT **E1-4** found that *GroundLink* can guide the users of the 3D editor. **E3** found: *"When I [look at feedforward on] the opacity, I can understand why it's like that, and the functionality is there, [and what action to take]."* **E4** focused on a more holistic aspect of the feedforward when thinking about it: *"It's a good idea [for a system to] give you an indication of where and to put your [element], especially in terms of visual [cues] and the changes [needed]."* Similarly, **E3,4** found filtration useful as it maps to contributors and/or scene elements. **E3** mentioned: *"There are a lot of times when you feel you remember someone mentioning something like the director or the camera person, and you couldn't*

[38] Reverse scored

EXPERT

quite remember what it was. It's nice that you can click on a contributor to see what exactly they wanted."

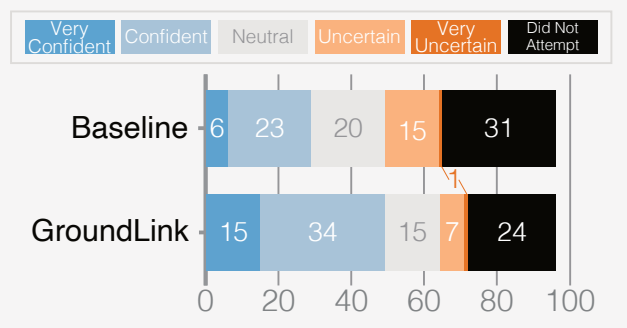


Figure 12: Cumulative confidence ratings across all comparative study participants.

COMPARATIVE STUDY

Consistently, the number of tasks that comparative study participants confidently attempted[39] was significantly higher (Figure 12) for *GroundLink* ($p = 0.014$). However, the number of tasks attempted between the two different conditions did not significantly differ ($p = 0.321$), though the mean favored *GroundLink*.

This modulation in confidence is also suggested in interviews. **P1-5,7-10** suggested that they were more confident in their edits when they were using *GroundLink*. **P3** mentioned: *"[With]* GroundLink *I am more confident. I have some clues or creative messages."* **P4** attributed the confidence to the feedforward: *"It gives me [positional constraints] and also a warning message."* **P8,9,11,12** also found that *GroundLink* is capable of guiding their edits. **P8** attributed this to the feedforward mechanism: *"There's more guidance [with* GroundLink*], there's more to the tasks and to the things that need to be done [inferred from the feedforward mechanism]."* For **P9**, they found the positional feedforward particularly helpful: *"I was moving the light (scene element). There were arrows of where you can move it to. I liked how [*GroundLink*] also highlights the things that are wrong. So [if the] opacity is wrong, it highlights them."* Similarly, many participants (**P3,7,8,11,12**) found the mapping between dashboard, meetings, and the 3D editor helpful in referencing scene elements, indicating that the referential links in realizing Dual Coding Theory might be functional in *GroundLink*. For **P7**, filtration combined with mapping was useful: *"I can now select the objects easily which I can interact with, from this objects filter, which is pretty cool because I can get my task and I can quickly identify this asset in this scene."*

OVERALL

The results show that providing feedforward and synchronization between 3D editor and dashboard can aid the formation of common ground and increase confidence and ease in 3D editing. In doing so, we found a clue in realizing Dual Coding Theory more effectively: the provision of referential link, and visualizing it in-situ on the editor application.

6.2.1 *Challenge: Visual Clutter and Information Overload.*

GroundLink provided more information, directly on the workspace, and inevitably this caused varying levels of visual clutter, confusion, and information overload.

[39] "Confident" or "Very Confident" - Top-2 box

E1,2 found the visual clutter can cause lack of efficiency in editing the 3D scene. **E1** attributed this to multiple blinking areas: *"Because there are all these green zones flashing right now, but the light (scene element) needed to be moved to a specific spot with all those zones flashing, it could mean that that light can go in any of those [areas]."*

The amount of information itself was overwhelming to **E1,2** as well. As part of it, **E2** attributed this to the extraneous language that AI summary includes: *"I'm not super impressed with the AI summary. It never answered the things I wanted it to answer. [...] The AI responses here are a little flowery."* Similarly, due to the amount of content provided, it caused some distraction for **E5**.

EXPERT

Most participants also experienced some difficulties in coping with the visual clutter it creates on the interface (**P1-5,8,10,11**). **P4** suggested similarly, but attributed the confusion to the overlaps of effects: *"I think the circles, which overlap, make it confusing. [...] Also, the shape of a circle might give me a sense of uncertainty."* **P6,10** found some cues on *GroundLink* confusing because they appeared to demand precision. As **P6** mentioned: *"I wasn't sure what* GroundLink *wanted although it was being very precise and trying to be precise. I felt frustrated."*

Similar to experts, **P5** found the details added to the AI summary excessive. As **P5** suggested: *"[AI summary of* GroundLink*] was not useful. It was very hand-wavy without any action. [Copilot] was more direct to the point."*

We found the comments on AI summary, though, are more nuanced. We found the exact opposite comments around Copilot's summaries. **P3,9** found Copilot's text too long (compared to *GroundLink*). **P3** suggested: *"I found [Copilot's summary] not intuitive enough and not straightforward enough. So I have to read line by line and there are so many lines there."* For **P12**, they found Copilot's summary not making sense: *"Honestly, Copilot was just giving me answers that did not make sense. And I felt if I really did follow [Copilot], it would be completely wrong."*

COMPARATIVE STUDY

The visual clutter, and the amount of information seems to cause varying levels of confusion and overwhelm. But as illustrated in the example of AI summary, it could be due to some preferences of users. Also, **P4**'s frustration on the uncertainty caused by circular shape of the positional feedforward ironically validates that uncertainty is communicated through *GroundLink*'s design as intended.

OVERALL

6.2.2 *Challenge: Perceiving* GroundLink *as Restricting Edits.*

As **E2** mentioned, participants did not bluntly follow the constraints visualized in *GroundLink*: *"Even if I say as the director, 'Oh, the opacity of the water should be 50%', I'm guessing. [...] And so I was making a creative decision."*

EXPERT

Similarly, **P12** was also exhibiting more conscious edits: *"I did go rogue for some [feedforward effects]. There was one decision specifically where I did not agree with the tool on how to place it. [If these disagreements happen, I would] have talked to my supervisors."* When users diverged from the restrictions visualized and surfaced, *GroundLink* respected those edits.

However, two participants **P6,10** found it limited their edits and reduced confidence. As **P10** suggested: *"I accomplished better, more things in Copilot. Whereas when I was working with* GroundLink*, I felt like I was doing something wrong because there were so many constraints being put. [...] I take more control [with Copilot]."*

COMPARATIVE STUDY

OVERALL

As such, it appears that we would need to find a way to balance the provision of extra information, and reducing the perception of excessive control. We discuss further about this topic in the discussion.

6.3 Overall Perception: Participants Preferred *GroundLink*

EXPERT

All experts (**E1-5**) were excited about the potential impact it may have on virtual production, as **E1** summarized: *"[GroundLink] allows teams to be a lot more in tune with one another. There's a lot of collaboration [in VP], and it is a lot easier [with* GroundLink*]. Information is right there. There's no second-guessing it because if the information here is wrong, then that's not on the artist or whoever's invoking these changes."* **E1-2, E4** particularly emphasized that the concept implemented in *GroundLink* never existed before, as **E4** suggests: *"[VP] doesn't have a bible application for communicating. [GroundLink] is something huge and something big and I see it has great potential."* **E5** mentioned that it might be able to simulate some in-person review: *"You should meet with some of the key people on set, because the moment you are in the volume, you get a different feeling of scale of things. [When] you have [*GroundLink*], you can simulate it quite nicely."*

Most expert participants (**E1**, **E3-5**) also suggested the generalizability of *GroundLink's* design concepts, as **E1** summarized: *"[*GroundLink's *concept] is something that's awesome for using in Unreal, Maya, or something [which involves] the 3D sense of the world."*

Experts also suggested the value of *GroundLink* in VP onboarding (**E1,2,4**) and education at colleges and universities (**E4**). **E4** suggested: *"It gives [students] that perception of how they can use and see on their computers and monitors and what's the perfect style."*

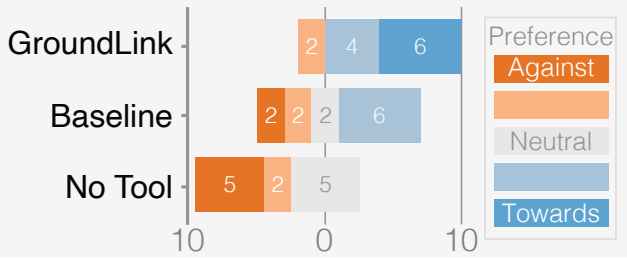


Figure 13: Result of the preference testing.

COMPARATIVE STUDY

Similar to experts' perception, most comparative study participants preferred working with *GroundLink* when working with meeting recordings for VP. 6/12 comparative study participants strongly preferred working with *GroundLink*, and 4/12 expressed slight preference toward it. This differs from Copilot, which only 6/12 participants slightly preferred working with, and none preferred working without any tool (i.e., just the video player). Only two participants (**P6, P10**) expressed slight preference against *GroundLink*, likely be due to the perceived restrictive nature of *GroundLink* (subsubsection 6.2.2).

7 Discussion

Our findings reveal how *GroundLink* reshapes common ground formation in virtual production by embedding past meeting decisions directly into the editing workflow. Here, we explore how these results extend collaborative cognition theories, examine the tensions between guidance and autonomy that emerged, and consider broader implications for collaboration systems in virtual production.

7.1 Forming Common Ground Through *GroundLink*

Our study reveals a noteworthy finding: Shared Mental Models (SMM) and Transactive Memory Systems (TMS), traditionally viewed as relatively independent cognitive frameworks [16, 121], might be simultaneously supported through unified interface design. *GroundLink*'s in-situ feedforward did not just help users know more (SMM) or understand who knows what (TMS). It created the perception that both were simultaneously supported through surfacing contextual information, in-situ.

This integration worked because *GroundLink* provided constituent information when the users needed it. When participants obtained positional constraints in the 3D editor while seeing who made those decisions, they would have built both shared understanding and transactive knowledge. The consistent results across comparative study participants and experts suggest this approach may help address an important challenge in collaborative creative work: the translation gap between discussion and implementation.

This integrative capability of *GroundLink* shows that users no longer need to choose between deep shared knowledge or efficient knowledge distribution. *GroundLink* suggests that appropriate interface design can support both cognitive systems through cohesive information provision.

> **Design Implication 1:** Remote meeting-based collaboration systems can provide constituent information for both shared mental models and transactive memory systems simultaneously, leveraging in-situ presentation, to realize both collaborative cognition systems.

7.2 Realizing Dual Coding Theory: The Important Role of Referential Links

While both *GroundLink* and the baseline provided verbal (meeting-mostly verbal) and non-verbal (3D scene) information separately, our results highlight that simply having multiple representations may not be sufficient. The important difference was the explicit referential linking that *GroundLink* provided between these modalities.

This finding extends Dual Coding Theory [93] by suggesting that in complex creative tasks, the cognitive challenge is not just processing different information types but maintaining coherence and connection between them. When **P7** was able to "quickly identify this asset" through filtered object selection, they were not accessing new information but leveraging pre-established connections by *GroundLink* that would otherwise require significant mental effort.

The seemingly trivial act of mapping verbal discussions to visual elements may have sizeable impacts on usability, suggesting that VP tools, or even collaborative creative tools more broadly, may need to consider the cognitive burden of maintaining these referential links manually. This also clarifies *GroundLink's* novelty relative to prior VP review/annotation systems, which often keep comments as static marks and/or decouple them from the tool where edits are performed. In contrast, *GroundLink* provides referential links between

meeting utterances and scene elements, making meeting-derived decisions actionable in the editor at the moment of manipulation, while preserving provenance back to the originating discussion segment.

Design Implication 2: Systems should prioritize automatic, persistent referential linking between verbal and non-verbal information sources, as the cognitive burden of maintaining these mappings manually may hinder task performance.

7.3 Design Friction as a Feature: Balancing Guidance and Autonomy

The experience of a minority of our participants reveals a fundamental tension in collaborative system design. Most participants found *GroundLink*'s constraints empowering. They felt more confident, possibly because the system provided boundaries. Yet **P6** and **P10** experienced these same constraints as restrictive, feeling they had lost control.

This tension became most salient through our atypical participant. **P6** did not follow the instructions (which were to "make edits according to what collaborators have discussed in meetings") despite repeated reminders. Interestingly, this case inadvertently validated our design approach. We posit that **P6**'s substantial performance difference (3/8 tasks with *GroundLink* vs. 8/8 with baseline) was not a failure but design friction [27] working as intended. **P6**'s admission of being "confident, but I understood nothing" with the baseline illustrates the danger of unconstrained editing in collaborative contexts.

However, this friction comes with a drawback: visual clutter. Most participants raised concerns about overwhelming visual feedback. Yet, this same clutter provided the very resistance necessary to prevent thoughtless edits. This suggests that the challenge is not eliminating visual complexity but calibrating it appropriately. **E1**'s suggestion of (task) priority-based filtering offers a path forward. Interfaces could adapt visual feedback intensity based on the importance of edits rather than displaying all constraints equally.

The fact that **P12** and **E2** successfully "went rogue" when they disagreed with suggestions demonstrates that *GroundLink* may achieve an important balance: it makes ignoring team decisions require deliberate choice rather than accidental oversight.

Design Implication 3: Collaborative editing systems should implement adaptive visual friction that scales with the importance of decisions, ensuring that overriding team decisions demands deliberate intent. However, this should simultaneously minimize unnecessary visual clutter for routine edits.

7.4 The Double-Edged Sword of Augmented Collaboration Tool

Expert concerns about junior professionals becoming too reliant on *GroundLink* resemble broader anxieties about the impact of generative AI on critical thinking capabilities [68]. **E1**'s worry that the system "holds your hand" too much suggests a risk of replacing rather than supporting collaborative skills and competencies.

GroundLink aims to redistribute rather than eliminate cognitive work. Users are still encouraged to exercise their own judgment, and the interface attempts to support these cognitive processes. The only difference provided by *GroundLink* is that users now simply need to focus on evaluating decisions rather than reconstructing them. The experience of participants overriding the system suggestions when they disagreed indicates that critical thinking remains active.

More intriguing is **E5**'s observation about feedback translation. In traditional VP workflows, harsh director feedback is diplomatically translated (by coordinators) before reaching artists. *GroundLink*'s mediated presentation might inadvertently solve this emotional labor problem, as the system, by default, removes unnecessary details in the summary view. This may potentially encourage more honest feedback, knowing it will be presented neutrally. Whether this leads to better or worse creative outcomes remains an open question.

The authoring challenge raised by experts points to a deeper issue: whose interpretation becomes the standard? The system currently requires manual authoring, and this process embeds subjective decisions about what constitutes an actionable insight versus mere discussion. This hidden editorial layer could significantly impact team dynamics in ways we do not yet understand. This editorial layer is important because shared understanding is not the same as agreement. *GroundLink* can help users see what their collaborators think, but in its current form it does not help resolve disagreements. Here, a suggestion from **E1** points to one possible direction: equipping each comment or change item with richer sharing features (subsubsection 6.1.2). Future implementations could integrate share features and opinion-aggregation mechanisms[40] so that disagreements are made visible. Such mechanisms could help mitigate overreliance on the system, and mitigate the risk that junior staff simply follow the feedforward suggestions.

In addition, future authoring systems could incorporate manual controls with adjustable weights, allowing teams to define shared "standard" comments or priorities. Once these are in place, much of the remaining system work becomes a classification problem: matching new data points to similar, previously labeled ones. This is a well-studied problem in machine learning, supporting feasibility [82, 113].

Taken together, these sharing features and adjustable controls could also serve as practical scaffolding for automating our Wizard-of-Oz pipeline. Concretely, a future system could log time-stamped edit events from Unity (object ID, property, old/new values) and align them with time-coded meeting transcripts (and cursor/whiteboard events, when available) using shared timestamps. A lightweight dialog-act classifier could then identify candidate decision/constraint utterances and convert them into structured comment items (e.g., target object, constraint/parameter, rationale, confidence, and source timestamp). Matching becomes an entity-linking and ranking step, which maps utterance to scene objects using object metadata (names/tags), then selects the best match. Finally, per-item sharing and editing can function as a human-in-the-loop confirmation step.

Designers could also stage cues more carefully. For example, surfacing them only at key moments such as before sign-off, or assign a severity measure to different violations so that only high-severity issues interrupt the creative flow. These strategies offer more concrete ways to balance the creativity-constraint trade-off.

[40] For example, reaction buttons or polls.

Design Implication 4: Systems mediating collaborative decisions need to make their interpretive/authoring layer transparent, allowing users to understand and question how meeting discussions become actionable constraints while preserving space for professional judgment and creative exploration.

7.5 Limitations and Future Work

These findings provide a strong basis for insights and discussion, though some limitations affect their generalizability. Our expert pool's homogeneity (all men, predominantly from our formative study) reflects industry demographics [34] but limits perspective diversity. The overlap with formative participants for the expert evaluation may have created positive bias toward solutions addressing their previously stated pain points. The user study included only a few participants, with short-term exposure, reducing statistical power, as well as limmiting generalizability. Given the exploratory nature of the study and the prioritization of avoiding Type II error (hence uncorrected p-values), results should be interpreted with caution. Our evaluation relies on qualitative reports and a customized questionnaire rather than validated SMM/TMS instruments. While being appropriate for our exploratory, domain-specific focus, it means our findings provide indicative, qualitative evidence of perceived SMM/TMS formation rather than formal, standardized measurements.

Our simplified task design enabled novice participation for the comparative study, but might not capture the full complexity of professional VP workflows involving hours-long meetings, numerous stakeholders, and high-stakes creative decisions. The absence of objective correctness measures, inherent to creative work, makes it difficult to determine whether *GroundLink* actually improves outcomes or merely increases confidence.

The authoring process remains unresolved. This work focused on the consumption experience of meetings, rather than authoring. Extracting actionable decisions from naturalistic meeting discourse for authoring remains unsolved, particularly given the needs for translation between filmmaker vision and technical implementation (**E5**).

Future work should deploy *GroundLink* in authentic virtual production environments over extended periods, with more participants to understand whether benefits persist or users develop workarounds, as well as for obtaining generalizable insights. Investigating impacts on upstream communication, such as how knowing feedback will be embedded might change meeting dynamics, could reveal unexpected system effects. Future work could also utilize validated SMM/TMS scales to strengthen the robustness and comparability of the results. Finally, the risk of homogenizing creative decisions across the industry deserves careful consideration, as these tools shape not just individual edits but collective creative processes.

As **E1**, **E3-5** suggested, design concepts considered in this work could generalize to other digital content creation, other domains of the 2D/3D editing[41], as well as beyond a workplace tool (e.g., for education). Our empirical findings suggest the transferable value lies in the mechanism of maintaining referential links between meeting evidence and work objects, so guidance remains traceable to its source. For example, BIM/CAD tools could surface meeting-derived coordination cues linked to the originating discussion, while educational tools could present recorded critique as shareable examples that students can review and revise as their exercises. While the current version of *GroundLink* focuses on Unity, the same meeting-indexing and in-situ feedforward concepts could also be realized as a family of interoperable add-ons that share a common Electron application for dashboard, allowing constraints and rationales derived from meetings to appear consistently across multiple contents creation tools. Future designers could utilize our generalizable design concepts in other domains, both in professional tool development and HCI research to further extend the reach of our contribution.

[41] E.g., coordinating structural and design changes in architectural BIM models, or propagating design-review decisions into CAD assemblies for manufacturing.

8 Conclusion

This paper introduced *GroundLink*, a Unity add-on for virtual production that embeds decisions from prior meetings directly into the 3D editing environment to support the formation of common ground. Across a comparative user study and an expert evaluation, participants reported clearer perceived shared understanding and awareness of who to consult, greater confidence, and easier scene editing experiences. These findings were also supported by some survey questions. Participants also raised concerns, such as the visual clutter, occasional perceptions of restricted agency, and potential risks of over-reliance. We presented four design implications to inform future work.

Acknowledgments

We acknowledge the use of Cursor to help us develop the prototype (https://cursor.com/en).

References

[1] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2013. YouMove: enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*. 311–320.

[2] Linda Argote and Paul Ingram. 2000. Knowledge Transfer: A Basis for Competitive Advantage in Firms. *Organizational Behavior and Human Decision Processes* 82, 1 (May 2000), 150–169. doi:10.1006/obhd.2000.2893

[3] Valentino Artizzu, Kris Luyten, Gustavo Rovelo Ruiz, and Lucio Davide Spano. 2024. Virgilites: multilevel feedforward for multimodal interaction in vr. *Proceedings of the ACM on Human-Computer Interaction* 8, EICS (2024), 1–24.

[4] Jeremy N Bailenson. 2021. Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. (2021).

[5] Tully Barnett, Jason Bevan, Cameron Mackness, and Zoë Wallin. 2024. *Beyond Virtual Production: Integrating Production Technologies* (1 ed.). Focal Press, London. doi:10.4324/9781003449492

[6] Dean C Barnlund. 1970. A TRANSACTIONAL MODEL OF COMMUNICATION. (1970).

[7] Olivier Bau and Wendy E Mackay. 2008. OctoPocus: a dynamic guide for learning gesture-based command sets. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*. 37–46.

[8] Andrea Bianchi, Zhi Lin Yap, Punn Lertjaturaphat, Austin Z Henley, Kongpyung Justin Moon, and Yoonji Kim. 2024. Inline Visualization and Manipulation of Real-Time Hardware Log for Supporting Debugging of Embedded Programs. *Proceedings of the ACM on Human-Computer Interaction* 8, EICS (2024), 1–26.

[9] Jeremy P Birnholtz, Thomas A Finholt, Daniel B Horn, and Sung Joo Bae. 2005. Grounding needs: achieving common ground via lightweight chat in large, distributed, ad-hoc groups. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 21–30.

[10] Helen Blair, Susan Grey, and Keith Randle. 2001. Working in film–employment in a project based industry. *Personnel review* 30, 2 (2001), 170–185.

[11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[12] Rob Brennan, Simon Quigley, Pieter De Leenheer, and Alfredo Maldonado. 2018. Automatic extraction of data governance knowledge from slack chat channels. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*. Springer, 555–564.

[13] Nathalie Bressa, Henrik Korsgaard, Aurélien Tabard, Steven Houben, and Jo Vermeulen. 2021. What's the situation with situated visualization? A survey and

perspectives on situatedness. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 107–117.
[14] Patrick D Bridge and Shlomo S Sawilowsky. 1999. Increasing physicians' awareness of the impact of statistics on research outcomes: comparative power of the t-test and Wilcoxon rank-sum test in small samples applied research. *Journal of clinical epidemiology* 52, 3 (1999), 229–235.
[15] Giuseppe Caggianese, Valerio Colonnese, and Luigi Gallo. 2019. Situated visualization in augmented reality: Exploring information seeking strategies. In *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 390–395.
[16] Janis A Cannon-Bowers, Eduardo Salas, and Sharolyn Converse. 1993. Shared mental models in expert team decision making. *Individual and group decision making: Current issues* 221 (1993), 221–46.
[17] Xinyue Chen, Nathan Yap, Xinyi Lu, Aylin Gunal, and Xu Wang. 2025. MeetMap: Real-Time Collaborative Dialogue Mapping with LLMs in Online Meetings. *Proceedings of the ACM on Human-Computer Interaction* 9, 2 (2025), 1–35.
[18] Yue Chen and Qin Gao. 2023. Visualizing timeline-anchored comments enhanced social presence and information searching in video-based learning. *Computer Applications in Engineering Education* 31, 5 (2023), 1306–1320.
[19] Liqi Cheng, Hanze Jia, Lingyun Yu, Yihong Wu, Shuainan Ye, Dazhen Deng, Hui Zhang, Xiao Xie, and Yingcai Wu. 2024. VisCourt: In-Situ Guidance for Interactive Tactic Training in Mixed Reality. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–14.
[20] Ed H Chi, Peter Pirolli, Kim Chen, and James Pitkow. 2001. Using information scent to model user information needs and actions and the Web. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 490–497.
[21] Hyunsung Cho, Matthew L. Komar, and David Lindlbauer. 2023. RealityReplay: Detecting and Replaying Temporal Changes In Situ Using Mixed Reality. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (Sept. 2023), 1–25. doi:10.1145/3610888
[22] Herbert H. Clark and Susan E. Brennan. 1991. *Grounding in communication*. American Psychological Association, 127–149. doi:10.1037/10096-006
[23] Clark, Herbert H. 1996. Communities, commonalities, and communication. *Rethinking linguistic relativity* 17 (1996), 324–355.
[24] Gregorio Convertino, Helena M Mentis, Mary Beth Rosson, John M Carroll, Aleksandra Slavkovic, and Craig H Ganoe. 2008. Articulating common ground in cooperative work: content and process. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1637–1646.
[25] Sven Coppers, Kris Luyten, Davy Vanacken, David Navarre, Philippe Palanque, and Christine Gris. 2019. Fortunettes: feedforward about the future state of GUI widgets. *Proceedings of the ACM on Human-Computer Interaction* 3, EICS (2019), 1–20.
[26] Karina Cortiñas-Lorenzo, Siân Lindley, Ida Larsen-Ledet, and Bhaskar Mitra. 2024. Through the looking-glass: Transparency implications and challenges in enterprise AI knowledge systems. *arXiv preprint arXiv:2401.09410* (2024).
[27] Anna L Cox, Sandy JJ Gould, Marta E Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design frictions for mindful interactions: The case for microboundaries. In *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*. 1389–1397.
[28] Richard L Daft and Robert H Lengel. 1986. Organizational information requirements, media richness and structural design. *Management science* 32, 5 (1986), 554–571.
[29] Robert J DeFillippi and Michael B Arthur. 1998. Paradox in project-based enterprise: The case of film making. *California management review* 40, 2 (1998), 125–139.
[30] Tom Djajadiningrat, Kees Overbeeke, and Stephan Wensveen. 2002. But how, Donald, tell us how? On the creation of meaning in interaction design through feedforward and inherent feedback. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*. 285–291.
[31] Doodle. 2019. *The Doodle State of Meetings Report 2019*. Technical Report. Doodle. https://doodle.com/en/resources/research-and-reports-/the-state-of-meetings-2019/
[32] Steven P Dow, Manish Mehta, Blair MacIntyre, and Michael Mateas. 2010. Eliza meets the wizard-of-oz: blending machine and human control of embodied characters. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 547–556.
[33] Sean E Ellis and Dennis P Groth. 2004. A collaborative annotation system for data visualization. In *Proceedings of the working conference on Advanced Visual Interfaces*. 411–414.
[34] Kim Elsesser. 2025. Gender Gaps Widen Behind The Scenes In 2024's Top-Grossing Films — forbes.com. https://www.forbes.com/sites/kimelsesser/2025/01/03/gender-gaps-widen-behind-the-scenes-in-2024s-top-grossing-films/. [Accessed 21-08-2025].
[35] Andreas Rene Fender and Christian Holz. 2022. Causality-preserving Asynchronous Reality. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–15. doi:10.1145/3491102.3501836
[36] Katherine Fennedy, Jeremy Hartmann, Quentin Roy, Simon Tangi Perrault, and Daniel Vogel. 2021. OctoPocus in VR: Using a dynamic guide for 3D mid-air gestures in virtual reality. *IEEE Transactions on Visualization and Computer Graphics* 27, 12 (2021), 4425–4438.
[37] Luis Garicano and Esteban Rossi-Hansberg. 2006. Organization and Inequality in a Knowledge Economy*. *Quarterly Journal of Economics* 121, 4 (Nov. 2006), 1383–1435. doi:10.1162/qjec.121.4.1383
[38] William W Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a resource for design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 233–240.
[39] Frederic Gmeiner, Kaitao Luo, Ye Wang, Kenneth Holstein, and Nikolas Martelaro. 2025. Exploring the Potential of Metacognitive Support Agents for Human-AI Co-Creation. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*. 1244–1269.
[40] Frederic Gmeiner, Nicolai Marquardt, Michael Bentley, Hugo Romat, Michel Pahud, David Brown, Asta Roseway, Nikolas Martelaro, Kenneth Holstein, Ken Hinckley, et al. 2025. Intent Tagging: Exploring Micro-Prompting Interactions for Supporting Granular Human-GenAI Co-Creation Workflows. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–31.
[41] Pascal Goffin, Tanja Blascheck, Petra Isenberg, and Wesley Willett. 2020. Interaction techniques for visual exploration using embedded word-scale visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[42] Tiago Joao Guerreiro, Daniel Medeiros, Daniel Mendes, Maurício Sousa, Joaquim A Jorge, Alberto Raposo, and Ismael HF dos Santos. 2014. Beyond Post-It: Structured Multimedia Annotations for Collaborative VEs.. In *ICAT-EGVE*. 55–62.
[43] Jocelyn Handy and Lorraine Rowlands. 2017. The systems psychodynamics of gendered hiring: Personal anxieties and defensive organizational practices within the New Zealand film industry. *Human Relations* 70, 3 (2017), 312–338.
[44] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
[45] Jeffrey Heer, Fernanda B Viégas, and Martin Wattenberg. 2009. Voyagers and voyeurs: Supporting asynchronous collaborative visualization. *Commun. ACM* 52, 1 (2009), 87–97.
[46] Jane Hoffswell, Arvind Satyanarayan, and Jeffrey Heer. 2018. Augmenting code with in situ visualizations to aid program understanding. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
[47] Siying Hu, Hen Chen Yen, Ziwei Yu, Mingjian Zhao, Katie Seaborn, and Can Liu. 2023. Wizundry: A cooperative wizard of oz platform for simulating future speech-based interfaces with multiple wizards. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–34.
[48] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 17–24.
[49] Kori Inkpen, Rajesh Hegde, Sasa Junuzovic, Christopher Brooks, John C Tang, and Zhengyou Zhang. 2010. AIR Conferencing: Accelerated instant replay for in-meeting multimodal review. In *Proceedings of the 18th ACM international conference on Multimedia*. 663–666.
[50] Nuwan Janaka, Chloe Haigh, Hyeongcheol Kim, Shan Zhang, and Shengdong Zhao. 2022. Paracentral and near-peripheral visualizations: Towards attention-maintaining secondary information presentation on OHMDs during in-person social interactions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–14.
[51] Karen A Jehn. 1995. A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative science quarterly* (1995), 256–282.
[52] Zipeng Ji, Pengcheng An, and Jian Zhao. 2025. ClassComet: Exploring and Designing AI-generated Danmaku in Educational Videos to Enhance Online Learning. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*. 552–575.
[53] Candace Jones. 1996. Careers in project networks: The case of the film industry. *The boundaryless career: A new employment principle for a new organizational era* 58 (1996), 75.
[54] Thomas Jung, Mark D Gross, and Ellen Yi-Luen Do. 2001. Space Pen: annotation and sketching on 3D models on the Internet. In *Computer Aided Architectural Design Futures 2001: Proceedings of the Ninth International Conference held at the Eindhoven University of Technology, Eindhoven, The Netherlands, on July 8–11, 2001*. Springer, 257–270.
[55] Thomas Jung, Mark D Gross, and Ellen Yi-Luen Do. 2002. Annotating and sketching on 3D web models. In *Proceedings of the 7th international conference on Intelligent user interfaces*. 95–102.
[56] Sasa Junuzovic, Kori Inkpen, Rajesh Hegde, Zhengyou Zhang, John Tang, and Christopher Brooks. 2011. What did I miss? In-meeting review using multimodal accelerated instant replay (AIR) conferencing. In *Proceedings of the sigchi conference on human factors in computing systems*. 513–522.
[57] Kadner, Noah. 2021. *THE VIRTUAL PRODUCTION FIELD GUIDE*. Vol. 2. Epic Games.
[58] Tobias Kauer, Derya Akbaba, Marian Dörk, and Benjamin Bach. 2024. Discursive patinas: Anchoring discussions in data visualizations. *IEEE transactions on visualization and computer graphics* (2024).

[59] Manolya Kavakli and Cinzia Cremona. 2022. The Virtual Production Studio Concept – An Emerging Game Changer in Filmmaking. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Christchurch, New Zealand, 29–37. doi:10.1109/VR51125.2022.00020

[60] John F Kelley. 1983. An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 193–196.

[61] Anjali Khurana, Michael Glueck, and Parmit K Chilana. 2024. Do I Just Tap My Headset? How Novice Users Discover Gestural Interactions with Consumer Augmented Reality Applications. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–28.

[62] Joseph Kim and Julie A. Shah. 2016. Improving Team's Consistency of Understanding in Meetings. *IEEE Transactions on Human-Machine Systems* 46, 5 (Oct. 2016), 625–637. doi:10.1109/THMS.2016.2547186

[63] Rebecca Krosnick, Fraser Anderson, Justin Matejka, Steve Oney, Walter S. Lasecki, Tovi Grossman, and George Fitzmaurice. 2021. Think-Aloud Computing: Supporting Rich and Low-Effort Knowledge Capture. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–13. doi:10.1145/3411764.3445066

[64] Werner Kunz and Horst WJ Rittel. 1970. Issues as elements of information systems. (1970).

[65] Benjamin Lafreniere, Parmit K Chilana, Adam Fourney, and Michael A Terry. 2015. These aren't the commands you're looking for: Addressing false feedforward in feature-rich software. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. 619–628.

[66] Austin Lee, Hiroshi Chigira, Sheng Kai Tang, Kojo Acquah, and Hiroshi Ishii. 2014. AnnoScape: remote collaborative review using live video overlay in shared 3D virtual workspace. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. 26–29.

[67] Benjamin Lee, Michael Sedlmair, and Dieter Schmalstieg. 2023. Design patterns for situated visualization in augmented reality. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2023), 1324–1335.

[68] Hao-Ping Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, and Nicholas Wilson. 2025. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI conference on human factors in computing systems*. 1–22.

[69] Joanne Leong, John Tang, Edward Cutrell, Sasa Junuzovic, Gregory Paul Baribault, and Kori Inkpen. 2024. Dittos: Personalized, Embodied Agents That Participate in Meetings When You Are Unavailable. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–28.

[70] Kyle Lewis. 2003. Measuring transactive memory systems in the field: scale development and validation. *Journal of applied psychology* 88, 4 (2003), 587.

[71] Haofeng Li, Cheng-Hung Lo, Arturo Smith, and Zhiyuan Yu. 2022. The development of virtual production in film industry in the past decade. In *International Conference on Human-Computer Interaction*. Springer, 221–239.

[72] Manling Li, Lingyu Zhang, Heng Ji, and Richard J Radke. 2019. Keep meeting summaries on topic: Abstractive multi-modal meeting summarization. In *Proceedings of the 57th annual meeting of the association for computational linguistics*. 2190–2196.

[73] Jian Liao, Adnan Karim, Shivesh Singh Jadon, Rubaiat Habib Kazi, and Ryo Suzuki. 2022. RealityTalk: Real-time speech-driven augmented presentation for AR live storytelling. In *Proceedings of the 35th annual ACM symposium on user interface software and technology*. 1–12.

[74] Feifan Liu and Yang Liu. 2008. Correlation between rouge and human evaluation of extractive meeting summaries. In *Proceedings of ACL-08: HLT, short papers*. 201–204.

[75] Tao Long, Kendra Wannamaker, Jo Vermeulen, George Fitzmaurice, and Justin Matejka. 2025. FeedQUAC: Quick Unobtrusive AI-Generated Commentary. *arXiv preprint arXiv:2504.16416* (2025).

[76] Ferruccio Mandorli, Harald E Otto, et al. 2024. Improving the learning experience within MCAD education: The provision of feedback on CAD model quality and dormant deficiency in real-time. *Computer-Aided Design and Applications* 21, 5 (2024), 739–758.

[77] David Maulsby, Saul Greenberg, and Richard Mander. 1993. Prototyping an intelligent agent through Wizard of Oz. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. 277–284.

[78] Lucid Meetings. 2022. How many meetings are there per day in 2022? (And should you care?) - The Lucid Meetings Blog. https://blog.lucidmeetings.com/blog/how-many-meetings-are-there-per-day-in-2022/

[79] Britta Meixner, Matthew Lee, and Scott Carter. 2017. Chat2Doc: From Chats to How-to Instructions, FAQ, and Reports. In *Proceedings of the 2017 ACM Workshop on Multimedia-based Educational and Knowledge Technologies for Personalized and Social Online Training*. 27–30.

[80] Microsoft. 2023. Will AI Fix Work? (2023).

[81] Bryan Min and Haijun Xia. 2025. Feedforward in Generative AI: Opportunities for a Design Space. doi:10.48550/arXiv.2502.14229 arXiv:2502.14229 [cs].

[82] Shervin Minaee, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, and Jianfeng Gao. 2021. Deep learning–based text classification: a comprehensive review. *ACM computing surveys (CSUR)* 54, 3 (2021), 1–40.

[83] Andreea Muresan, Jess McIntosh, and Kasper Hornbæk. 2023. Using feedforward to reveal interaction possibilities in virtual reality. *ACM Transactions on Computer-Human Interaction* 30, 6 (2023), 1–47.

[84] Gabriel Murray, Steve Renals, and Jean Carletta. 2005. Extractive summarization of meeting recordings. (2005).

[85] Mukesh Nathan, Mercan Topkara, Jennifer Lai, Shimei Pan, Steven Wood, Jeff Boston, and Loren Terveen. 2012. In case you missed it: Benefits of attendee-shared annotations for non-attendees of remote meetings. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. 339–348.

[86] Karin Niemantsverdriet and Thomas Erickson. 2017. Recurring Meetings: An Experiential Account of Repeating Meetings in a Large Organization. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (Dec. 2017), 1–17. doi:10.1145/3134719

[87] Ikujiro Nonaka. 1998. The Knowledge-Creating Company. In *The Economic Impact of Knowledge*. Elsevier, 175–187. doi:10.1016/B978-0-7506-7009-8.50016-1

[88] Donald A Norman. 1988. *The psychology of everyday things*. Basic books.

[89] Nels Numan, Gabriel Brostow, Suhyun Park, Simon Julier, Anthony Steed, and Jessica Van Brummelen. 2025. CoCreatAR: Enhancing authoring of outdoor augmented reality experiences through asymmetric collaboration. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–22.

[90] Gary M Olson and Judith S Olson. 2000. Distance matters. *Human–computer interaction* 15, 2-3 (2000), 139–178.

[91] Olson, Judith S and Olson, Gary M. 2014. Common Ground. In *Working Together Apart: Collaboration over the Internet*. Springer, 39–42.

[92] Heather L O'Brien, Paul Cairns, and Mark Hall. 2018. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies* 112 (2018), 28–39.

[93] Allan Paivio. 1991. Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie* 45, 3 (1991), 255.

[94] Gun Woo Park, Payod Panda, Lev Tankelevitch, and Sean Rintel. 2024. CoExplorer: Generative AI Powered 2D and 3D Adaptive Interfaces to Support Intentionality in Video Meetings. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–10.

[95] Gun Woo Park, Payod Panda, Lev Tankelevitch, and Sean Rintel. 2024. The CoExplorer Technology probe: A generative AI-powered adaptive interface to support intentionality in planning and running video meetings. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 1638–1657.

[96] Gun Woo Park, Anthony Tang, and Fanny Chevalier. 2024. JollyGesture: Exploring Dual-Purpose Gestures and Gesture Guidance in VR Presentations. In *Proceedings of the 50th Graphics Interface Conference*. 1–14.

[97] Soya Park and Chinmay Kulkarni. 2023. Retrospector: Rapid collaborative reflection to improve collaborative practices. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (Sept. 2023), 1–20. doi:10.1145/3610084

[98] Henning Pohl. 2024. Body-Based Augmented Reality Feedback During Conversations. *Proceedings of the ACM on Human-Computer Interaction* 8, MHCI (2024), 1–22.

[99] Henning Pohl and Kasper Hornbæk. 2024. Integrated Calculators: Moving Calculation into the World. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 343–355.

[100] Virgile Rennard, Guokan Shang, Julie Hunter, and Michalis Vazirgiannis. 2023. Abstractive meeting summarization: A survey. *Transactions of the Association for Computational Linguistics* 11 (2023), 861–884.

[101] Yvonne Rogers and Judi Ellis. 1994. Distributed Cognition: an alternative framework for analysing and explaining collaborative working. (1994).

[102] Hauke Sandhaus and Eva Hornecker. 2018. A WOZ study of feedforward information on an ambient display in autonomous cars. In *Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 90–92.

[103] Nazmus Saquib, Rubaiat Habib Kazi, Li-Yi Wei, and Wilmot Li. 2019. Interactive body-driven graphics for augmented video performance. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[104] Kevin Sheppard, Stanislav Khrapov, Gábor Lipták, mikedeltalima, Rob Capellini, alejandro cermeno, Hugle, esvhd, Snyk bot, Alex Fortin, JPN, Matt Judell, Weiliang Li, Austin Adams, jbrockmendel, M. Rabba, Michael E. Rose, Nikolay Tretyak, Tom Rochette, UNO Leo, Xavier RENE-CORAIL, Xin Du, and Burak Çelik. 2022. *bashtage/arch: Release 5.3.1*. doi:10.5281/zenodo.6684078

[105] Joongi Shin, Michael A Hedderich, AndréS Lucero, and Antti Oulasvirta. 2022. Chatbots facilitating consensus-building in asynchronous co-design. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–13.

[106] Rajinder S Sodhi, Brett R Jones, David Forsyth, Brian P Bailey, and Giuliano Maciocci. 2013. BeThere: 3D mobile collaboration with spatial input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 179–188.

[107] Nicole Sultanum, Anastasia Bezerianos, and Fanny Chevalier. 2021. Text visualization and close reading for journalism with storifier. In *2021 IEEE Visualization Conference (VIS)*. IEEE, 186–190.

[108] Jon Swords and Nina Willment. 2024. The emergence of virtual production – a research agenda. *Convergence: The International Journal of Research into New Media Technologies* 30, 5 (Oct. 2024), 1557–1574. doi:10.1177/13548565241253903

[109] Jon Swords and Nina Willment. 2024. 'It used to be fix-it in post production! now it's fix-it in pre-production': how virtual production is changing production networks in film and television. *Creative Industries Journal* (Nov. 2024), 1–17. doi:10.1080/17510694.2024.2430025

[110] John Tang, Jennifer Marlow, Aaron Hoff, Asta Roseway, Kori Inkpen, Chen Zhao, and Xiang Cao. 2012. Time travel proxy: using lightweight video recordings to create asynchronous, interactive meetings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3111–3120.

[111] Miini Teng, Martin Yu, Weina Jin, Dae Woo Hwang, David Stitt, Cristina Conati, Giuseppe Carenini, Tal Jarus, and Thalia Field. 2024. Human-computer interactions and compassionate healthcare: A Wizard of Oz study using a self-administered AI-assisted cognitive assessment. *medRxiv* (2024), 2024–12.

[112] Michael Terry and Elizabeth D Mynatt. 2002. Side views: persistent, on-demand previews for open-ended tasks. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*. 71–80.

[113] Sergios Theodoridis and Konstantinos Koutroumbas. 2006. *Pattern recognition*. Elsevier.

[114] Simon Tucker, Ofer Bergman, Anand Ramamoorthy, and Steve Whittaker. 2010. Catchup: a useful application of time-travel in meetings. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 99–102.

[115] Jandre J Van Rensburg, Catarina M Santos, Simon B de Jong, and Sjir Uitdewilligen. 2022. The five-factor perceived shared mental model scale: a consolidation of items across the contemporary literature. *Frontiers in psychology* 12 (2022), 784200.

[116] Jo Vermeulen and Kris Luyten. 2013. Crossing the Bridge over Norman's Gulf of Execution: Revealing Feedforward's True Identity. (2013).

[117] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17 (2020), 261–272. doi:10.1038/s41592-019-0686-2

[118] Alex Waibel, Michael Bett, Michael Finke, and Rainer Stiefelhagen. 1998. Meeting browser: Tracking and summarizing meetings. In *Proceedings of the DARPA broadcast news workshop*. 281–286.

[119] Dakuo Wang, Haoyu Wang, Mo Yu, Zahra Ashktorab, and Ming Tan. 2022. Group chat ecology in enterprise instant messaging: How employees collaborate through multi-user chat channels on slack. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–14.

[120] Jixuan Wang, Jingbo Yang, Haochi Zhang, Helen Lu, Marta Skreta, Mia Husić, Aryan Arbabi, Nicole Sultanum, and Michael Brudno. 2022. PhenoPad: building AI enabled note-taking interfaces for patient encounters. *NPJ digital medicine* 5, 1 (2022), 12.

[121] Daniel M Wegner. 1987. Transactive memory: A contemporary analysis of the group mind. In *Theories of group behavior*. Springer, 185–208.

[122] Wesley Willett, Yvonne Jansen, and Pierre Dragicevic. 2016. Embedded data representations. *IEEE transactions on visualization and computer graphics* 23, 1 (2016), 461–470.

[123] Zhenyu Xu, Hailin Xu, Zhouyang Lu, Yingying Zhao, Rui Zhu, Yujiang Wang, Mingzhi Dong, Yuhu Chang, Qin Lv, Robert P Dick, et al. 2024. Can large language models be good companions? An LLM-based eyewear system with conversational common ground. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 2 (2024), 1–41.

[124] Susan Zwerman and Jeffrey A. Okun. 2023. *The VES Handbook of Virtual Production* (1 ed.). Focal Press, New York. doi:10.4324/9781003366515

A Expert Interview Participant Details

PID (Formative-Interview)	Role	Years of VP Experience	Number of Past VP Projects	*PID* (Formative-Survey)	*PID* (Expert Evaluation)	Country	Age	Gender
FI1	VP Supervisor	5	8	**F2**	**E1**	US	54	Man
FI2	VP Artist	2	3	**F10**	*DNP*	South Korea	41	Man
FI3	VP Artist	5	30	**F15**	**E4**	Canada	35	Man
FI4	Compositor	10	12	**F17**	**E3**	Canada	28	Man
FI5	VP Lead/Researcher	8	15	**F22**	*DNP*	Canada	38	Man
FI6	VP Supervisor	5	4	**F5**	*DNP*	Canada	46	Man
DNP	VP Researcher	4	10	**F23**	**E2**	Canada	40	Man
DNP	Executive Producer	4	10	*DNP*	**E5**	US	*DNR*	Man

Table 1: Participant information for the semi-structured, expert interviews. *PID*: Participant ID, *DNP*: Did Not Participate, *DNR*: Did Not Report.

B Formative Study - Survey

ID	Question	Response Type
1	Which parts of the virtual production workflow are you mostly involved in?	Text
2	What is your experience with virtual production workflows?	Text
3	What are the main steps in your virtual production workflow? (Is it more waterfall-like or iterative process between phases? What tools do you use in each step? Were those Remote/In-person/On-site?)	Text
4	How often do you collaborate with others? (For each group) - Staff from different roles - Staff in the same role	5-Point Likert (Never-Always)
5	What are the most common issues you encounter in virtual production?	Text
6	How do you record and resolve these issues? In terms of: Frequency; Recording method (e.g., Jira, Word document); Resolution method (e.g., Zoom meeting, VR meeting, email).	Text
7	Do you use Augmented Reality (AR) systems in your virtual production?	Yes/No
8	When do you use AR systems and why?	Text
9	Do you review 3D scenes using AR technology? Describe its usefulness or problems.	Text
10	Have you onboarded staff(s) to a 3D scene created by your team?	Yes/No
11	[If yes] Describe how you onboarded others to a scene. (e.g., meeting format, devices used, software used for onboarding)	Text
12	Have you been onboarded to a 3D scene?	Yes/No
13	[If yes] Describe how you were onboarded to a scene. (e.g., meeting format, devices used, software used for onboarding)	Text
14	How often do onboarding sessions to a scene occur?	7-Point Likert (Daily-Never)+Other
15	Rate the effectiveness of existing onboarding methods as: (For each case) - A person being onboarded - A person onboarding others	5-Point Likert (Very Problematic-Very Good)+N/A

Table 2: Survey questionnaire used for the formative study.

C Formative Study - Semi-Structured Interview Questions

Please think about your virtual production experiences in general.

- How would you describe your virtual production experiences? (Compared to other production practices?)
- Do you collaborate with others? (Frequency, context; follow-up from survey responses)
  - How do you collaborate (tools, methods, etc.)?
  - Are there any challenges in existing collaborative practices? (Follow-up from the survey)
  - Were there any instances when you were missing information from other collaborators?
    - How did you resolve this issue?
  - Were there any instances when you were not sure about the intents of other collaborators?
    - How did you resolve this issue?
  - Were there instances where you had to redo some of the work others had done, or others modified your work without knowing your intent?
    - How did you resolve this issue?
- Do you review past meeting videos or issue notes (e.g., from version control or production management tools like Flow)?
  - How do you review the notes or videos?
  - What works well with these? Are there any challenges?

Please think about an instance when you onboarded a staff member to a 3D scene you have been working on.

- Could you walk me through how you introduced the 3D scene?
  - Which part of the overall workflow was the project in when you introduced the new scene to the other person?
  - Were there any constraints you established during the creation of the 3D scene (e.g., object should not be placed here, lighting should be this way)?
    - (If yes) What were those constraints?
    - (If yes) How problematic were those constraints if violated?
    - (If yes) How did you communicate those constraints?
    - (If yes) How frequently, and how substantially, did collaborators violate these constraints?
  - Are there other types of information you try to communicate to collaborators?
  - How easy or difficult is it to communicate that information? Why?
    - How long does a new collaborator usually take to get used to the scene?
    - Are those new collaborators in the same role as you, or not?
    - Do those scenes require frequent collaboration? Cross-functional collaboration?

Now, let's think about the opposite scenario.

- Do you get introduced to a new 3D scene as part of your job?
  - (If yes) How did other collaborators introduce the scene to you?
    - When was it? What was the situation like?
    - Any success stories? Or any instance of an introduction that did not work well?
    - When you were introduced to a scene, were there instances when you performed well or poorly while collaborating with others?
  - Do you think there could have been aspects that, if implemented, would have helped you get used to the scene faster?

D Comparative Study - Semi-Structured Interview Questions

- Describe one change you made in the scene and what information led you to make that change.
- Did you ever worry that you might undo someone else's work or might not consider the decisions made by others? What helped to reassure you?
- How confident are you that your final edits will be accepted? Why?
- Were there any points that were the most helpful?
- Were there any points that were the most confusing or unhelpful?
- Do you think accessing the context from different projects might be useful?
  - Under what circumstances?
  - What kind of information?
- Any suggestions for improvements?

E Expert Evaluation - Semi-Structured Interview Questions

- In what ways do you think the effects happening on the Unity editor or the inspector help or deter your work?
  - Why?
  - What are the opportunities or challenges you could foresee from it?
- There was a separate panel for providing a summarized meeting detail. In what ways do you think it would help or hinder the completion of your edits?
  - When you interact with one panel, how would you interact with the other panel?
    - How would the system help or hinder you from performing those tasks?
- There was a filter tool with an AI summary. In what ways would you use filtration and AI summary?
  - In what ways do you think this will be helpful or unhelpful?
- What do you think are the strengths and weaknesses of the tool?
- We are envisioning using the tool for more junior virtual production professionals. In what ways do you think the tool can help or hinder their completion of the task?
  - In what ways do you think they will know or not know the knowledge established within the team?
  - In what ways do you think they will have confidence in the edits that they performed?
  - How would the usage of the tool help or hinder the user in figuring out who to contact if they have questions about previous edits being done?
- Do you have any suggestions for improvements?

F Questionnaires

ID	Relates To	Question
1	SMM	I understood why earlier collaborators made the design choices I saw on this screen.
2	SMM	I was aware of the constraints or conventions that applied when I edited the scene.
3	SMM	While working, I felt "on the same page" with the existing team.
4	TMS	I could easily tell who I would need to consult about a specific asset or decision.
5	TMS	I trusted that the information I saw about prior decisions was accurate and credible.
6	TMS	My edits fit smoothly with what the team had already done.
7	PC	I worried that my changes might conflict with earlier decisions.*

Table 3: Post-task questionnaire used to quantify participants' subjective, qualitative experiences relating to theories. *SMM*: Shared Mental Model, *TMS*: Transactive Memory System, *PC*: Perceived level of conflict. *: reverse scored.

ID	Task Description
1	Boat needs to be moved and edited to be appearing like an illusion.
2	Tent needs to be relocated.
3	Rock should be moved: it is not usable by the actors.
4	The sub camera needs to show the boat coming behind Alice and Bob.
5	A log should be floating on the river, and more food items need to be placed around the scene.
6	The light source should appear more low-angled.
7	The paddle should be relocated.
8	The main camera needs to capture wider area, while being located closer to the actors.

Table 4: List of tasks for the river scene.

ID	Task Description
1	Relocate and adjust the mushroom table.
2	Adjust the sub camera. It is currently a placeholder only.
3	Adjust the main directional light.
4	Move the rock: left of the tent, and color should follow the style guidelines.
5	Relocate the view-blocking tree (Spruce 6).
6	Bring the campfire closer to the tent (entrance).
7	Place the log around campfire.
8	Place decorative mushrooms in the scene (use Mushroom Large).

Table 5: List of tasks for the forest scene.

Questionnaire/ Measure	**ID**	*W*	*p*	**Baseline Mean**	**Baseline 95% CI**	***GroundLink* Mean**	***GroundLink* 95% CI**
NASA-TLX	Mental Demand	17.0	0.507	4.17	[3.42, 4.83]	3.92	[3.17, 4.67]
(Lower: Better)	Physical Demand	19.0	0.374	1.83	[1.33,2.33]	2.25	[1.58, 3.00]
	Temporal Demand	34.5	0.733	4.42	[3.42, 5.33]	4.17	[3.33, 5.00]
	Performance	13.0	0.042	4.92	[3.42, 5.33]	3.33	[2.58, 4.17]
	Effort	30.0	0.519	4.25	[3.92, 4.58]	3.92	[3.17, 4.75]
	Frustration	11.0	0.168	4.25	[3.50, 4.92]	3.25	[2.42, 4.17]
UES-SF	FA-S1	12.0	0.209	3.08	[2.50, 3.67]	3.75	[3.17, 4.33]
(Higher: Better)	FA-S2	0.0	0.024	3.33	[2.67, 4.00]	4.08	[3.75, 4.42]
	FA-S3	10.5	0.077	3.42	[2.92, 3.92]	4.25	[3.75, 4.67]
	PU-S1	13.0	0.133	2.50	[2.00, 3.00]	3.25	[2.58, 3.83]
	PU-S2	24.0	0.718	3.00	[2.25, 3.67]	3.17	[2.50, 3.75]
	PU-S3	16.0	0.232	2.58	[2.08, 3.17]	3.08	[2.58, 3.58]
	AE-S1	5.5	0.140	2.75	[2.17, 3.33]	3.33	[2.83, 3.75]
	AE-S2	17.0	0.886	3.17	[2.50, 3.75]	3.17	[2.75, 3.58]
	AE-S3	28.0	0.642	2.92	[2.25, 3.58]	3.17	[2.75, 3.58]
	RW-S1	6.0	0.165	3.50	[3.08, 3.83]	4.00	[3.58, 4.33]
	RW-S2	0.0	0.011	2.83	[2.25, 3.42]	4.08	[3.83, 4.33]
	RW-S3	6.0	0.317	3.75	[3.17, 4.17]	4.08	[3.83, 4.33]
Post-task	1	9.0	0.026	2.75	[2.33, 3.17]	3.58	[3.17, 4.00]
(Table 3)	2	5.0	0.036	2.42	[2.00, 2.92]	3.42	[2.83, 4.00]
(Higher: Better)	3	15.0	0.191	2.58	[2.08, 3.08]	3.17	[2.75, 3.58]
	4	7.0	0.032	2.42	[1.92, 2.92]	3.67	[2.92, 4.33]
	5	4.0	0.015	2.67	[2.17, 3.17]	3.92	[3.33, 4.42]
	6	14.0	0.156	2.92	[2.50, 3.33]	3.42	[3.00, 3.83]
	7	5.0	0.062	2.17	[1.58, 2.83]	3.08	[2.58, 3.58]
Post-task Confidence	Number of Tasks Completed	18.0	0.321	5.42	[4.75, 6.17]	6.00	[5.17, 6.83]
(Higher: Better)	Perceived Confidence (Top-2 box)	2.0	0.014	2.42	[1.58, 3.33]	4.08	[3.00, 5.00]

Table 6: List of all the test statistics and confidence intervals.