



The CoExplorer Technology Probe: A Generative AI-Powered Adaptive Interface to Support Intentionality in Planning and Running Video Meetings

Gun Woo (Warren) Park
warren@dgp.toronto.edu
Microsoft Research
Cambridge, United Kingdom
University of Toronto
Toronto, Canada

Lev Tankelevitch
lev.tankelevitch@microsoft.com
Microsoft Research
Cambridge, United Kingdom

Payod Panda*
payod.panda@microsoft.com
Microsoft Research
Cambridge, United Kingdom

Sean Rintel
serintel@microsoft.com
Microsoft Research
Cambridge, United Kingdom

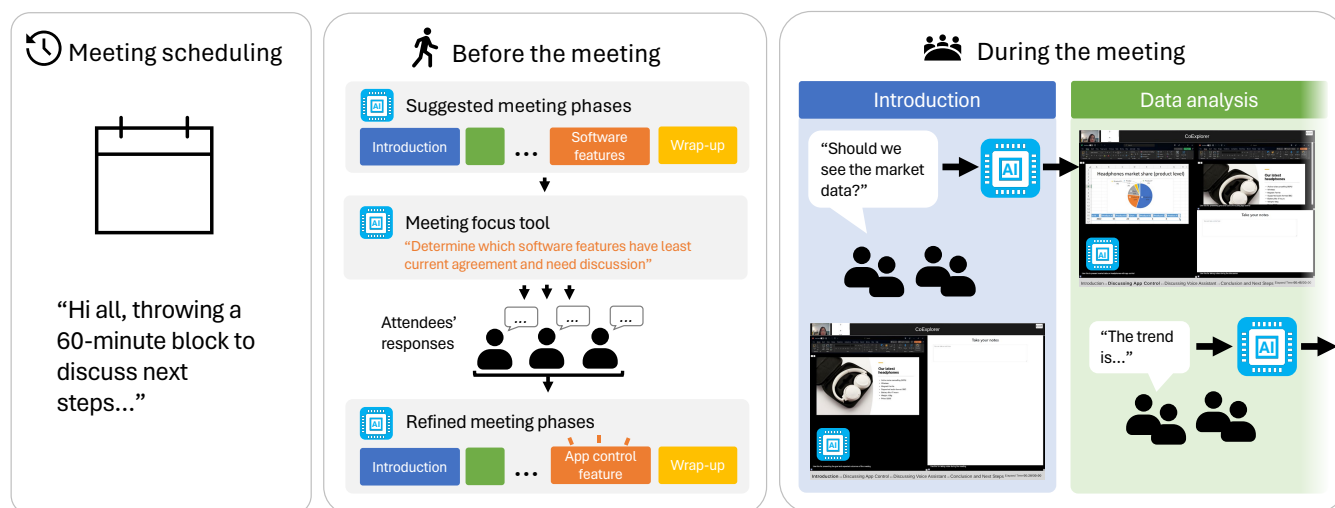


Figure 1: CoExplorer uses Generative AI to reduce the effort of meeting intentionality. From a natural language calendar invitation, GenAI is used to generate a meeting goal, a suggested list of meeting phases and their associated activities, and a meeting focus tool to help find specific needs for discussion. From attendee responses to that tool, GenAI refines the list of meeting phases. When the meeting is underway, GenAI is used to monitor the talk and either suggest or follow users' transitions between phases to fit the trajectory of the meeting. For each phase, GenAI generates apps and optimizes their position and size to fit the meeting phase.

*Corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
DIS '24, July 01–05, 2024, IT University of Copenhagen, Denmark
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0583-0/24/07
<https://doi.org/10.1145/3643834.3661507>

ABSTRACT

Effective meetings are effortful, but traditional videoconferencing systems offer little support for reducing this effort across the meeting lifecycle. Generative AI (GenAI) has the potential to radically redefine meetings by augmenting intentional meeting behaviors. CoExplorer, our novel adaptive meeting prototype, preemptively generates likely phases that meetings would undergo, tools that allow capturing attendees' thoughts before the meeting, and for each phase, window layouts, and appropriate applications and files. Using CoExplorer as a technology probe in a guided walkthrough, we studied its potential in a sample of participants from a global technology company. Our findings suggest that GenAI has the

potential to help meetings stay on track and reduce workload, although concerns were raised about users' agency, trust, and possible disruption to traditional meeting norms. We discuss these concerns and their design implications for the development of GenAI meeting technology.

CCS CONCEPTS

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

KEYWORDS

video meetings, effectiveness, effort, design, adaptive user interface, windowing system, speech recognition; intent recognition, technology probe

ACM Reference Format:

Gun Woo (Warren) 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 *Designing Interactive Systems Conference (DIS '24), July 01–05, 2024, IT University of Copenhagen, Denmark*. ACM, New York, NY, USA, 20 pages. <https://doi.org/10.1145/3643834.3661507>

1 INTRODUCTION

Video meetings have enabled a new era of distributed work, but this has not gone hand-in-hand with improved meeting effectiveness [28, 67, 70], and may have introduced additional fatigue [5]. Evolving from the telephony technology paradigm [52, 59], videoconferencing systems have largely focused on creating a connection between endpoints and representing video and audio on a canvas. While research has explored broadening the technological support for person, reference, and task spaces [13], scheduling [17], agendas [35], moderation [3], and even decision-making [50], improving holistic meeting effectiveness by reducing the effort of planning and running meetings remains an unmet design challenge. For example, calendar interfaces make it easy to schedule a meeting, but very few require the user to specify a meeting *goal*, and videoconferencing interfaces tend to simply provide a place for repeating the text of invitations, without making use of that information as relevant to the interface design or experience [91].

“Generative AI” (GenAI) is a generic term encompassing all end-user tools whose technical implementations include a generative model based on deep learning [88]. GenAI has the potential to solve these problems and open up a new intentionality-driven design paradigm for video meetings. Commercial videoconferencing systems such as Zoom Companion AI [61] and Microsoft Teams Meeting Copilot [69] are already building GenAI features into their meeting systems. However, these just scratch the surface of how GenAI could help teams have more intentional meetings and overcome long-standing ineffectiveness.

This paper reports on a study of the design opportunities and challenges in this space through CoExplorer, a GenAI-powered technology probe (see section 3). CoExplorer represents some of the core capabilities of a holistic goal-driven adaptive system for planning and running meetings. CoExplorer: (1) uses a meeting invitation to generate tentative meeting goal and phases, (2) uses the

meeting goal to generate a meeting focus tool, (3) suggests an appropriate set of applications and their layout for phases (detected in real-time) during a meeting, and (4) analyzes meeting conversation to fine-tune the phase structure automatically.

We conducted a user study involving 26 participants from a global technology company. In the hour-long study, participants were given a guided tour of CoExplorer, frequently engaging with the researcher to discuss the pros and cons of design concepts across the meeting lifecycle. We inquired about their views on various system elements, including their benefits, drawbacks and potential issues. Our thematic analysis of the interviews found that many participants valued the system's ability to facilitate attendee alignment on meeting purpose, and to automatically choose relevant applications or files based on the current meeting phase. However, they also emphasized the importance of verifying the system's choices and appreciated prompts for confirmation at appropriate intervals. These findings highlight an underlying conflict between ease-of-use and higher agency, and we discuss strategies to balance these tensions in our design implications. Our contributions are:

- CoExplorer, a GenAI-powered meeting system that optimizes task space in response to meeting goal-relevant input and immediate activity context,
- Findings from a user study that sheds light on the perceived merits, utility, and concerns of our novel system
- Design insights for future creators interested in developing adaptive windowing systems
- A set of system and user prompts that could facilitate the development of a similar system

2 BACKGROUND AND RELATED WORK

2.1 Meetings and effectiveness

Planning and preparing for an effective meeting takes effort. Much effort goes into appropriate scheduling given different preferences and relevance of attendees [17, 38, 73, 93]. However, the larger problem is whether the meeting should occur at all and whether it has clear goals. Clear goals are imperative for effective meetings [8, 36], but setting goals takes effort and time of an organizer and potentially attendees [11], which many aspects of organizational culture may work against. There is often a lack of understanding of the kind of tasks best suited to meetings [9, 65, 68, 83]. Even when goals are set, meeting preparation is often sacrificed due to time pressures [30, 37].

Running and participating in an effective meeting also takes effort. Agendas are a valuable tool in structuring meetings [11, 21], as long as they do not become rote or a crutch [84]. One important aspect is managing the agenda while also allowing for flexibility necessary to incorporate new ideas, disagreement etc. [56, 58, 84, 106]. Agendas are often a source of tension in terms of how many topics will be included and who wants those topics raised, which leads to the question of which topics must be discussed versus which could be resolved asynchronously before the meeting to condense the agenda and leave time for discussion of differences or productive conflict [35, 71, 83]. While the agenda provides a topical structure to manage, the other effort of running a meeting is the work of managing the meeting's *phases* [25]. Phases are organized locally in the moment, and may be synonymous with

agenda items, be more granular than an agenda item, or encompass several agenda items. Phases are activity-oriented, often involving the use of multiple resources, and different phases may require different resources. Here the effort is in ensuring that the relevant participants are focused on the relevant materials at the right time.

Post-meeting effort starts during the meeting in capturing notes and toward the end capturing action items or next steps [54, 74, 75, 84]. Significant research has gone into reducing this effort through automatic generation of meeting summaries, highlights, and action items [4]. If a meeting has unclear goals and/or not been well-run, post-meeting effort may be higher or another meeting may be required, creating a vicious circle of effort [84].

In this paper we focus on planning and running meetings as the two most crucial points of intervention. The effort of planning and running meetings apply in both unmediated and mediated situations, but video-mediated meetings have added layers that either do not support the effort or even exacerbate it. There is a need for adaptive video meeting systems to reduce the effort of planning by helping turn goals into meaningful action phases for the meeting, and then starting and flexibly changing phases to reduce the effort of running and participating in the meeting.

2.2 Using AI to Improve the Effectiveness and Adaptability of Video Meetings

AI interventions have been proposed as a solution to many kinds of meeting problems, such as to reduce the effort of *pre-meeting* work such as automating meeting scheduling [17] and agenda item voting [11, 34, 35]. Research has also focused on improving decision-making *during the meeting* itself. The field of Group Support Systems (GSS) was very large in the 1980s through early 2000s [50]. De Vreede et al. [23], report that the most common reason for failure in meetings relying on Group Support Systems was "Goal[s] poorly defined by the process owner"—yet ironically, direct technological support for expressing goals was often lacking in such systems. Since then, attention has shifted to how AI agents can facilitate various aspects of meetings, often focused on automatically detecting people's roles [7, 108], useful actions [66], and social dynamics [10] in meetings, and on issues of inclusion and participation (e.g. [32, 41, 57, 76]) than directing the meeting's purpose. To the latter point, Kim and Shah [55] showed that a model that automatically assesses consistency of understanding across meeting participants and suggests topics for review has potential to improve team understanding. Aseniero et al. [3] showed that a system enabling participants to provide real-time meeting feedback via a back-channel can improve their engagement and awareness. Other work showed that automatic detection of different types of conversations [116], emotions [117], bodily cues [16], and environmental factors [20] in meetings can predict aspects of meeting effectiveness. For reducing *post-meeting* effort, significant research has gone into automatic generation of meeting summaries, highlights, and action items, [4, 54, 74, 75], as well as metrics [21, 43] and post-meeting dashboards that attempt to provide feedback on meeting effectiveness [86, 87]. GenAI has sparked a surge in commercial meeting assistance tools. Zoom Meeting AI Companion ¹ and Microsoft

Teams Meeting Copilot ² build AI tools directly into videoconferencing, while Read ³ and Fellow ⁴ are examples of AI services that can integrate with a range of videoconferencing applications.

Both the prior research and commercial tools show promise, but they have shortcomings if we consider the overall problems of ineffective meetings and meeting fatigue. First, they concentrate on features to improve specific pain points, usually because there is a clear technical path to a solution (e.g. automated agenda creation). However, in some cases these features *add* to effort, such as providing flexibility that needs to be manually managed (e.g. using a meeting focus tool [35] or choosing window layouts [45]). More importantly, they miss the bigger picture of improving the *intentionality* of meetings across the meeting lifecycle—everyone in a meeting knowing why they are there, what it is for, and what they need to discuss. In sum, the research gap we explore is how to design a system that reduces the overall effort of meeting intentionality through a goal-driven approach to *focusing* the communication, *refinement*, and *moderation* of purpose-driven effort, and *eliminating* technical overhead in organizing, moderating, sharing documents, and arranging the user interface.

2.3 Adaptive User Interfaces

Adaptive user interfaces adapt to varying user needs either at the time of setup and/or dynamically during operation. By adapting to in-the-moment user needs, adaptive interfaces reduce cognitive load by improving navigability of information [1, 12, 31, 33, 64, 94, 105]. One of the simplest adaptive interfaces is context-dependent adaptive windowing systems. These enable easy retrieval of groups of tools needed in the moment, either as groups of windows per task [95, 103, 104] or time [46]. Co-Activity Manager [44, 45] combines traditional window configuration task-switching with the ability to share a particular configuration with collaborators. In such systems, users manually choose window configurations, although sometimes the placement of components may be optimized automatically [81]. AI has obvious application for optimizing adaptive windowing of relevant resources to fit with context and/or device affordances [79, 97, 98].

In videoconferencing contexts, the layout of interfaces has received attention in specific areas. The issue of whether video adds value to audio at all is a classic problem, and a recent study has found that video of content is important for all business needs video of content is needed but not necessarily video of people [96]. When people are considered to be important, much attention has been paid to exotic solutions that preserve eye gaze, usually by warping the image of the eyes [49] but sometimes by changing the interface itself [107]. Beyond gaze, preserving spatiality has also been explored [42, 102]. End-users also configure their window layouts to suit their working preferences as best they can given the limited flexibility of videoconferencing applications [6], and some classes of end-users (e.g. no- and low- vision [85, 101]) find that the videoconferencing interface was never designed to suit their needs.

However, *adaptive interfaces* have had limited exploration. One long-standing limitation in videoconferencing has been how to

¹<https://news.zoom.us/zoom-ai-companion/>

²<https://support.microsoft.com/en-us/copilot-teams>

³<https://www.read.ai/>

⁴<https://fellow.app/>

enable remote users to see enough of one another or an environment to make sense of a task in context [53]. Research has investigated various approaches to fluid group formation [47], moving video around to improve reference-space capabilities [40] and automated view direction or focus logic [110, 113].

We have found very little on adaptive approaches applied to the agenda or phases of meetings, nor to the display and arrangement of task-oriented materials in video meetings. One example is a 1995 attempt at adapting interface layouts to work contexts. The "dynamically adaptive multi-disciplinary workstation" (DAMDW) [15] changed its display of tools at the outset of a telemedicine consultation to suit a physician's specialization. However, it did not change thereafter and could not adapt to context changes in the moment. In 2003, Antunes and Costa [2] investigated whether genre analysis could be used not just to categorize meetings and activities, but encoded into an interface to reflect the needs of a meeting. They found that the genre concept was useful to clarify the organizational context of meetings and that the systems they tested supported recurrent work practices without undue constraint. However, writing in 2003, they also found that doing the work manually was not a scalable approach: "we currently cannot envisage a way to generalize so many different meetings occurring in organizations." More recently, VisPoll [18] enables an education video streamer to set up areas of a visual stream that users may interact with, which can be used to enable a large audience to provide visual input to questions asked about the stream, and the streamer, in turn, to manage their cognitive load by seeing the aggregate of those audience inputs in a manageable visual fashion. However, this system does not use AI and requires effort of the streamer to decide on the visual polling to be used.

GenAI systems using Large Language Models (LLMs) are well suited to meeting interventions because they are very good at performing using natural language inputs. Text is ideal for pre-meeting requirements, and during the meeting a Voice User Interface (VUI) is well-suited to driving textual representation into language models [22]. Proactive VUIs [90, 109, 115] can support time-critical events, which is essential for recognizing how well meeting talk matches meeting needs. Researchers have utilized LLMs as a tool for prototyping decision-making algorithms in interactive system prototypes [63, 82], but less so in adaptive meeting systems. One example is CrossTalk [112], which harnesses an intelligent substrate to prompt context-appropriate controls such as screen sharing or to present pertinent information. Another inspirational is Co-activity Manager [45], which combines a context-dependent windowing system, a shared task space, and critically, an intelligent adaptive interface.

We speculate that the sparsity of adaptive interface exploration for meetings has to do with the difficulty of balancing flexibility with the manual work required to actually make meaningful changes on the fly, especially if that distracts users from carrying out the actual tasks of the meeting. GenAI holds the promise of reducing this manual load [80]. That being said, using GenAI to enable such capabilities brings its own challenges, among which is that predictions are non-deterministic. This creates risks of users being unable to achieve tasks as expected or in repeatable ways [29].

2.4 Research questions

In sum, video meetings are essential for collaborative work and decision-making, but have been largely stuck in a one-size-fits-all container optimized for connection without much intentionality. The status quo leaves goal setting, agenda creation, and scene setting to people who are often too busy to prepare before the meeting. It also leaves meeting facilitation and resource management to people who are already under cognitive load making decisions about the content and potentially experiencing videoconferencing fatigue based on the interface. To that end, this paper uses a technology probe to explore how GenAI may offer more holistic support for video meetings. We explore three research questions about the opportunities and challenges that users face with a video meeting system that uses GenAI:

- RQ1. What are users' perceptions of GenAI driving a meeting's purpose for a collaborative meeting of teammates?
- RQ2. What are users' perceptions of GenAI surfacing implicit needs and resources in meetings, and the potential effects on participation and effectiveness?
- RQ3. What are users' perceptions of human-on-the-loop interactivity in GenAI video meeting systems?

These questions drive the design concepts that we implement in our technology probe, as laid out in subsection 3.1, our findings in section 5, and our discussion in section 6.

3 DESIGNING COEXPLORER

We began by choosing a meeting type that would benefit from GenAI-augmented help with planning and running a meeting, getting motivation from existing work e.g., [80]. We landed on meetings of cross-functional product teams in the technology industry, which often face struggles of coordination [48], especially at the start of projects when they need to form plans about which they have different opinions and stakes, and for which the resources are scattered across different storage locations and apps. We then developed a fictional meeting scenario emphasizing the need for effective decision-making. In this scenario, the team's current product, the Strata Headphones 2, is lagging behind competitors in market share, and so the team (hardware and software developers, designers, and researchers) needs to decide on the feature set that will increase market share of a new version. The team has files such as the list of design, hardware, and software features that can be included, the price for implementing them, competitor information, and the current product's specification sheet.

3.1 Design Strategies for GenAI interventions

After defining the scenario and design concepts, we considered where CoExplorer would provide the most meaningful interventions in the meeting lifecycle. One early decision was to leave the post-meeting stage out of this probe, partially because there has been significant work on meeting summarization and dashboards already (as reviewed e.g., [4, 54, 74, 75, 86, 87]) and partially because, as above, the level of post-meeting effort is dependent on the planning and running of the meeting. As such, we concentrated on the pre- and during meeting phases. Figure 2 below identifies the primary pain points in these phases (initiating focused discussion; resource management for each phase), and forms of intervention

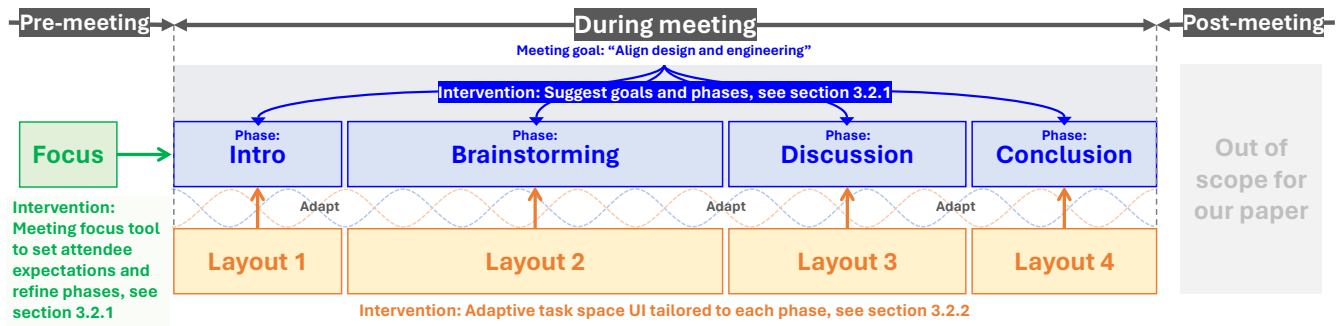


Figure 2: This framework shows the pre-, during, and post- meeting stages and identifies pain points (initiating focused discussion; resource management for each phase), and suggests points of intervention (meeting focus tool; adaptive phase definitions; layouts tailored to each phase).

(meeting focus tool; adaptive phase definitions; layouts tailored to each phase). Figure 2 can be read in conjunction with Figure 1, which shows a timeline of how meeting attendees would experience CoExplorer. We used this framework to craft a narrative of how CoExplorer would provide interventions in the scenario.

Design Strategy 1: Clarify underlying needs and available resources. The clear definition of objectives, agendas, preparatory materials, and task-related resources is an essential aspect of efficient meeting preparation and facilitation, notably in virtual settings [19, 60]. Yet, due to time constraints, both meeting planners and participants may cut corners in laying out pre-meeting preparations [92]. During the meeting, while *explicit* items on the agenda direct the discourse, the meeting actually often transitions through *implicit* phases. These phases may align with agenda points but can also encompass several points or be subdivided into finer details. Moreover, essential files and applications are often linked to each action-oriented phase. The implicit nature of these phases makes accessing appropriate resources at the opportune moment burdensome. This design strategy explores the challenges and potential benefits of GenAI in identifying implicit phases as they arise during the meeting, and using them to drive a specific arrangement of resources and the meeting’s intended progression.

Design Strategy 2: Incorporate collective feedback to shape meeting objectives. Agendas are typically set by meeting organizers, but should also reflect the priorities of other attendees [100], yet current videoconferencing systems do not easily allow for collaborative crafting of meeting objectives. Participants’ contributions can be integrated through mechanisms like voting systems to create or fine-tune the agenda [35]. This design strategy explores the potential of GenAI to produce a tool to assimilate varying viewpoints, focusing the conversation on reconciling differences.

Design Strategy 3: Manage the system through a Human-on-the-Loop (HOTL) methodology. HOTL characterizes human-machine interactions in which the automated system mainly allows humans to *abort* the machine’s decisions, once initiated by the system [77]. This is in contrast to *Human-in-the-Loop* (HITL), where the system requires user approval *prior* to executing an action. This grants automated systems more independence and limits the number of prompts to users, which is compatible with a meeting context where participants often have lower capacity to process information

outside the interpersonal interactions of meetings [78]. Introducing HOTL in meeting systems has been impeded by inadequate predictive performance in natural language models, but with the launch of GPT3.5, there has been noticeable improvement in accuracy across a spectrum of fields [14, 51]. Nonetheless, to foster and sustain the trust of users, HOTL systems must clear a higher threshold than the often-preferred human-*in-the-loop* methodology in Human-Computer Interaction [27, 77, 82, 112]. This design strategy explores the difficulties and possibilities of building trust with users through HOTL, and how this might influence user perceptions of both efficiency and adaptability.

Overall implementation: We implemented CoExplorer using Unity and the GPT-3.5 and GPT-4 GenAI models. We utilized LiveKit for real-time video communication to provide cross-platform compatibility. With CoExplorer, we wanted to explore the extent to which GenAI could generate appropriate meeting interfaces—accordingly, we implement our design strategies primarily through designing GenAI assistants that generated the visual interfaces satisfying our requirements. For implementation details, we include all our system prompts in Appendix B, and specific details for the various strategies in the section below.

3.2 How CoExplorer Facilitates Meetings

3.2.1 Formulating the Initial Phases of the Meeting. The meeting organizer distributes a meeting invitation, and as the meeting time approaches, for each attendee CoExplorer outlines the meeting’s goals (Figure 4A, top left) and describes the reasoning for those goals (Figure 4A, bottom left). It also suggests phases anticipated for the meeting (Figure 4A, left, shown as segmented bars (see Appendix A, Figure 6, left). Details for each phase include the name of the phase, the expected duration, and the pertinent activities for that phase. Below we expand on how the system implements the design strategies above. The strategies are not implemented in the order presented, and one strategy might be implemented multiple times.

Implementing Design Strategy 1: Inspired by the successful use of chaining in LLMs [111], we posited that given a brief meeting invitation text, GenAI could identify both the purpose of the meeting, and a tailored list of applications for each phase. We emphasize our focus is on exploring the potential applications

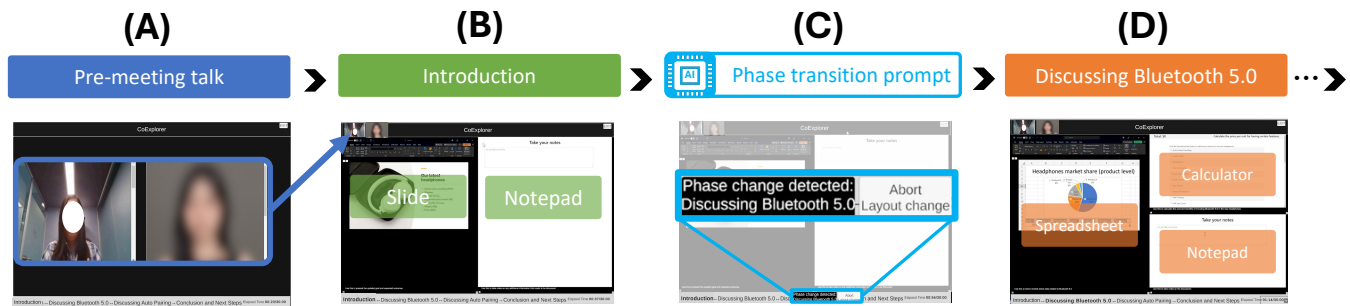


Figure 3: Upon recognition of a phase transition, CoExplorer notifies users and modifies the display to accommodate the new phase.

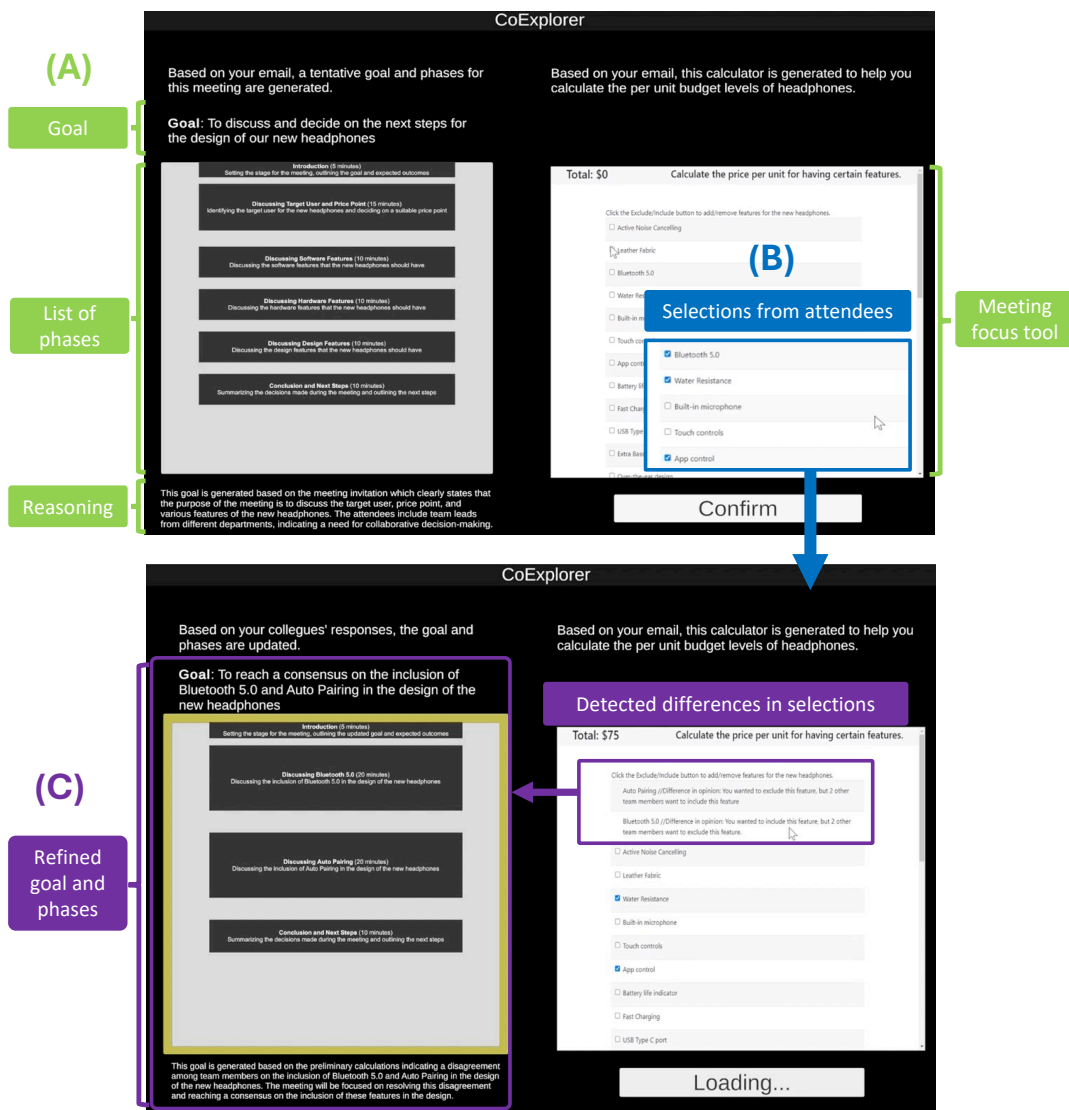


Figure 4: (A) Sequence of phases and tool for determining meeting focus. (B) Participants in the meeting employ the focus tool to select preferred features. (C) CoExplorer uses the aggregated preferences to adjust the meeting's objective and flow.

enabled by GenAI’s capabilities, and thus formal evaluation of this capability is out of scope for this paper. GenAI is tasked with elaborating on the provided invitation text and compiling a list of the phases of the meeting, complete with titles and descriptions. We then guide GenAI to structure these details into a coherent list that includes explanations for the choices it has made, crafting a narrative that can both inform the CoExplorer system and persuade users of the validity of the decisions. To achieve this, we initialise GenAI requests with system prompts, bolstered by examples, to ensure the output is clear and actionable (see Appendix B.1 for the system prompts for phase generation. This preliminary information generated by GenAI is fixed at the beginning and does not undergo refinement during the meeting, intentionally designed to preclude any confusion that might arise from changes during the meeting.

Implementing Design Strategy 2: CoExplorer also primes participants to consider their own requirements and those of the team by generating a discussion initiator. In this scenario, the discussion initiator takes the form of a Meeting Focus Tool that lets meeting participants assess the implications of including specific features in a product from their role perspective and voice their preferences on key features. CoExplorer utilizes GenAI to generate the Meeting Focus Tool—please see Appendix B.3 for the system prompts we used for GPT-4 to generate the tool. The Meeting Focus Tool is displayed in Figure 4B as CoExplorer generates it. Once all preferences are communicated, as seen in Figure 4B, CoExplorer synthesizes these varied responses to refine both the meeting’s objectives and phases. CoExplorer presents the revised information as shown in Figure 4C (see Appendix A, Figure 6, right), which narrows the team’s focal points for discussion to areas with the most divergent views. Areas of divergence are chosen for discussion because live meetings are most suited to dynamic discussions, including productive conflict, while areas of agreement can be handled asynchronously or set aside for later [99]. In this scenario, the GenAI system proposed that the tool take the form of a feature ranking aid (see Appendix B, Figure 7 for the initial raw aid generated, and Figure 6 (B) and (C) for the final version). GenAI’s versatility could enable the tool to alternately manifest as a chatbot or questionnaire.

During our experiments, we observed that GenAI is indeed proficient at devising a tool for this need. To enable a later evaluation, we opted to employ a pre-generated version of the tool rather than generating it on-the-fly. This decision was made with the intention of ensuring a consistent experience for all users. Upon utilizing the pre-generated tool, we noted certain aspects of the user interface, such as the color scheme used for highlighting selected buttons, could be misinterpreted. Consequently, we implemented manual adjustments to the pre-generated version to refine the prototype for better clarity and functionality.

3.2.2 Dynamic Window Management. Upon examination of the updated phases, the team commences the meeting. CoExplorer curates the necessary documents and applications for each phase, generating an ideal layout for the display. Initially, participants engage in a social introduction phase, where their video feeds are maximized (Figure 3A). As the meeting progresses from casual conversation to its formal agenda, CoExplorer detects the shift to a new phase, specifically the project introduction. The project manager outlines the problem being addressed.

Implementing Design Strategy 1: To process spoken dialogue, we segmented speech into discrete utterances based on pauses. For each utterance, a transcript was produced using the Microsoft Azure Speech API. These transcripts were then provided to GenAI, which determined the pre-identified phase of the meeting the utterance pertained to. If GenAI’s prediction indicated a new phase that differed from CoExplorer’s current phase, the user interface was updated to reflect the new phase, shifting to the window layout associated with it. This layout switch was based on a pregenerated list of window configurations established at the start of the meeting.

In organizing the layout of multiple windows, we employed a tiling approach. This decision was influenced by our experimentation with GenAI’s ability to generate window sizes and positions (see subsection B.2 for the system prompts we used to generate layouts). While GPT-4 was adept at creating well-fitted window layouts, GPT-3.5 — the version available to us during prototyping — fell short in this capacity. Consequently, a freeform window layout was deemed unsuitable. Thus, we established that a tiling window layout optimally utilized the available screen space within CoExplorer.

Implementing Design Strategy 3: At this juncture, CoExplorer asks all participants on whether they want to halt the transition (Figure 3C). If no objections occur, CoExplorer adapts the screen layout to fit this more work-focused phase: video sizes are reduced, with a PowerPoint presentation occupying the left side of the display, and a collaborative notepad on the right (Figure 3B).

In the scenario through which participants were guided, a hardware engineer begins discussing the importance of Bluetooth 5.0 for the headphones. This is taken up by a software engineer who points out the challenges in supporting this feature through software. Once more, CoExplorer senses a phase shift (requesting a confirmation on whether to proceed with the change; Figure 3C). With no opposition, CoExplorer transitions to a “Discussing Bluetooth 5.0” phase (Figure 3D). The PowerPoint is replaced with an Excel sheet on the left side, a calculator for the Meeting Focus Tool atop the right side, and the notepad downsized to the bottom right, guiding participants back to the highlighted contentious topic. A full meeting would proceed this way to a decision and discussion of next steps. Due to time limitations, we did not complete the full meeting scenario in CoExplorer. Current GenAI systems such as Microsoft Copilot are able to detect and outline action items from a transcript, and thus a future CoExplorer-like system could use this in conjunction with adaptive windowing to place action items under the video of the relevant person.

4 USER STUDY METHODS

We organized 26 study sessions as controlled simulations designed to accurately capture experiences at the appropriate moments. One participant joined a researcher in each of these sessions. The researcher introduced the scenario and interacted with the participant via a semi-structured interview and think-aloud session. This approach enabled us to gather more detailed insights related to our research questions than we could have obtained through conventional dyadic or triadic studies.

4.1 Participants

We aimed to capture a wide range of experiences with CoExplorer, and thus recruited participants with a variety of experiences, levels of expertise, and roles. Participants were 26 employees drawn from a large global technology company. Employees were recruited through a combination of convenience sampling, snowballing, and batch emails, ensuring a diverse profile by age, region, and gender (Age: 27% 18-29, 54% 30-44, 19% 45-59; Region: 69% UK (including EU), 27% US, 4% Canada; 18 males and 8 females). Of 26 participants, 23% of the participants had the job role title “Principal”, and 42% had “Senior”. Participants included diverse roles such as Project Manager, Cloud Advocate, Software Engineer, IT Service Manager, Hardware Engineer, and Design Researcher. After grouping roles into five broad categories, we had 12% designers, 4% hardware engineers, 31% project managers, 23% researchers, and 31% software engineers.

4.2 System

Participants used their personal laptops with a Windows executable of CoExplorer. The study proceeded with a Node.js server and a CoExplorer Windows executable which synchronized the state, regulated video communication and had a browser-like windowing system. The web page of ChatGPT was manually overlaid with UIKit⁵. The adaptations to the client application (including connectivity with the client application) were manually added for this user study version of CoExplorer to prevent context size overflow (because we forced ChatGPT to generate 30+ feature-price pairs.). The interactive phases were manually initiated by the researcher to prevent misclassification of phases based on user responses to interview questions during the study. To ensure a consistent starting point for all participants, the initial list of phases and a Meeting Focus Tool were pre-generated using ChatGPT. These decisions were made to standardize the CoExplorer experience for the user study.

4.3 Study Process

Each study session lasted approximately one hour and was conducted via Microsoft Teams. Our scenario (discussed above) was chosen from among several that we piloted (e.g. predicting how public policies might affect property prices, and devising designs for a new keyboard) as the most widely relatable to our participant pool. The following steps illustrates the user study:

- (1) Participants chose a role in the scenario closest to their actual work experience (options included Program Manager, Software Engineer, Hardware Engineer, Designer, and Researcher). This role categorization assisted in prompting realistic experiences.
- (2) Participants were introduced to the meeting scenario and CoExplorer. To simulate the meeting lifecycle, participants pasted a calendar invitation into CoExplorer (§3).
- (3) Participants experienced the initial phase list presentation and responded on the focus tool to express their thoughts about the meeting. To demonstrate CoExplorer’s ability to

highlight differences in opinion among participants, we assigned contrasting opinions to the other simulated attendee roles. Based on differences, phase list refinements happened and the system displayed the refined phase list to participants.

- (4) Participants experienced the simulated meeting. This meeting included social introduction phase within a meeting setting, and a minimum of two different phase transitions. They interacted with the system and discussed the value, challenges, and necessary human involvements regarding each facet of CoExplorer. To stimulate authentic feedback, the participants were kept unaware of certain implementation details, such as the source of files and applications.

Participants’ insights were sought after the study session. Each participant received a gift card as thanks.

4.4 Analysis

The data collected from the semi-structured interviews and think-aloud sessions were transcribed using Microsoft Teams and later verified by the researchers. This data underwent a twofold analysis process entailing closed coding for investigating attitudes and perceptions towards specific concepts, such as the system choosing files to share, followed by iterative open coding to uncover rationales behind attitudes or perceptions concerning other parts of the prototype.

5 FINDINGS

5.1 Participants’ current meeting experiences

Participants in our sample frequently participated in remote meetings; 22 daily, nine exceeding ten hours weekly. Except for P12 and P18, all engaged in multi-disciplinary meetings.

Meeting Structure. Over half of our participants (17 of 26) mentioned using predefined agendas for their meetings. P11 stated this was important to their meetings: “*We always have an agenda and a flow... super important.*”. However, 11 participants reported also experiencing non-linear meetings, which was predominantly common among senior roles, like people with ‘Principal’ (4 of 6) in their title, as well as particular roles like project managers (5 of 8) and designers (2 of 3). P22 details this non-linearity: “*very few meetings follow the linear path that one had hoped, and most often you would start with setting up the context and people would have more interest or questions in certain parts of the overall context than others.*”

Eight participants reported meetings to often exceed set time, while nine indicated punctuality, mainly attributed to back-to-back meeting schedules. Opinions on meeting durations were divided, with 12 reporting satisfaction and nine reporting dissatisfaction. Key challenges included biased coverage of topics (P6, P13, P17, P20, P23), insufficient pre-meeting preparation (P15, P19, P20), and running over time (P12, P13, P19, P22). Proposed solutions included pre-meeting voting (P8, P13) and pre-reads (P4, P12, P15, P26).

Multi-window usage. Most participants (22 of 26) used multi-window setups for work, consciously arranging the windows, as described by P10: “*My communication tools... on one screen, my main tools... on the main screen*”. Despite the popularity of multi-windows in work, 18 participants considered it unsuitable for meetings due

⁵<https://getuikit.com>

to difficulties in screen sharing and inconsistent resolutions, emphasizing setup inefficiency (P7, P9, P10) and issues with varying display sizes (P8, P9, P17, P20). These demonstrate a mismatch between user desire and technological support resulting in infrequent use of multi-window setup during meetings.

5.2 RQ1: Using team input to drive meeting purpose

CoExplorer provided ways for meeting organizers to capture initial opinions from attendees. This has potential to democratize the process of affecting the meeting flow, ensuring that voices are heard while maintaining efficiency.

Almost all participants (24 of 26) found capturing meeting attendees' preferences in advance to be helpful, appreciating the ability to align and prioritize discussion topics. P10 mentioned: *"a lot of the discussion often talks about the differences I guess. So it would be good to see where everybody's aligned and where everyone's not aligned."* Over half of our participants (17 of 26) wanted this tool well in advance of the meeting to treat it as a preparatory exercise. Additionally, six participants suggested adding more ways to express opinions, such as text boxes, to make the focus tool more flexible.

Just over half of participants (14 of 26) appreciated the structured and automatic democratization of meeting flow, as it helped ensure that discussions were focused and avoided repeating topics which everyone agreed upon. Participants recognized the potential to save time and delve into more in-depth conversations. P18 mentioned, *"I think that's [the concept of the meeting focus tool] a really good idea. (...) We apparently all agree that this [item on the list] is something that we want. And so we can discuss the other two items more in depth."* Eight participants suggested no need for a human mediator for the meeting flow refinement process. As P2 mentioned: *"I think the system can automatically monitor it. Why use human?"* Additionally, eight participants suggested that the meeting phase list should be refined during the meeting as well. They desired real-time phase changes based on meeting outcomes and time management considerations, as P23 mentioned: *"So I would really appreciate if it could change it [the phase list during the meeting] and they could also tell me like you have this much time left you can talk about this."*

Six participants further emphasized the value of capturing initial thoughts and facilitating collaboration, when asked for the overall advantages of CoExplorer. P19 mentioned: *"The biggest win was the creation and synthesis of data into a way that could be consumed in a unique space for collaboration. (...) I loved how the discussion could be reflected in the agenda and the shifts around that and the ability to manage that and even to do collaboration on that. I thought that those were sort of really, really interesting parts and I think it's very interesting to think about how one can manage."*

5.2.1 Challenge: Embraced role of organizer. Contrary to our design, just over half of our participants (14 of 26) wanted the meeting phase refinement results to go through the meeting organizer. These participants believed that the organizer would have a better understanding of the meeting dynamics and should have the final say on suggestions. P8 mentioned, *"The person who's chairing the meeting should be able to accept, reject, or modify any suggestions because they might know better or they might like the suggestion or*

not." Participants suggested additional roles for organizers, such as confirming with participants or scheduling follow-up meetings, to enhance meeting management.

5.2.2 Challenge: Disruptive refinements. Although many participants appreciated the phase list refinement feature, six opposed pre-meeting refinements, and eight were against during-meeting refinements. Participants highlighted concerns about the potential disruption caused by refined phase lists. Almost half of our participants (11 of 26) worried that the phase list could become too focused, disregarding important topics or overwriting the organizer's agenda. P21 expressed their concerns, stating, *"Before the meeting happens, if this (refined phase list) is showing up as if this is replacing what I have manually set as the agenda, then without seeing that, that's strange. Right? Why would it just get overwritten? My stuff gets overwritten."*

During meetings, P17 was apprehensive about system-driven changes leading to dominance or biased conversation control: *"If someone has a dominant personality, they can take over the meeting. That might be made even worse if someone was speaking a lot and then the phases changed based on that."* P4 and P9 raised concerns about refining phases during meetings without the right stakeholders present. P9 mentioned, *"I wouldn't want the phases to change in the meeting because again, the meeting is called with specific participants. If you had a phase change during the meeting, would the right stakeholders be there? I would rather it take the outcomes or the output of the meeting and then suggest further meetings."*

5.3 RQ2: Surfacing implicit needs and resources

CoExplorer allows participants to structure their meetings without adding additional tasks to their existing meeting routines through automated phase suggestion and refinements.

Seven participants appreciated the easy structuring capability provided by CoExplorer. P10 mentioned, *"So from just a line of text to get a basic structure is very quick, and having someone then able to tweak that is really, really useful."*, and P11 *"I really, really like the first feature that you show me how using an AI model based on a meeting description can generate the agenda items. That's super cool. The second one, the previous input meeting where participants select some topics [in the meeting focus tool] and based on that, adjust the agenda. Those two were mind blowing. Well, those two for me was, oh my God, this is a great way to pick up something big."* Participants' comments emphasized that CoExplorer showed potential to improve the following effectiveness issues:

- **Productivity of meetings** (P10, P16, P22, P25): P22 said *"Too many meetings just don't achieve anything for anyone just because the meeting organizer did not prepare the agenda did not even know what the agenda should be. The flow should be what things to take care of, how to modulate the discussion, how to move from a phase to phase. Hard stuff. So support is definitely needed."*
- **Clarity of meeting goals** (P8): *"Much better clarity on whether the meeting goals are being met"*
- **Efficiency of meetings** (P10, P23): P23 said *"I think I can speed up things because whenever I do meetings, I'm like, OK, give me a second. Let me find that document. Ohh yeah, here's the relevant document or something like that."*

Four participants also valued CoExplorer's ability to assist meeting organizers in driving the meeting. Such a system is particularly helpful for meeting organizers given the tendency for people to impose most responsibility on the organizers to be the primary drivers of meetings. People saw it as a useful tool for meeting orchestration, especially for organizers who may struggle with managing complex cross-functional meetings, as mentioned by P22: *"I think what you're trying to do here makes a lot of sense. And in my experience, I would not assume that more than 20% of people who hosted meetings like these feel comfortable and confident of orchestrating these tricky conversations."*

5.3.1 Challenge: Capturing nuances. Participants expressed concerns about the system's ability to detect nuances in the files and conversations they create. Eight participants mentioned having multiple versions of files while working and struggled to see how the system could accurately infer which version to use (e.g., private vs. public versions, variations on ideas, and iterative refinements). They relied on visual representation or location cues to differentiate files. P5 stated, *"I don't really number the versions really well. I'm trying to do like V1, V2. But then the way I remember which files to share and which files not to share is by remembering their location. So, I need that visual representation of where each file is."*

Participants also stressed the challenge of system detection and reaction to subtle nuances in the conversations during meetings. Some participants (P3, P4, P6, P7, and P22) believed that the system should consider the emotions and thoughts of individual attendees when refining phases. However, participants acknowledged the difficulty of quantifying and addressing these nuanced aspects. P22 questioned whether it was feasible or desirable for an AI system to index on emotions and undertones, stating, *"Alongside the rational content of the discussion there are some undertones inevitably here and as an orchestrator we try and respond to that. I don't know if that can be or even should be replicated in an AI system here. Maybe we are better off just getting rid of all those emotions and undertones. Maybe, maybe not."*

5.3.2 Challenge: Excessive structuring of the meeting. While many participants appreciated the provision of structure by CoExplorer, some (P1, P4, P9, P12, P20, and P22) voiced concerns about excessive structuring. P9 expressed concerns about the impact of imposed time limits: *"It rather limits the discussion or can limit the discussion. If we say, hey, discussing target users but no more than 15 minutes, this is never enough to discuss target users. What if we end up spending 30 minutes on it? Do we then not talk about hardware features?"* P22 suggested reducing phase level detail: *"Rather than going into six different phases and names and descriptions of that, that felt like it is putting too tight a structure already. Instead, if it stayed at the level of just doing three different phases broadly and one liner descriptions, that's a good enough starting point I felt."* P12 noted that overly structured meetings would not align with their company's norms. P12 said, *"We probably wouldn't pay too much attention. ... There may be organizations where meeting culture is much more regimented and they do look to things like this to structure their agenda."* Overall, these thoughts from participants indicate the value and challenge of non-linear meetings. Conversations often go off on tangents and some people value this either implicitly (e.g., P4: *"I think it could create very linear meetings."*), implying that linear meetings might be

undesirable) or explicitly (e.g., P1: *"If that meeting is structured, then people might feel pressured to only stick to the point"*). This suggests the importance of balancing meeting structure and flexibility in meetings.

5.4 RQ3: Merits and challenges of implementing Human-on-the-Loop (HOTL) interactions

We used the HOTL interaction paradigm to manage changes in phases, and the windowing system to minimize distractions and information overload during meetings. Participants had mixed responses to the various aspects of HOTL design, and we identified some common challenges.

Ten participants expressed excitement about the system's capability of curating shared applications and files. P4 appreciated the system's ability to improve document discoverability, allowing participants to easily access relevant resources during meetings: *"One of the most difficult things in meetings is, do we have a document around this... So having that as resources that people can get to easily is really vital in a meeting."* P4, P17, and P24 mentioned that system could auto-populate the windows without verifying the window 'sharing' decision with any attendees. Furthermore, almost half of our participants (12 of 26) found that the system showing content in the shared task space was not intrusive.

Participants also provided input on the search scope for curation. A slim majority (15 of 26) favored curating work-related files, such as those within a shared work cloud, as they assumed team members would have common access permissions. In contrast, indexing files from the internet (4 of 26), local systems (6 of 26), or anywhere available (2 of 26) were less preferred options. Additionally, over half of our participants (17 of 26) mentioned trusting the curated contents. However, many participants linked their trust to conditions, such as their continued experience with the system (P2, P21, P23, P25) or the ability to see previews of the chosen files (P3, P4, P10, P12, P26).

The ability of the system to automatically add windows was positively received by almost half of our participants (12 of 26; 7 were neutral). Participants found it particularly useful for quickly retrieving and sharing information during meetings, thereby making meetings more efficient. P16 mentioned: *"I think in terms of getting information out quickly, rather than having to look for it all, it would be useful. I feel like a lot of meetings they're a big time sink. It's just people switching in between apps."* Six participants also appreciated the concept of the system resizing and relocating windows, finding it beneficial for organizing and managing their workspace during meetings.

However, participants were less inclined to have the system remove windows. Eight expressed that they did not want the system to close windows, but most (20 of 26) mentioned that they expected the system to allow for window recall in a format similar to the last opened window interface in web browsers or a hamburger menu.

5.4.1 Challenge: Agency. Using the HOTL paradigm, we only allowed users to *abort* actions performed by the system. This raised concerns among participants, who desired a more traditional human-in-the-loop approach to retain agency.

Most (21 of 26) participants mentioned that the system should consult humans when showing any file or application on the shared

task space. Additionally, nine participants believed that similar checks should be performed for windowing system operations, and over half (16 of 26) said that checks should be performed for phase changes.

Participants recognized the potential distractions that such prompts could cause during meetings. They emphasized the need to reduce the frequency of prompts and suggested offloading these prompts as a pre-meeting exercise. P6 highlighted the necessity of reducing the frequency of prompts: *“You don’t want to present and then you have something popping up quite regularly for some reason. That could be distracting.”*, as well as P13: *“I wouldn’t want to be in a meeting and then be notified 20 times that we’re now moving to a next slide.”* Six participants suggested the creation of a “Meeting choreography”, where the meeting organizer reviews and decides on the files, applications, and phases that should be prepared before joining the meeting. Similarly, nine participants proposed pre-defining the scope of the search for curation, such as putting files into a certain folder or tagging files in advance of the meeting.

5.4.2 Challenge: Algorithmic aversion/mistrust. Participants expressed concerns about the ability of the system to generate content accurately. Nearly every participant expressed significant skepticism and mistrust regarding the actual content generated by CoExplorer without seeing the system err (except P20 who mentioned that some of the suggestions given by CoExplorer were not suitable). This suggests of a pre-existing aversion towards, or mistrust of, automated systems. As P9 summarized, *“My experience with meeting technology is when people walk in and they click go and it works, they are super happy. The minute something doesn’t work, they are super unhappy, and there’s no middle ground.”* Participants emphasized the importance of the system understanding the nuances and context correctly, without displaying irrelevant screens or inappropriately transitioning the phases, based on their experience with meeting technology. In particular, seven participants criticized the system’s inability based on their existing experiences, sometimes even before using the system, rather than evaluating the system’s performance during the study. P3 based their view of the system on their experience with ChatGPT: *“So a lot of these multidisciplinary meetings are quite sensitive. They’re quite delicate balance of understanding each other’s perspective so people don’t get hurt.(...) Expecting ChatGPT to come up with all these subtle nuances and understand these nuances, it’s slightly unrealistic because that’s never going to happen.”* P5 also expressed concern regarding an algorithm-generated decision: *“I feel like if this is auto generated, I might not even know if this is something that we’re gonna talk about in this particular meeting for now.”* These concerns highlighted the need for the system’s content generation to be dependable and aligned with participants’ expectations.

6 DISCUSSION

Our research provides empirical early evidence for design concepts (subsection 3.1) that can enhance the productivity of meetings. There is value in GenAI’s ability to extract and reformulate goal-direction from existing information to support intentionality in meeting planning and execution. This includes helping organizers highlight meeting goals and adjust plans to address inter-attendee issues before the meeting, and surfacing the potential implicit phase

structure of the meeting in advance, while adapting it to in-the-moment activities. Our design concepts were perceived to be more useful when larger quantities of resources and participants needed management, and when the meetings were intended to be more structured than non-linear. Participants gave us insight into circumstances where GenAI might not be beneficial, and their overall concerns about the place of and trust in GenAI for augmenting meeting intentionality.

6.1 Trust in AI systems: balancing system autonomy with user trust

Participants appreciated our intention to equip meeting attendees with an intelligent system. However, despite the potential benefits the study shows, there is a general reluctance among meeting attendees to trust AI systems which do more than merely give suggestions. This resonates with existing literature on algorithmic aversion within decision-making fields [26]. Participants wanted decisions made by the system to be validated by the meeting organizer. Even though CoExplorer did not make mistakes in predictions for most participants, we still noted a marked mistrust toward CoExplorer. This issue surfaced in the initial stages of the user study when interaction with the system was limited, and lasted throughout.

Designers need to address this mistrust to integrate more advanced intelligent systems into meetings. The majority of participants and existing literature suggest a human-in-the-loop approach [82, 112], highlighting the need for human intervention and oversight. However, frequent interventions can overload meeting attendees with information, reducing their capacity to process it. Further, constant system prompts could be disruptive and distracting to the meeting process, as recognized by our findings. Even sharing a single file could involve reviewing multiple pieces of information—location, permissions, format—which makes it pivotal to balance the trade-off between information provision and disruption. Future investigations could focus on identifying the threshold of disruption for informational prompts—how much information is it safe for these prompts to have, and how frequently can they be shown to be effective?

Design implication 1: A balance between keeping meeting attendee effort low during the meeting, while offering more agency and establishing trust, could be reached by offloading some of the human-in-the-loop tasks to periods of low engagement—prior to the meeting, during breaks, or when certain attendees are less engaged. Example human-in-the-loop prompts are shown in Figure 5.

6.2 Structure in meetings: linearity, stages, and approach

6.2.1 Linearity of meetings. Our findings suggest that both linear and non-linear meetings are common for our study participants, leading to mixed responses about surfacing the implicit phase structure in meetings. Some participants valued the structure that CoExplorer imposed on meetings by surfacing meeting goals and phases, helping to clarify attendee perceptions in advance of meetings, and supporting organizers in driving the meeting. This was perceived

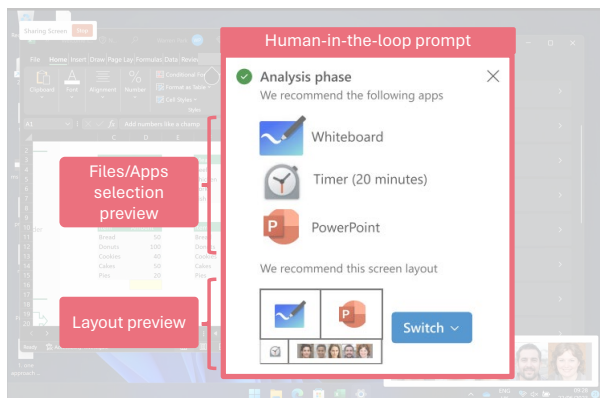


Figure 5: An example human-in-the-loop design for verifying app selections

to improve meeting clarity and therefore productivity. However, some participants expressed concerns about the excessive structure imposed by CoExplorer, worrying that it may limit the natural spontaneity and flexibility of conversations. Considering these individual perceptions and preferences, it may be crucial for meeting support systems to capture the general consensus on the desired level of structure and adapt the workflow accordingly, rather than adopting a one-size-fits-all approach. Reconciling the variations in the group consensus and the individual preferences is a direction that needs further research. Systems could also help people decide on how much structure may be needed depending on their goals. It also points out the value in developing ways to adaptively adjust the level of structure during meetings. For instance, when people are moving through phases smoothly, the suggestion of structure could be more restrained, whereas if the meeting is being side-tracked, structure could be suggested in a stronger way by tracking the progress of the meeting through an intelligent system.

Design implication 2: The design of intelligent meeting systems should consider ways to address the tension between structured, linear meetings enabling intentionality and clarity, and less structured, non-linear meetings, enabling flexibility. This includes capturing attendees' preferences about structure as well as helping decide how much structure may be beneficial for their goals. Variations in need for structure during a given meeting should also be considered to adapt the interface.

6.2.2 Meeting stages and broader context. Our study underscores the strengths of intelligent systems in creating a trajectory of intentionality across the meeting lifecycle. This can help workers focus on making meetings matter by augmenting effort rather than hoping that unfocused effort during the scheduled meeting time will have good results.

An interesting avenue for expansion might be giving the systems capability to track progress beyond the context of individual meetings, and tapping into the larger context of team projects. For instance, the meeting focus tool or phase generation could adapt to the stage the project is in (e.g., initial conceptual design, prototyping, or pre-launch). This might have several benefits. For instance,

a system that tracks an entire project could prevent the team from wasting time rehashing decisions that were already made. We could bolster user interface support with features such as 'graying out' previously discussed topics to signal they should not be revisited. Further research may illuminate how meetings can best use GenAI to ensure they are run more efficiently and coherently.

Design implication 3: The meeting life cycle extends much further in both directions (before and after meetings), and designing interactions for these stages might make attendees more intentional throughout the meeting life cycle. For instance, the meeting focus tool might be most beneficial when shown to attendees well in advance of the meeting. We should also consider designing experiences that adapt beyond just the scheduled meeting, such as the context of the larger project within which the meeting is taking place.

6.3 Task load and agency: striking a balance between organizers and attendees

Our design intention for CoExplorer was to reduce the effort and task load for meeting attendees and organizers by using automation. However, participants' feedback indicates the need for some human intervention, with most pointing to the meeting organizer or chair as the appropriate person to validate the decisions made by the AI system. For example, participants suggested more human-in-the-loop design elements or more previews for organizers, which provides information on differing options and opinions among attendees. The implication is that the organizer, as an obvious target, is responsible for resolving these differences before the meeting begins. Implementing such a system, then, reduces the task load and effort for meeting attendees at the expense of higher task load for the meeting organizer. This might be mitigated by ensuring that the added tasks for the organizer happen during low engagement times, as discussed in Design Implication 1 above.

These implied tasks are a consequence of participants' perceived necessity for human moderation of decisions made by GenAI systems. Indeed, some participants saw the current design as limiting their agency during meetings. CoExplorer's phase suggestions were perceived as a rigid task list that devalued their contributions, and would prefer such a system to be a "suggester" rather than a "dictator". This suggests a tension between the simultaneous desire for low task load (e.g., fewer verification prompts) and high agency (e.g., more human moderation). This tension merits further investigation.

Design implication 4: There's a delicate balancing act between task-load and agency that an intelligent meeting system must manage. Systems should act as an assistant—reducing task-load and enhancing agency—not a decision-maker.

6.4 Balancing efficiency and sociality

Per RQ1 (using team input to drive meeting purpose), many of our participants appreciated that the meeting focus tool could help steer meeting activity to the topics that needed most discussion, in this case based on differences of opinion that needed to be resolved. This is clearly a step in the right direction towards effective meetings, but there are other factors at play. First, many meetings have a social

component or may be entirely about sociality, such as establishing or maintaining collegiality [9]. To the extent that a meeting is the right venue for sociality, privileging only task efficiency with no time or highly constrained time for small talk at the beginning and end of meetings [114] may be counter-productive. It might not only reduce team engagement but also reduce serendipitously raising important issues. Further, if all meetings focus only on points of difference, this might lead to problems with team morale, making it seem as if the team disagrees more than is actually the case, and/or leading to a dread of meetings because they always involve the difficult work of resolving difficulties. There may be value in teams first considering their points of agreement to establish common ground and trust for dealing with disagreement.

As such, even if resolving disagreements is the most important value that a meeting has, either the meeting focus tool or some other aspect of the meeting focus and/or phase creation of the GenAI needs to be guided with meta-prompts to balance sociality and agreement with tasks and disagreement. Gonzales-Diaz et al. [39] propose a version of of this balance in their Transitions concept. They propose an interface that literally changes before, during, and after a meeting. Pre- and post- meeting interfaces are simple blank spaces that afford free-form spatialized talk, while the during-meeting phase changes to look more like a traditional meeting space, in their case an auditorium, which is also a literal visual indication of space for ‘real work’ and a literal focusing mechanism of an audience on a presenter. They note that their design of the visual look and user experience of “making space for social time” needs significant work, and this might be quite different for different teams. As such, this would be another area in which GenAI’s ability to make use of user preferences might create very different-looking meeting interfaces for different teams while following the guiding principle. The precise balance of effectiveness and sociality will, of course, require future exploration.

Design implication 5: The design of intelligent meeting systems should consider ways to balance effectiveness with sociality. This must preserve the social and ritual value of meetings when it is needed. It must also ensure that the drive for effectiveness does not create a cascade of social problems that undermine the very effectiveness that is aimed for, such as balancing agreement and disagreement so that attendees can trust that conflict will be productive.

6.5 Limitations

We acknowledge that our study has some limitations. Our participants were recruited from a single company. This choice was made to maintain a consistent meeting culture across our participants while comparing differing perceptions towards the same prototype. Future research could broaden the scope to include multiple varied companies and sectors with differing meeting practices.

Our participant pool was not gender-balanced (female participants were only 31% of the total, none were non-binary). The 31% female proportion is unfortunately similar to the average in the technology sector in 2023 [24], but future research should seek greater gender balance.

Since confidentiality issues prevented participants from bringing their actual work files into the study, our guided walk-through used

fictitious documents. Future GenAI design research will, of course, benefit from incorporating real-world data, documents, meetings, etc., which would allow participants to evaluate the benefits or drawbacks with more relevance to their work.

We also acknowledge that our study method did not validate the generated phases and support tools, nor was it a comparative evaluation of GenAI creating an interface versus an interface created by a person, or GenAI running meetings versus meetings run by people. These limitations are a natural consequence of choosing the guided walk-through method. In common with the Participatory Prompting method [89], rather than speculating on fictional GenAI capabilities, our technology probe walk-through allowed users to engage with a working example of a system that was not yet possible to build to a fidelity necessary for plausible evaluation. For example, as of the time of writing, GenAI is not proficient at abstract reasoning (e.g., see [62, 72]), which limited the kinds of meeting phases that our implementation of CoExplorer could detect. Given the early state of GenAI in this usage area, our approach allowed us to capture rich insights into users’ reactions to GenAI interventions at various stages of the meeting lifecycle, and what challenges and opportunities that might bring.

Finally, this research did not extensively explore a range of design options for meeting interfaces, as this study focused on users’ reactions to the meaning and logic of GenAI driving interface changes, not on evaluating or optimizing the look and feel of the system. We settled on the current design based on a pilot study of the clarity of the logic for users, but we could have experimented with more ways to present and organize different phases and tools through co-design sessions or focus group studies. We leave such investigations to future work.

7 CONCLUSION

We designed CoExplorer, a GenAI video meeting system that aims to promote intentionality while reducing the effort of participating in a meeting. We used CoExplorer as a technology probe, utilizing specific design concepts that provoked strong responses from our study participants. We delved into possible user concerns regarding the automation of meeting planning duties and surfacing implicit needs and resources during a meeting. Our research further highlighted the difficulties in maintaining user control amidst the automation of decision-making aspects. Notably, we discerned a possible algorithmic aversion or mistrust that could potentially hinder the success of future systems. Based on these findings, we discussed the challenges in depth and proposed design implications linked to our results. Additionally, we also provide the prompts we used with an LLM in order to generate the UI and responses in the Appendix. Our results and design implications suggest next steps for creating GenAI-mediated meeting systems, and the Appendix helps other researchers recreate and build upon our system.

ACKNOWLEDGMENTS

We thank Ava Scott for helpful discussions, and the study participants for volunteering their time. We acknowledge the use of ChatGPT by OpenAI on the implementation of the prototype, as well as proofreading the text.

REFERENCES

- [1] Khalid Al-Omar and Dimitrios Rigas. 2009. A user performance evaluation of personalised menus. In *2009 Second International Conference on the Applications of Digital Information and Web Technologies*. IEEE, London, United Kingdom, 104–109. <https://doi.org/10.1109/ICADIWT.2009.5273847>
- [2] Pedro Antunes and Carlos J. Costa. 2003. From Genre Analysis to the Design of Meetingware. In *Proceedings of the 2003 ACM International Conference on Supporting Group Work* (Sanibel Island, Florida, USA) (GROUP '03). Association for Computing Machinery, New York, NY, USA, 302–310. <https://doi.org/10.1145/958160.958209>
- [3] Bon Adriel Aseniero, Marios Constantinides, Sagar Joglekar, Ke Zhou, and Daniele Quercia. 2020. MeetCues: Supporting online meetings experience. In *2020 IEEE Visualization Conference (VIS)*. IEEE, 236–240.
- [4] Sumit Asthana, Sagih Hilleli, Pengcheng He, and Aaron Halfaker. 2023. Summaries, Highlights, and Action items: Design, implementation and evaluation of an LLM-powered meeting recap system. <https://doi.org/10.48550/arXiv.2307.15793> arXiv:2307.15793 [cs].
- [5] Jeremy N. Bailenson. 2021. Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue. *Technology, Mind, and Behavior* 2, 1 (Feb. 2021). <https://doi.org/10.1037/tmb0000030>
- [6] Karolina Balogova and Duncan Brumby. 2022. How Do You Zoom?: A Survey Study of How Users Configure Video-Conference Tools for Online Meetings. In *Proceedings of the 1st Annual Meeting of the Symposium on Human-Computer Interaction for Work* (Durham, NH, USA) (CHIWORK '22). Association for Computing Machinery, New York, NY, USA, Article 5, 7 pages. <https://doi.org/10.1145/3533406.3533408>
- [7] Satanjeev Banerjee and Alexander Rudnicky. 2004. Using simple speech-based features to detect the state of a meeting and the roles of the meeting participants. In *Eighth International Conference on Spoken Language Processing*.
- [8] Henning Bang, Synne L. Fuglesang, Mariann R. Ovesen, and Dag Erik Eilertsen. 2010. Effectiveness in top management group meetings: The role of goal clarity, focused communication, and learning behavior: Effectiveness in top management meetings. *Scandinavian Journal of Psychology* 51, 3 (Jan. 2010), 253–261. <https://doi.org/10.1111/j.1467-9450.2009.00769.x>
- [9] Rachel Bergmann, Sean Rintel, Nancy Baym, Advait Sarkar, Damian Borowiec, Priscilla Wong, and Abigail Sellen. 2022. Meeting (the) Pandemic: Videoconferencing Fatigue and Evolving Tensions of Sociality in Enterprise Video Meetings During COVID-19. *Computer Supported Cooperative Work (CSCW)* (Nov. 2022). <https://doi.org/10.1007/s10606-022-09451-6>
- [10] Indrani Bhattacharya, Michael Foley, Ni Zhang, Tongtao Zhang, Christine Ku, Cameron Mine, Heng Ji, Christoph Riedl, Brooke Foucault Welles, and Richard J Radke. 2018. A multimodal-sensor-enabled room for unobtrusive group meeting analysis. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*. 347–355.
- [11] Ana Cristina Bicharra Garcia, John Kunz, and Martin Fischer. 2004. Cutting to the Chase: Improving Meeting Effectiveness by Focusing on the Agenda. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) (CSCW '04). Association for Computing Machinery, New York, NY, USA, 346–349. <https://doi.org/10.1145/1031607.1031664>
- [12] Peter Brusilovsky. 1996. Methods and techniques of adaptive hypermedia. (1996).
- [13] Bill Buxton. 2009. Mediaspace – Meaningspace – Meetingspace. In *Media Space 20 + Years of Mediated Life*, Steve Harrison (Ed.). Springer, London, 217–231. https://doi.org/10.1007/978-1-84882-483-6_13
- [14] Lingjiao Chen, Matei Zaharia, and James Zou. 2023. How is ChatGPT's behavior changing over time? <http://arxiv.org/abs/2307.09009> arXiv:2307.09009 [cs].
- [15] W.J. Chimiak, R. Rainer, and J. Cook. 1995. An adaptive multi-disciplinary telemedicine system. In *Proceedings of the Fourth International Conference on Image Management and Communication (IMAC 95)*. 244–249. <https://doi.org/10.1109/IMAC.1995.532593>
- [16] Jun-Ho Choi, Marios Constantinides, Sagar Joglekar, and Daniele Quercia. 2021. KAIROS: Talking heads and moving bodies for successful meetings. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*. 30–36.
- [17] Andy Chun, Hon Wai, and Rebecca Y. M. Wong. 2003. Optimizing agent-based meeting scheduling through preference estimation. *Engineering Applications of Artificial Intelligence* 16, 7 (Oct. 2003), 727–743. <https://doi.org/10.1016/j.engappai.2003.09.009>
- [18] John Joon Young Chung, Hijung Valentina Shin, Haijun Xia, Li-yi Wei, and Rubaiat Habib Kazi. 2021. Beyond Show of Hands: Engaging Viewers via Expressive and Scalable Visual Communication in Live Streaming. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 109, 14 pages. <https://doi.org/10.1145/3411764.3445419>
- [19] Melissa A. Cohen, Steven G. Rogelberg, Joseph A. Allen, and Alexandra Luong. 2011. Meeting design characteristics and attendee perceptions of staff/team meeting quality. *Group Dynamics: Theory, Research, and Practice* 15, 1 (March 2011), 90–104. <https://doi.org/10.1037/a0021549>
- [20] Marios Constantinides, Sanja Šćepanović, Daniele Quercia, Hongwei Li, Ugo Sassi, and Michael Eggleston. 2020. ComFeel: Productivity is a matter of the senses too. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–21.
- [21] Ross Cutler, Yasaman Hosseinkashi, Jamie Pool, Senja Filipi, Robert Aichner, Yuan Tu, and Johannes Gehrke. 2021. Meeting Effectiveness and Inclusiveness in Remote Collaboration. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 1–29. <https://doi.org/10.1145/3449247>
- [22] Ritwik Dasgupta. 2018. *Voice User Interface Design: Moving from GUI to Mixed Modal Interaction*. Apress, Berkeley, CA. <https://doi.org/10.1007/978-1-4842-4125-7>
- [23] Gert-Jan De Vreede, Robert M. Davison, and Robert O. Briggs. 2003. How a silver bullet may lose its shine. *Commun. ACM* 46, 8 (Aug. 2003), 96–101. <https://doi.org/10.1145/859670.859676>
- [24] Emily Dean. 2023. 60+ Women in Tech Statistics You Need to Know in 2023 – Closing the Gaps – techopedia.com. <https://www.techopedia.com/women-in-tech-statistics>
- [25] Arnulf Deppermann, Reinhold Schmitt, and Lorenza Mondada. 2010. Agenda and emergence: Contingent and planned activities in a meeting. *Journal of Pragmatics* 42, 6 (June 2010), 1700–1718. <https://doi.org/10.1016/j.pragma.2009.10.006>
- [26] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. 2014. Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err. (2014), 13.
- [27] Rui Dong, Zhicheng Huang, Ian Iong Lam, Yan Chen, and Xinyu Wang. 2022. WebRobot: Web Robotic Process Automation using Interactive Programming-by-Demonstration. <http://arxiv.org/abs/2203.09993> arXiv:2203.09993 [cs].
- [28] 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/>
- [29] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). Association for Computing Machinery, New York, NY, USA, 278–288. <https://doi.org/10.1145/3025453.3025739> event-place: Denver, Colorado, USA.
- [30] Sayed M. Elsayed-Elkhouly, Harold Lazarus, and Volville Forsythe. 1997. Why is a third of your time wasted in meetings? *Journal of Management Education* 16, 9 (Jan. 1997), 672–676. <https://doi.org/10.1108/02621719710190185> Publisher: MCB UP Ltd.
- [31] Leah Findlater, Karyn Moffatt, Joanna McGrenere, and Jessica Dawson. 2009. Ephemeral adaptation: the use of gradual onset to improve menu selection performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Boston MA USA, 1655–1664. <https://doi.org/10.1145/1518701.1518956>
- [32] Szu-Wei Fu, Yaran Fan, Yasaman Hosseinkashi, Jayant Gupchup, and Ross Cutler. 2022. Improving meeting inclusiveness using speech interruption analysis. In *Proceedings of the 30th ACM International Conference on Multimedia*. 887–895.
- [33] Krzysztof Z. Gajos, Katherine Everitt, Desney S. Tan, Mary Czerwinski, and Daniel S. Weld. 2008. Predictability and accuracy in adaptive user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Florence Italy, 1271–1274. <https://doi.org/10.1145/1357054.1357252>
- [34] Ana Cristina Bicharra Garcia, John Kunz, and Martin Fischer. 2003. Meeting Details: Methods to instrument meetings and use agenda voting to make them more effective. In *meeting of the Center for Integrated Facility Engineering, Stanford (no. TR147)*.
- [35] Ana Cristina Bicharra Garcia, John Kunz, and Martin Fischer. 2005. Voting on the agenda: the key to social efficient meetings. *International Journal of Project Management* 23, 1 (Jan. 2005), 17–24. <https://doi.org/10.1016/j.ijproman.2004.05.003>
- [36] Jennifer L. Geimer, Desmond J. Leach, Justin A. DeSimone, Steven G. Rogelberg, and Peter B. Warr. 2015. Meetings at work: Perceived effectiveness and recommended improvements. *Journal of Business Research* 68, 9 (Sept. 2015), 2015–2026. <https://doi.org/10.1016/j.jbusres.2015.02.015>
- [37] Paul Geyer. 2020. Do directors prepare well for board meetings? *Governance Directions* 72, 6 (Aug. 2020), 272–276. <https://doi.org/10.3316/informit.263991955412630> Publisher: Governance Institute of Australia.
- [38] Michael Gibbs, Friederike Mengel, and Christoph Siemroth. 2021. Work from Home & Productivity: Evidence from Personnel & Analytics Data on IT Professionals. <https://doi.org/10.2139/ssrn.3843197>
- [39] Carlos Gonzalez Diaz, John Tang, Advait Sarkar, and Sean Rintel. 2022. Making Space for Social Time: Supporting Conversational Transitions Before, During, and After Video Meetings. In *2022 Symposium on Human-Computer Interaction for Work (CHIWORK 2022)*. Association for Computing Machinery, New York, NY, USA, 1–11. <https://doi.org/10.1145/3533406.3533417>

- [40] Jens Emil Grønbaek, Banu Saatci, Carla F. Griggio, and Clemens Nylandstedt Klokmoose. 2021. MirrorBlender: Supporting Hybrid Meetings with a Malleable Video-Conferencing System. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 451, 13 pages. <https://doi.org/10.1145/3411764.3445698>
- [41] Zixuan Guo and Tomoo Inoue. 2019. Using a Conversational Agent to Facilitate Non-Native Speaker's Active Participation in Conversation. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI EA '19). Association for Computing Machinery, New York, NY, USA, 1–6. <https://doi.org/10.1145/3290607.3313075>
- [42] Jörg Hauber, Holger Regenbrecht, Mark Billinghurst, and Andy Cockburn. 2006. Spatiality in Videoconferencing: Trade-Offs between Efficiency and Social Presence. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* (Banff, Alberta, Canada) (CSCW '06). Association for Computing Machinery, New York, NY, USA, 413–422. <https://doi.org/10.1145/1180875.1180937>
- [43] Yasaman Hosseinkashi, Lev Tankelevitch, Jamie Pool, Ross Cutler, and Chinmaya Madan. 2024. Meeting effectiveness and inclusiveness: large-scale measurement, identification of key features, and prediction in real-world remote meetings. arXiv:2304.00652 [cs.HC]
- [44] Steven Houben, Jakob E Bardram, Jo Vermeulen, Kris Luyten, and Karin Coninx. 2013. Activity-centric support for ad hoc knowledge work: a case study of co-activity manager. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013).
- [45] Steven Houben, Jo Vermeulen, Kris Luyten, and Karin Coninx. 2012. Co-activity manager: integrating activity-based collaboration into the desktop interface. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. ACM, Capri Island Italy, 398–401. <https://doi.org/10.1145/2254556.2254633>
- [46] Donghan Hu and Sang Won Lee. 2020. Scrapbook: Screenshot-based Bookmark for Effective Curation of Digital Resources. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology*. ACM, Virtual Event USA, 46–48. <https://doi.org/10.1145/3379350.3416181>
- [47] Erzhen Hu, Md Aashikur Rahman Azim, and Seongkook Heo. 2022. FluidMeet: Enabling Frictionless Transitions Between In-Group, Between-Group, and Private Conversations During Virtual Breakout Meetings. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 511, 17 pages. <https://doi.org/10.1145/3491102.3517558>
- [48] Avan R. Jassawalla and Hemant C. Sashittal. 2000. Cross-Functional Dynamics in New Product Development. *Research-Technology Management* 43, 1 (2000), 46–49. <https://doi.org/10.1080/08956308.2000.11671331>
- [49] Jason Jerald and Mike Daily. 2002. Eye Gaze Correction for Videoconferencing. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications* (New Orleans, Louisiana) (ETRA '02). Association for Computing Machinery, New York, NY, USA, 77–81. <https://doi.org/10.1145/507072.507088>
- [50] Starr Roxanne Hiltz and Jerry Fjermestad. 2000. Group Support Systems: A Descriptive Evaluation of Case and Field Studies. *Journal of Management Information Systems* 17, 3 (2000), 115–159. <https://doi.org/10.1080/07421222.2000.11045657>
- [51] Douglas Johnson, Rachel Goodman, J Patrinely, Cosby Stone, Eli Zimmerman, Rebecca Donald, Sam Chang, Sean Berkowitz, Avni Finn, Eiman Jahangir, Elizabeth Scoville, Tyler Reese, Debra Friedman, Julie Bastarache, Yuri Van Der Heijden, Jordan Wright, Nicholas Carter, Matthew Alexander, Jennifer Choe, Cody Chastain, John Zic, Sara Horst, Isik Turker, Rajiv Agarwal, Evan Osmundson, Kamran Idrees, Colleen Kiernan, Chandrasekhar Padmanabhan, Christina Bailey, Cameron Schlegel, Lola Chambliss, Mike Gibson, Travis Osterman, and Lee Wheless. 2023. *Assessing the Accuracy and Reliability of AI-Generated Medical Responses: An Evaluation of the Chat-GPT Model*. preprint. In Review. <https://doi.org/10.21203/rs.3.rs-2566942/v1>
- [52] Jon Crowcroft. 1997. Supporting Videoconferencing on the Internet. In *Video-Mediated Communication*, Kathleen E. Finn, Abigail J. Sellen, and Sylvia B. Wilbur (Eds.). 519–540.
- [53] Sasa Junuzovic, Kori Inkpen, Rajesh Hegde, and Zhengyou Zhang. 2011. Towards Ideal Window Layouts for Multi-Party, Gaze-Aware Desktop Videoconferencing. In *Proceedings of Graphics Interface 2011* (St. John's, Newfoundland, Canada) (GI '11). Canadian Human-Computer Communications Society, Waterloo, CAN, 119–126.
- [54] Vaiva Kalnikaitė, Patrick Ehlen, and Steve Whittaker. 2012. Markup as you talk: establishing effective memory cues while still contributing to a meeting. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (CSCW '12). Association for Computing Machinery, New York, NY, USA, 349–358. <https://doi.org/10.1145/2145204.2145260>
- [55] Joseph Kim and Julie A Shah. 2016. Improving team's consistency of understanding in meetings. *IEEE Transactions on Human-Machine Systems* 46, 5 (2016), 625–637.
- [56] David J. Kocsis, Gert-Jan de Vreede, and Robert O. Briggs. 2015. *Designing and Executing Effective Meetings with Codified Best Facilitation Practices*. Cambridge University Press, 483–503. <https://doi.org/10.1017/CBO9781107589735.021>
- [57] Steffi Kohl, André Calero Valdez, and Kay Schröder. 2023. Using Speech Contribution Visualization to Improve Team Performance of Divergent Thinking Tasks. In *Proceedings of the 15th Conference on Creativity and Cognition* (Virtual Event, USA) (C&C '23). Association for Computing Machinery, New York, NY, USA, 319–324. <https://doi.org/10.1145/3591196.3596824>
- [58] Kiron Koshy, Alison Liu, Katharine Whitehurst, Buket Gundogan, and Yasser Al Omran. 2017. How to hold an effective meeting. *International Journal of Surgery Oncology* 2, 5 (June 2017), e22–e22. <https://doi.org/10.1097/IJ9.0000000000000022>
- [59] Robert E. Kraut and Robert S. Fish. 1997. Prospects for Video Telephony. In *Video-Mediated Communication*, Kathleen E. Finn, Abigail J. Sellen, and Sylvia B. Wilbur (Eds.). 541–561.
- [60] Liana Kreamer, George Stock, and Steven Rogelberg. 2021. Optimizing Virtual Team Meetings: Attendee and Leader Perspectives. *American Journal of Health Promotion* 35, 5 (June 2021), 744–747. <https://doi.org/10.1177/08901171211007955e>
- [61] Lactetia Taylor. 2023. Zoom introduces Zoom AI Companion — available at no additional cost with paid Zoom user accounts. <https://news.zoom.us/zoom-ai-companion/>
- [62] Martha Lewis and Melanie Mitchell. 2024. Using counterfactual tasks to evaluate the generality of analogical reasoning in large language models. *arXiv preprint arXiv:2402.08955* (2024).
- [63] Jiannan Li, Mauricio Sousa, Karthik Mahadevan, Bryan Wang, Paula Akemi Aoyagui, Nicole Yu, Angela Yang, Ravin Balakrishnan, Anthony Tang, and Tovi Grossman. 2023. Stargazer: An Interactive Camera Robot for Capturing How-To Videos Based on Subtle Instructor Cues. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–16. <https://doi.org/10.1145/3544548.3580896>
- [64] Justin Matejka, Tovi Grossman, and George Fitzmaurice. 2013. Patina: dynamic heatmaps for visualizing application usage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris France, 3227–3236. <https://doi.org/10.1145/2470654.2466442>
- [65] Matt Martin. 2020. The State of Meetings in 2020 | Clockwise. <https://www.getclockwise.com/blog/the-state-of-meetings-in-2020>
- [66] Moira McGregor and John C Tang. 2017. More to meetings: challenges in using speech-based technology to support meetings. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 2208–2220.
- [67] 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/>
- [68] Microsoft. [n.d.]. In the Office, It's All About Moments That Matter. <https://www.microsoft.com/en-us/worklab/in-the-office-it-is-all-about-moments-that-matter>
- [69] Microsoft. 2023. Get started with Copilot in Microsoft Teams meetings - Microsoft Support. <https://products.support.services.microsoft.com/en-us/office/get-started-with-copilot-in-microsoft-teams-meetings-0bf9dd3c-96f7-44e2-8bb8-790bedf066b1>
- [70] Microsoft. 2023. Work Trend Index | Will AI Fix Work? <https://www.microsoft.com/en-us/worklab/will-ai-fix-work>
- [71] Shaila M. Miranda and Robert P. Bostrom. 1993. The Impact of Group Support Systems on Group Conflict and Conflict Management. *Journal of Management Information Systems* 10, 3 (1993), 63–95. <https://doi.org/10.1080/07421222.1993.11518011>
- [72] Melanie Mitchell, Alessandro B. Palmari, and Arseny Moskvichev. 2023. Comparing Humans, GPT-4, and GPT-4V on abstraction and reasoning tasks. *arXiv preprint arXiv:2311.09247* (2023).
- [73] Lillio Mok, Lu Sun, Shilad Sen, and Bahareh Sarrafzadeh. 2023. Challenging but Connective: Large-Scale Characteristics of Synchronous Collaboration Across Time Zones. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 611, 17 pages. <https://doi.org/10.1145/3544548.3581141>
- [74] Thomas P. Moran, Leysia Palen, Steve Harrison, Patrick Chiu, Don Kimber, Scott Minneman, William van Melle, and Polle Zellweger. 1997. "I'll get that off the audio": a case study of salvaging multimedia meeting records. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (CHI '97). Association for Computing Machinery, New York, NY, USA, 202–209. <https://doi.org/10.1145/258549.258704>
- [75] William Morgan, Pi-Chuan Chang, Surabhi Gupta, and Jason M. Brenier. 2009. Automatically detecting action items in audio meeting recordings. In *Proceedings of the 7th SIGDIAL Workshop on Discourse and Dialogue* (SigDIAL '06). Association for Computational Linguistics, USA, 96–103.
- [76] Philipp Müller, Michael Xuelin Huang, and Andreas Bulling. 2018. Detecting low rapport during natural interactions in small groups from non-verbal behaviour. In *23rd International Conference on Intelligent User Interfaces*. 153–164.
- [77] Saeid Nahavandi. 2017. Trusted Autonomy Between Humans and Robots: Toward Human-on-the-Loop in Robotics and Autonomous Systems. *IEEE Systems*,

- Man, and Cybernetics Magazine* 3, 1 (Jan. 2017), 10–17. <https://doi.org/10.1109/MSMC.2016.2623867>
- [78] Eyal Ofek, Shamsi T Iqbal, and Karin Strauss. 2013. Reducing Disruption from Subtle Information Delivery during a Conversation: Mode and Bandwidth Investigation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013).
- [79] Antti Oulasvirta, Niraj Ramesh Dayama, Morteza Shiripour, Maximilian John, and Andreas Karrenbauer. 2020. Combinatorial Optimization of Graphical User Interface Designs. *Proc. IEEE* 108, 3 (2020), 434–464. <https://doi.org/10.1109/JPROC.2020.2969687>
- [80] Gun Woo (Warren) Park, Payod Panda, Lev Tankelevitch, and Sean Rintel. 2024. CoExplorer: Generative AI Powered 2D and 3D Adaptive Interfaces to Support Intentionality in Video Meetings. (2024). <https://doi.org/10.1145/3613905.3650797>
- [81] Seonwook Park, Christoph Gebhardt, Roman Rädle, Anna Feit, Hana Vrzakova, Niraj Dayama, Hui-Shyong Yeo, Clemens Klokmoose, Aaron Quigley, Antti Oulasvirta, and Otmar Hilliges. 2018. AdaM: Adapting Multi-User Interfaces for Collaborative Environments in Real-Time. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14. <https://doi.org/10.1145/3173574.3173758> arXiv:1803.01166 [cs].
- [82] Kevin Pu, Rainey Fu, Rui Dong, Xinyu Wang, Yan Chen, and Toví Grossman. 2022. SemanticOn: Specifying Content-Based Semantic Conditions for Web Automation Programs. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. ACM, Bend OR USA, 1–16. <https://doi.org/10.1145/3526113.3545691>
- [83] Qatalog. [n.d.]. Killing Time Report | language.work. <https://language.work/research/killing-time-at-work/>
- [84] Steven G. Rogelberg. 2019. *The Surprising Science of Meetings: How You Can Lead Your Team to Peak Performance*. Oxford University Press, New York.
- [85] Jazz Rui Xia Ang, Ping Liu, Emma McDonnell, and Sarah Coppola. 2022. “In This Online Environment, We’re Limited”: Exploring Inclusive Video Conferencing Design for Signers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI ’22). Association for Computing Machinery, New York, NY, USA, Article 609, 16 pages. <https://doi.org/10.1145/3491102.3517488>
- [86] Samiha Samrose, Daniel McDuff, Robert Sim, Jina Suh, Kael Rowan, Javier Hernandez, Sean Rintel, Kevin Moynihan, and Mary Czerwinski. 2021. Meeting-coach: An intelligent dashboard for supporting effective & inclusive meetings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [87] Samiha Samrose, Ru Zhao, Jeffery White, Vivian Li, Luis Nova, Yichen Lu, Mohammad Rafayet Ali, and Mohammed Ehsan Hoque. 2018. CoCo: Collaboration Coach for Understanding Team Dynamics during Video Conferencing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 160 (jan 2018), 24 pages. <https://doi.org/10.1145/3161186>
- [88] Advait Sarkar. 2023. Will Code Remain a Relevant User Interface for End-User Programming with Generative AI Models?. In *Proceedings of the 2023 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software* (<conf-loc>, <city>Cascadia</city>, <country>Portugal</country>, </conf-loc>) (Onward! 2023). Association for Computing Machinery, New York, NY, USA, 153–167. <https://doi.org/10.1145/3622758.3622882>
- [89] Advait Sarkar, Ian Drosos, Robert DeLine, Andy Gordon, Carina Negreanu, Sean Rintel, Jack Williams, and Ben Zorn. 2023. Participatory prompting: a user-centric research method for eliciting AI assistance opportunities in knowledge workflows. In *Proceedings of the 34th Annual Conference of the Psychology of Programming Interest Group* (PPIG 2023).
- [90] Maria Schmidt, Wolfgang Minker, and Steffen Werner. 2020. USER ACCEPTANCE OF PROACTIVE VOICE ASSISTANT BEHAVIOR. *Konferenz Elektronische Sprachsignalverarbeitung* (2020).
- [91] Ava Elizabeth Scott, Lev Tankelevitch, and Sean Rintel. 2024. Mental Models of Meeting Goals: Supporting Intentionality in Meeting Technologies. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (CHI ’24). Association for Computing Machinery, New York, NY, USA.
- [92] Ava Elizabeth Scott, Lev Tankelevitch, and Sean Rintel. 2024. Mental Models of Meeting Goals: Supporting Intentionality in Meeting Technologies. (2024). <https://doi.org/10.1145/3613904.3642670>
- [93] Sandip Sen and Edmund H. Duffee. 1998. A Formal Study of Distributed Meeting Scheduling. *Group Decision and Negotiation* 7, 3 (May 1998), 265–289. <https://doi.org/10.1023/A:1008639617029>
- [94] Anil Shankar, Sushil J. Louis, Sergiu Dascalu, Linda J. Hayes, and Ramona Houmanfar. 2007. User-context for adaptive user interfaces. In *Proceedings of the 12th international conference on Intelligent user interfaces*. ACM, Honolulu Hawaii USA, 321–324. <https://doi.org/10.1145/1216295.1216357>
- [95] Greg Smith, Patrick Baudisch, George Robertson, Mary Czerwinski, Brian Meyers, Daniel Robbins, and Donna Andrews. 2003. GroupBar: The TaskBar evolved. (2003).
- [96] Willem Standaert, Steve Muylle, and Amit Basu. 2021. How shall we meet? Understanding the importance of meeting mode capabilities for different meeting objectives. *Information & Management* 58, 1 (2021), 103393. <https://doi.org/10.1016/j.im.2020.103393>
- [97] Zinovia Stefanidi, George Margetis, Stavroula Ntoa, and George Papagiannakis. 2022. Real-Time Adaptation of Context-Aware Intelligent User Interfaces, for Enhanced Situational Awareness. *IEEE Access* 10 (2022), 23367–23393. <https://doi.org/10.1109/ACCESS.2022.3152743>
- [98] Åsne Stige, Efraxia D. Zamani, Patrick Mikalef, and Yuzhen Zhu. 2023. Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda. *Information Technology & People* ahead-of-print, ahead-of-print (Jan. 2023). <https://doi.org/10.1108/ITP-07-2022-0519>
- [99] Barbara J. Streibel. 2003. *The manager’s guide to effective meetings*. McGraw-Hill, New York.
- [100] Jan Svennevig. 2012. The agenda as resource for topic introduction in workplace meetings. *Discourse Studies* 14, 1 (Feb. 2012), 53–66. <https://doi.org/10.1177/1461445611427204>
- [101] John Tang. 2021. Understanding the Telework Experience of People with Disabilities. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 30 (apr 2021), 27 pages. <https://doi.org/10.1145/3449104>
- [102] John C. Tang, Kori Inkpen, Sasa Junuzovic, Keri Mallari, Andrew D. Wilson, Sean Rintel, Shiraz Cupala, Tony Carbury, Abigail Sellen, and William A.S. Buxton. 2023. Perspectives: Creating Inclusive and Equitable Hybrid Meeting Experiences. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 351 (oct 2023), 25 pages. <https://doi.org/10.1145/3610200>
- [103] Craig Tashman. 2006. WindowScape: a task oriented window manager. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*. ACM, Montreux Switzerland, 77–80. <https://doi.org/10.1145/1166253.1166266>
- [104] Craig Tashman and W. Keith Edwards. 2012. WindowScape: Lessons learned from a task-centric window manager. *ACM Transactions on Computer-Human Interaction* 19, 1 (March 2012), 1–33. <https://doi.org/10.1145/2147783.2147791>
- [105] Kashyap Todi, Gilles Bailly, Luis Leiva, and Antti Oulasvirta. 2021. Adapting User Interfaces with Model-based Reinforcement Learning. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–13. <https://doi.org/10.1145/3411764.3445497>
- [106] Ambreen Usmani, Rehana Rehman, Shazia Babar, and Azam Afzal. 2012. Impact of structured meetings on the learning of faculty members. 26, 03 (2012).
- [107] Roel Vertegaal, Ivo Weevers, Changuk Sohn, and Chris Cheung. 2003. GAZE-2: Conveying Eye Contact in Group Video Conferencing Using Eye-Controlled Camera Direction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) (CHI ’03). Association for Computing Machinery, New York, NY, USA, 521–528. <https://doi.org/10.1145/642611.642702>
- [108] Alessandro Vinciarelli, Fabio Valente, Sree Harsha Yella, and Ashtosh Sapru. 2011. Understanding social signals in multi-party conversations: Automatic recognition of socio-emotional roles in the AMI meeting corpus. In *2011 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 374–379.
- [109] Jing Wei, Tilman Dingler, and Vassilis Kostakos. 2021. Understanding User Perceptions of Proactive Smart Speakers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 4 (Dec. 2021), 1–28. <https://doi.org/10.1145/3494965>
- [110] Wolfgang Weiss, Rene Kaiser, and Manolis Falelakis. 2014. Orchestration for Group Videoconferencing: An Interactive Demonstrator. In *Proceedings of the 16th International Conference on Multimodal Interaction* (Istanbul, Turkey) (ICMI ’14). Association for Computing Machinery, New York, NY, USA, 72–73. <https://doi.org/10.1145/2663204.2669624>
- [111] Tongshuang Wu, Michael Terry, and Carrie Jun Cai. 2022. AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>New Orleans</city>, <state>LA</state>, <country>USA</country>, </conf-loc>) (CHI ’22). Association for Computing Machinery, New York, NY, USA, Article 385, 22 pages. <https://doi.org/10.1145/3491102.3517582>
- [112] Haijun Xia, Tony Wang, Aditya Gunturu, Peiling Jiang, William Duan, and Xiaoshuo Yao. 2023. CrossTalk: Intelligent Substrates for Language-Oriented Interaction in Video-Based Communication and Collaboration. <https://doi.org/10.1145/3586183.3606773> arXiv:2308.03311 [cs].
- [113] Lining Yao, Anthony DeVincenzi, Anna Pereira, and Hiroshi Ishii. 2013. FocalSpace: Multimodal Activity Tracking, Synthetic Blur and Adaptive Presentation for Video Conferencing. In *Proceedings of the 1st Symposium on Spatial User Interaction* (Los Angeles, California, USA) (SUI ’13). Association for Computing Machinery, New York, NY, USA, 73–76. <https://doi.org/10.1145/2491367.2491377>
- [114] Michael A. Yoerger, Kyle Francis, and Joseph A. Allen. 2015. So Much More than “Chit-Chat”: A Closer Look at Premeeting Talk. In *The Cambridge Handbook of Meeting Science*, Joseph A. Allen, Nale Lehmann-Willenbrock, and Steven G. Rogelberg (Eds.). Cambridge University Press, Cambridge, 153–174. <https://doi.org/10.1017/9781107311111.008>

- //doi.org/10.1017/CBO9781107589735.008
- [115] Nima Zargham, Leon Reicherts, Michael Bonfert, Sarah Theres Voelkel, Johannes Schoening, Rainer Malaka, and Yvonne Rogers. 2022. Understanding Circumstances for Desirable Proactive Behaviour of Voice Assistants: The Proactivity Dilemma. In *Proceedings of the 4th Conference on Conversational User Interfaces*. ACM, Glasgow United Kingdom, 1–14. <https://doi.org/10.1145/3543829.3543834>
 - [116] Ke Zhou, Marios Constantinides, Luca Maria Aiello, Sagar Joglekar, and Daniele Quercia. 2021. The Role of Different Types of Conversations for Meeting Success. *IEEE Pervasive Computing* 20, 4 (2021), 35–42.
 - [117] Ke Zhou, Marios Constantinides, Sagar Joglekar, and Daniele Quercia. 2022. Predicting meeting success with nuanced emotions. *IEEE Pervasive Computing* 21, 2 (2022), 51–59.

A APPENDIX A: PHASE GENERATION AND REFINEMENTS (HIGHER RESOLUTION)

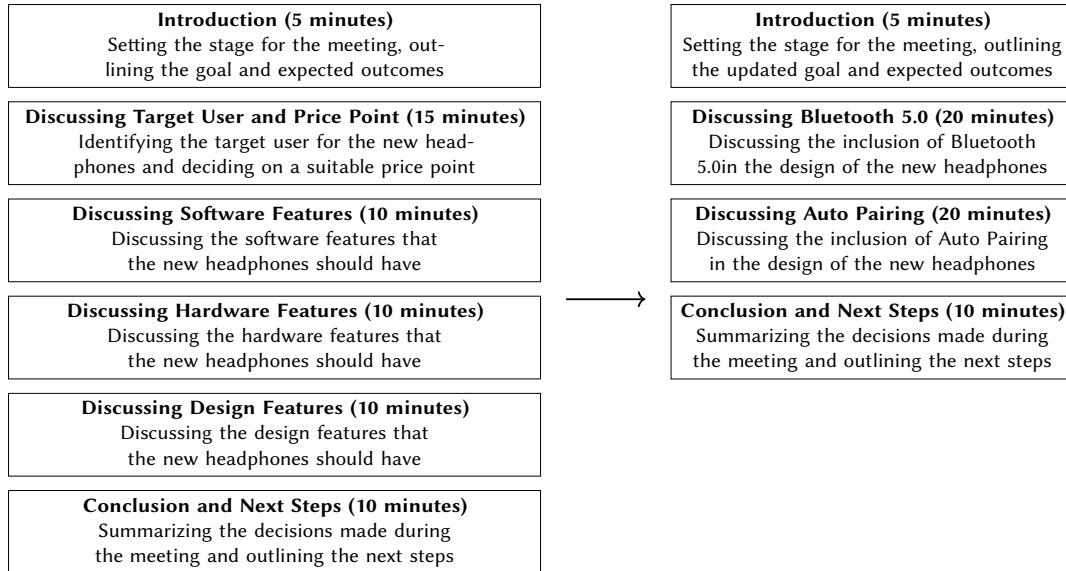


Figure 6: List of generated initially generated phases (left) and the refined phases (right). After refinement, the phases became more focused and realistic for the given timeframe.

B APPENDIX B: EXAMPLE PROMPTS

We attach versions of the example prompts for GPT 3.5-Turbo that we used for the implementation of the CoExplorer. CoExplorer generated prompts automatically based on the templates. During this, CoExplorer optimized the prompts as per the user inputs, and prompted the language model multiple times to obtain the results in the correct format. The prompts given is one instantiation of output prompts that the CoExplorer was producing.

B.1 Phase Generation

B.1.1 Example system prompt.

[No prose]

[Output only JSON]

Do not write normal text

You are a JSON generator which converts meeting agenda text into a more descriptive agenda description. You always need to have an introduction phase at the beginning.

B.1.2 Example user prompt.

Please breakdown the following meeting agenda that someone has sent in email into meeting phases that we would need to go through. We have a 60 minute meeting scheduled. Based on the given information, give a goal of the meeting (goal), as well as the explanation on why you chose the goal (exp). And also give the phase definition in a list (pi). Each phase definition should include: (1) Phase title (pt), (2) Phase description (pd) which should include a sub goal of a phase, (3) Behaviors to be encouraged (be), (4) Behaviors to be discouraged (bd), (5) priority (p), (6) amount of time allocation (t), (7) direction (d) (i.e., is it an iterative phase or directional phase). Please only respond in JSON with each element needed as a key within a phase i.e., if we have two phases, [{"pt": "xxx", "pd": "xxx", "be": ["xxx", "yyy"], "bd": ["xxx", "yyy"], "p": "high", "t": 2, "d": "iterative"}, {"pt": "yyy", "pd": "zzz", "be": ["ttt", "kk"], "bd": ["lll", "mmm"], "p": "low", "t": 8, "d": "directional"}]. So the overall JSON to export is {"goal": "xxx", "pi": [<phase definitions>], "exp": "xxx"}. Explanation should start by saying this goal is generated... or

similar. Please use the full 60 minutes. Here is the meeting invitation:

(The email invitation would be attached.)

Refinement could be done using the same script, but by attaching the refinement scenario. This script can be generated through code, and does not have to be detailed.

B.2 Layout Generation

B.2.1 Example system prompt.

[No prose]

[Output only JSON]

Do not write normal text

You are a helpful assistant who creates screen layout that has appropriate apps that are most helpful for users to complete the task successfully. Respond only in JSON following the format. Example format:

```
[{"PhaseTitle": "xxx", "timer": n, "programList":
  [{"name": "yyy", "description": "zzz"},
   {"name": "kk", "description": "lll"}]},
 {"PhaseTitle": "xxx", "timer": n, "programList":
  [{"name": "yyy", "description": "zzz"},
   {"name": "kk", "description": "lll"}]}].
```

Strictly follow this format. n is integer, and

programList.name should either be a name of a program in a program list given or a URL. programList.

description is where you should put a extremely concise reason why you chose that program. e.g.,

Use this for presenting agenda; Use this for viewing relevant budget data

B.2.2 Example user prompt.

I will give you the list of phases in a meeting in JSON format. Each phase in JSON is defined with several keywords.

"pt" represents phase title, "pd" represents phase description, "be" represents behaviours to be encouraged, "bd" represents the behaviours to be

discouraged, "p" represents priority (high, medium, low), "t" represents recommended duration of time for the phase, and "d" represents directionality (directional i.e., cannot be returned, and should be preceded by a certain phase or iterative i.e., can be transitioned to this whenever).

You need to generate what kind of programs are needed for helping goals of each phase (defined by the description) to be met the most efficiently. You can generate a list of 1-5 program name/URL, and the sequence of generation will affect where they are being placed, and size.

Therefore, you need to be sensible about ordering so that important

programs can be shown with the bigger sizes. The rule is as follows: If you have one program on the list, that would be full screen. If you have two programs, it would be one on the left half (first program on the list), and one on the right half. If you have three programs, the right-hand-side panel will split in half, creating two small panels at the top and the bottom.

Four programs mean the left-hand-side panel will also be split. Five programs mean two equally sized panel at the top, and three equally sized panel at the bottom.

The ordering in the list will be used to place program to panels in a clockwise ordering (top left panel is the first panel). Here is the list of programs available, and if the program that you want is not listed, please generate a URL for the program that you need instead of the program name. Please feel free to give Bing Search URL with the search term filled, and generate at least one URL:

List of programs could take arbitrary formats.

B.3 Meeting Focus Tool

B.3.1 Example system prompt for a specialized calculator generation.

You need to generate a HTML page, and only HTML+CSS

+JavaScript based page as a response which allows me to calculate the total value to have the a list of features for a given scenario. Each feature needs to have an "include" button

(green tick) or "exclude" button (red cross), and you need to calculate the total at the end,

when the submit is clicked. You should not show any prices including itemised ones before this.

You need to generate at least 30 features (features should be unique and descriptive. No Feature 26 or something likethat), and incorporate that into the page, those are relevant for the scenario. Assign random prices for each feature. The HTML page needs to incorporate all the features that you generated embedded (i.e., no ... or "many more features here" etc.). No prose. No add more features here etc. You need to list all the features on the HTML.

B.3.2 Example user prompt.

Designer, software engineer, hardware engineer, PM, marketing expert, and a researcher are gathering to hink about what features that a new headphone product that they release might have. It could be electronics feature such as active noise cancelling (95%) or it could be about materials (leather etc.).

This user prompt may not occasionally result in explicit features being stated but rather like “<!--Here will be inserted many more features-->”. Such cases can be detected, and additional user prompt could be given, for instance:

It does not show 30+ features

This generates the following working web page:

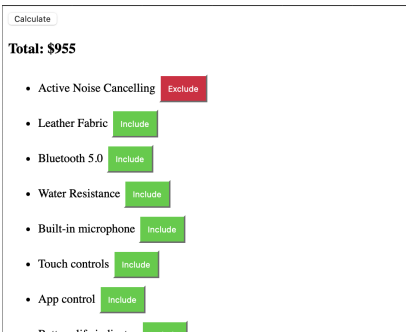


Figure 7: Example raw ranking aid app generated by ChatGPT as a meeting focus tool

Received 8 February 2024; accepted 18 April 2024