DesignMemo: Integrating Discussion Context into Online Collaboration with Enhanced Design Rationale Tracking

BOYU LI, The Hong Kong University of Science and Technology (Guangzhou), China
LINJIE QIU, The Hong Kong University of Science and Technology (Guangzhou), China
DUOTUN WANG, The Hong Kong University of Science and Technology (Guangzhou), China
QIANXI LIU, The Hong Kong University of Science and Technology (Guangzhou), China
RYO SUZUKI, University of Colorado Boulder, USA
MINGMING FAN, The Hong Kong University of Science and Technology (Guangzhou), China and The Hong Kong University of Science and Technology, China
ZEYU WANG*, The Hong Kong University of Science and Technology (Guangzhou), China and The Hong Kong University of Science and Technology, China

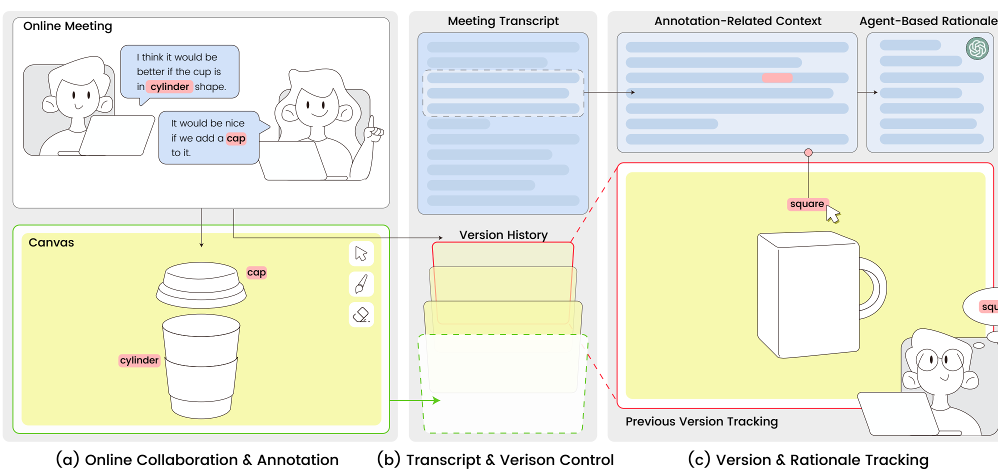


Fig. 1. Overview of DesignMemo. DesignMemo integrates online meetings with a shared design canvas and a GPT-empowered agent to support collaborative design. (a) It automatically transcribes the discussions and generates relevant design annotations. (b) As users submit their work, the system stores both the transcripts and design histories. In the future, users can review past design rationales by querying transcripts linked to specific design elements. The system informed by these transcripts can dynamically provide context-aware responses to enhance decision-making.

*Corresponding author.

Authors' Contact Information: Boyu Li, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, blibr@connect.hkust-gz.edu.cn; Linjie Qiu, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, lqiu250@connect.hkust-gz.edu.cn; Duotun Wang, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, dwang866@connect.hkust-gz.edu.cn; Qianxi Liu, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China, qliu930@connect.hkust-gz.edu.cn; Ryo Suzuki, University of Colorado Boulder, Boulder, Colorado, USA, qliu930@connect.hkust-gz.edu.cn; Mingming Fan, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China and The Hong Kong University of Science and Technology, Hong Kong, China, mingmingfan@ust.hk; Zeyu Wang, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China and The Hong Kong University of Science and Technology, Hong Kong, China, zeyuwang@ust.hk.

Remote collaborative design has become increasingly popular, but current design tools often overlook the importance of contextual communication during synchronized design activities, which is critical for understanding the rationale and decisions behind design choices. In this paper, we introduce DesignMemo, a proof-of-concept system that integrates the verbal context of remote discussions into visual design history tracking. The system automatically labels the visual elements with an annotation, which is linked to a certain transcript of the meeting, so that the user can easily recall the context of the visual design by clicking the element. The system also integrates an LLM agent for annotation-oriented summarization based on global context tracking, so users can quickly follow the rationale of the design without reading the lengthy transcript. Our user study with 24 participants suggests that the ability to track communication context makes the iterative design process smoother and more efficient.

CCS Concepts: • **Human-centered computing** → *Synchronous editors.*

Additional Key Words and Phrases: Collaborative Design, Context Tracking, Annotation-oriented Summarization

ACM Reference Format:
Boyu Li, Linjie Qiu, Duotun Wang, Qianxi Liu, Ryo Suzuki, Mingming Fan, and Zeyu Wang. 2025. DesignMemo: Integrating Discussion Context into Online Collaboration with Enhanced Design Rationale Tracking. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article 398 (November 2025), 32 pages. https://doi.org/10.1145/3757579

1 Introduction

Collaborative design in remote settings has become increasingly important in today's creative workflows. For example, tools like Miro [1] and Mural [2] facilitate more efficient collaboration by enabling synchronized design processes. While existing collaborative design tools offer features like real-time editing, commenting, and voice chat to help distributed teams stay aligned, they often fall short in integrating visual design activities with ongoing discussions. Tools such as Zoom [59] and Voov [32]facilitate communication but lack mechanisms to connect conversation content directly to the evolving design artifacts. Without such integration, designers often find it difficult to recall design rationale from previous discussion in real time, as they must focus on both participating in meetings and actively working on design tasks. Although automatic transcription features are available, capturing the link between specific parts of the conversation and corresponding design versions or elements is still a time-consuming process. Designers have to manually search through lengthy transcripts to find the relevant parts and match them to visual elements, and even read the surrounding conversation to understand the rationale based on the broader context. Such a laborious process makes it difficult for users to recall the context behind critical design decisions, leading to uncertainty during iteration and reduced efficiency when modifying designs. Therefore, there is a pressing need for collaborative systems that tightly integrate the design canvas with meeting logs, enabling users to effortlessly trace design rationale and maintain continuity across design sessions.

To address these challenges, we introduce DesignMemo, a system that integrates discussion context recall into remote collaborative design, thus enhancing the experience of reviewing design history. Based on the insights derived from our formative study (Section 3), we design DesignMemo to facilitate remote online discussions and collaborative concept designs by providing a comprehensive trace of the discussion contexts. We use concept sketching as our primary experimental

ACM 2573-0142/2025/11-ART398
https://doi.org/10.1145/3757579

medium due to its widespread use in the early stages of design processes [22, 49], suitability for collaborative design patterns [42], and ease of tracking in the design process [40]. Our system introduces two key features:

- DesignMemo leverages a large language model (LLM) to automatically extract annotations from real-time meeting conversations, facilitating efficient annotation of design elements across versions. This process empowers users to directly link dialogue with specific design canvas elements, thereby seamlessly aligning conversations with design components.
- DesignMemo aids users in deciphering design rationales within extensive transcripts. By clicking on annotations linked to specific design elements, users can revisit the underlying rationale. They can either access the original conversational context or employ an LLM agent to generate an automatic summary based on the transcript and selected annotations.

In our system, meeting transcripts from collaborative design sessions provide contextual guidance for the LLM agent. The agent not only generates annotations for real-time tagging of design content but also facilitates the retrieval of the underlying design rationale. Specifically, when a user selects an annotation linked to a specific design component, the agent utilizes the full transcript—integrating the annotation's timestamp and semantic context—to extract pertinent dialogue segments and generate a concise summary of the underlying design rationale.

Compared with previous systems, such as Lopez et al. [16] which rely on manual, text-centric documentation of design rationale, our approach addresses key limitations in real-time collaboration. In those systems, users must manually type and organize notes during fast-paced discussions, detracting from creative engagement, and later face tedious navigation to recall rationale. In contrast, our system automates the process, enabling instant recording and retrieval of relevant discussions through intuitive canvas interactions and replacing manual note-taking with seamless action-driven capture. Furthermore, while recent works [3, 8, 51]-especially MeetVis [43]-have explored rapid context recall in meetings by summarizing spoken audio, these approaches treat design artifacts as static resources. This focus on discussion content alone disconnects dialogue from the dynamic evolution of design elements, such as changes in UI components post-feedback. DesignMemo uniquely bridges this gap through bidirectional integration: users can anchor discussion clips directly to evolving design elements, and selecting any element reveals its linked discussion history. Unlike conventional meeting-centric tools, this binding not only captures the emergence of decisions but also illustrates how design rationale informs specific iterative changes, thereby tightly coupling collaborative dialogue with actionable design context.

Finally, we present a comparative study involving DesignMemo and a baseline setup to evaluate the usability and effectiveness of our proposed designs. The results demonstrate that DesignMemo empowers users to continuously document and revisit prior design considerations throughout the design process, enriching their contextual understanding and facilitating more informed subsequent iterations. We conclude that the concepts of context record and recall modules for history retrieval have the potential to be applied to more complex design scenarios, such as 3D modeling [37] and augmented reality conversation [21].

In summary, this paper makes the following contributions:

- A formative study identifying how the motivation behind design influences the iterative design process and providing key considerations for integrating discussion context into collaborative design history reviews.
- The concepts of contextual annotation linking design content and discussion, leveraging transcripts as contextual information for LLM agents to track design rationale.

- A web-based platform, DesignMemo, that implements our designs and is evaluated through quantitative and qualitative measures to distill key insights into the recording and recalling of discussion context within a collaborative design system.

2 Related Work

2.1 Retrospective Meeting

HCI and CSCW research has investigated many techniques for tracking historical data, particularly during and following collaborative meetings[23, 35, 51]. Conventionally, a designated individual is responsible for documenting meeting notes to facilitate easy retrieval. In the early 2000s, Klemmer et al. [24] explored digital tracing by taking snapshots of various states during collaborative web editing sessions. Recent developments in video conferencing platforms, such as Zoom [59] and Tencent Meeting [32], have offered practical features that automatically record audio and video and transcribe speech through AI-powered recognition to distill important discussions into paragraphs. To go further, MeetCues [3] automatically generates an interactive summary page that highlights notable moments and key audio snippets to engage in online meetings. When presented in a readable interface that aligns with the conversation flow, MeetScript develops low-effort contextualized user interactions and incorporates filtering mechanisms into live transcripts to help users locate important information [8]. These mechanisms include an interaction history panel and an auto-disappearing feature based on user inputs. The core concept of these works is to reduce interruptions and distractions during meetings [46].

However, with regard to remote meetings for collaborative design, the design content is often disconnected from the discussion records, despite their close interrelation. To identify design considerations for a specific element, users must sift through the chronological transcripts to find the relevant details, which can be both tedious and time-consuming. In a recent interview, the concept of “Meeting Bridges” was introduced to enhance connectivity in asynchronous collaboration [51]. But the specific form these “Bridges” will take is still undefined. Our formative study identifies the design considerations of such traceable representation. Our system employs contextual annotations as a “bridge” to bind design and rationale as a whole, accelerating the process of context information retrieval to improve efficiency in both synchronous and asynchronous collaboration.

2.2 The Tracking of Design History and Rationale

Efficient recording and tracking in iterative design have long been a key focus in HCI. Previous work introduced intuitive representations for tracking purposes. For example, Chronicle [13] used a series of probes to filter design versions over time. Chen et al. [7] employed a Directed Acyclic Graph (DAG) to trace back image editing operations. Some other work implemented the Git-like version control feature to record the creative design history [9, 47, 58]. However, all the above systems for recording and tracking design history overlook the track of design rationale.

Design rationale is more than a record of the design process; it evolves alongside the product to better grasp the reasons behind design [18, 33]. Tracking design rationale is crucial for maintaining coherence, supporting decision-making, and acting as a memory for designers and stakeholders [10, 34]. Traditional co-located design collaboration often uses artifacts such as decision cards [15] and tangible labels [20] to improve the traceability of the design rationale. With the rise of remote collaboration via the Internet, collaborative systems have enabled interactive tracking of design rationale through interactive means. For instance, Helaba integrates sticky labels to clarify design decisions during discussions [14]. Other systems employed shared design spaces, such as DR graphs [6] and shared design documents [11], to facilitate note-taking and tracking.

While these systems are promising for documenting design rationale, they may still challenge users to capture detailed information during spontaneous meetings. Recalling past content can also be difficult after long intervals. Leveraging an LLM agent, DesignMemo addresses these issues by analyzing the global context of previous meetings and summarizing local context through interaction on annotations, thus reducing efforts to track design rationale.

2.3 Co-Design Systems in Remote Collaboration

Co-design involves the collaborative creativity of designers and non-design-trained participants [41]. Its significance has been explored in designers' communication with diverse groups, including children [53, 54], breastfeeding mothers [52], and individuals with disabilities [5, 12, 44]. Evidence shows that creativity increases when designers incorporate ideas from end-users who offer perspectives beyond their expertise [27].

In particular, remote co-design is proven to be an efficient and practical design method to bridge geographical and time zone disparities [17], utilizing video calls [45], social media [39], and dedicated platforms for both synchronous and asynchronous collaboration. For instance, Lee et al. [25] introduced a model based on improvisation theories to help researchers and designers manage unexpected situations. Walsh et al. [50] developed DisCo, a tool for distributed asynchronous co-design with children and adults. More recently, Ayobi et al. [4] investigated using shared computational notebooks to connect technicians and medical stakeholders.

Unlike previous research, our paper does not limit the focus to a specific task or user group. Instead, we aim for broader applicability. Our system enables users to uncover design rationales that may not be immediately apparent from the design content itself, helping them better understand the perspectives of others and thus promoting collaborative design among diverse participants.

3 Formative Study

To explore the difficulties caused by insufficient traceability of information during the design iteration process, we conducted a formative study focusing on two primary groups: designers and stakeholders. Designers often struggle to retain design rationale over time, especially in fast-paced workflows involving complex and evolving design content. Prior work [16, 38, 43, 51] shows that capturing and recalling information from meetings and the design process imposes a significant cognitive and logistical burden, making it difficult for designers to balance creative tasks with the ongoing effort of documenting and managing rationale. Stakeholders, on the other hand, are typically not design experts but still play a crucial role in the project. Their decisions can significantly influence the design's direction and outcome. However, without a clear understanding of the design rationale, stakeholders may misinterpret the design's purpose, leading to communication breakdowns or decisions that conflict with the intended design goals. By ensuring that stakeholders have a comprehensive understanding of the design logic, we aim to facilitate better decision-making and collaboration, ultimately contributing to the success of the project.

To investigate how contextual information is captured and maintained during remote design collaboration, we formulated the following research questions (RQ):

- **RQ1:** How do designers and stakeholders currently document and manage design rationale in remote, collaborative design workflows?
- **RQ2:** What challenges do they face in accessing, interpreting, and using design rationale during iterative design processes?

3.1 Method

3.1.1 Participants. We recruited a total of 12 participants with prior experience in collaborative design projects. The participant pool included 6 designers (D1–D6) and 6 stakeholders (S1–S6). Participants were recruited through professional networks and screened using a brief questionnaire to ensure relevant experience with remote design collaboration and tool usage. We selected participants who had recently (within the past year) taken part in remote design projects involving iterative decision-making, to ensure recency and relevance of their reflections. Designers were individuals who regularly engage in remote, collaborative, and iterative design activities. Their roles included UI/UX designers, product designers, and service designers. All designers had more than two years of professional experience and were familiar with common collaborative design tools such as Figma, Miro, and Zoom. Stakeholders were project managers, clients, or team members from non-design roles (e.g., marketing, engineering) who had previously participated in design review meetings and contributed to design-related decisions. Although they did not have formal design training, they frequently engaged with designers during remote sessions and influenced the direction of design outcomes. Participants were selected to represent both the creative and decision-making sides of the design process. This diverse composition allowed us to capture a range of perspectives on how discussion context is recorded, accessed, and understood during remote collaboration.

3.1.2 Procedure. We conducted semi-structured interviews with each participant over video conferencing platforms such as Zoom [59] or Google Meet. Each session lasted approximately 45–60 minutes. Before the interview, participants were asked to recall a recent collaborative design project they had been involved in and to describe their role in that project. The interview was divided into three subsections:

- Workflow and Tool Usage: Participants were instructed to describe how they typically conducted remote collaborative design sessions, including tools used for sketching, discussion, and documentation.
- Challenges in Context Retention: We asked participants about their experiences in capturing and recalling design rationale, involving how they managed meeting notes, transcripts, or design change tracking.
- Reflections and Ideal Support: Participants were encouraged to reflect on what kinds of tool support would help them better access and organize design context in future projects.

3.2 Challenges

Based on the analysis of participant interviews, we identified four recurring challenges that impact how design context is recorded, retrieved, and utilized throughout the design process.

Users found it cumbersome to record the content of the meeting and add notes to the design. Participants expressed that the process of documenting meeting discussions and attaching notes to design artifacts was time-consuming and disjoint. D5 noted: "*I often have to jump between tools to write notes during meetings and then manually link them back to the designs later, which is tedious and error-prone.*" This disconnect can reduce the efficiency of collaboration and lead to incomplete documentation of the design rationale, affecting the overall quality of the design.

Users face difficulties locating relevant discussions of a specific design component in entire records. Although many users rely on audiovisual recordings and transcripts to document design discussions, finding specific design rationales corresponding to a particular design component is often cumbersome. As D7 noted, "*When I want to understand why a specific change was made, I have to search through long and disorganized transcripts, which is time-consuming and frustrating.*"

Although current algorithms can extract key points from meeting minutes, they often fail to align these insights with the specific details of the design process, and controlling the granularity of the information presented remains challenging.

Users often feel that accessing design rationale is disconnected from their ongoing design activities. In the design process, designers frequently review past meeting records or related materials to understand what changes were made to a design and the reasoning behind them. As D4 explained, "*I need quick access to why a certain decision was made, but having to leave my design space to find this information really breaks my focus.*" The issue becomes even more pronounced in fast-paced, collaborative environments where context changes frequently, and retrieving the right information at the right time is crucial. Many designers express frustration at the difficulty of locating relevant records, especially when the system offers limited filtering options or lacks intelligent support to present context in an unobtrusive manner. This friction can cause delays in decision-making and lead to misinterpretations of design history, ultimately affecting the efficiency and coherence of the design process.

Users are struggling to smoothly compare and view design consideration iterations. Designers often struggle to keep track of how a project has evolved over time, particularly in collaborative settings where multiple iterations and changes occur frequently. As D3 expressed, "*It can be really difficult to see how things have progressed and what the key turning points were in the design.*" Without a clear visual representation of iterations, users are left piecing together fragmented information from various sources, leading to confusion and potential miscommunication among team members. This lack of visibility into the design history makes it harder to understand the rationale behind decisions and to align current work with past iterations, ultimately hindering collaboration and the overall design process.

3.3 Design Consideration

C1: Users should not be interrupted from their discussions and design work by creating context records. Develop an automated recording system that captures discussion content using technologies like voice recognition or screen activity tracking, minimizing the need for manual input. This system should operate through a non-intrusive interface that updates notes in the background or displays brief summaries only when users pause their work, ensuring designers remain focused without constant task switching. Additionally, the system should exhibit context awareness by intelligently recognizing discussion topics and key moments, automatically generating content cues, and tagging important points, further reducing the need for manual tagging and enhancing the overall workflow.

C2: Users should have context records that correspond to specific design elements. Implement element-specific tagging that allows users to link context records, such as notes or discussions, directly to specific design elements like components or features. This tagging should be integrated into an interactive design canvas, where users can easily view, manage, and retrieve context records associated with each element. By interacting with the canvas, users can seamlessly access relevant information tied to individual parts of the design, enhancing both organization and usability.

C3: Users should be able to easily access context records to enhance their design process. Ensure seamless integration of context records within the design environment, allowing users to access them directly from their workspace without disrupting their workflow. This can be achieved through features like in-context tooltips, pop-up windows, or sidebars that provide immediate access to relevant information. To enhance usability, implement an intuitive access mechanism such as a searchable database or advanced filtering system, enabling users to efficiently locate records by keywords, tags, or design elements. Additionally, develop a contextual agent that intelligently understands user queries or actions, retrieving and presenting relevant context records to support ongoing design tasks.

C4: Users should have a clear visual representation of design iterations to facilitate understanding and collaboration. Develop a unified dashboard that consolidates all design iterations along with their associated context records, providing an overview of design progress and changes. This dashboard should feature an interactive timeline where users can click on specific points to access detailed information about each iteration, including metadata such as dates and descriptions. By using this timeline, users should be able to directly understand the reasoning behind design decisions and track the evolution of the project, facilitating collaboration and deeper comprehension.

4 DesignMemo

Based on the above-summarized insights, we developed DesignMemo, a proof-of-concept collaborative design system integrated with video conferencing. Implementing features according to the design considerations can rapidly validate different concepts and explore their relevance in helping users understand the discussion context of the concept sketch design, paving the way for effectively enhancing the collaborative design process.

To illustrate how DesignMemo supports collaborative design, we describe its core functionalities from two aspects. First, we explain how the system assists users in capturing and organizing critical contextual information during the design process (Section 4.1). Second, we show how the system enables users to revisit and retrieve this information to reconstruct the design rationale (Section 4.2).

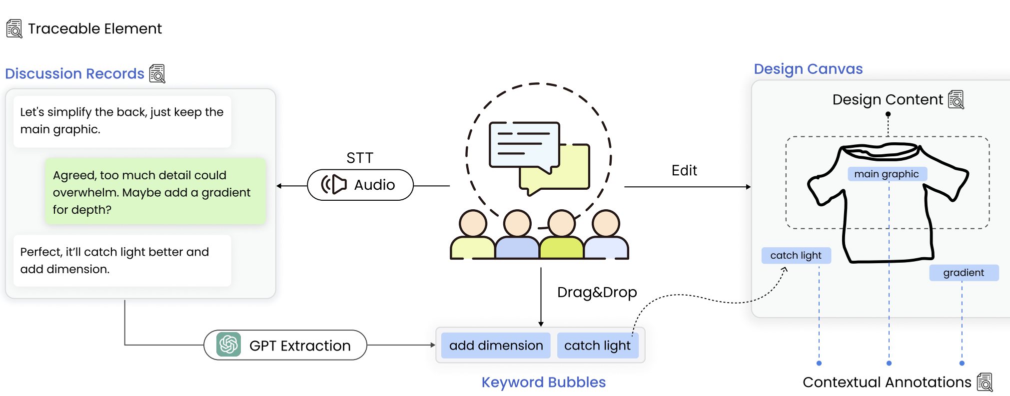


Fig. 2. Traceable Information: discussion records, design content, and contextual annotations. STT stands for "Speech to Text".

4.1 Traceable Information in the Design Process

In our investigation of remote collaboration, we identified several key types of traceable information: discussion records, design content, and contextual annotations, as shown in Fig. 2. Each of these elements encapsulates the logical progression from discussion to design within the collaborative process. By capturing this information, we aim to facilitate the ability of designers to efficiently trace back to the original decisions and the underlying rationale, thereby ensuring that the logic and context behind design choices are readily accessible for future reference.

4.1.1 Discussion Records. In remote collaborative design processes, all verbal discussions among participants are transcribed and recorded as transcripts. These transcripts capture the content of design discussions, which is crucial for understanding the context behind design decisions. By continuously recording and storing these discussions in a database, we ensure that all verbal interactions are preserved and can be recalled for future reference, maintaining the context and rationale behind the design.

4.1.2 Design Content. We also include the content of the design itself as traceable information. This means that as users iterate on their designs, multiple versions are created and stored. Specifically, since we focus on concept sketches in our design system, every stroke drawn by the users is meticulously recorded in the database. In doing so, we capture the evolution of the design process, allowing users to revisit any stage of their work and to understand the progression from one iteration to the next. This granular level of tracking ensures that every detail is preserved, enabling a comprehensive review and understanding of the design development over time.

4.1.3 Contextual Annotations. Contextual annotations act as a vital bridge between design discussions and the design content, facilitating the seamless integration of dialogue and design. In collaborative design processes, discussions frequently reference specific elements of the design, for instance, some strokes in a concept sketch. However, a significant challenge lies in linking the temporal flow of discussion transcripts with the spatial-oriented design content. To overcome this, we drew inspiration from traditional design annotations, enhancing them to enable querying discussion records related to associated design elements.

Automatic Annotation Generation. In the traditional design process, annotations typically require manual input, forcing designers to interrupt their creative flow to take notes or mark up relevant elements. While this manual annotation is important for collaborators to understand the design, it can disrupt the cognitive flow, as frequent pauses to take notes can fragment the thought process and hinder creativity. Additionally, the time and cognitive effort needed for manual annotation can reduce overall efficiency, potentially causing delays and diminishing creative output.

To alleviate the cognitive overhead of manual annotation during collaborative sketching, we developed an automated annotation generation pipeline that identifies and extracts meaningful keywords from verbal design discussions in real-time. This pipeline consists of three main stages:

- **Speech Transcription:** We utilize the Web Speech API to transcribe live audio from video conferences into text. A pause detection algorithm monitors speech intervals, identifying natural sentence boundaries and topic shifts, thereby ensuring semantic completeness and reducing noise.
- **Utterance Segmentation and Keyword Extraction:** Segmented sentences are analyzed using a prompt-engineered GPT model. This model extracts concise, contextually relevant design keywords or phrases (e.g., *"add gradient"*, *"simplify back"*, *"emphasize contrast"*) without requiring manual input, thus preserving the flow of the design process.
- **Annotation Visualization:** The extracted keywords are displayed as annotation bubbles within the user interface, facilitating immediate reference and application during collaborative sessions.

Associating Annotations with Design Elements. Once annotations are generated, they appear as floating keyword bubbles. These candidate annotations are visible to all participants in the collaborative session, allowing any user to review, select, and drag relevant keywords onto the shared design canvas. This lightweight interaction enables team members to externalize important discussion parts and link them to corresponding visual elements in real time. When a user drops an annotation onto the canvas, its spatial position is recorded. While the system does not explicitly bind annotations to individual strokes or components, the placement itself indicates the intended association. Detailed interaction designs can be found in Section. 5.1.3. These spatially anchored annotations remain persistent across the design session and are visible to all collaborators, supporting shared understanding and traceable design rationale. In Section 4.2.3, we describe how these annotations further serve as entry points for retrieving related discussion context. More technical details are provided in Appendix C.1.1.

4.2 History Review

Version history plays a crucial role in collaborative design by allowing teams to track changes and understand the evolution of their design content. Unlike traditional versioning systems, our approach incorporates discussion context into the version history. This enables users not only to revisit previous design content but also to understand the reasoning and discussions that shaped those design decisions. The system integrates both synchronous and asynchronous design management to meticulously document each iteration, thereby supporting future version reviews. When users submit a design version, the system captures not only the design content but also real-time discussion records, providing rich contextual information that clarifies the rationale behind key design decisions. This ensures that even in asynchronous design scenarios, collaborators who were not part of the live discussions can access and review the detailed conversation history. They can gain insights into the design rationale, understand the evolution of ideas, and contribute effectively to ongoing work, thereby maintaining continuity and coherence in the design process.

4.2.1 Commit and Reset. The Commit function allows users to save the current state of the design, specifically the concept sketch along with its contextual annotations, as a fixed version. Additionally, the transcript is recorded, capturing the discussion of the iteration process in detail. This operation is essential to mark significant milestones or key decision points in the design process. By committing a version, designers ensure that they can easily return to this state in the future for reference or further development.

The Reset function enables users to roll back to a previous design version. This feature is particularly valuable during experimental design phases, where multiple iterations are explored. Resetting to an earlier version allows designers to abandon less successful paths without losing progress on more promising ideas.

4.2.2 Interactive History Review. The visualization of the design iteration is crucial for tracking the history of design decisions and understanding how the design has developed and changed gradually. We offer a comprehensive history review feature, organized along a timeline, to facilitate an in-depth review of the design process. Every time a version is submitted, DesignMemo will store it in a database and render it in a timeline view. Each version is displayed as a miniature design canvas, allowing the user to interactively view the design content and examine the details of each design iteration. Furthermore, users can seamlessly trace the contextual evolution of each design iteration directly within the history review, leveraging our innovative annotation-based context recall feature, which will be detailed subsequently.

4.2.3 Annotation-based Context Recall. Effectively recording and managing design iterations in collaborative environments presents a significant challenge, especially when teams engage in both synchronous and asynchronous collaboration [58]. The challenge lies in maintaining a coherent and accessible record of the design's evolution, which is often complicated by the fragmented nature of communication across different time zones and work schedules. As users iterate through multiple

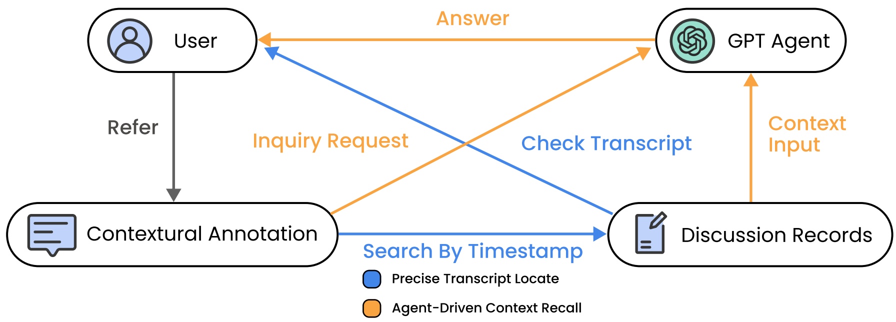


Fig. 3. Annotation-based Context Recall: Precise Transcript Locate or Agent-Driven Context Recall.

design versions, the original rationale behind certain decisions can become obscured. When these reasons are not properly documented, it can disrupt the logical continuity of subsequent designs. We refer to communication in the remote collaborative design process as contextual information, which includes the reasons behind key decisions. Contextual information is crucial because it provides the necessary background for future iterations to build on previous work effectively. Without this information, team members may struggle to understand the reasoning behind earlier design choices. This lack of clarity can lead to misinterpretations, redundant efforts, or even the reversal of critical decisions.

Our history system leverages the contextual annotations we proposed earlier, enabling users not only to revisit previous design content but also to access the discussion context linked to specific design elements. Although users can view all discussion records associated with a target version to gain a comprehensive understanding of the design logic at that time, reviewing the entire transcript can be time-consuming and challenging when they are only interested in the reasoning behind a specific part of the design. To address this, we introduce two methods to help users recall context more efficiently through annotations: precision transcript search and agent-driven context recall, as illustrated in Fig. 3.

Precise Transcript Search. To gain insight into the rationale behind a design, it is intuitive to revisit the discussions that occurred during its creation. Since users create annotations to highlight specific content during the design process, these annotations can be leveraged to facilitate targeted searches. By recording the timestamp of each annotation in the database, we can identify and retrieve the corresponding conversation in the transcript that is closest to the annotation's creation time. The conversation is then presented to the user. Additionally, to provide a broader context, users can browse the surrounding communication records, moving forward or backward from the identified conversation.

Agent-Driven Context Recall. When trying to understand the rationale behind specific design decisions, reviewing entire conversation transcripts may place a heavy cognitive load on users. To address this issue, we propose an LLM-powered agent that automatically recalls and summarizes relevant design rationales based on user queries. Our system leverages contextual annotations linked to specific design elements. When an annotation is selected, the system retrieves the corresponding timestamp and locates the surrounding conversation segments—both local context (i.e., transcripts shortly before and after the annotation) and global context (i.e., the full transcript of the design session). These retrieved transcripts are then used to construct a tailored prompt for the large language model. By feeding this structured context into the LLM, the agent can generate concise and relevant explanations that help users recall the reasoning behind past design decisions. This process reduces the effort required to trace rationale manually and improves the clarity and accessibility of collaborative design histories. Prompt used for LLM are in the Appendix C.2.

4.3 Implementation

DesignMemo is a web-based application developed using React for the front end and Agora for real-time video and audio transmission. The backend is built with JavaScript and utilizes MySQL for data storage and version management. Communication between the front-end and back-end is handled via the HTTP protocol, while the shared design canvas, used for sketching, relies on WebSocket for communication. For speech recognition, the Web Speech API is employed, and the GPT API is used to analyze transcribed text for generating annotations automatically. The application has been deployed on the public Internet, allowing access for users across the Internet and facilitating user study. More details are in the Appendix C.

5 Interactions and Interfaces

The interface design requires careful integration of remote meetings, the design canvas, and version history. As shown in Fig. 4, the DesignMemo interface is divided into three views, allowing users to switch between them: a)Whiteboard: The main interface for collaborative design, including a design canvas for creating concept sketches, a video conference area that displays the camera feed of each participant, and a potential annotation generation area. b) Transcript. A dialog box that displays real-time speech recognition for ongoing conversations. c) History Management. This view includes options to commit and reset the current design content, along with interactive exploration of past design iterations.

Based on this interface, we need to design web-based interactions with two primary goals: First, to simplify the process of recording contextual information, especially when adding contextual annotations; and Second, to enhance the management and review of version history, enabling users to quickly revisit the discussions that informed past design decisions.

5.1 Collaborative Design

5.1.1 Video Conference. Videoconferencing is essential to establish a virtual synchronous presence in the collaborative design process [36], and DesignMemo fully supports this by enabling video communication in real time between collaborators. Positioned at the top of the interface, the video stream allows participants to see each other via a webcam and engage in live audio discussions. As the conversation progresses, DesignMemo transcribes the dialogue in real time, recording it in a dialog box that can be reviewed at any time. The transcript is displayed in a separate interface to avoid disrupting the design workflow.

5.1.2 Shared Sketch Canvas. In this paper, we focus on concept sketching as our primary experimental subject. Concept sketches are foundational in the early stages of most design processes, supporting a wide range of preliminary tasks and enabling rapid iteration. To explore this, we implement a collaborative sketch interface with standard interactions on a shared canvas. For basic usage, users can employ the pencil tool to draw strokes and utilize the eraser to collaboratively remove strokes on the shared canvas.

5.1.3 Annotation on Design. The annotations not only provide supplementary textual information to the design but also serve as critical links to the discussion records, as described in Section 4.2.3. The system is designed to support the automatic generation of candidate annotations, offering users a selection to lower the workload. During discussions, the system automatically extracts key information from the dialogue, displaying it as keyword bubbles above the design canvas. These bubbles remain visible for approximately 10 seconds, minimizing disruption to the ongoing

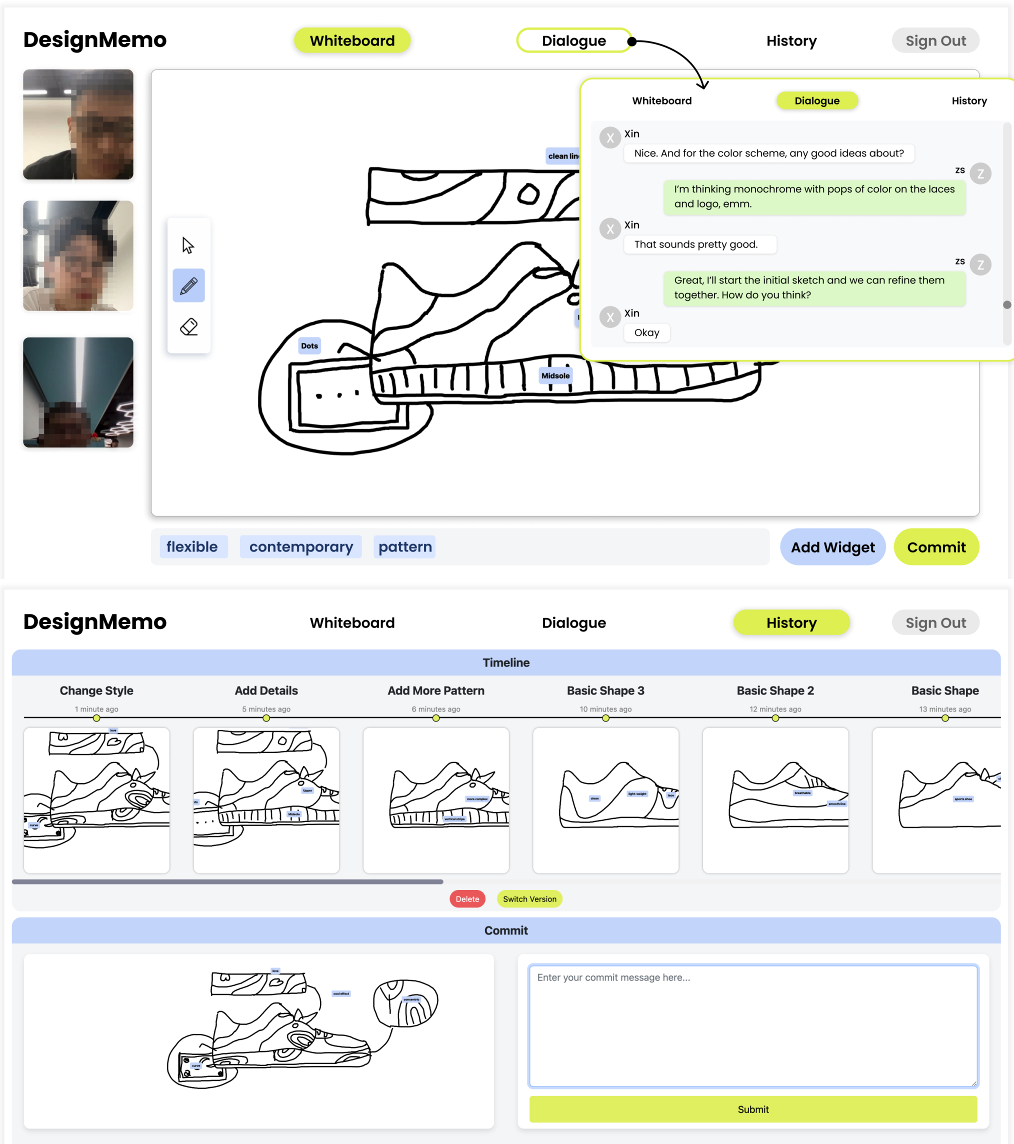


Fig. 4. Three main views in DesignMemo: Whiteboard, Dialogue, and History.

conversation. As the dialogue progresses, these bubbles will appear and disappear, allowing users to disregard those that are less pertinent. If users identify a keyword as valuable for annotating the design, they can easily drag and drop the bubble onto the relevant area corresponding to the sketch. These annotations can be repositioned freely within the whiteboard, and the system also supports manual input of the annotations.

When users complete a prototype design iteration, they can submit the current version to the version history system. To facilitate browsing and recalling the design process, each version is displayed on small canvases arranged chronologically in an interactive gallery view. These canvases behave similarly to the main design canvas but with restricted editing capabilities, allowing only content viewing.

For instance, when users conclude a design iteration and intend to archive it for subsequent phases, they can transition to the history interface. The interface displays the current design, including both the sketch and annotations. Users can provide a summary message to accompany the iteration before submission. Following submission, a new canvas is added to the timeline, depicting the design along with the associated commit message and timestamp. The timeline visually catalogs the submitted versions, allowing users to interact with these canvases by dragging, viewing, or zooming in and out. To revert to a previous version, users can select the desired version, click the toggle button, and resume editing on the main whiteboard. The active version is highlighted on the timeline with a green border.

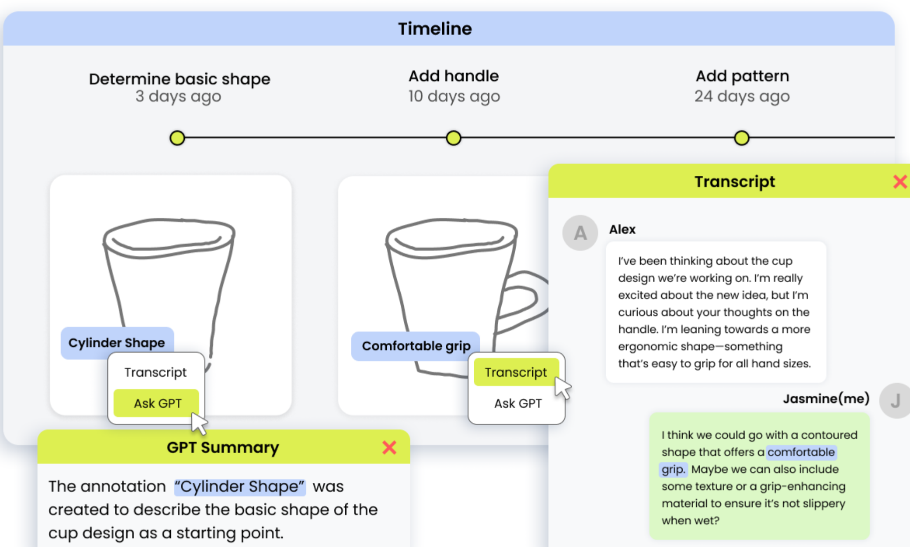


Fig. 5. Retrospection of design and consideration by transcript and agent summary.

5.2 Retrospective Design Rationale

Designers, throughout the long-term design process, will go through multiple iterations of versions, each carrying the design considerations and the reasons for the improvement of a specific period. As new requirements arise from stakeholders, designers may revisit earlier versions to extract and integrate specific design elements. However, over time, the rationale behind these iterations can fade, making it difficult to fully understand past design decisions. Traditional methods rely on meeting records to assist in understanding design changes, but this is not only time-consuming but also inefficient. Therefore, a more intuitive and natural way to view the reasons that correspond to each version of the design is needed.

DesignMemo enables designers to seamlessly navigate the design evolution, reviewing both content and rationale across versions and time without switching between iterations. By simply right-clicking on any annotation linked to a design component, two key interaction options are presented. First, users can opt to view the conversation records from the time the change was made. When selected, the system opens a dialog box that navigates directly to the relevant transcript, allowing users to review the discussions that influenced the design decisions. Alternatively, users can choose the summarized explanation of the agent. This feature takes advantage of both the global and local context to generate a concise summary of the design rationale, providing a quick overview without the need to sift through lengthy transcripts.

6 Evaluation

In this section, we conducted a within-subjects experiment to explore how DesignMemo enhances active participation and understanding during video conferences. Specifically, we examined how the system aids in the iterative prototyping process across meetings held at different times by providing contextual information about the discussions.

6.1 Participants

A total of 24 participants (11 females and 13 males), with an average age of 26.1 (SD=3.58), were recruited from local universities and companies. Their professional backgrounds included but were not limited to design, programming, sculpture, and photography. All participants noted that regular meetings and collaborations with others were an integral part of their daily routines. Upon completion of the study, each participant was compensated with $20 for their involvement.

6.2 Study Setup

DesignMemo was deployed on a remote server that participants connected to using the Chrome web browser. Participants were required to have a camera, a microphone, and a computer to participate in the study tasks. The baseline condition was informed by previous work [8] and current practices identified in our formative study. It adopted a standard remote collaboration setup that combined Zoom with Figma, supporting transcript recording and version control features. To avoid introducing unnecessary variables, participants were limited to Figma's basic sketching and text tools, ensuring that editing functions were consistent under two conditions. In each condition, groups collaboratively brainstormed ideas for a designated design task as shown in Fig. 11. To ensure fairness, each group completed comparable tasks of similar difficulty across both systems. For example, one group brainstormed initial designs for two websites using both the DesignMemo and the baseline setups.

6.3 Procedure

24 participants were divided into eight groups, each comprising three members. They were instructed to conduct experiments through remote meetings under two conditions: 1) Using DesignMemo; 2) Using Zoom Meeting and Figma. During experiments, two participants participated in both brainstorming sessions, while the third was mandated to partake only in the second retrospective discussion session. In total, it took around 60 minutes for each group to finish our study under each condition. All the experimental sessions were video recorded.

Introduction and Training Session. Participants were first informed of the purpose of the study and instructed to acquaint themselves with the operational aspects of the DesignMemo system, thereby ensuring their proficiency in performing sketch editing and utilizing traceable annotation and version features. It is worth noting that this preliminary step was deemed unnecessary when using Zoom and Figma, as all participants had already demonstrated familiarity with the aforementioned tools in prior engagements.

First Brainstorming Meeting. In both conditions, participants spent the first 5 minutes reading the task, such as designing a cup and preparing a project presentation. Then, they were instructed to join a 15-minute open discussion on the given topic and develop preliminary concept sketches.

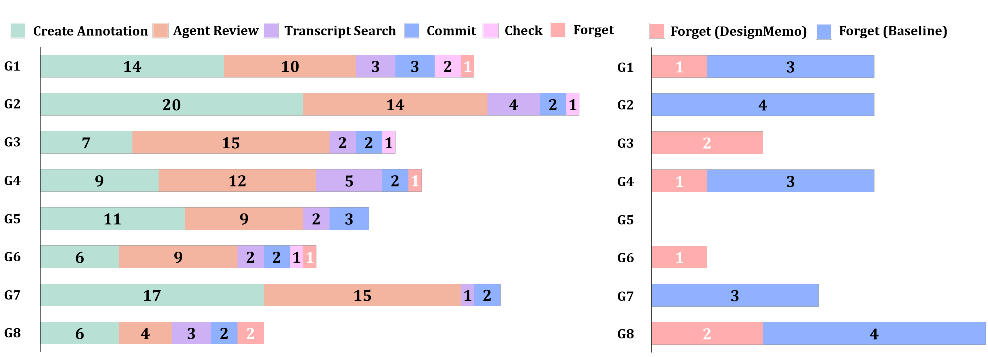
Second Retrospective Meeting. This discussion occurred one week after the initial one. The third participant first spent around five minutes reviewing the task and content of the first meeting. The person was asked to act as a stakeholder, reviewing the previous design and providing suggestions. Subsequently, all participants engaged in a collaborative discussion lasting 15 minutes, during which they worked to refine the preliminary sketches into a final, consensus-driven design.

Post-Study Survey and Interview. After the above two meetings, all three participants in the same group were asked to rate their experiences with DesignMemo compared to using Zoom and Figma on a 5-point Likert scale (1-strongly disagree, 5-strongly agree). The survey includes a five-point questionnaire based on their previous experimental experience and semi-structured interviews to gather detailed qualitative feedback on DesignMemo.

Data analysis. We systematically analyzed each group's interaction behaviors under both the DesignMemo and baseline conditions by manually coding screen recordings. As shown in Fig. 8, we identified and labeled eight distinct action types during the user study in DesignMemo, whereas the baseline setting supports only three action types. *Draw* refers to participants sketching on the canvas. *Recall* indicates efforts to remember content from the previous meeting, while *Forget* means the participants fail to recall something or achieve a false recall in the second meeting. For annotation-based interactions, *Create* represents either dragging an automatic annotation onto the canvas or manually texting an annotation. The actions *Agent Review* and *Transcript Search* reflect context-recall strategies—leveraging agent-driven summaries and direct navigation in the transcript, respectively. Regarding version control, *Commit* denotes creating a new design version in DesignMemo, and *Check* refers to reviewing previous versions through the history interface.

6.4 Results and Findings

We illustrate the count of interaction actions in DesignMemo and compare the *Forget* action count between the DesignMemo and Baseline settings in Fig. 6. To understand how users interact in two conditions, we present one group's activity of DesignMemo and baseline in Fig. 8, which shows the interaction actions over time.



(a) Frequency of actions in DesignMemo (b) Forget action under two conditions

Fig. 6. Action visualization. (a) The number of actions taken by the groups across two meetings using DesignMemo. (b) A comparison of the number of *Forget* actions across two meetings under the DesignMemo and baseline conditions.

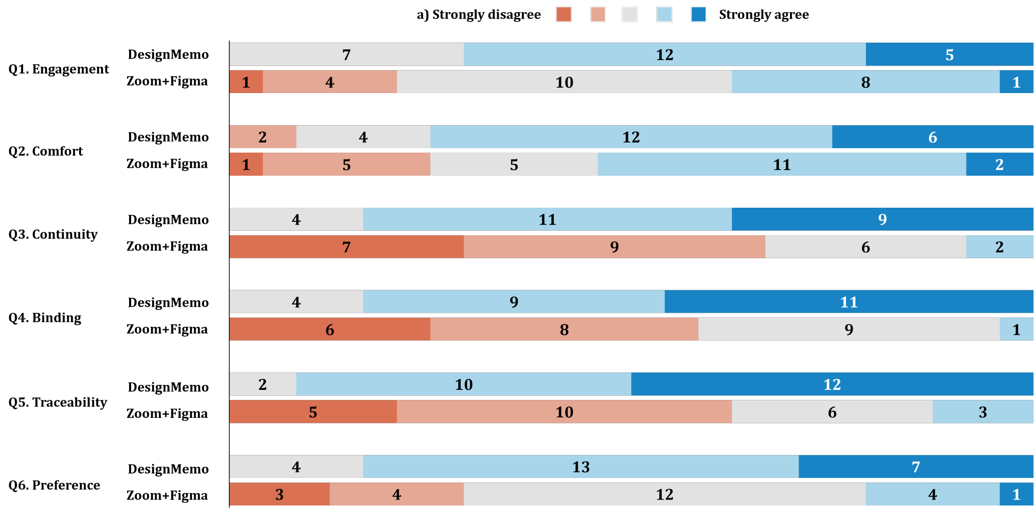


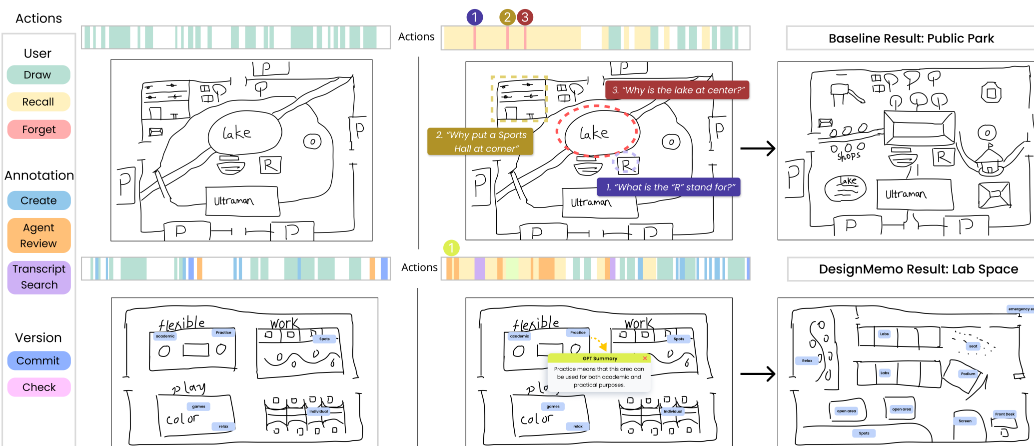
Fig. 7. User ratings for DesignMemo and a baseline setup with Zoom and Figma.

6.4.1 Effectiveness and Usability in Real-Time Conferencing. The users first emphasized how DesignMemo effectively records discussions during collaborative design sessions and links them to specific design elements.

Automatic shared annotation generation enhanced the participation in real-time meetings. The recognition of keyword annotation in DesignMemo was appreciated by participants for its ability to reduce the cognitive load of balancing active participation with note-taking. Groups are active to create annotations (mean = 11.25, SD = 4.94) in design meetings. P2 commented, *"The automatic annotations align well with things in my heart, exempting me from taking notes manually, greatly saving my time."* As shown in Fig. 7, Participants (17/24) agreed that DesignMemo engaged them well in design collaboration. Nearly half of them (11/24) mentioned that the automatic identification of critical keywords helps stay in sync with the conference proceedings, especially, such a method facilitates the collective capturing of key points, decisions, and action items in real time, reducing the loss of important information. As P4 noted, "*In general cases, the system always reminds me to consider which keywords are meaningful and how they relate to our design.*" The sharing of keywords also fostered a sense of collaboration and mutual understanding, as highlighted by P7: "*It is interesting and convenient that we drag the same word into our canvas, indicating that we are on the same side.*" Overall, DesignMemo promotes an automatic and shared approach to group participation, making users more engaged and discussions smoother.

Contextual annotations enhanced the ability to recognize the binding of design and consideration across the timeline. Many Participants (20/24) agreed that contextual annotations consistently bind the design and rationale, informing how users perceive and engage with evolving creative processes. For those who proposed their ideas, these emerging keywords often prompt participants to associate their comments with the design elements during on-the-fly meetings. This helps alleviate concerns about losing track of critical design elements and their underlying thoughts. As P7 noted, *"Having such items set up here really puts me at ease because it can remind me anytime about what's been going on with the design."* During the second meeting, the third stakeholder often presented viewpoints that differed from those of the designers. Users think the contextual annotations effectively record diverse opinions and bind them to a cohesive design. As

P1 noted, *"I can check out all the different items people are showing. They provide contrasting ideas around a specific design, helping me check what's happening with the designs at previous conferences."*



(a) Session history for the first meeting (b) Session history for the second meeting (c) Final results of the second meeting

Fig. 8. Participant Workflow. We present Group 1st's user study session under both the Baseline condition (Top) and the DesignMemo condition (Bottom). For each meeting, we visualize a timeline of actions using color coding. We also include the design outcomes discussed during the sessions, along with selected questions and actions that occurred throughout the meetings.

6.4.2 Effectiveness and Usability in Post-hoc Recollection. Recognizing the benefits of DesignMemo in the real-time meeting, the results also highlighted the advantages of design rationale tracking:

Multiple versions of discussion records enabled iterative and In-depth thinking among various design versions. From the results shown in Fig. 7, DesignMemo receives higher scores in maintaining continuity across multiple meetings than the baseline setup. Participants appreciated the ability to directly view and interact with different design versions in the history review graph, offering a straightforward way to identify points of interest. P8 commented, *"It is nice that the visualization of multiple versions effectively grasps the entire design iteration process. I can quickly skim through and then closely examine the parts I don't understand."* Such visualization aids designers in forging an overall comprehension of the iterative design process, guiding them from a broad understanding of design thinking to detailed insights on specific design elements. Another participant highlighted the importance of preserving partial design results and promoting in-depth consideration in a focus design element, stating, *"I often switch between various design perspectives. This feature allows me to freely explore the currently inspired design elements without worrying about losing ideas in other branches."* Overall, the multiple versions tracking in DesignMemo was seen as a forward in remote design meetings, providing an efficient way to store previous results that are easy for all participants to recall. Although participants acknowledged the potential of version control, our evaluation for such function was a bit limited to the lack of a long-term design process as groups did not commit (mean = 2.25, SD = 0.433) and check (mean = 0.63, SD = 0.70) the previous versions frequently.

Automatic annotation summary facilitated users to understand the local context. Many participants (12/24) strongly agreed with the annotation's ability to enhance traceability when two meetings are disconnected. During the second meeting, participants heavily relied on the *Agent*

Review function to recall previous discussion content, a preference also supported by quantitative data showing higher usage of *Agent Review* (mean = 11.0, SD = 3.54) compared to *Transcript Search* (mean = 2.75, SD = 1.19). Such contextual annotation involves completed history records on specific design elements, which is one unique benefit brought by DesignMemo. P9 praised the accuracy of the agent-driven summary in those context-enriched substrates. commented, "*I love the annotation items because this (contextual annotation) makes accessing history so simple and so accurate.*" Compared to the lengthy and informal transcript of an entire meeting, participants widely appreciate the contextual annotation, describing it as *"the medium between design and human thought."* This feature significantly improved the efficiency of collaborative design by enabling precise access to relevant information. P11 noted, "*I am not willing to sift through the whole video or translate; I simply want to seek specific information.*" And his collaborator, P10, added "*It is valuable for me to click on such items to get the specific descriptions. Otherwise, recalling content in a large space would be difficult, no one wants to search for a needle in a haystack every time.*" As shown in Fig. 6 (b), groups were more prone to forgetting previous meeting details or providing inaccurate recollections in the baseline setup (mean = 2.38, SD = 1.50) compared to the DesignMemo setup (mean = 0.63, SD = 0.70). This difference can be largely attributed to the high accuracy of summaries generated by the *Agent Review* action. User ratings for traceability and binding (Fig. 7) further reinforce the effectiveness of contextual recall. In our semi-structured interviews, nearly all participants confirmed that the recalled content was appropriate and relevant. While two groups noted minor inaccuracies during design tasks involving an academic poster and a conference website, these issues stemmed from discussions of recently introduced research ideas not covered in the model's training data. To our knowledge, such limitations could be addressed by leveraging a more up-to-date LLM with access to online information [29]. Thus, we believe that DesignMemo provided a promising way to make working with multiple meetings in design easier and more productive. It offers new and exciting possibilities for collaborative design in an iterative process.

6.4.3 Technical and Design Limitations of Current System. Through our user study, we identify limitations of the current implementation.

Errors in transcription and reduced readability. Our user study revealed two key issues in the use of spoken dialogue transcripts: transcription errors that affect machine understanding, and unstructured speech that hinders human comprehension. First, we observed that errors, as P17 noted "*It records my unclear pronunciation and identifies a wrong sentence meaning.*" While human readers are able to infer the intended meaning from context, such errors can confuse computational agents and result in incorrect outputs. In contrast, the second issue lies in the inherently unstructured nature of spoken language. This makes it difficult for human readers to follow the conversation, especially when there are topic shifts, interruptions, or filler words. Interestingly, despite this lack of structure, large language models are still capable of extracting relevant information from these transcripts to a certain extent [48]. To mitigate these issues, future improvements could focus on enhancing transcription accuracy through context-aware speech recognition models and improving human readability via automatic summarization or discourse reorganization techniques.

Agent can not see the design content. In the semi-structured interview, we heard that participants occasionally find some annotations are irrelevant to the design content in canvas, as explained by P13, "*We have talked about this, but we didn't draw this idea out.*" This disconnect between verbal discussion and visual output can lead to confusion in interpreting the transcript, especially when ideas are discussed but not visually documented. While participants who attended the meeting can often resolve these ambiguities based on shared context, stakeholders who were not present may struggle to identify such issues unless they explicitly ask questions. This issue can be mitigated to a large extent by implementing a vision large language model to track the design process by meeting transcripts and screen recordings.

6.5 Design Implications

Based on the above findings, we distill several implications for the future collaborative design system in remote collaboration:

Collaborative design system should record not only the iterations of the design content but also the motivations behind those designs. While some research focuses on tracking the version history of design content, such as the necessity of design history in creative workflows [47], the use of branches in design histories [9], and the visualization of branches in collaborative settings [58], merely reviewing iterations does not provide a complete understanding of the design process. Designers often lose sight of the rationale behind earlier decisions, leading to confusion or misalignment. This is because people often forget the purpose of the design, leading to confusion. To maintain the continuity of the iteration design process across different times, collaborative design systems should consider effectively storing the thinking behind the design decisions. This could be done through notes, annotations [57], or direct integration with decision logs that explain the reasoning behind design changes. This feature ensures that key decisions, design choices, and annotations from previous meetings are easily accessible, reducing redundancy, preventing vision drift, and maintaining team alignment over multiple sessions.

Collaborative design system should link context information with specific design elements. Transcripts generated through speech-to-text methods are often cluttered with irrelevant information, making it difficult to extract meaningful insights. Some research has attempted to address this issue. For instance, MeetCues focuses on extracting subtle cues from meetings [3], MeetScript provides tools for manually refining or tagging transcripts [8], and Meeting Bridges shifts the task of organizing meeting content to post-meeting asynchronous activities [51]. Nevertheless, when we shift to remote collaborative design, it remains challenging to locate the rationale behind particular design elements within transcripts. Therefore, future collaborative design systems should allow users to query design rationale directly from the design canvas. The contextual annotation method proposed in this paper offers a solution by linking timestamps to design elements, enabling users to click on relevant sections to instantly access the associated conversation.

Collaborative design system that integrates an assistant agent should take advantage of discussion records. Many existing studies on AI-assisted design tools often lack contextual information in previous and ongoing meetings for agents to provide accurate and relevant assistance [28, 31]. This absence of context leads to AI tools offering generic suggestions that may not align with the specific needs of a project [51]. Although modern LLMs are incredibly powerful and possess extensive prior knowledge, they struggle to provide effective assistance when they lack a specific context on a design project. Our user study shows that LLM can utilize transcripts to understand contextual information and provide more task-oriented support. Providing the agent with transcripts in chronological order not only helps it grasp the design background but also tracks the evolution of the design as the process unfolds. Therefore, when incorporating LLM agents into design systems, do not overlook the importance of transcripts as a key source of information.

Collaborative design system should carefully examine content across multiple levels of detail when recalling design context. Existing studies on collaborative meetings have not adequately differentiated between varying levels of detail in presenting information. This lack of granularity control often forces users to choose between overly detailed records or overly simplified summaries, such as manually marking important content during real-time speech transcription [23], or adding extensive annotations to transcripts [30], limiting their ability to efficiently demonstrate relevant information. Different levels of detail in design rationale offer distinct benefits. In DesignMemo, we support users with three types of information to help them recall the previous design: annotation, agent summaries, and origin transcript. Annotations are just limited keywords created by the user to label the design. While transcripts provide detailed information that can help users understand the design process by reviewing conversations, they can be time-consuming and may impose a high cognitive workload. On the other hand, summaries from an agent are more concise and easier to understand, but their accuracy and detail may not always meet user expectations. Therefore, the future system should provide the rationale for the design at various levels of detail, offering users multiple options to suit their needs, and balancing between comprehensive transcripts and concise summaries.

7 Limitations and Future Work

7.1 Limitations

Despite the promising results and potential benefits, our approach faces several limitations that could affect its overall performance and applicability in real-world scenarios.

Lack of Evaluation for long-term, large team, and various design iteration processes. Due to practical constraints, the user studies are conducted in controlled experiments that may differ from real-world design. First, the user study only covers two iterations of the design process with a gap of just one week, which might not reflect the iterative and incremental nature of design work that unfolds over extended periods in real-life situations that last several months. Second, we conducted the studies with a small sample size of only two to three participants presenting certain limitations. However, in larger teams with more significant numbers of participants, there is a higher potential for discovering new insights that might not be evident in smaller groups. Third, our participants represent only a limited range of design backgrounds, including game scene design, architectural design, sculpture design, and interaction design. However, there may be many other design domains that we haven't included. Since DesignMemo is intended for general designers, we must evaluate a broader range of design tasks to ensure the system's versatility and effectiveness across different design fields.

Lack of consideration for asynchronous design practices. Our study does not address asynchronous design processes, which are common in real-world settings where designers often work individually across different times [25]. In practice, it is typical for a designer to continue iterating on a concept sketch produced during a prior collaborative session. Such asynchronous workflows often involve fragmented engagement with past discussions, where designers seek specific rationale rather than relying on direct conversations with teammates. While our system supports accessing and reviewing design rationale at any time, our user study did not include scenarios where participants independently revisited prior discussions or decisions outside of synchronous collaboration. As a result, we have yet to examine how effectively the system supports individual reflection, delayed feedback, or staggered design activities in asynchronous workflows.

Lack of in-the-wild evaluation. Our study was conducted in a controlled lab setting, which may not fully capture the complexity and unpredictability of real-world design environments. In practice, design processes are often shaped by contextual factors such as team dynamics, organizational constraints, and evolving profit goals [51]. These factors may influence how rationale is captured, interpreted, and reused over time. While our study offers valuable insights into the system's potential, its artificial setting limits our understanding of how the tool performs in longitudinal, organically evolving design workflows. A more extensive in-the-wild deployment is needed to evaluate the system's effectiveness in supporting design rationale across varied, messy, and dynamic real-world conditions.

7.2 Future Work

As we look toward future developments, there are several avenues for enhancing the capabilities of DesignMemo to better support diverse design tasks and improve user experience.

Exploring the role of multi-Modal context in collaborative design support. Our current system relies solely on conversational transcripts as input for the conversational agent. However, collaborative design frequently involves references to external materials—such as papers, sketches, or web pages [55]—that provide essential context for understanding design rationale. In our study, we observed cases where the lack of integrated multi-modal context led to reduced agent relevance and user frustration. This points to a broader research opportunity: understanding how access to and alignment across diverse modalities (e.g., text, images, sketches, documents) affects the effectiveness of agent-supported collaboration [19]. Future work could investigate what types of contextual information are most critical for different design stages, how they should be represented to the agent, and how to balance completeness with cognitive load in multi-modal agent interactions.

Investigating rationale tracking across diverse design domains and modalities. While our study focuses on collaborative concept sketching, the underlying rationale tracking mechanism may generalize to a broader range of design practices, such as character design, animation [26], or architectural planning. Future research could explore how the capture and retrieval of rationale functions across these varied domains, particularly where design artifacts are more spatially complex or temporally extended. Additionally, emerging platforms such as augmented and virtual reality introduce new forms of spatial interaction and embodiment [56]. These modalities raise open research questions about how design rationale can be embedded in 3D environments, how users access past decisions through spatial cues, and how such embedded representations affect collaborative understanding and creative flow. Studying these interactions could deepen our understanding of how rationale tracking operates in immersive, multi-modal settings.

Understanding the cognitive and collaborative impacts of design rationale availability. Future research may investigate how the availability of explicit design rationale affects designers' cognitive processes and team collaboration over time. In collaborative settings, rationale can serve as a cognitive anchor to help team members recall prior decisions, align on shared goals, and avoid redundant discussions. However, it may also constrain creative exploration if earlier justifications are treated as fixed constraints [34]. Little is known about how designers interpret, trust, or reinterpret rationale authored by others, especially in asynchronous workflows. Further studies could examine when and how rationale is most effectively accessed, how it influences design continuity and decision quality, and how it mediates communication in evolving team configurations.

8 Conclusion

In this work, we have introduced DesignMemo, a proof-of-concept system that integrates the verbal context of remote discussions into visual design history tracking. The system automatically generates potential annotations for the design according to the audio, which is linked to a certain transcript of the meeting, so that the user can easily recall the context of the design by clicking the annotation. The system also integrates agent summarization based on global and local context information supported by the transcript, so that the user can quickly understand the rationale behind a specific design part. We initially conducted a formative study for in-depth design considerations, interviewing designers and stakeholders who had gained extensive experience in the design process. Our user study simulated two iterations in a collaborative design process that included 16 designers and 8 stakeholders. Through user study, we distill several implications for the future design of collaborative design systems, emphasizing the significance of integrating transcripts into the design history review. We envision that our concept of utilizing transcripts as contextual information for both agents and DesignMemo will inspire future work in the design process, with potential applications across various collaborative design domains, including product development, architectural planning, and interactive media design.

Acknowledgments

This work was partially supported by Guangzhou Basic Research Scheme #2024A04J4229 and Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things #2023B1212010007. We would also thank Red Bird MPhil Program at HKUST(GZ) for their general support.

References

[1] 2022. Miro. https://miro.com/
[2] 2022. Mural. https://mural.co/
[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)*. 236–240. https://doi.org/10.1109/VIS47514.2020.00054
[4] Amid Ayobi, Jacob Hughes, Christopher J Duckworth, Jakub J Dylag, Sam James, Paul Marshall, Matthew Guy, Anitha Kumaran, Adriane Chapman, Michael Boniface, and Aisling Ann O'Kane. 2023. Computational Notebooks as Co-Design Tools: Engaging Young Adults Living with Diabetes, Family Carers, and Clinicians with Machine Learning Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 300, 20 pages. https://doi.org/10.1145/3544548.3581424
[5] Filip Bircanin, Margot Brereton, Laurianne Sitbon, Bernd Ploderer, Andrew Azaabanye Bayor, and Stewart Koplick. 2021. Including Adults with Severe Intellectual Disabilities in Co-Design through Active Support. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 486, 12 pages. https://doi.org/10.1145/3411764.3445057
[6] Rob Bracewell, Ken Wallace, Michael Moss, and David Knott. 2009. Capturing Design Rationale. *Computer-Aided Design* 41, 3 (2009), 173–186. https://doi.org/10.1016/j.cad.2008.10.005 Computer Support for Conceptual Design.
[7] Hsiang-Ting Chen, Li-Yi Wei, and Chun-Fa Chang. 2011. Nonlinear Revision Control for Images. *ACM Trans. Graph.* 30, 4, Article 105 (2011), 10 pages. https://doi.org/10.1145/2010324.1965000
[8] Xinyue Chen, Shuo Li, Shipeng Liu, Robin Fowler, and Xu Wang. 2023. MeetScript: Designing Transcript-based Interactions to Support Active Participation in Group Video Meetings. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 347 (2023), 32 pages. https://doi.org/10.1145/3610196
[9] Kathy Cheng, Phil Cuvin, Alison Olechowski, and Shurui Zhou. 2023. User Perspectives on Branching in Computer-Aided Design. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 371 (Oct. 2023), 30 pages. https://doi.org/10.1145/3610220
[10] Nigel Cross. 2001. Chapter 5 - Design Cognition: Results from Protocol and Other Empirical Studies of Design Activity. In *Design Knowing and Learning: Cognition in Design Education*, Charles M. Eastman, W. Michael McCracken, and Wendy C. Newstetter (Eds.). Elsevier Science, Oxford, 79–103. https://doi.org/10.1016/B978-008043868-9/50005-X
[11] Peter Dalsgaard and Kim Halskov. 2012. Reflective Design Documentation. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*. Association for Computing Machinery, New York, NY, USA, 428–437. https://doi.org/10.1145/2317956.2318020
[12] Maitraye Das, Thomas Barlow McHugh, Anne Marie Piper, and Darren Gergle. 2022. Co11ab: Augmenting Accessibility in Synchronous Collaborative Writing for People with Vision Impairments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 196, 18 pages. https://doi.org/10.1145/3491102.3501918
[13] Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2010. Chronicle: Capture, Exploration, and Playback of Document Workflow Histories. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. Association for Computing Machinery, New York, NY, USA, 143–152. https://doi.org/10.1145/1866029.1866054
[14] Marisela Gutierrez Lopez, Mieke Haesen, Kris Luyten, and Karin Coninx. 2015. Helaba: A System to Highlight Design Rationale in Collaborative Design Processes. In *Cooperative Design, Visualization, and Engineering*, Yuhua Luo (Ed.). Springer International Publishing, Cham, 175–184.
[15] Marisela Gutierrez Lopez, Gustavo Rovelo, Mieke Haesen, Kris Luyten, and Karin Coninx. 2017. Capturing Design Decision Rationale with Decision Cards. In *Human-Computer Interaction - INTERACT 2017*, Regina Bernhaupt, Girish

Dalvi, Anirudha Joshi, Devanuj K. Balkrishan, Jacki O'Neill, and Marco Winckler (Eds.). Springer International Publishing, Cham, 463–482.

[16] Marisela Gutierrez Lopez, Gustavo Rovelo Ruiz, Kris Luyten, Mieke Haesen, and Karin Coninx. 2018. Re-thinking Traceability: A Prototype to Record and Revisit the Evolution of Design Artefacts. In *Proceedings of the 2018 ACM International Conference on Supporting Group Work (GROUP '18)*. Association for Computing Machinery, New York, NY, USA, 196–208. https://doi.org/10.1145/3148330.3148334

[17] Dean M. G. Hargreaves and Toni Robertson. 2012. Remote Participatory Prototyping Enabled by Emerging Social Technologies *(PDC '12)*. Association for Computing Machinery, New York, NY, USA, 25–28. https://doi.org/10.1145/2348144.2348153

[18] John Horner and Michael E. Atwood. 2006. Design Rationale: the Rationale and the Barriers. In *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles (NordiCHI '06)*. Association for Computing Machinery, New York, NY, USA, 341–350. https://doi.org/10.1145/1182475.1182511

[19] Hanyao Huang, Ou Zheng, Dongdong Wang, Jiayi Yin, Zijin Wang, Shengxuan Ding, Heng Yin, Chuan Xu, Renjie Yang, Qian Zheng, et al. 2023. ChatGPT for shaping the future of dentistry: the potential of multi-modal large language model. *International Journal of Oral Science* 15, 1 (2023), 29.

[20] Robert J. K. Jacob, Hiroshi Ishii, Gian Pangaro, and James Patten. 2002. A Tangible Interface for Organizing Information Using A Grid. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. Association for Computing Machinery, New York, NY, USA, 339–346. https://doi.org/10.1145/503376.503437

[21] Shivesh Jadon, Mehrad Faridan, Edward Mah, Rajan Vaish, Wesley Willett, and Ryo Suzuki. 2024. Augmented Conversation with Embedded Speech-Driven On-the-Fly Referencing in AR. *arXiv preprint arXiv:2405.18537* (2024).

[22] Ben Jonson. 2005. Design Ideation: The Conceptual Sketch in The Digital Age. *Design Studies* 26, 6 (2005), 613–624. https://doi.org/10.1016/j.destud.2005.03.001

[23] Vaiva Kalnikaitunde, 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* (Seattle, Washington, USA) *(CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 349–358. https://doi.org/10.1145/2145204.2145260

[24] Scott R. Klemmer, Michael Thomsen, Ethan Phelps-Goodman, Robert Lee, and James A. Landay. 2002. Where Do Web Sites Come From? Capturing and Interacting with Design History. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/503376.503378

[25] Kung Jin Lee, Wendy Roldan, Tian Qi Zhu, Harkiran Kaur Saluja, Sungmin Na, Britnie Chin, Yilin Zeng, Jin Ha Lee, and Jason Yip. 2021. The Show Must Go On: A Conceptual Model of Conducting Synchronous Participatory Design With Children Online. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 345, 16 pages. https://doi.org/10.1145/3411764.3445715

[26] Boyu Li, Linping Yuan, Zhe Yan, Qianxi Liu, Yulin Shen, and Zeyu Wang. 2024. AniCraft: Crafting Everyday Objects as Physical Proxies for Prototyping 3D Character Animation in Mixed Reality. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 99, 14 pages. https://doi.org/10.1145/3654777.3676325

[27] Yan Li, Jian Wang, Xianglong Li, and Wu Zhao. 2007. Design Creativity in Product Innovation. *The International Journal of Advanced Manufacturing Technology* 33 (01 2007), 213–222. https://doi.org/10.1007/s00170-006-0457-y

[28] Long Ling, Xinyi Chen, Ruoyu Wen, Toby Jia-Jun Li, and RAY LC. 2024. Sketchar: Supporting Character Design and Illustration Prototyping Using Generative AI. *Proc. ACM Hum.-Comput. Interact.* 8, CHI PLAY, Article 337 (Oct. 2024), 28 pages. https://doi.org/10.1145/3677102

[29] Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, et al. 2024. Deepseek-v3 Technical Report. *arXiv preprint arXiv:2412.19437* (2024).

[30] Ching (Jean) Liu, Chi-Lan Yang, Joseph Jay Williams, and Hao-Chuan Wang. 2019. NoteStruct: Scaffolding Note-taking while Learning from Online Videos. 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/10.1145/3290607.3312878

[31] Fang Liu, Junyan Lv, Shenglan Cui, Zhilong Luan, Kui Wu, and Tongqing Zhou. 2024. Smart "Error"! Exploring Imperfect AI to Support Creative Ideation. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 121 (April 2024), 28 pages. https://doi.org/10.1145/3637398

[32] Tencent Holdings Ltd. 2018. Voov Meeting. https://voovmeeting.com/. Accessed: 2024-07-11.

[33] A. MacLean, R. M. Young, and T. P. Moran. 1989. Design Rationale: The Argument Behind The Artifact. *SIGCHI Bull.* 20, SI (mar 1989), 247–252. https://doi.org/10.1145/67450.67497

[34] Helena M. Mentis, Paula M. Bach, Blaine Hoffman, Mary Beth Rosson, and John M. Carroll. 2009. Development of Decision Rationale in Complex Group Decision Making. In *Proceedings of the SIGCHI Conference on Human*

Factors in Computing Systems (CHI '09). Association for Computing Machinery, New York, NY, USA, 1341–1350. https://doi.org/10.1145/1518701.1518904

[35] Karin Niemantsverdriet and Thomas Erickson. 2017. Recurring Meetings: An Experiential Account of Repeating Meetings in a Large Organization. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 84 (Dec. 2017), 17 pages. https://doi.org/10.1145/3134719

[36] Roy Pea and Robb Lindgren. 2008. Video Collaboratories for Research and Education: An Analysis of Collaboration Design Patterns. *IEEE Transactions on Learning Technologies* 1, 4 (2008), 235–247. https://doi.org/10.1109/TLT.2009.5

[37] Cecil Piya, Vinayak , Senthil Chandrasegaran, Niklas Elmqvist, and Karthik Ramani. 2017. Co-3Deator: A Team-First Collaborative 3D Design Ideation Tool. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6581–6592. https://doi.org/10.1145/3025453.3025825

[38] William C Regli, Xiaochun Hu, Michael Atwood, and Wei Sun. 2000. A survey of design rationale systems: approaches, representation, capture and retrieval. *Engineering with computers* 16 (2000), 209–235.

[39] Lill Francis Miranda Reyes and Sisse Finken. 2012. Social Media as A Platform for Participatory Design. In *Proceedings of the 12th Participatory Design Conference: Exploratory Papers, Workshop Descriptions, Industry Cases - Volume 2 (PDC '12)*. Association for Computing Machinery, New York, NY, USA, 89–92. https://doi.org/10.1145/2348144.2348173

[40] P.A Rodgers, G Green, and A McGown. 2000. Using Concept Sketches to Track Design Progress. *Design Studies* 21, 5 (2000), 451–464. https://doi.org/10.1016/S0142-694X(00)00018-1

[41] Elizabeth B.-N. Sanders and Pieter Jan Stappers. 2008. Co-creation and the New Landscapes of Design. *CoDesign* 4, 1 (2008), 5–18. https://doi.org/10.1080/15710880701875068

[42] James A. Self. 2019. Communication through Design Sketches: Implications for Stakeholder Interpretation During Concept Design. *Design Studies* 63 (2019), 1–36. https://doi.org/10.1016/j.destud.2019.02.003

[43] Yang Shi, Chris Bryan, Sridatt Bhamidipati, Ying Zhao, Yaoxue Zhang, and Kwan-Liu Ma. 2018. MeetingVis: Visual Narratives to Assist in Recalling Meeting Context and Content. *IEEE Transactions on Visualization and Computer Graphics* 24, 6 (2018), 1918–1929. https://doi.org/10.1109/TVCG.2018.2816203

[44] Laurianne Sitbon. 2018. Engaging IT Students in Co-Design with People with Intellectual Disability. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188620

[45] Geertje Slingerland, Maria Murray, Stephan Lukosch, John McCarthy, and Frances Brazier. 2022. Participatory Design Going Digital: Challenges and Opportunities for Distributed Place-Making. *Computer Supported Cooperative Work (CSCW)* 31 (07 2022). https://doi.org/10.1007/s10606-022-09438-3

[46] Seoyun Son, Junyoug Choi, Sunjae Lee, Jean Y Song, and Insik Shin. 2023. It is Okay to be Distracted: How Real-time Transcriptions Facilitate Online Meeting with Distraction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 64, 19 pages. https://doi.org/10.1145/3544548.3580742

[47] Sarah Sterman, Molly Jane Nicholas, and Eric Paulos. 2022. Towards Creative Version Control. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 336 (Nov. 2022), 25 pages. https://doi.org/10.1145/3555756

[48] Katrin Tomanek, Jimmy Tobin, Subhashini Venugopalan, Richard Cave, Katie Seaver, Jordan R. Green, and Rus Heywood. 2024. Large Language Models As A Proxy For Human Evaluation In Assessing The Comprehensibility Of Disordered Speech Transcription. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 10846–10850. https://doi.org/10.1109/ICASSP48485.2024.10447177

[49] Barbara Tversky. 2002. What Do Sketches Say about Thinking?. In *2002 AAAI Spring Symposium, Sketch Understanding Workshop, Stanford University, AAAI Technical Report SS-02-08*, Vol. 148. 151.

[50] Greg Walsh, Allison Druin, Mona Leigh Guha, Elizabeth Bonsignore, Elizabeth Foss, Jason C. Yip, Evan Golub, Tamara Clegg, Quincy Brown, Robin Brewer, Asmi Joshi, and Richelle Brown. 2012. DisCo: A Co-Design Online Tool for Asynchronous Distributed Child and Adult Design Partners. In *Proceedings of the 11th International Conference on Interaction Design and Children (IDC '12)*. Association for Computing Machinery, New York, NY, USA, 11–19. https://doi.org/10.1145/2307096.2307099

[51] Ruotong Wang, Lin Qiu, Justin Cranshaw, and Amy X. Zhang. 2024. Meeting Bridges: Designing Information Artifacts that Bridge from Synchronous Meetings to Asynchronous Collaboration. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 35 (apr 2024), 29 pages. https://doi.org/10.1145/3637312

[52] Chelsea-Joy Wardle, Mitchell Green, Christine Wanjiru Mburu, and Melissa Densmore. 2018. Exploring Co-design with Breastfeeding Mothers. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3174056

[53] Cara Wilson, Margot Brereton, Bernd Ploderer, and Laurianne Sitbon. 2019. Co-Design Beyond Words: 'Moments of Interaction' with Minimally-Verbal Children on the Autism Spectrum. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15.

https://doi.org/10.1145/3290605.3300251

[54] Julia Woodward, Zari McFadden, Nicole Shiver, Amir Ben-hayon, Jason C. Yip, and Lisa Anthony. 2018. Using Co-Design to Examine How Children Conceptualize Intelligent Interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3174149

[55] 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. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 60, 16 pages. https://doi.org/10.1145/3586183.3606773

[56] Lin-Ping Yuan, Boyu Li, Jindong Wang, Huamin Qu, and Wei Zeng. 2024. Generating Virtual Reality Stroke Gesture Data from Out-of-Distribution Desktop Stroke Gesture Data. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 732–742. https://doi.org/10.1109/VR58804.2024.00093

[57] Amy X. Zhang and Justin Cranshaw. 2018. Making Sense of Group Chat through Collaborative Tagging and Summarization. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 196 (Nov. 2018), 27 pages. https://doi.org/10.1145/3274465

[58] Lei Zhang, Ashutosh Agrawal, Steve Oney, and Anhong Guo. 2023. VRGit: A Version Control System for Collaborative Content Creation in Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 36, 14 pages. https://doi.org/10.1145/3544548.3581136

[59] Inc. Zoom Video Communications. 2021. Zoom. https://zoom.us/. Accessed: 2024-07-10.

A Concept Sketch Description

Figure 9 support some examples of concept sketch.

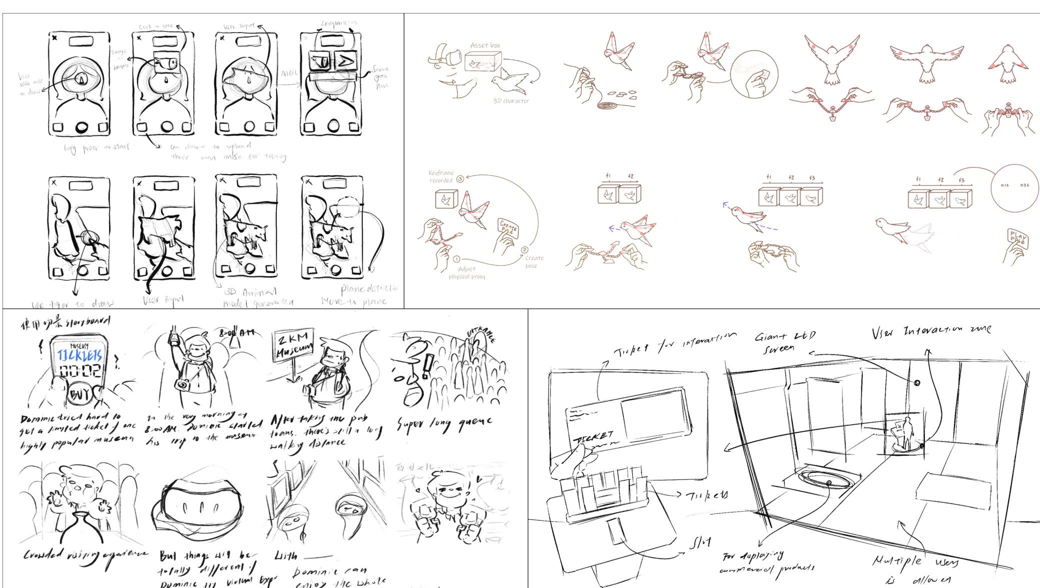


Fig. 9. Examples of our research focus: concept sketch.

B Formative Study Detail

B.1 Interview Questions

The following is a selected list of guiding questions used in our semi-structured interviews. These questions were not intended to be a fixed questionnaire; rather, they served as prompts to facilitate open-ended conversations. The actual questions asked varied slightly depending on the participant's role (designer or stakeholder) and the context of the discussion. Follow-up questions were asked when appropriate to further probe individual experiences or clarify specific responses.

The interview sessions were structured around three main themes: (1) Workflow and Tool Usage, (2) Challenges in Context Retention, and (3) Reflections and Ideal Tool Support.

1. Workflow and Tool Usage

- Can you describe a recent remote collaborative design project you participated in?
- What tools did you use to support collaboration during that project (e.g., for meetings, sketching, documentation)?
- How did you typically conduct design meetings or discussions remotely?

2. Challenges in Context Retention

- How did you or your team document key decisions made during discussions?
- What difficulties have you encountered when trying to recall why certain design decisions were made?
- How do you usually access past design rationale? What do you find frustrating or helpful about the current process?

3. Reflections and Ideal Tool Support

- What features or tools do you wish you had to better track and retrieve design rationale?
- If a tool could automatically capture meeting content and associate it with your design work, what would be important to you?
- How would you envision an ideal system that supports better understanding of past design discussions?

B.2 Representative Interview Quotes Supporting Key Challenges

This appendix provides selected verbatim quotes from participant interviews that support the four primary challenges (**C1–C4**) identified in Section 3. These quotes illustrate the participants' lived experiences and reflect the recurring patterns that informed our thematic coding.

Challenge 1: Difficulty in Recording and Linking Notes (C1).

> *"It's hard to take notes while also presenting or sketching. I usually have to choose one or the other."* (D1)
>
> *"I open Notion in one tab, Zoom in another, and Figma in a third. I always forget which note belongs to which screen discussion."* (S2)

These quotes highlight the cognitive overload of manually documenting meetings, and the disconnect between design artifacts and corresponding rationale.

Challenge 2: Difficulty Locating Relevant Discussion (C2).

> *"We record everything, but good luck finding that one comment from last week's call."* (D3)
>
> *"Even if I remember what we said, scrolling through transcripts with no tags is a nightmare."* (D5)

These observations underscore the lack of fine-grained indexing or element-level linkage in current tooling.

Challenge 3: Disconnection Between Design and Rationale (C3).

> *"I often want to check why we chose a design. But the decision info is always somewhere else—maybe Slack, maybe a doc."* (D4)

> *"Switching away from the design canvas to find the reason behind a change breaks my rhythm."* (D2)

This aligns with the need for rationale access without leaving the design workspace.

Challenge 4: Lack of Visibility into Iteration History (C4).

> *"Sometimes I forget which version had what feature, and there's no easy way to compare."* (S6)

> *"The design evolved a lot, but we couldn't really explain how or why it changed to the client."* (D6)

This emphasizes the lack of clear visual histories and the importance of embedding context into version review.

C System Implementation

C.1 System Architecture

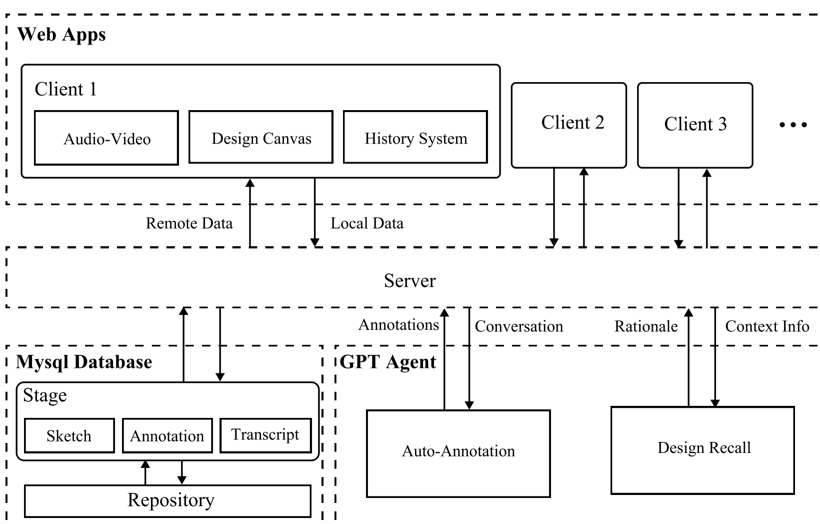


Fig. 10. System architecture of DesignMemo.

C.1.1 Annotation Creation. This appendix provides technical details on the annotation generation pipeline used in DesignMemo, including real-time speech processing, utterance segmentation, GPT-based keyword extraction, and the spatial linking of annotations to design elements on the canvas.

Pipeline. The annotation system follows a four-stage pipeline, implemented using a combination of browser-native APIs and server-side processing:

(1) **Real-time Speech Transcription.** Audio from the video conferencing stream is transcribed in real time using the browser-native `Web Speech API`, which outputs a continuous stream

of partial and final transcript segments. We listen for ‘onresult‘ events to retrieve finalized utterances for downstream processing.

(2) **Utterance Segmentation.** A final segment is assumed to correspond to a semantically complete utterance, typically finalized after ∼500–800ms of silence. Each finalized utterance is stored in a temporary buffer and immediately passed to the annotation module.

(3) **Keyword Extraction with GPT.** When a finalized utterance is received, it is asynchronously sent to a GPT-4 API endpoint. We use a structured zero-shot prompt to instruct the model to extract at most 3 semantically meaningful, design-relevant keywords. The call is made with temperature set to 0.3 to promote deterministic and focused outputs. The model is accessed via an internal wrapper that includes automatic response validation and duplicate filtering.

(4) **Annotation Display.** Extracted keywords are rendered as draggable DOM elements (bubbles) in a designated UI tray below the transcript window. Each bubble has a TTL (time-to-live) of 10 seconds, after which it fades out unless interacted with. Users can drag a bubble to any position on the 2D design canvas, triggering a ‘drop‘ event that stores the annotation’s content, (x, y) coordinate, and timestamp in the backend database.

Prompt Engineering for Annotation Generation. We use the following system prompt and query structure to ensure concise and domain-relevant outputs:

```
SYSTEM: You are a helpful assistant in a collaborative design tool.
Given a sentence from a design discussion, identify up to 3 short
keywords or phrases that describe visual intent, style suggestions,
or design changes. Be concise. Only return a comma-separated list of
meaningful keywords. If nothing useful can be extracted, return an
empty list.

USER: "We should add more contrast to the background and maybe simplify
the text layout."
ASSISTANT: contrast, simplify text
```

The model is queried via OpenAI’s API using the `GPT-4` engine.

Associating Annotations with Canvas Elements. The system uses a lightweight spatial mapping approach to associate annotations with specific design content, without requiring an explicit object hierarchy. The association mechanism works as follows:

- Each design canvas is a 2D coordinate plane, rendered as an HTML5 `<canvas>` element. Design strokes and annotation bubbles are both layered on the same coordinate system.
- When a user drags a keyword bubble onto the canvas, a ‘drop‘ event is triggered. This records the current (x, y) position of the annotation relative to the canvas dimensions.
- The annotation object is stored as a JSON blob containing:
  - `content`: extracted keyword or phrase,
  - `position`: `(x, y)` coordinate,
  - `timestamp`: the associated utterance time,
  - `canvasVersionID`: the current version of the canvas.
- Although there is no explicit binding between an annotation and a semantic object (e.g., a specific stroke), the spatial locality allows users to implicitly encode this association by placement. In the UI, hovering over or clicking an annotation highlights its local region for visual clarity.

This design choice provides flexibility while avoiding the need for object-level annotation binding, making the system more compatible with freeform sketches or canvas drawings without structured metadata.

C.2 Annotation-based Rationale Recall

To help users revisit the reasoning behind past annotations, we implement a rationale recall feature based on timestamp alignment. When a user selects an annotation, the system queries the meeting transcript using the annotation's associated `timestamp` to extract the relevant conversational context.

This context window—typically a few utterances before and after the annotation time—is then passed to GPT-4 via OpenAI's API, using a system prompt designed to extract concise rationale statements. An example prompt is as follows:

```
SYSTEM: You are a helpful assistant embedded in a collaborative design tool. Given a user-created annotation and its surrounding context from a collaborative design session, your task is to generate one concise sentence that explains the design rationale behind the annotation—that is, why this design suggestion or decision was made. You will be provided with: The annotation itself; The transcript excerpt surrounding the time the annotation was created (local context); The full transcript of the session (global context), for broader understanding if needed. Focus on the reasoning or motivation behind the annotation. Your output should help someone quickly understand the intent of the design decision without reading the entire conversation. Only output one sentence.

USER: {Full Transcript} : {Annotation} : {Transcript excerpt around timestamp}

ASSISTANT: {One-sentence rationale}
```

The resulting rationale is displayed alongside the annotation in the interface, allowing users to quickly understand or recall the original design motivation. This approach supports fluid traceability without requiring users to manually write or tag rationale during live collaboration.

D Demographic Information

Table 1 shows the demographic information of participants in our study.

E DesignMemo Examples

Figure 11 shows some design examples of DesignMemo.

ID	Age	Gender	Profession	Experience (Design)
P1	21-30	M	VR Researcher, Game Developer	5-10 years
P2	21-30	W	Web Engineer, UI/UX Designer	2-5 years
P3	30-40	M	Web Engineer	2-5 years
P4	21-30	W	Artist, Animator	5-10 years
P5	21-30	W	UI/UX Designer, Sketch	2-5 years
P6	21-30	M	AI Researcher, VR Researcher	2-5 years
P7	30-40	M	Artist, Sculptor	10-15 years
P8	30-40	W	Artist	10-15 years
P9	30-40	M	Architect	10-15 years
P10	21-30	M	AI Researcher	2-5 years
P11	21-30	M	Game Developer, Web Engineer	5-10 years
P12	21-30	M	AI Researcher, Web Engineer	2-5 years
P13	21-30	W	Photographer, Art Director	5-10 years
P14	21-30	W	3D Designer	5-10 years
P15	31-40	M	Sketch, Illustrator	10-15 years
P16	21-30	M	Product Manager	5-10 years
P17	21-30	M	AI Researcher	5-10 years
P18	21-30	W	UI/UX Designer	2-5 years
P19	21-30	W	Studio Designer, Photographer	5-10 years
P20	21-30	W	Game Designer	5-10 years
P21	21-30	W	Designer, Technical Artist	5-10 years
P22	21-30	M	AI Researcher	0-2 years
P23	21-30	M	Web Engineer	2-5 years
P24	21-30	W	Product Manager	2-5 years

Table 1. The demographic information of participants in our study

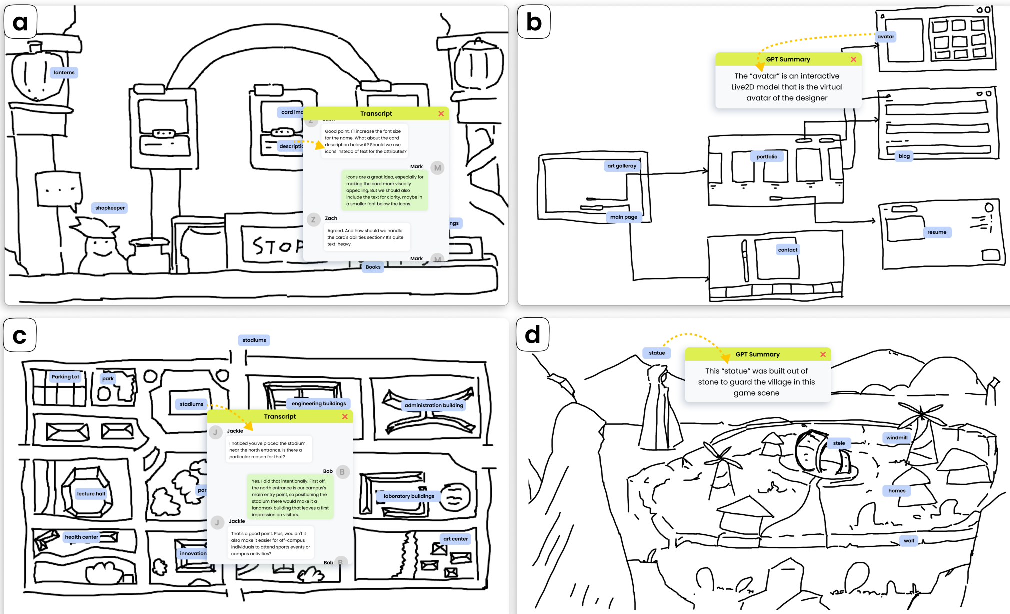


Fig. 11. Design examples using DesignMemo. (a) 2D Store page for the card game. (b) Designer website. (c) Campus plan. (d) 3D village scene for an RPG game.