InsightBridge: Enhancing Empathizing with Users through Real-Time Information Synthesis and Visual Communication

Junze Li
The Hong Kong University of Science and Technology
Hong Kong, China
jlijj@connect.ust.hk

Yue Zhang*
Shenzhen University
Shenzhen, China
zhangyue2020@email.szu.edu.cn

Chengbo Zheng
The Hong Kong University of Science and Technology
Hong Kong, China
cb.zheng@connect.ust.hk

Dingdong Liu
The Hong Kong University of Science and Technology
Hong Kong, China
dliuak@connect.ust.hk

Zeyu Huang
The Hong Kong University of Science and Technology
Hong Kong, China
zhuangbi@connect.ust.hk

Xiaojuan Ma†
The Hong Kong University of Science and Technology
Hong Kong, China
mxj@cse.ust.hk

Abstract

User-centered design necessitates researchers deeply understanding target users throughout the design process. However, during early-stage user interviews, researchers may misinterpret users due to time constraints, incorrect assumptions, and communication barriers. To address this challenge, we introduce InsightBridge, a tool that supports real-time, AI-assisted information synthesis and visual-based verification. InsightBridge automatically organizes relevant information from ongoing interview conversations into an empathy map. It further allows researchers to specify elements to generate visual abstracts depicting the selected information, and then review these visuals with users to refine the visuals as needed. We evaluated the effectiveness of InsightBridge through a within-subject study (N=32) from both the researchers' and users' perspectives. Our findings indicate that InsightBridge can assist researchers in note-taking and organization, as well as in-time visual checking, thereby enhancing mutual understanding with users. Additionally, users' discussions of visuals prompt them to recall overlooked details and scenarios, leading to more insightful ideas.

CCS Concepts

• **Human-centered computing → User interface toolkits**; • **Computing methodologies → Artificial intelligence**.

Keywords

User-centered Design, Human-AI Collaboration, Empathy Map, Visual Communication

*This work was done during Yue Zhang's internship at the Hong Kong University of Science and Technology.
†Corresponding author

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713640

ACM Reference Format:
Junze Li, Yue Zhang, Chengbo Zheng, Dingdong Liu, Zeyu Huang, and Xiaojuan Ma. 2025. InsightBridge: Enhancing Empathizing with Users through Real-Time Information Synthesis and Visual Communication. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3706598.3713640

1 Introduction

User-centered design (UCD) is an iterative, empathetic, and contextualized design process that places the needs, preferences, and requirements of target users at the forefront in every stage of the design and development lifecycle [1, 22, 46, 71]. The typical steps of the UCD process include understanding the user context, specifying user requirements, ideating and prototyping potential design solutions, and continuously evaluating and refining these solutions based on user feedback [9, 22, 37]. This process is often resource-intensive, necessitating significant time, effort, and financial investment [46, 71]. If designers fail to demystify diverse user contexts and needs through thorough and comprehensive user research, especially in domains where the design scenarios are highly personalized, their subsequent design solutions can lead to poor user satisfaction, decreased usability, and ultimately, higher costs due to redesigns and lost business opportunities. UCD practitioners thus employ various methods to help externalize and articulate knowledge gathered about a particular user or user group and remove bias in the process. Empathy mapping [49, 53, 66] is one such method widely adopted from the early stages of UCD.

UCD practitioners typically conduct empathy mapping by synthesizing their knowledge of users into a visual map [28]. Such knowledge could be in the form of user data collected from ongoing activities such as interviews and workshops or practitioners' understanding of users [22, 36, 40]. It is typically represented as sticky notes or digital equivalents, spatially arranged on a canvas according to pre-defined key themes, such as user thoughts, feelings, behaviors, and pain points [28, 49, 66]. In this way, the resulting empathy map intuitively visualizes the user perspective and supports practitioners in immersing into the users' view as well as uncovering the design gaps [15, 40, 77]. However, developing deep empathy by creating and communicating with an empathy map is not easy, especially during an interaction with target users. Designers may fail to capture all relevant details and/or misinterpret users' expressions and needs when summarizing them into sticky notes organized on the empathy map, potentially due to biases, incorrect assumptions, or miscommunication during user interviews [2, 19, 69, 70, 78]. This can result in an empathy map that superficially reflects the true user experience, leading to design decisions that are misaligned with the actual user needs. A typical solution to such an issue is to re-engage with the target users in later design iterations. For example, Davidoff et al. proposed to conduct speed dating sessions with intended users, during which designers present visual artifacts such as storyboards that illustrate their understanding of usage scenarios and design concepts to gather immediate feedback [22]. Despite the effectiveness of such a re-engagement process in deepening the practitioner's understanding of users, this approach requires extra time and effort to develop the design concepts and craft storyboards. If the interpretation of user needs is wrong in the first place, significant efforts would be wasted [20, 51, 60]. Thus, means to facilitate UCD practitioners' empathy-building with users while creating an empathy map are still desired.

In this paper, we introduce InsightBridge, an AI-empowered tool that augments the process of empathizing with users during interviews by real-time, automatic creation of single-user empathy maps and generation of line drawing style visual abstracts of user context to facilitate mutual understanding as shown effective in prior work [32, 59]. To be more specific, InsightBridge can process conversational speech between researchers[1] and users as their interview goes on, extract and organize information relevant to the themes on a given empathy map template into sticky notes. Then, researchers could select the sticky notes that need further confirmation or elaboration from users, and InsightBridge will generate a visual abstract depicting the corresponding contextual information for users to review. InsightBridge can update and refine the visual abstracts instantly based on users' feedback to support further discussions.

We integrate InsightBridge into the Miro platform[2], which is widely used by UX researchers to collect and brainstorm design ideas. To evaluate its efficacy on enhancing researchers' empathy with users, this work aims to explore the following two research questions through a comparative user study:

RQ1 How well can InsightBridge assist researchers in synthesizing information from users' responses into the empathy map and generating contextualized visual abstracts for validation and communication? Whether and how does each feature of InsightBridge support these functions?

RQ2 How well can researchers trigger users to provide more detailed and insightful user context by leveraging each feature of InsightBridge to facilitate mutual understanding in empathy development? How do they use each feature of InsightBridge?

To answer these questions, we recruited 16 pairs of researchers and users to conduct regular user interviews. The study was carried out in a within-subject manner, in which each researcher interviewed with paired user with and without the assistance of InsightBridge regarding two design topics respectively. We counterbalanced the order of these two conditions and the two design topics by using Latin Square [31]. Based on the analysis of study outcomes, we demonstrated that InsightBridge can synthesize multifaceted information into the empathy map reliably and accurately, and generate contextualized visual abstracts efficiently. Meanwhile, the process of visual checking and adjustment trigger users to recall overlooked details, provide more insightful feedback, and further facilitate researchers' understanding and empathizing with users.

The main contribution of this work are threefold:

- InsightBridge, a novel system to enhance researcher's empathy with target users in the user interview.
- An experiment discovering the efficacy of InsightBridge in real-time information synthesis and visual communication to strength the mutual understanding between researchers and user.
- Design implications on how to extend the use of InsightBridge and how to involve convergent and divergent thinking into the development of empathy-enhancing tools.

2 Related Work

2.1 Empathy in User-Centered Design

UCD entails centering the design process around the user needs, which generally starts from "knowing the users" and requires empathy of designers towards the target users [1, 77]. Design researchers have proposed various methods to facilitate researchers' empathic analysis of users [40, 77]. Wright and McCarthy [77] highlight three groups of empathic techniques. The first group of methods involves researchers having direct contact with the users. Researchers might adopt ethnographic methods, immerse themselves into the detailed contexts of the users, and perform observations to form an understanding of users [72, 77]. Conducting interviews with the users is another direct-contact method to promote empathy [55].

The second group of methods is creating a narrative around the target user group [77] (or communication techniques [40]). For example, researchers can create a persona to characterize the user group, specifying their goals, needs, and encountered challenges [8, 18]; or they can opt for scenario-based design [15], creating authentic usage scenarios to inspire the inquiry of user needs. The creation of such narratives is often based on user raw data or can be from direct contact with users. For example, the empathy map, characterizing users' thoughts, feelings, goals, etc. spatially on canvas, typically results from a synthesis of researchers' interviews with the users [49, 53, 66]. Another group of methods facilitates researchers' imagination of the users through, for example, role-playing. Experience prototyping introduced by Buchenau and Suri [12] is a typical method of this group, which promotes researchers' empathy by situating researchers themselves as the target users.

It is worth noting that promoting empathy does not mean "becoming the user", but refers to understanding the user from researchers' own perspective [40, 77]. Kouprie and Visser [40] frame the empathy process into a four-phrase model, including discovery, immersion, connection, and detachment. While the first two phases focus on embracing the users' perspectives, the latter two phases require researchers to reflect upon their own experiences to develop

[1] In the following text, we use the term "researcher" to refer to researchers who conduct user-centered design.
[2] https://miro.com/

novel design knowledge and further guide design concept formulation. In this process, the divergence of users' and researchers' views could lead to challenges in empathy, such as loss of the nuances of user experience [55] and incorrect assumptions by researchers [2]. To this end, this paper aims to bridge this gap by leveraging AI to generate boundary objects (*i.e.,* visual abstracts) to smooth the synthesis and development of both researchers' and users' perspectives during the interview.

2.2 Support Communication beyond Words

Speech is the main modality in human-human conversations. However, speech alone may limit the transmission of information and understanding of each other [32, 59]. One of the most famous theories in the field of interpersonal communication, the media richness theory [23], suggests that the richness of the medium's message affects the effectiveness of communication. Therefore, researchers tried to adopt diverse materials and devices to support communication in different scenarios. For example, the card is a common medium to foster communication between designers and target users [75]. Beck *et al.* proposed the instant card technique for UCD which enables formalized development of scenarios with target users by editing and modifying instant cards [3]. Curtis *et al.* attempted to use smartwatches to support the communication needs of people with language impairment aphasia [21]. Wood *et al.* conducted robot-mediated interviews to investigate the effect of humanoid robot for supporting interview conversations [76]. Efficient note-taking assistance can also support communication. Haliburton *et al.* investigated the use of microphone, walking stick, and highlighting buttons to help people take notes during walking meetings [33]. Among the vast possibilities beyond words, the most prevalent choice of modality is visual. It is supported by the dual coding theory [54], which suggests that there are two basic channels through which humans process information: the verbal channel and the visual channel. Through the combination of both verbal and visual information, people can speed up the processing of information and perceive the information more sufficiently and thoroughly [54]. Existing research has extensively explored the effectiveness of visuals in supporting communication. The Visual Captions system developed by Liu *et al.* helps audiences understand and remember complex or unfamiliar concepts mentioned by the speakers [45]. It anticipates the "visual intent" of a conversation and proactively suggests appropriate visual materials in real time. The effectiveness of communication is especially important in an educational setting because it can directly affect the effectiveness of learning. In teaching tool operation skills, visual communication allows students to better understand the instructor's instructions and more accurately replicate actions compared to verbal communication [65]. Visual's supportive role in communication is even hard to replace on some special occasions. For people with language barriers, like nonverbal autistic individuals, visual aids can be adjusted in elements like button size and color to match their motor and cognitive skills, making communication easier for users with varying needs [24]. The findings and design considerations in these previous work inspire and support the design of InsightBridge's real-time information synthesis (verbal channel) and visual communication (visual channel) functions.

2.3 Support Communication with AI

As an emerging field, a growing body of research is dedicated to exploring the potential of AI to support interpersonal communication and improve the quality of communication. Existing research suggests that AI has demonstrated positive effects in facilitating communication in different scenarios and among different groups [43, 44, 63, 81].

AI can help reduce the barriers of communication among people with different backgrounds and contexts. To combat generation gaps, Liu *et al.* leveraged generative AI to propose more expressive, creative, and efficient family communication through arts therapy [44]. Similarly, LegacySphere promotes more open and inclusive dialog on life philosophies with AI-powered avatars, role-playing, and storytelling designs [63]. In healthcare, a field with substantial communication conflicts, AI interventions can help caregivers and beneficiaries communicate their feelings and reach mutual understanding [61]. AI's automated symptom monitoring and conversational agent system can also help cancer patients communicate with their doctors on a regular basis, which reduces lags and barriers to information delivery [79].

User research is another area that desires communication support for researchers and users, but little research has been done to apply the surging capability of generative AI to facilitating communication therein. Therefore, InsightBridge is developed to facilitate user interviews during the UCD process. To keep a natural communication environment without a high burden of additional devices, *e.g.,* VR devices in LegacySphere [63], InsightBridge is integrated into the Miro whiteboard platform which is widely used by UCD researchers [25, 42, 47]. Meanwhile, to keep the fluency of on-going interviews, InsightBridge provides real-time functions without long delay and specific preparation of AI-generated contents [44, 63]. The flexible combination of the sticky notes in the empathy map and the visual abstracts in the visual area of InsightBridge enhances the verbal and visual communication channels at the same time, which is essential of interpersonal communication but not realized in previous work [54, 61, 79]. The APIs provided by OpenAI can process and generate multi-modal data. However, it is only developed for humans communicating and interacting with AI. In our case, InsightBridge is a tool designed for supporting human-human communication by leveraging OpenAI APIs with special mechanisms and design considerations.

3 System Design

In this section, we introduce our InsightBridge system, which leverages Large Language Models (LLMs) to generate visuals that facilitate communication and understanding between researchers and users during the completion of empathy maps. InsightBridge is a plugin developed in Miro, a widely-used diagramming whiteboard platform for empathy-building, due to the quick note-taking and ease of information organization [25, 42, 47]. Miro enables collaboration through elements like text, sticky notes, and lines. Its collaborative capability, developer-friendly architecture, and rich set of elements and templates (including empathy map templates) make it the ideal platform for developing InsightBridge.

Referring to previous research that shows real-time generation of dialog-related visual content can enhance communication [45],

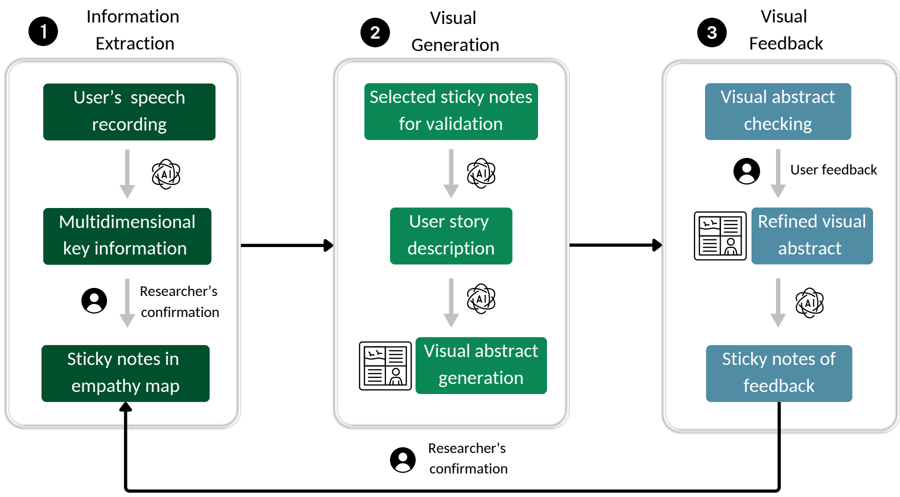


Figure 1: The workflow of InsightBridge. We leverage LLMs to support the three main functions of InsightBridge, including information extraction, visual generation, and visual feedback. The actions done by human or AI are annotated respectively.

we incorporate the generation and modification of visual abstracts into the workflow of user interview and implement in InsightBridge (Figure 1). We expect the generated visuals to serve as key boundary objects [67], that aggregate perspectives from both researchers and users, supporting focused discussion around them, and finally augment and facilitate researcher-user communication. Specifically, InsightBridge first helps researchers extract and organize key points from users' responses by generating sticky notes (Figure 1, 1). When researchers identify key points they want to visualize, InsightBridge enters the second stage: generating a visual abstract (Figure 1, 2). After the visual is generated, InsightBridge gathers user feedback to create new sticky notes and visual abstracts (Figure 1, 3). These three stages cycle until the interview is finalized. Finally, InsightBridge analyzes pain points and gain points from all sticky notes and conversation history, again in the form of generating sticky notes.

3.1 Information Extraction

During typical user interviews, researchers must carefully listen to user responses while quickly identifying and recording key information. Following this, a best practice is to structure the collected data into a summative representation (*e.g.,* an empathy map) to validate insights with users and identify design opportunities [7, 41]. However, manually noting down key information is already labor-intensive, adding further complexity to the subsequent synthesis and re-engagement with users during the interview. To address these challenges and enhance communication with timely user feedback, InsightBridge automatically extracts and organizes key points from user responses, streamlining the process and reducing the cognitive load on researchers.

When the user begins answering questions, the researcher presses the **"Record + Generate Sticky Notes"** button to start recording. The system automatically records the response and transcribes audio files to text using OpenAI's API[3]. Then these texts are fed as part of the prompts to ChatGPT 4o model, which has been provided with the contexts on each dimensions of the empathy map. For example, when the user says, "I am worried about whether remote toys might have some sharp edges, as these sharp edges could harm my cat", InsightBridge generates a sticky note recording the user's feeling of worry and categorizes it under the *Feels* dimension. This categorization can be expressed as $d_f = C(K)$, where d_f stands for the *Feels* dimension, C represents categorization, and K represents the key points.

InsightBridge is highly flexible because the content of the generated sticky notes is fully editable, allowing researchers to adjust as needed. Researchers can also delete unnecessary sticky notes, or if none of the generated notes are satisfactory, they can add an empty one from Miro's toolbar. In other words, the sticky notes are content-editable, removable, and addible, providing researchers with ample flexibility to collaborate with the AI.

3.2 Visual Abstract Generation

The key enhancement InsightBridge brings to researcher-user communication is the real-time generation of highly relevant and thought-provoking visual abstracts during the interview. Leveraging the text-to-image capabilities of LLMs, InsightBridge creates visuals based on the sticky notes from the synthesized empathy map (Figure 1, 1). Notably, InsightBridge does not generate images from all the sticky notes but instead requires researchers to select specific combinations. The timing of visual generation is also decided by researchers based on the progress of interviews, to avoid their distraction from irrelevant and unimportant visual abstracts. While the selection and decision process may add some cognitive load, we argue that it is essential as it allows researchers to emphasize their priorities, ensuring the generated images are more aligned with their interests – an important factor in the empathy process [40, 77].

InsightBridge uses the DALL·E 3 model [58] to generate images from key points in selected sticky notes. While DALL·E 3 produces

[3]https://platform.openai.com/docs/guides/speech-to-text

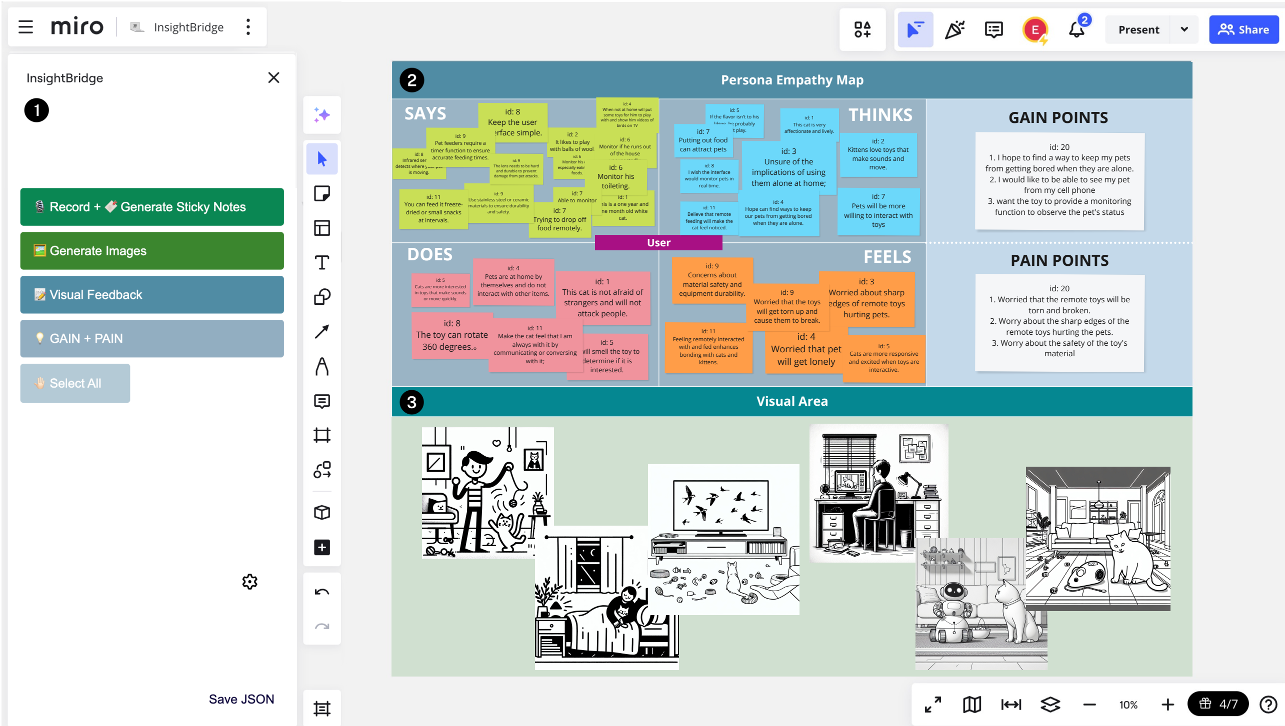


Figure 2: The user interface of InsightBridge, which includes three main views: manipulation panel (1), persona empathy map area (2) for information synthesis, and visual area (3) for visual communication. Some sample sticky notes with key information and generated visual abstracts are provided in the corresponding area.

high-fidelity, detail-rich images that enhance visual appeal, such images might consist of excessive details that diverge from the user's real experience and backfire to hinder user immersion. To address this, our prompting strategy focuses on generating minimalist-style images, primarily as black-and-white line drawings, ensuring the visuals remain abstract and generalized. This process can be summarized as $V_s = D(K, S)$, where V_s is the visual abstract, D is the DALL·E 3 model, and S defines the style which is black-and-white, line-drawing-like, and minimalistic. For example, when a researcher selects a sticky note with the content "My kitten likes to play with plush toys" that needs alignment and presses the button of **"Generate Images"**, InsightBridge produces a simple black-and-white sketch resembling a comic strip, using clean lines to depict the kitten and the plush toy. This minimalist style allows users to build their own mental context based on the visual.

Moreover, our prompting strategy employs the chain-of-thought (CoT) technique [74] by instructing the LLMs to first describe the user story with concrete scenes and relevant elements based on the selected key points, then generate the visual abstracts from the user story. The rationale behind this CoT design is to mimic and align with the typical approach researchers use to create narratives (*e.g.,* scenarios, storyboards) from user data [15, 22, 41, 77]. We anticipate that the narrative-centered images produced through this prompting strategy will stimulate users with concrete scenes, encouraging them to provide more meaningful feedback.

3.3 Visual Feedback

Even if researchers can ask follow-up questions to confirm with users, it is challenging to ensure that researchers and users fully align in their understanding of specific concepts and scenarios [2, 48]. To minimize misunderstandings, InsightBridge incorporates users in the creation of visual abstracts by encouraging them to provide feedback. After showing the visual abstract to users, researchers can press the **"Visual Feedback"** button to immediately collect user feedback on the visual abstract. InsightBridge transcribes the user's feedback into text, which is simultaneously used for generating new sticky notes and updating the visual abstract. For example, in Figure 2, when discussing pet toys, the visual abstract initially depicted a robot-shaped interactive toy based on the user's input. After checking it, the user suggested that "the toy would be better shaped like a mouse, as it would attract the cat's interest more." In response, InsightBridge generated a sticky note with the suggestion, "I wish the toy was in the shape of a mouse" and updated the visual abstract by changing the toy's shape from a robot to a mouse. The visual abstract provides a more intuitive way of presenting the content of use than words. It acts as a bridge, allowing researchers and users to quickly compare it with their own internal understanding, guiding them towards a shared alignment.

3.4 Interaction Flow

This section introduces the interaction flow of InsightBridge by demonstrating how it helps researchers organize user interviews. Alice, a user experience researcher, aims to understand pet owners' opinions and needs regarding pet toys for remote interaction. She finds a target user and begins the interview.

3.4.1 Recording and organizing the user's responses. When the user starts answering a question, Alice presses the **"Record + Generate Sticky Notes"** button, and InsightBridge automatically records the response. Once the user finishes, Alice presses the button again to stop recording, and InsightBridge generates a sticky note containing the key points of the user's response. The sticky note is then automatically categorized into the relevant frames corresponding to the four dimensions of the empathy map.

3.4.2 Generating and modifying visual abstract. After asking a few questions, Alice notices three sticky notes capturing the user's product requirements: "want a pet toy that interacts with cats", "want to be able to control it remotely", and "want to be able to walk around". Although Alice has a general idea of what the toy might look like, she is unsure how closely her vision aligns with the user's. She selects these three sticky notes (or, if needed, clicks the **"Select All"** button to select all the sticky notes) and then clicks **"Generate Images"** to create a visual abstract. InsightBridge generates a visual in the Visual Area, depicting a white cat and a robot-shaped interactive toy. When the user begins commenting on the visual abstract and suggesting modifications, Alice presses the **"Visual Feedback"** button to record the feedback. Once the user finishes, Alice presses the button again, and InsightBridge generates feedback sticky notes, categorizes them into the empathy map, and updates the visual abstract based on the feedback.

3.4.3 Analyzing the interview. After completing all the questions on the interview outline, Alice faces an abundance of key points. She wants to quickly analyze the gain points and pain points in order to confirm them with the user on the spot. She clicks the **"GAIN + PAIN"** button. InsightBridge selects all the sticky notes and analyzes them to extract gain points and pain points. The points are recorded on sticky notes, which are pasted in the corresponding areas of the empathy map.

4 Evaluation

To evaluate the efficiency and usefulness of InsightBridge , we recruited participants to conduct research-oriented interviews in a within-subject design. In this section, we introduce the details of experiment procedures and the corresponding measurements.

4.1 Study Procedures

In each session of our study, we invited one participant as the researcher and the other as the user to conduct a user interview. Following the requirements of our institution's IRB, we recruited 16 participants as researchers (7 males, 7 females, and 2 non-binary/third gender; age range 19-27, *M*=21.63, *S.D.*=1.93). These researchers have backgrounds in Human-Computer Interaction, UI/UX, and Product Design, and possess abundant experience in communicating and empathizing with target users during the UCD process. For the Baseline condition comparing with the use of InsightBridge, we allow researchers to choose the tool they are most familiar with, including paper and pen, tablet, or laptop. We install the original Miro software on the tablet and laptop. Both the paper and electronic empathy map template are provided. Researchers can decide whether and how to take notes or sketch drafts based on their common practice. In this within-subject study, we select the following two design topics for each system condition: 1) Pet toy: "*Designing an Interactive Pet Toy with Remote Control to Facilitate Human-Pet Communication*"; 2) Fitness training: "*Designing an Interactive Fitness Device with Daily Object Recognition to Enhance Human-Environment Interaction for Physical Training*". Since these two topics are highly related to users' personal experiences and preferences, it is important to empathize with users thoroughly as the initial step of the UCD iteration. By adopting *Latin Square*, four combinations are formed to counterbalance the experiment:

- Pet toy (Baseline) - Fitness training (InsightBridge)
- Fitness training (Baseline) - Pet toy (InsightBridge)
- Pet toy (InsightBridge) - Fitness training (Baseline)
- Fitness training (InsightBridge) - Pet toy (Baseline)

To pair with each researcher, we recruited 16 participants as target users (5 males, 11 females; age range 21-31, *M*=25.25, *S.D.*=2.88) for the user interviews. All the users raise pets and have fitness training habits.

During the onboarding phase of our study, we introduce the study workflow to both the researchers and users. Specifically, we provide researchers with a tutorial to familiarize them with InsightBridge and its functions. To reduce the burden on researchers, keep a smooth interview process and control the confounding factors, we set up a list of 10 questions regarding each design topic for researchers' reference. The aim and focus of questions are: 1) Q1-Q2: warm up and collect background information; 2) Q3-Q5: inquire about current practice and challenges; 3) Q6-Q10: collect user needs and requirements. Researchers can flexibly modify or add follow-up questions based on their pace and interview habits. After each researcher-user pair completes an experiment condition, they are required to complete a questionnaire separately. Finally, we conduct semi-structured interviews by asking open-ended questions to researchers and users independently to collect their feedback on their interview experiences.

4.2 Measurements

We evaluate the performance of InsightBridge by analyzing participants' self-ratings in the questionnaires and their feedback to the final semi-structured interviews. All questions in the questionnaire are rated with a 7-point Likert scale, with 1 denoting strongly disagree and 7 denoting strongly agree.

4.2.1 Measurements for Researchers. In the questionnaires, we focus on the following four dimensions:

Task Load. Conducting user interviews is a complex process that requires researchers to pay attention to the ongoing discussion as well as user emotions and behavior while ensuring the fluency and accuracy of question-asking. We hypothesize that InsightBridge 's features, such as information extraction and visual abstract generation, could potentially help with the interview process and lower the task load. We use the NASA Task Load Index [35]. It includes six dimensions: 1) mental demand; 2) physical demand; 3) temporal demand; 4) task completeness; 5) effort to complete the task; and 6) any frustration that occurred.

System Usability. We would like to explore the usability of the features in InsightBridge in real interview contexts. We adopt the standard System Usability Scale (SUS) tool to measure the system usability in five aspects [11]: 1) easy to learn; 2) easy to use; 3) compatible with the ongoing task; 4) overall usefulness; 5) future use.

Researchers' View of Mutual Understanding. The alignment between researchers' understanding and users' actual thoughts and feelings is a crucial factor for empathizing with users. We leverage measurements of mutual understanding from previous work to propose five aspects in our study [34, 80]: 1) the effectiveness of communication with users, 2) the ease of asking users questions, 3) the ability of users to provide insightful feedback, and 4) the accuracy of understanding user needs.

Assessment of InsightBridge 's Features The above three dimensions are measured by researchers using both the Baseline system and InsightBridge for statistical comparison. Regarding the specific functions of InsightBridge , we evaluate the quality of the following five features: 1) extracting key information; 2) organizing key information; 3) checking sticky notes carefully; 4) generation of visual abstracts; 5) adjustment of visual abstracts.

4.2.2 Measurements for Users. Since most system interactions occur on the researcher side, the measurements of the users' experience are mainly their self-report mutual understanding during the interview in the following five dimensions: 1) the effectiveness of communication with researchers, 2) the ease of answering questions, 3) the comfort in sharing ideas, 4) the impact of researchers' operations on the device on communication, and 5) the accuracy of understanding users' needs.

4.2.3 Semi-structured Interviews. For the semi-structured interviews, we prepared open-ended questions to assess researchers' views on InsightBridge 's performance (*e.g.,* "Did you find many manual edits necessary for the automatic note-taking and visual-checking results? Why?"), its effectiveness in eliciting user feedback (*e.g.,* "Was user feedback on the generated visuals insightful? Why?"), and how it supports a better understanding of users (*e.g.,* "Did any new understanding about users emerge from the interviews?"). Similar questions were posed to users, with language tailored to their perspective (*e.g.,* "What are your thoughts on the generated visuals?" in relation to InsightBridge 's performance).

We conducted a thematic analysis [10] of the semi-structured interview data. Two authors independently reviewed the interview recordings and generated initial codes. They met frequently to discuss and resolve any discrepancies, leading to the formation of high-level themes. We triangulated the qualitative analysis with participants' self-rating scores to strengthen the qualitative analysis quality.

5 Results

In this section, we present our findings on the impacts of implementing InsightBridge in user interviews from the researchers' and users' perspectives respectively.

5.1 RQ1: InsightBridge for Information Synthesis and Visual Communication

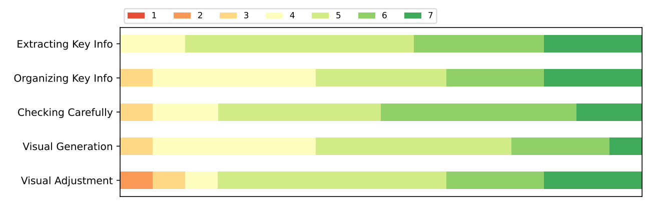


Figure 3: Researchers' self-ratings of the satisfaction degree regarding each key feature of InsightBridge in 7-point Likert scale. Score 7 denotes that researchers strongly agree on the usefulness of each feature, and score 1 denotes that they strongly disagree.

To analyze the effect and impact of each feature of InsightBridge, we let researchers who mainly interact with InsightBridge rate their satisfaction degree of each feature in a 7-point Likert scale, and the result is shown in Figure 3. Overall, the researchers show a positive attitude towards each feature. And in the following subsections, we will represent and interpret detailed findings of whether and how does each feature of InsightBridge support the information synthesis and visual communication functions based on the analysis of interaction data and qualitative feedback.

5.1.1 Efficient and Reliable Information Synthesis. During the researchers' use of InsightBridge, they need to check the automatically generated sticky notes in each dimension of the empathy map and possibly modify these notes when they found any discrepancies with their own understanding. We summarized the number of finalized sticky notes and researchers' modification operations in Table 1. Researchers' self-rating results of the note-taking function are shown in Figure 3. While most researchers confirmed that they checked the generated notes carefully (as shown in Figure 3 "Checking carefully"), we found that they made few modifications in the dimensions of *Does*, *Gains* and *Pains*. This implies that they are generally satisfied with the extracted key information in these three dimensions, which also echos the relatively positive ratings of "Extracting key information" and "Organizing key information" in Figure 3.

Table 1: The note-taking results of each researcher in each quadrant of the empathy map. We summarize the average number of finalized sticky notes, the average number of modifications (editing and deleting), and the corresponding modification rate.

Dimension	Avg. # Sticky Notes	Avg. # Modifications	Modification Rate
Says	14.06	3.00	21.34%
Thinks	15.25	2.38	15.61%
Does	9.69	0.25	2.58%
Feels	5.44	0.81	14.89%
Gains	3.44	-	-
Pains	4.19	-	-

On the other hand, for the dimensions of *Thinks* and *Feels*, researchers have modified around 15% of the generated sticky notes. The key points in *Says* were modified most, which accounts for 21.34% of all the sticky notes. Interestingly, researchers did not consider the real-time modification a huge burden during the ongoing interview. For instance, researcher R6 mentioned that "*Even if there are errors [in the generated sticky notes], they are not the kind of errors in the content that are misleading, but there may be some pronunciation problems that lead to inaccurate voice recognition, which is easy to pick out at a glance*". Meanwhile, the automatic information extraction and organization features process all the conversations between researchers and users, which prevents researchers from omitting key information in particular aspects. For example, researcher R3 mentioned that "*Considering my personal note-taking preferences, I tend to pay more attention to the dimension of Thinks, and sometimes ignore the pieces of information in Feels dimension. But this tool [InsightBridge] can remind me that users' expression of emotions and feelings is equally important*".

5.1.2 Visual Checking Behavior. All the participating researchers tried the visual generation and adjustment features during their implementation of InsightBridge. On average, each researcher chose to generate 6.25 visual abstracts, and for each checkpoint, they generated 2.33 visuals. We also analyzed the temporal patterns of researchers' visual checking behavior by tracking when they attempted to generate visual abstracts and how many visuals were generated in each phase of the interview. As shown in Figure 4, we discovered that most researchers preferred to generate visual abstracts and check with users during the phase of asking and collecting user needs, especially after finishing all the interview questions. Researcher R4 said that "*If the user mentions a more scenario-based concept or a more difficult-to-imagine scenario, it will make me want to generate a visual abstract. In addition, after collecting enough information, I choose to make a final confirmation*".

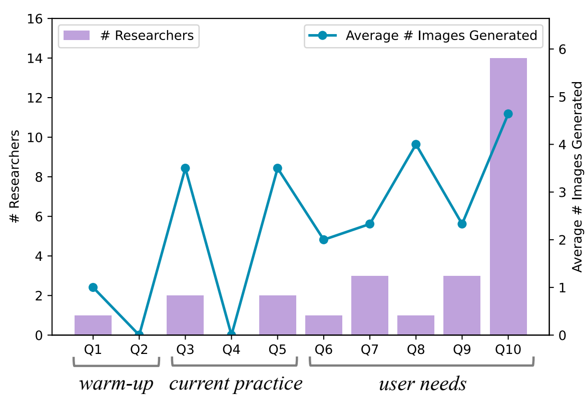


Figure 4: The timing of researchers to do visual checking, and the number of generated visuals at each checkpoint. We track the interview process through three phases, *warm-up*, *current practice* and *user needs*.

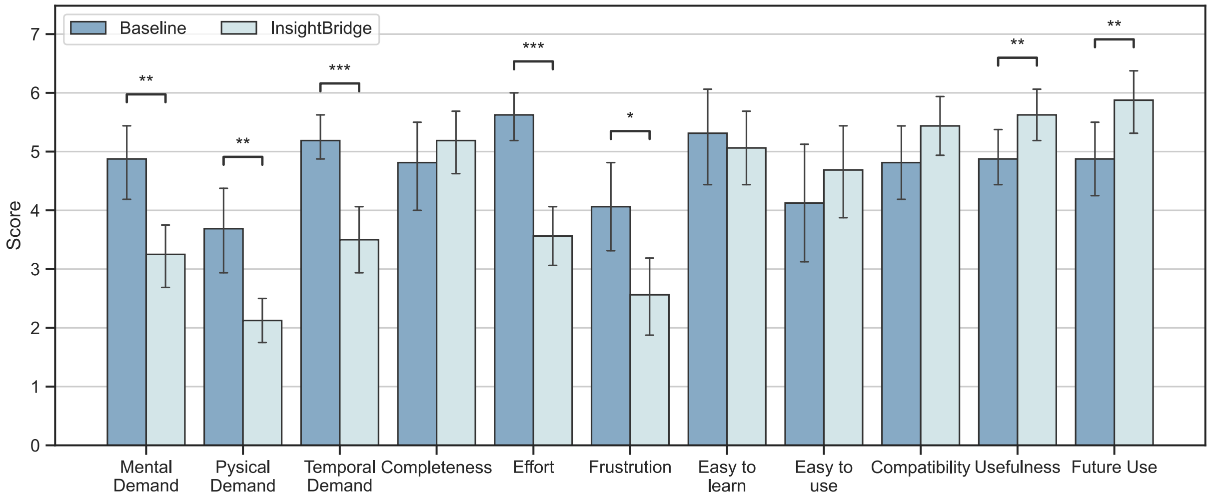
Based on the researchers' satisfaction degree of visual generation and adjustment features shown in Figure 3, we found that most researchers confirmed the usefulness of these visual-related features for visual communication. It can broaden and enrich the available medium when researchers communicate with users. As researcher R7 shared, "*Originally, I am not very good at sketching and hand-drawing, so I rarely try the visual-based method to communicate with users. However, this tool [InsightBridge] allows me to overcome the limitations of my personal ability and use more diverse mediums to promote our communication*". Regarding the timing and procedure to trigger the visual generation feature, researcher R8 commented that "*I think it is a good idea to let us [researchers] first select some of the sticky notes that we want to further confirm and discuss, and then generate the corresponding visuals, since this makes the generated visuals representative and focused, without causing many hallucination issues*". However, researcher R10 held an opposite attitude and said that "*To ensure the smoothness of the interview, sometimes I may not have enough time to select the sticky notes which want to align with users further, especially during the second half of the interview*".

5.1.3 Overall Task Load and Usability Comparison. As described in Section 4.2, the researchers rated the level of task load and overall system usability. Before comparing the task load and usability between the Baseline system and InsightBridge, we firstly performed a Mann-Whitney test to check if researchers' ratings under each design topic have a significant difference. And there is no significant difference ($p > 0.050$ for all the questions) detected, which excludes the design topic as a confounding factor. Then, we conducted the Wilcoxon signed-rank statistical test to explore whether the rating difference between two systems is significant, which is presented in Figure 5.

Regarding the task load, we found that researchers' task load while using InsightBridge is significantly lower than that while using the Baseline system in the dimension of mental demand ($p = 0.002$), physical demand ($p = 0.001$), temporal demand ($p = 0.001$), effort ($p = 0.001$), and frustration ($p = 0.014$). The feature of automatic speech recognition and note-taking is considered as an effective factor to lower the researchers' burden and cognitive load. For instance, researcher R4 commented that "*If I take notes by myself, I need to spend a lot of time typing, especially for recording detailed expressions from users, and I also have to make a lot of repetitive decisions [whether to record or not]. When using this system [InsightBridge], I can save a lot of physical and mental energy, and have more eye contact with users*". For the comparison of system usability, there is no significant difference in the aspects of "easy to learn", "easy to use", and compatibility. It reflects that even though more features are designed in InsightBridge, researchers can still learn and use InsightBridge easily and involve it in the interview process smoothly. Moreover, researchers thought InsightBridge is more useful ($p = 0.009$) and expressed more willingness to use InsightBridge ($p = 0.010$) in the future significantly.

5.2 RQ2: InsightBridge for Facilitating Mutual Understanding and Insightful Feedback

5.2.1 Mutual Understanding. As described in Section 4.2, both the researchers and users rated the diverse aspects of mutual understanding. Similar to the analysis pipeline in Section 5.1.3, we firstly excluded the design topic as a confounding factor after running the Mann-Whitney test. And then the Wilcoxon signed-rank test



Baseline	4.88(1.31)	3.69(1.58)	5.19(0.75)	4.81(1.52)	5.62(0.81)	4.06(1.57)	5.31(1.74)	4.12(2.00)	4.81(1.33)	4.88(1.02)	4.88(1.31)
Insight-Bridge	3.25(1.18)	2.12(0.89)	3.50(1.21)	5.19(1.11)	3.56(1.03)	2.56(1.36)	5.06(1.34)	4.69(1.62)	5.44(1.03)	5.62(0.96)	5.88(1.09)
W	112.00	88.50	102.00	32.00	117.50	88.00	40.50	31.50	18.00	10.00	7.00
p-value	**0.002**	**0.001**	**0.001**	0.301	**0.001**	**0.014**	0.263	0.289	0.089	**0.009**	**0.010**
eff. size	**0.87**	**0.95**	**0.94**	0.18	**0.96**	**0.68**	0.23	0.19	0.45	**0.74**	**0.79**

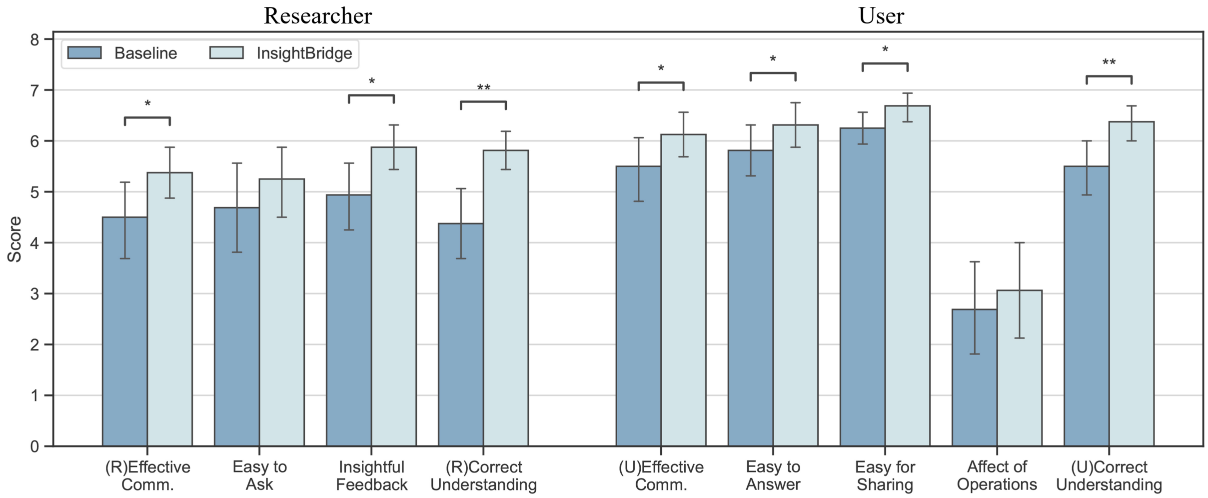
Figure 5: Statistical comparison of researchers' self-ratings of task load and usability between the Baseline system and InsightBridge. Error bars indicate 95% confidence intervals around the mean values. We applied the Wilcoxon signed-rank test, where the statistic value W and p-value (*:p < .050, **:p < .010, ***:p < .001) are reported. The effect size calculated using RBC is also shown.

was conducted to measure the statistical difference of participants' level of mutual understanding when using the Baseline system and InsightBridge as shown in Figure 6.

From the researchers' perspective, they consider using InsightBridge can support higher effectiveness of communication (p = 0.045). Meanwhile, users' feedback are more insightful (p = 0.013) and the understanding of user needs is more accurate (p = 0.006) with the assistance of InsightBridge. Nevertheless, there is no significant difference shown in the aspect of ease of asking questions. In general, InsightBridge can promote researcher's communication with users and empathize with users better through correct understanding of users' expressions and ideas. As researcher R1 mentioned, "*At that time, I did not quite understand what the user was saying [about the potential usage scenario], and my understanding of what he talked about was very vague. However, this tool [InsightBridge] generated a pretty good and contextualized image. Based on the discussion of the generated image, my understanding of what the user was describing would be much clearer*". Moreover, the contextualized environment depicted in the generated visual abstracts can trigger researchers to confirm and discuss with users, which strength the alignment of researchers' perceived scenarios and users' described scenarios. For instance, Figure 7 represents two manually created drafts, which mainly focus on the device without showing the surrounding environments. Some researchers mentioned the importance of confirming user's context. For example, researcher R15 said that "*[Under the design topic of fitness training] I originally assumed that comfortable training only requires enough equipment, but the generated images made me realize that the air conditioning, lighting, and mirrors in the environment are all important. This made me further inquire about the user's needs for these environmental factors which really enhances my understanding*".

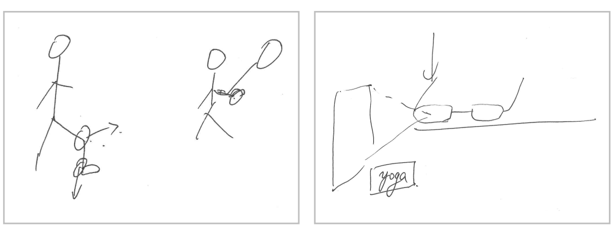
5.2.2 Insightful Feedback. The checking and refinement of generated visual abstracts can facilitate users' more insightful feedback. Based on the analysis of user's feedback and discussions triggered by visual abstracts, we summarize three type of facilitation: contextualization, instantiation, and brainstorming. The representative visual samples under each design topic are shown in Figure 8.

Contextualization The concrete and contextualized scenario presented in the visual abstract serves as a prompt for users to think of the relation between the design target and the potential surrounding people or objects in the environment. For instance, user U7 mentioned that "*The picture shows the environment of a living room, so I thought, besides on the floor, where else can I use such a pet toy? Then there are a lot of clutter on the floor, so I thought that I should set aside an area for the cat to play with the toy without messing up the whole room. Even if it [scenario shown in visual abstracts] may be different from the actual settings in my house, the process of checking it can actually help me come up with more needs*"



	(R)Effective Comm.	Easy to Ask	Insightful Feedback	(R)Correct Understanding	(U)Effective Comm.	Easy to Answer	Easy for Sharing	Affect of Operations	(U)Correct Understanding
Baseline	4.50(1.67)	4.69(1.82)	4.94(1.39)	4.38(1.45)	5.50(1.26)	5.81(1.11)	6.25(0.68)	2.69(1.89)	5.50(1.10)
Insight-Bridge	5.38(1.02)	5.25(1.53)	5.88(0.96)	5.81(0.83)	6.12(0.89)	6.31(0.87)	6.69(0.60)	3.06(2.05)	6.38(0.72)
W	14.00	33.50	8.00	7.00	3.00	3.00	3.00	16.50	7.00
p-value	**0.045**	0.204	**0.013**	**0.006**	**0.019**	**0.033**	**0.032**	0.255	**0.010**
eff. size	**0.58**	0.26	**0.76**	**0.82**	**0.83**	**0.79**	**0.79**	0.27	**0.79**

Figure 6: Statistical comparison of researchers' and users' self-ratings of mutual understanding between the Baseline system and InsightBridge. Error bars indicate 95% confidence intervals around the mean values. We applied the Wilcoxon signed-rank test, where the statistic value W and p-value (*:p < .050, **:p < .010, ***:p < .001) are reported. The effect size calculated using RBC is also shown.



Researcher R10 Researcher R11

Figure 7: Researchers' manually created visual abstracts when using the Baseline system. (left: a protective gear for real-time gesture feedback; right: a pair of glasses to play teaching videos.)

(Figure 8 (2)). Similarly, user U3 recalled that "*The gym I used to go to is always crowded and noisy, so the audio-based feedback from the device may be not applicable*" (Figure 8 (1)). In some cases, the personal traits of the character generated in visual abstracts may not be fully consistent with users'. For example, there was a male athlete generated in the visual abstract for the user U8, but U8 is a female user. Then, U8 said that "*This reminds me that many gym facilities are not female-friendly. For example, the weights of some anaerobic equipment are too high and difficult to adjust them. Also, the size of some equipment is not suitable for females to fix and providing protection*".

Instantiation When users refer to some particular design concepts and ideas, InsightBridge can guide them to instantiate the vague concepts to avoid the superficial expression of user needs and requirements. For example, both the users U3 and U6 mentioned that cell phone can be the possible device to support fitness training and communication with pets. However, they did not provide enough details of how to interact with cell phones before checking the generated visuals (Figure 8 (3) and (4)). Specifically, researcher R6 mentioned that "*Here, a mobile phone is plugged into the robot vacuum cleaner. This seems a bit nonsense at first, but then the user [U6] says that it can be possible. The user considers that she can hold a replica of the phone remotely, and the state of the robot is the same as the state of the phone. In this way, she can check her cat's status and control the robot remotely*". Another example is that the image (3) of Figure 8 triggered the user U6 to think of the difficulty of operating cell phones during exercise. Then the user U6 expressed the need for a bigger interactive screen to display the real-time feedback of the gesture and physical activity data.

Brainstorming Based on the users' personalized information provided and the discussions during the iteration of visual refinements, InsightBridge can facilitate users' divergent thinking to brainstorm some novel applications. In Figure 8, image (5) is an interactive human exoskeleton to monitor the fitness training process, and image (6) is a Transformers toy with remote control for

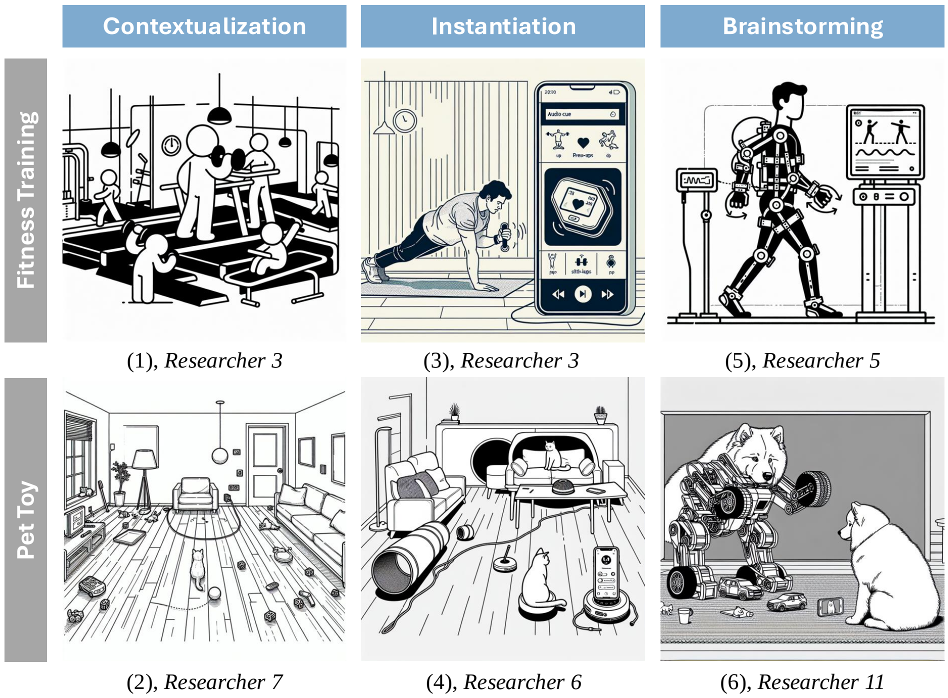


Figure 8: Samples of generated visual abstracts by InsightBridge from each type of facilitation under each design topic which elicit users' deep and insightful feedback. The researcher ID under each visual abstract indicates the corresponding image was generated during the interview conducted by this researcher.

dogs. These brainstorming-oriented generated visuals not only assist researchers in confirming the concrete format of novel concepts but also encourage users to provide more insights. For instance, researcher R5 mentioned that "*When users refer to some creative ideas, the tool [InsightBridge] can generate corresponding images with high-quality. These images can help users further give more rationale and details. It is also a promotion for the subsequent design process*".

6 Discussion

In this work, we aim to design a novel AI-powered system to enhance researchers' understanding of user needs and trigger users to provide insightful feedback during user interviews. Based on the analysis of participants' interaction data, questionnaire ratings, and qualitative feedback in Section 5, we summarized our main findings:

- The key point extraction and organization features can realize reliable information synthesis through prompt and accurate speech understanding, flexible editing, and comprehensive organization.
- The visual generation and adjustment features can support efficient visual communication due to the stable generation constrained by human preferences, and convenient visual refinement mechanisms.
- By leveraging the information synthesis and visual communication functions, InsightBridge can facilitate mutual understanding between researchers and users and trigger users to provide more detailed and insightful user needs.

In this section, we further discuss how well InsightBridge achieves our design goals, the generalizability of InsightBridge, design considerations, limitations, and future work.

6.1 Common Ground Building Based on Visual Abstracts

The visual abstract generated by InsightBridge can establish a common ground between researchers and users, aligning with the Common Ground Theory proposed by Clark and Brennan [17]. This theory posits that shared understanding and knowledge in common between communicators provide a basis for effective and collaborative communication. By transforming verbal expressions into visual representations, InsightBridge creates a tangible, shared reference point for discussion, which allows both parties to literally "see" the same thing, mitigating misinterpretations that typically arise from purely verbal exchanges [48, 50, 56]. Moreover, the iterative visual checking loop provides a series of shared references to reduce users' burden of recalling details and recognizing misalignment, which further elicits deeper, more insightful user feedback. Leveraging visual abstracts to establish common ground can be adopted in any scenario where abstract ideas need to be concretized for mutual understanding. For instance, visual abstracts could facilitate communication between medical professionals and patients about patients' physical status and living environment during the consultation phase before treatment in the healthcare domain [52]. Due to the hallucination of LLMs and potential biases in the generated contents [39], the character in AI-generated visual abstracts may not reflect the true personal traits of users, especially for the marginalized and minoritarian participants. On the one hand, InsightBridge supports the loop of visual generation and adjustment to rectify such inconsistency conveniently. On the other hand, such misalignment of personal traits might remind participants of contrasting personal experiences and needs, which echoes the "I know it when I see it" phenomenon [57] and corresponds to the feedback given by user U8 in Section 5.2.2. However, researchers need to carefully balance participant comfort and the reliability of collected data during the visual abstract checking process, especially in sensitive settings [73].

6.2 InsightBridge Facilitating Convergent and Divergent Thinking

The design and interaction flow of InsightBridge embodies both the convergent and divergent thinking in UCD [27, 30, 62]. Specifically, the note-taking function can be considered as a convergent process. It synthesizes complex conversations into structured, categorized insights, focusing the vast array of user responses into manageable and intuitive data points as sticky notes. Similarly, the visual checking phase, where researchers select specific sticky notes for visual generation, further narrows the focus to significant and intricate user scenarios, exemplifying convergent thinking. Conversely, the subsequent visual feedback and refinement session introduces a divergent thinking process. As users provide feedback and collaboratively refine the visual abstracts with AI, overlooked details and ideas could emerge. In other words, the process of users reflecting upon the visual abstract is essentially a form of divergent thinking, which triggers them to express more and deeper needs. In this way, the researcher's understanding of the target user could be further expanded. The hallucination issue of LLMs brings uncertainty to the divergent thinking process. If the generated contents are continuously unsatisfied after several rounds of iteration, the users may feel distracted or even upset. On the contrary, some hallucinated or irrelevant generated contents can also steer users to diverse aspects and perspectives, which is shown in the Brainstorming type of facilitation in Section 5.2.2. Finally, the prompted and brainstormed user feedback is integrated into the empathy map again as another convergent process to increase the depth and breadth of user needs and context. To this end, we believe InsightBridge serves as a novel yet effective boundary object in the researcher-user communication to facilitate both convergent and divergent thinking, as well as iteration in these two ways of thinking [5].

6.3 Extend the Use of Generated Empathy Maps and Visuals

While our user study has demonstrated the effectiveness and usefulness of InsightBridge in supporting user interviews in the early stage of the design process, we believe it can also be extended and utilized in more stages of the UCD process. The key lies in, through generating visual abstracts, InsightBridge is capable of establishing an empathetic channel in researcher-user communication, thus could potentially support the general "knowing the users" activity in the UCD process. For example, during prototyping, the system could assist in creating quick visual mockups based on verbal descriptions of the design concepts, allowing for rapid iteration and user feedback [13]. Meanwhile, InsightBridge's ability to facilitate communication between researchers and users could be leveraged in co-design sessions, where visual abstracts could serve as a common language for collaborative ideation [14, 68]. By extending InsightBridge's capabilities across the user-centered design lifecycle, designers could benefit from enhanced communication, visualization, and insight generation throughout the entire process, potentially leading to more user-aligned and innovative design outcomes.

Besides extending InsightBridge to support more researcher-user communication scenarios, we also see opportunities to improve communication within research teams through circulating the generated and confirmed empathy maps and visuals. In our experiment, to shape a deep understanding of each target user initially, we adopt the single-user empathy map as the template. However, InsightBridge can be implemented to assist research teams in aggregating separate one-user empathy maps into a comprehensive multiple-user empathy map, representing a group perspective of multiple users [6, 26]. Converged and synthesized information in the empathy map provides a basis for diverging research team's thinking and collaboration.

6.4 Design Considerations

6.4.1 Overcoming researchers' existing perspectives. Some previous work indicate biases from researchers could be introduced in the empathy-building process through, for example, legitimating researchers' inherent perspectives [38] or displacing the first-hand experiences of minoritarian users [4]. The feature of automatically synthesizing information into the empathy map can potentially break researchers' existing perspectives and biases, as reflected by the feedback from researcher R3 in Section 5.1.1. Future designers should consider leveraging AI's capability to synthesize diverse perspectives, aiming to minimize bias. Additionally, while not investigated in this paper, potential bias inherent in AI should also be carefully examined to ensure it does not replace users' perspectives.

6.4.2 Balancing agency in human-AI collaboration. In the current mechanism of visual abstract generation, the timing and the seeds (*i.e.,* selected sticky notes) are fully controlled by researchers. The rationality of giving such an agency to researchers lies in allowing researchers to lead the communication directions and avoiding overwhelming researchers with too many AI-generated irrelevant visuals. However, considering the researcher R10's feedback in Section 5.1.2, there is still room to increase AI's autonomy in visual generation for communication, such as automatically selecting meaningful sticky notes based on the researchers' preferences and habits, users' voice tone of expressed scenarios, and the progress of the on-going interview. It might also help researchers overcome the two critiques of empathy-building mentioned in Section 6.4.1. In some cases, *e.g.,* the conversation is stuck or users face communication barriers, AI could generate some visuals automatically to inspire and prompt user thinking. Based on researcher R7's feedback in Section 5.1.2, the AI-powered visual generation shows the potential of overcoming the limitation of researchers' personal abilities and enriching researchers' choice of methods for specific study scenarios. At the same time, the priming effect introduced by AI intervention should be considered in designing the human-AI collaboration mode. For example, the AI-generated contents, *i.e.,* sticky notes and visual abstracts should keep a reasonable size to avoid overly affecting researchers' decision-making process. The confidence level and other explainable factors could be further provided for researchers' reference and consideration [64].

6.4.3 Personalizing to different communication styles. User interviews or other design activities are highly unstructured, and researchers may have different communication styles and strategies [29]. AI's intervention in the communication without considering the cultural and other demographic characteristics of the user groups might lead to discomfort of researchers and users or misunderstandings [16], potentially hindering effective collaboration and empathy-building. Thus, the workflow of InsightBridge should be fine-tuned to consider the cultural and other background factors and to allow researchers to intervene in the AI-empowered interview process promptly and flexibly. Meanwhile, the order and layout of the sticky notes and visual abstracts in the corresponding areas could also be tailored to researchers' communication habits, such as the bottom-up communication or the communication with circular logic.

6.5 Limitations and Future Work

Our work has the following limitations. First, the researchers (aged 19-27) and users (aged 21-31) recruited in our experiment are young adults, which may bring some potential bias into the findings. The researchers' familiarity with the operation of electronic devices, and users' perception of generated visuals vary across participants from different age groups. In future work, exploring how InsightBridge benefits researchers and users over 30 will be an effective step to enhance its generalizability. Second, the two pre-selected design topics may limit the generalizability of our findings in other areas of design. The usability and efficacy of InsightBridge can be tested more comprehensively in diverse design scenarios. Moreover, in our experiment setting, each researcher only interviews one user on each topic. The long-term observation of researchers' interview behaviours can further deepen the analysis of InsightBridge's effectiveness. Furthermore, as the design process may involve a group of users, we aim to conduct a user study with a larger scale of participants in different scenarios in the future, *e.g.,* online meetings, to evaluate the InsightBridge's performance and efficacy when many more participants are engaged.

7 Conclusion

In our work, we introduce InsightBridge , a novel interactive system to enhance researchers' empathy with users during user interviews. Empowered by LLMs, InsightBridge can extract and organize key points from ongoing conversations into the empathy map. Researchers can select elements in the empathy map to generate corresponding visual abstracts. Then users check visuals, provide insights, and refine visuals as needed. Compared to the Baseline system, a within-subject study demonstrates that InsightBridge can assist researchers in reliable information synthesis and useful generated visual abstracts. The visual checking and refinement session with users can trigger more detailed and insightful user feedback and facilitate mutual understanding between two parties. We further sum up design implications and limitations to inspire future study about empathy enhancement in user-centered design.

Acknowledgments

This work is partially supported by the Research Grants Council of the Hong Kong Special Administrative Region, China under Grant No. C6044-23G. We thank anonymous reviewers for their valuable feedback and our study participants for their great contribution and effort. Lastly, we appreciate Jiaxiong Hu and Runhua Zhang's insightful discussions.

References

[1] Chadia Abras, Diane Maloney-Krichmar, Jenny Preece, et al. 2004. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications* 37, 4 (2004), 445–456.

[2] Mohammed Alhamadi, Omar Alghamdi, Sarah Clinch, and Markel Vigo. 2022. Data Quality, Mismatched Expectations, and Moving Requirements: The Challenges of User-Centred Dashboard Design. In *Nordic Human-Computer Interaction Conference* (Aarhus, Denmark) *(NordiCHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 11, 14 pages. https://doi.org/10.1145/3546155.3546708

[3] Elke Beck, Marianna Obrist, Regina Bernhaupt, and Manfred Tscheligi. 2008. Instant card technique: how and why to apply in user-centered design. In *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008* (Bloomington, Indiana) *(PDC '08)*. Indiana University, USA, 162–165.

[4] Cynthia L. Bennett and Daniela K. Rosner. 2019. The Promise of Empathy: Design, Disability, and Knowing the "Other". In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300528

[5] Mark Bergman, Kalle Lyytinen, and Gloria Mark. 2007. Boundary objects in design: An ecological view of design artifacts. (2007).

[6] Eva Bittner and Omid Shoury. 2019. Designing automated facilitation for design thinking: A chatbot for supporting teams in the empathy map method. (2019).

[7] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. *Qualitative HCI research: Going behind the scenes.* Morgan & Claypool Publishers.

[8] Stefan Blomkvist. 2002. Persona–an overview. *Retrieved November* 22, 2004 (2002), 15.

[9] S. Bodker. 1999. Scenarios in user-centred design-setting the stage for reflection and action. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*, Vol. Track3. 11 pp.–. https://doi.org/10.1109/HICSS.1999.772892

[10] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis.* American Psychological Association.

[11] J Brooke. 1996. SUS: A quick and dirty usability scale. *Usability Evaluation in Industry* (1996).

[12] Marion Buchenau and Jane Fulton Suri. 2000. Experience prototyping. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques.* 424–433.

[13] Reinhard Budde, Karlheinz Kautz, Karin Kuhlenkamp, and Heinz Züllighoven. 1990. What is prototyping? *Information Technology & People* 6, 2/3 (1990), 89–95.

[14] Ingrid Burkett. 2012. An introduction to co-design. *Sydney: Knode* 12 (2012).

[15] John M Carroll. 1997. Scenario-based design. In *Handbook of human-computer interaction.* Elsevier, 383–406.

[16] Kaiping Chen, Anqi Shao, Jirayu Burapacheep, and Yixuan Li. 2024. Conversational ai and equity through assessing gpt-3's communication with diverse social groups on contentious topics. *Scientific Reports* 14, 1 (2024), 1561.

[17] HH Clark. 1991. Grounding in Communication. *Perspectives on Socially Shared Cognition/American Psychological Association* (1991).

[18] Alan Cooper. 1999. *The inmates are running the asylum.* Springer.

[19] Victor Philip Cornet, Tammy Toscos, Davide Bolchini, Romisa Rohani Ghahari, Ryan Ahmed, Carly Daley, Michael J Mirro, and Richard J Holden. 2020. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. *JMIR mHealth and uHealth* 8, 7 (2020), e17703.

[20] Catherine Courage and Kathy Baxter. 2005. *Understanding your users: A practical guide to user requirements methods, tools, and techniques.* Gulf Professional Publishing.

[21] Humphrey Curtis and Timothy Neate. 2023. Watch Your Language: Using Smartwatches to Support Communication. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 51, 21 pages. https://doi.org/10.1145/3597638.3608379

[22] Scott Davidoff, Min Kyung Lee, Anind K. Dey, and John Zimmerman. 2007. Rapidly exploring application design through speed dating. In *Proceedings of the 9th International Conference on Ubiquitous Computing* (Innsbruck, Austria) *(UbiComp '07)*. Springer-Verlag, Berlin, Heidelberg, 429–446.

[23] Alan R. Dennis and Susan T. Kinney. 1998. Testing Media Richness Theory in the New Media: The Effects of Cues, Feedback, and Task Equivocality. *Information Systems Research* 9, 3 (Sept. 1998), 256–274. https://doi.org/10.1287/isre.9.3.256

[24] Travis Dow, Pratishtha Pratishtha, Lorans Alabood, Vikram K. Jaswal, and Diwakar Krishnamurthy. 2024. Towards an Augmented Reality Agent to Support Communication for Nonspeaking Autistic People. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems (CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 338, 8 pages. https://doi.org/10.1145/3613905.3651063

[25] Luce Drouet, Wo Meijer, Aisling Ann O'Kane, Aneesha Singh, Thiemo Wambsganss, Andrea Mauri, and Himanshu Verma. 2023. The EmpathiCH Workshop: Unraveling Empathy-Centric Design. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 366, 7 pages. https://doi.org/10.1145/3544549.3573796

[26] Bruna Ferreira, Williamson Silva, Edson Oliveira, and Tayana Conte. 2015. Designing Personas with Empathy Map.. In *SEKE*, Vol. 152.

[27] Jonas Frich, Midas Nouwens, Kim Halskov, and Peter Dalsgaard. 2021. How Digital Tools Impact Convergent and Divergent Thinking in Design Ideation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 431, 11 pages. https://doi.org/10.1145/3411764.3445062

[28] Sarah Gibbons. 2018. Empathy Mapping: The First Step in Design Thinking. https://www.nngroup.com/articles/empathy-mapping/. Accessed: 2018-01-14.

[29] Vijai N Giri. 2006. Culture and communication style. *The Review of Communication* 6, 1-2 (2006), 124–130.

[30] Gabriela Goldschmidt. 2014. *Linkography: unfolding the design process.* Mit Press.

[31] David A Grant. 1948. The latin square principle in the design and analysis of psychological experiments. *Psychological bulletin* 45, 5 (1948), 427.

[32] Steve Greenspan, David Goldberg, David Weimer, and Andrea Basso. 2000. Interpersonal trust and common ground in electronically mediated communication. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* (Philadelphia, Pennsylvania, USA) *(CSCW '00)*. Association for Computing Machinery, New York, NY, USA, 251–260. https://doi.org/10.1145/358916.358996

[33] Luke Haliburton, Natalia Bartłomiejczyk, Albrecht Schmidt, Paweł W. Woźniak, and Jasmin Niess. 2023. The Walking Talking Stick: Understanding Automated Note-Taking in Walking Meetings. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 431, 16 pages. https://doi.org/10.1145/3544548.3580986

[34] Johannes AM Harmsen, RMD Bernsen, Ludwien Meeuwesen, Duane Pinto, and MA Bruijnzeels. 2005. Assessment of mutual understanding of physician patient encounters: development and validation of a mutual understanding scale (MUS) in a multicultural general practice setting. *Patient Education and Counseling* 59, 2 (2005), 171–181.

[35] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[36] Justin L Hess and Nicholas D Fila. 2016. The manifestation of empathy within design: findings from a service-learning course. *CoDesign* 12, 1-2 (2016), 93–111.

[37] Karen Holtzblatt, Jessamyn Burns Wendell, and Shelley Wood. 2004. *Rapid contextual design: a how-to guide to key techniques for user-centered design.* Elsevier.

[38] Lilly Irani. 2019. *Chasing innovation: Making entrepreneurial citizens in modern India.* Princeton University Press.

[39] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of Hallucination in Natural Language Generation. *ACM Comput. Surv.* 55, 12, Article 248 (March 2023), 38 pages. https://doi.org/10.1145/3571730

[40] Merlijn Kouprie and Froukje Sleeswijk Visser. 2009. A framework for empathy in design: stepping into and out of the user's life. *Journal of Engineering Design* 20, 5 (2009), 437–448.

[41] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research methods in human-computer interaction.* Morgan Kaufmann.

[42] Meira Levy and Irit Hadar. 2024. Learning to empathize with users through design thinking in hybrid mode: Insights from two educational case studies. *Journal of Systems and Software* 207 (2024), 111831.

[43] Junze Li, Changyang He, Jiaxiong Hu, Boyang Jia, Alon Y Halevy, and Xiaojuan Ma. 2024. DiaryHelper: Exploring the Use of an Automatic Contextual Information Recording Agent for Elicitation Diary Study. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 818, 16 pages. https://doi.org/10.1145/3613904.3642853

[44] Di Liu, Hanqing Zhou, and Pengcheng An. 2024. "When He Feels Cold, He Goes to the Seahorse"—Blending Generative AI into Multimaterial Storymaking for Family Expressive Arts Therapy. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 118, 21 pages. https://doi.org/10.1145/3613904.3642852

[45] Xingyu "Bruce" Liu, Vladimir Kirilyuk, Xiuxiu Yuan, Alex Olwal, Peggy Chi, Xiang "Anthony" Chen, and Ruofei Du. 2023. Visual Captions: Augmenting Verbal Communication with On-the-fly Visuals. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 108, 20 pages. https://doi.org/10.1145/3544548.3581566

[46] Ji-Ye Mao, Karel Vredenburg, Paul W. Smith, and Tom Carey. 2005. The state of user-centered design practice. *Commun. ACM* 48, 3 (mar 2005), 105–109. https://doi.org/10.1145/1047671.1047677

[47] Andrea Mauri, Yen-Chia Hsu, Marco Brambilla, Ting-Hao Kenneth Huang, Aisling Ann O'Kane, and Himanshu Verma. 2022. Empathy-Centric Design At Scale. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 75, 6 pages. https://doi.org/10.1145/3491101.3503744

[48] Rose McCabe and Patrick GT Healey. 2018. Miscommunication in doctor–patient communication. *Topics in cognitive science* 10, 2 (2018), 409–424.

[49] Áurea Hiléia da Silva Melo, Luis Rivero, Jonathas Silva dos Santos, and Raimundo da Silva Barreto. 2020. EmpathyAut: an empathy map for people with autism. In *Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems* (Diamantina, Brazil) *(IHC '20)*. Association for Computing Machinery, New York, NY, USA, Article 45, 6 pages. https://doi.org/10.1145/3424953.3426650

[50] C David Mortensen and Carter M Ayres. 1997. *Miscommunication.* Sage.

[51] Elizabeth D Mynatt, A-S Melenhorst, A-D Fisk, and Wendy A Rogers. 2004. Aware technologies for aging in place: understanding user needs and attitudes. *IEEE Pervasive Computing* 3, 2 (2004), 36–41.

[52] Lucille ML Ong, Johanna CJM De Haes, Alaysia M Hoos, and Frits B Lammes. 1995. Doctor-patient communication: a review of the literature. *Social science & medicine* 40, 7 (1995), 903–918.

[53] Alexander Osterwalder. 2010. Business model generation: A handbook for visionaries, game changers, and challengers.

[54] Allan Paivio. 1990. *Mental representations: A dual coding approach.* Oxford university press.

[55] Dev Patnaik and Robert Becker. 1999. Needfinding: the why and how of uncovering people's needs. *Design Management Journal (Former Series)* 10, 2 (1999), 37–43.

[56] Alexandra Paxton, Jennifer M Roche, Alyssa Ibarra, and Michael K Tanenhaus. 2021. Predictions of miscommunication in verbal communication during collaborative joint action. *Journal of Speech, Language, and Hearing Research* 64, 2 (2021), 613–627.

[57] Andrew Polaine. 2005. The flow principle in interactivity. In *ACM International Conference Proceeding Series*, Vol. 123. 151–158.

[58] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-Shot Text-to-Image Generation. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 8821–8831. https://proceedings.mlr.press/v139/ramesh21a.html

[59] Fraser JM Reid and Susan E Reed. 2007. Conversational grounding and visual access in collaborative design. *CoDesign* 3, 2 (2007), 111–122.

[60] E Frank Ritter, D Gordon Baxter, and F Elizabeth Churchill. 2014. *Foundations for designing user-centered systems: What system designers need to know about people.* Springer.

[61] John Rudnik, Sharadhi Raghuraj, Mingyi Li, and Robin N. Brewer. 2024. CareJournal: A Voice-Based Conversational Agent for Supporting Care Communications. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 526, 22 pages. https://doi.org/10.1145/3613904.3642163

[62] Mark A Runco. 1991. The evaluative, valuative, and divergent thinking of children. *The Journal of Creative Behavior* 25, 4 (1991), 311–319.

[63] Chenxinran Shen, Joanna Mcgrenere, and Dongwook Yoon. 2024. LegacySphere: Facilitating Intergenerational Communication Through Perspective-Taking and Storytelling in Embodied VR. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 119, 16 pages. https://doi.org/10.1145/3613904.3641923

[64] Donghee Shin. 2021. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies* 146 (2021), 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

[65] Cassandre Simon, Manel Boukli Hacene, Samir Otmane, and Amine Chellali. 2024. Study of communication modalities to support teaching tool manipulation skills in a shared immersive environment. *Comput. Graph.* 117, C (mar 2024), 31–41. https://doi.org/10.1016/j.cag.2023.09.011

[66] Waralak Vongdoiwang Siricharoen. 2021. Using empathy mapping in design thinking process for personas discovering. In *Context-Aware Systems and Applications, and Nature of Computation and Communication: 9th EAI International Conference, ICCASA 2020, and 6th EAI International Conference, ICTCC 2020, Thai Nguyen, Vietnam, November 26–27, 2020, Proceedings 9*. Springer, 182–191.

[67] Susan Leigh Star and James R Griesemer. 1989. Institutional ecology,translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social studies of science* 19, 3 (1989), 387–420.

[68] Marc Steen. 2013. Co-design as a process of joint inquiry and imagination. *Design issues* 29, 2 (2013), 16–28.

[69] Khai N. Truong, Gillian R. Hayes, and Gregory D. Abowd. 2006. Storyboarding: an empirical determination of best practices and effective guidelines. In *Proceedings of the 6th Conference on Designing Interactive Systems* (University Park, PA, USA) *(DIS '06)*. Association for Computing Machinery, New York, NY, USA, 12–21. https://doi.org/10.1145/1142405.1142410

[70] Priyan Vaithilingam, Ian Arawjo, and Elena L. Glassman. 2024. Imagining a Future of Designing with AI: Dynamic Grounding, Constructive Negotiation, and Sustainable Motivation. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 289–300. https://doi.org/10.1145/3643834.3661525

[71] Karel Vredenburg, Ji-Ye Mao, Paul W. Smith, and Tom Carey. 2002. A survey of user-centered design practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI '02)*. Association for Computing Machinery, New York, NY, USA, 471–478. https://doi.org/10.1145/503376.503460

[72] Jacqueline H Watts. 2008. Emotion, empathy and exit: Reflections on doing ethnographic qualitative research on sensitive topics. *Medical sociology online* 3, 2 (2008), 3–14.

[73] Jenny Waycott, Greg Wadley, Stefan Schutt, Arthur Stabolidis, and Reeva Lederman. 2015. The Challenge of Technology Research in Sensitive Settings: Case Studies in 'ensitive HCI'. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction* (Parkville, VIC, Australia) *(OzCHI '15)*. Association for Computing Machinery, New York, NY, USA, 240–249. https://doi.org/10.1145/2838739.2838773

[74] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems* 35 (2022), 24824–24837.

[75] Christiane Wölfel and Timothy Merritt. 2013. Method card design dimensions: A survey of card-based design tools. In *Human-Computer Interaction–INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part I 14*. Springer, 479–486.

[76] Luke Jai Wood, Kerstin Dautenhahn, Austen Rainer, Ben Robins, Hagen Lehmann, and Dag Sverre Syrdal. 2013. Robot-mediated interviews-how effective is a humanoid robot as a tool for interviewing young children? *PloS one* 8, 3 (2013), e59448.

[77] Peter Wright and John McCarthy. 2008. Empathy and experience in HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 637–646.

[78] 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

[79] Ziqi Yang, Xuhai Xu, Bingsheng Yao, Jiachen Li, Jennifer Bagdasarian, Guodong Gao, and Dakuo Wang. 2024. "I Wish There Were an AI": Challenges and AI Potential in Cancer Patient-Provider Communication. arXiv:2404.13409 [cs.HC] https://arxiv.org/abs/2404.13409

[80] Yanzhen Yue, Xiaojuan Ma, and Zhenhui Jiang. 2014. Share your view: impact of co-navigation support and status composition in collaborative online shopping. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3299–3308. https://doi.org/10.1145/2556288.2557143

[81] Chengbo Zheng, Dakuo Wang, April Yi Wang, and Xiaojuan Ma. 2022. Telling stories from computational notebooks: Ai-assisted presentation slides creation for presenting data science work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–20.