Investigating “Touch and Talk” for Blind and Low Vision People: Science Communication Assistance Through Exploring Multiple Tactile Objects

Ayaka Tsutsui*
University of Tsukuba
Tsukuba, Japan
Miraikan – The National Museum of Emerging Science and Innovation
Tokyo, Japan
ayakatsutsui1128@digitalnature.slis.tsukuba.ac.jp

Xiyue Wang*
Miraikan – The National Museum of Emerging Science and Innovation
Tokyo, Japan
wang.xiyue@lab.miraikan.jst.go.jp

Hironobu Takagi
IBM Research - Tokyo
Tokyo, Japan
takagih@jp.ibm.com

Chieko Asakawa
Miraikan – The National Museum of Emerging Science and Innovation
Tokyo, Japan
IBM Research
Yorktown, USA
chiekoa@us.ibm.com

Abstract

Tactile exploration is extremely important for blind and low vision (BLV) users to understand concepts. Although Interactive 3D Models that integrate modalities such as audio and vibration have been developed to create self-directed tactile exploration experiences, they depend on pre-defined commands and fixed interaction flows, which limit opportunities for adaptive guidance. In this study, we examine how interactive dialogue and temporal dynamics can enhance tactile learning experiences, particularly in the context of science communication where BLV users frequently encounter abstract and spatially complex topics. We conducted interviews with 22 tactile guidance experts to identify effective explanation techniques and communication strategies. We then employed a technology probe that combines multiple tactile models with a voice-based “Touch and Talk” system, using a Wizard-of-Oz approach with 10 BLV participants. The experiment revealed strategies that support understanding and foster curiosity. Based on findings, we propose a set of design implications aimed at supporting BLV users in autonomously exploring complex scientific content.

CCS Concepts

• **Human-centered computing → Empirical studies in accessibility**.

*Both authors contributed equally to this research.

This work is licensed under a Creative Commons Attribution 4.0 International License.
ASSETS '25, Denver, CO, USA
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0676-9/2025/10
https://doi.org/10.1145/3663547.3746373

Keywords

Interactive 3D models (I3M), Tactile graphics, 3D model, Accessibility, Communication, Visually impaired

ACM Reference Format:
Ayaka Tsutsui, Xiyue Wang, Hironobu Takagi, and Chieko Asakawa. 2025. Investigating “Touch and Talk” for Blind and Low Vision People: Science Communication Assistance Through Exploring Multiple Tactile Objects . In *The 27th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '25), October 26–29, 2025, Denver, CO, USA.* ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3663547.3746373

1 INTRODUCTION

Tactile exploration is essential for blind and low vision (BLV) individuals to acquire spatial knowledge [34, 54, 73] and understand concepts [28, 29, 44]. However, tactile objects alone have been shown to be insufficient for conveying complex, scientific, or conceptual information [15, 16, 41, 69]. Consequently, supplementary explanations by human guides or pre-recorded audio guides are required during tactile exploration. While human guides can provide adaptive real-time feedback, they may compromise user autonomy [62]. In contrast, pre-recorded audio guides lack flexibility and do not address individual interests or exploration patterns [14, 38].

To overcome these limitations, Interactive 3D models (I3Ms) that integrate multiple modalities such as audio [18, 26, 27], vibration [25, 48, 58], and even visual cues for BLV users have been developed [32, 46]. These I3Ms aim to provide a more self-directed learning experience by linking tactile interaction with real-time feedback. However, most existing I3Ms rely on pre-defined commands or fixed interaction flows, leaving little room for BLV users to ask questions, clarify concepts, or receive adaptive guidance [30, 55].

Recent research has attempted to address these issues using various approaches. Shi et al. analyzed gestures and hand movements to inform the design of I3Ms, but their use of static models left actual interactive exploration unexamined [67]. Working with BLV teachers, another I3M was created to let users ask questions on demand, yet the timing and strategies for those queries remain unclear [65]. A subsequent study combined tap controlled audio labels with a conversational interface for natural language queries, but it too stopped short of exploring how users navigate multiple distinct models within a single theme [60].

Although these studies demonstrate that combining tactile interaction with conversational cues can promote user agency, they often focus on fixed tasks or constrained dialogue structures. They rarely address how to support open-ended tactile exploration across multiple tactile representations or how to adapt explanations in a customizable manner.

To address these gaps, this study investigates how interactive dialogue and temporal dynamics can enhance tactile learning experiences. Our work focuses specifically on tactile exploration within the context of interactive science communication for BLV learners. Science communicators often present these learners with abstract and spatially complex topics which, when combined with adaptive and multimodal explanations, can stimulate curiosity and enhance learning [11, 59].

Our goal is to develop systems that support BLV users not only in understanding scientific content but also in determining when, how, and what kind of information they receive—enabling not only comprehension but also the formulation of new questions and deeper conceptual inquiry. To guide this investigation, we posed the following research questions (RQ):

- RQ1: What are the current status, challenges, and opportunities in conveying scientific concepts to BLV individuals, particularly in terms of content, timing, and communication strategies?
- RQ2: Based on the findings of RQ1, how can we design an assistive tool that supports science communication through tactile exploration, enabling both structured understanding and free exploration?

In Study 1, we investigated RQ1 by conducting semi-structured interviews with 22 tactile guidance experts including educators and science museum professionals to identify effective explanation techniques and communication strategies. Our findings highlight several key elements for tactile exploration: the use of multiple models, adaptive communication strategies, and balanced control, timing, and correction during explanation.

Based on these insights, in Study 2 we developed a technology probe that combined multiple tactile models representing scientific themes with a Wizard-of-Oz voice interface simulating a "Touch and Talk" system. To investigate RQ2, we conducted an experiment with 10 BLV participants who explored scientific themes by freely switching between storytelling narration and question-based interactions and receiving directional guidance. This experiment revealed recurring interaction patterns and challenges in model manipulation, as well as strategies that support understanding and foster curiosity. Based on these findings, we propose six design implications for future tactile systems that focus on adaptivity, timing, multimodal feedback, and personalization.

In conclusion, our study complements existing I3M research which primarily focuses on customizing tactile output or integrating passive audio labels by examining interactive dialogue and temporal dynamics that are essential for a fully integrated, self-directed learning experience. By supporting BLV users in engaging in tactile exploration based on intrinsic motivation, our approach is expected to ultimately promote the evolution of self-directed STEM learning. Our contributions include:

- Revealing effective explanation techniques and communication strategies through semi-structured interviews with 22 tactile guidance experts from diverse backgrounds.
- Investigating interaction patterns, challenges, and opportunities in BLV users' engagement with a voice-based "Touch and Talk" system through a Wizard-of-Oz study using multiple tactile models.
- Developing design implications for next-generation assistive technologies that enable BLV users to autonomously explore, understand, and generate new questions about scientific content.

2 RELATED WORK

2.1 Accessible Tactile Content for BLV Users

To enable BLV individuals to understand complex concepts, the effectiveness of accessible tactile content has been demonstrated across various fields, including everyday life [50], mobility [81], and education [9, 24, 40, 49, 64].

Tactile content can be broadly classified into three formats: 2D, 2.5D, and 3D [57]. 2D content is essentially flat and is used for tactile picture books [70], embossed paper-based tactile maps [20], and a wide range of educational materials [56, 72]. Recent advances in embossing technologies are further expanding the design space of 2D tactile media, allowing finer detail and more varied textures to be produced on standard substrates [75, 76]. 2.5D content, which is utilized for reliefs [82] and representations of buildings [83], associates each position with a single height value [33] and is frequently employed in depicting buildings and artworks. 3D content enables the depiction of rich textures and fine details and is used in museums and galleries to increase opportunities for BLV users to physically interact with exhibits [2, 3].

However, some tactile materials incorporate braille or legends to facilitate understanding for BLV users. Such additions, though, may reduce the amount of tactile information that can be directly perceived by touch [7, 37, 39, 66]. Moreover, since not all BLV users are able to read braille, incorporating it does not necessarily accommodate every user [10, 53]. Understanding these limitations is essential for designing tactile content that balances informational clarity with unimpeded tactile exploration.

2.2 Development of Interactive 3D Models (I3Ms)

To move beyond static labels, Interactive 3D Models (I3Ms) integrate audio, haptics, and other modalities directly into tactile models. Toolkits add voice labels to 3D-printed models [37, 66], while others pair audio with haptic feedback on complex forms [78, 79], making updates easier and enabling freer exploration [47]. Further enhancing interactivity, systems combine audio, vibration, and gesture recognition to broaden access for non-braille users [8] and increase the information conveyed [45, 71].

Researchers have also embedded clear audio explanations into tactile content [35, 42], including gesture-based audio guides in museums [55], camera-triggered audio labels [17], and QR code activated narrations on tactile diagrams [4, 23]. Navigation aids using combined tactile and audio feedback [1, 12, 77, 80] and wearable pen-type or ring-type devices [19, 74] further demonstrate how physical models and interaction technology can merge. Despite their promise, these solutions remain largely static and do not tailor support to individual knowledge or exploration styles [21, 43, 59].

2.3 Tactile Exploration Using I3Ms

I3Ms are now entering mainstream education. Recent studies have introduced modular, voice-guided assembly systems [51, 52] compared an interactive 3D model with a DIY interactive 2D tactile map of a complex building [63] and surveyed the adoption of 3D printing in classrooms [36]. Together, these works underscore the growing importance of tactile tools in both formal and informal learning settings.

Despite these advances, the exploration strategies of BLV users remain under-examined. Shi et al. analysed gestures and hand movements across multiple static models, revealing demand for audio cues, buttons, and gesture support, yet leaving dynamic exploration unaddressed [67]. Reinders et al. first conducted a Wizard-of-Oz study that combined natural-language queries with audio and haptic output, then evaluated a fully functional multimodal I3M; participants preferred the integration of touch, vibration, and conversation [59, 60]. However, because these studies focused on one model at a time, the question of when and how users should switch modes or navigate multiple related models particularly those differing in more than scale remains unresolved.

Building on this momentum, we investigate the design requirements for interactive dialogue and temporal dynamics that allow BLV users to move fluidly among multiple models while maintaining self-directed engagement with complex scientific themes. Ultimately, our goal is to foster intrinsically motivated tactile exploration and advance self-directed STEM learning for BLV audiences.

3 STUDY 1: INTERVIEW WITH EXPERTS WHO PROVIDES GUIDANCE DURING TACTILE EXPLORATION

To gain a deep understanding of the current state of tactile guidance for BLV individuals, identify challenges, and pinpoint areas for improvement, we conducted semi-structured interviews with 22 experts from diverse fields, all with extensive experience supporting BLV users' tactile exploration. By involving specialists from various domains, we gathered a broad range of perspectives not tied to any single use case. The interviews focused specifically on the materials used for scientific themes, the timing of explanations, and the communication strategies employed.

3.1 Participants

We recruited 22 participants with diverse backgrounds (10 males, 12 females, average age = 44.25, SD = 13.25). P14 and P15 reported their ages in their 40s and 50s; thus, the average age was calculated based on the responses of 20 participants. Among the participants, 7 work as science communicators at science museums, 1 is a tactile museum staff, 8 are teachers at schools for the blind (teaching subjects include 1 science, 3 acupuncture, 1 independent living skills, 1 mathematics, and 2 English), 5 are staff at BLV support facilities, and 1 is a researcher specializing in accessibility. Notably, P10 has low vision, while P20 and P22 are congenitally blind.

All participants have experience in providing explanations or guidance to BLV users using at least one type of tactile content, whether 2D tactile graphics, 3D models, or actual objects with experience ranging from 1 to 34 years (average year = 9.5). Demographic information, occupations, and years of experience for all participants are summarized in Table 3 in the Appendix.

3.2 Methods

Interviews were conducted via Zoom[1] and lasted 60 minutes per session. At the beginning of each interview, we obtained participants' consent for the interview and for recording the session. Two of the authors participated in each interview session, with one serving as the facilitator and the other taking notes. Initially, we gathered background information on the participants, including their prior experiences in tactile instruction. The interview sessions were then divided into three main categories of questions: (1) Tactile content used for conveying concepts, (2) Instructional approaches and correction during tactile exploration, and (3) Strategies for communication and adaptations based on learner needs. Detailed questions for each category are provided in Table 4 in the Appendix. According to the guidelines of the research institution, ethical pre-evaluation or permission was exempted for this study.

3.3 Data Collection and Analysis

The interview recordings were transcribed using Notta[2]. We then conducted an iterative thematic analysis [6] to identify themes. First, two authors independently analyzed the data and performed open coding [13]. Next, we discussed the codes and resolved conflicts, such as missing codes or disagreements.

Once the authors reached an agreement, we transferred the codes into digital sticky notes on Miro[3]. Using the affinity diagramming technique [5], we grouped the codes into themes. This process resulted in the identification of five key themes. It is important to note that the number and composition of themes in a thematic analysis can vary according to factors such as the volume of data and the specific research focus [6].

3.4 Findings

3.4.1 *Use of Diverse Tactile Materials and Disassemblable Models.*

All participants suggested that combining a variety of tactile materials and models to explain a single concept greatly enhances BLV users' conceptual understanding and promotes an integrated comprehension of tactile information. For example, six participants (P8–P14) reported creating custom teaching materials using items such as cotton, milk cartons, felt, and additional tactile diagrams

[1] https://zoom.us/
[2] https://www.notta.ai/en
[3] https://miro.com/en

when existing 3D models or tactile graphics did not provide sufficient detail. This finding is consistent with previous research [31].

In particular, seven participants (P1, P3, P5–P7, P20, and P22) strongly endorsed the use of disassemblable models. It was noted that disassemblable models allow an object to be understood from both an external and internal perspective. P9 remarked, *"For items like fish that cannot be understood without examining the interior, a disassemblable model is ideal,"* emphasizing that tactile exploration of internal structures is crucial for independent living.

Additionally, seven science communicators (P1–P7) reported that combining a 2D overview diagram of the International Space Station (ISS) with detailed 3D models and full-scale models proved especially effective for conveying complex topics. This approach stimulates users' imagination and enhances comprehensive understanding, an effect that verbal explanations alone cannot achieve. Moreover, the sequence in which multiple tactile contents are explored was frequently adjusted to meet individual needs. For example, P12 noted *"Student who struggle with mental imagery, we have them explore 3D models instead of 2D."* This two-step approach was found to be particularly effective.

Integrating multiple tactile sources helps BLV users deepen their understanding of content and concepts while also prompting new questions during exploration. However, five teachers (P10–P15) noted that there is currently no technology available to support the use of multiple models. For example, P15 stated, *"It would be extremely convenient if there were a technology that could automatically explain each model while learners use multiple models, this could be especially effective in group lessons."* Similarly, P12 commented, *"Totally blind children often require extensive explanations, whereas low-vision children, who still have a relatively preserved field of view, can sometimes become bored. I would be very pleased if there were a tool that could combine explanations with a question-based approach to further enhance knowledge during those moments."* These responses suggest that a system capable of providing both detailed explanations and interactive questions when using multiple models would be highly beneficial.

3.4.2 Exploration Strategies: Balancing and Switching Between User-Led and Guide-Led Approaches.

In the context of tactile exploration for BLV users, it was found that facilitators employed both user-led and guide-led approaches, switching between them based on the user's level of understanding and the specific context.

The user-led approach involves the facilitator initially providing only a general overview of the model, after which the user is encouraged to explore the tactile content freely and spontaneously. As users ask questions during their exploration, the facilitator responds accordingly, helping to deepen the user's understanding. 7 participants (P2, P6, P9, P14, P18, P19 and P21) emphasized that excessive correction or premature intervention by the facilitator tends to diminish the user's motivation and comprehension. For example, P9 stated, *"Explaining everything deprives the learner of opportunities to cultivate their own curiosity,"* and P21 warned that *"Excessive intervention can foster a negative attitude toward tactile engagement."* These participants highlighted the importance of minimal intervention that allows for independent exploration and spontaneous formation of questions.

In contrast, the guide-led approach is characterized by a predefined, systematic sequence of explanations that the user follows while exploring the tactile materials. This method helps reduce cognitive load and enables learners to build their understanding step by step, from an overview to more detailed content. For instance, P12 suggested that starting with a 3D model explanation and then transitioning to more challenging 2D tactile diagrams allowed for comprehensive understanding. P22, a researcher with visual impairment, stated, *"Exploring freely alone does not lead to understanding. I need to touch while listening to a story-based explanation to achieve deeper comprehension."* Similarly, 5 participants (P5, P6, P9, P17 and P20) reported that beginning with a summary of the overall content and then proceeding to storytelling-based explanations effectively facilitated knowledge retention.

3.4.3 Challenges in Determining the Timing of Explanations.

All participants emphasized the need to adjust the timing of explanations flexibly, depending on the user's level of understanding, pace of tactile exploration, and degree of engagement. However, determining the appropriate timing was found to be challenging, particularly when interacting with BLV users for the first time. P11 noted, *"When a student was exploring a map used for mobility training, their hand had stopped moving, so I began to speak. But the student told me not to explain at that moment because they were trying to visualize the map in their mind."* This illustrates the difficulty in judging when to intervene with an explanation. P14 stated, *"The student I regularly work with always makes the same facial expression when they don't understand, so I don't find it difficult to judge the right moment to explain. But with students I am not familiar with, it would be more difficult."* P15 noted, *"When they furrow their brows, I can tell they're lost at least with the students I work with regularly"*. These responses highlight that familiarity with the user affects the ability to determine appropriate timing.

Tactile museum and science museum staff, who often interact with BLV visitors for the first time, reported particular difficulty in reading individual reactions accurately. P4 explained, *"In a school setting, we might be able to judge from facial expressions or hand movements, but at a science museum, it's usually our first time meeting the visitor, so it's harder. That's why we try to observe participants closely and match the pace of our explanations to theirs."* These findings underscore that the timing of explanations depends heavily on user reactions and levels of engagement, requiring facilitators to make sensitive, context-aware decisions.

3.4.4 Correction Methods in Explanation.

When guiding BLV users in tactile exploration using tactile content, all participants reported having had experiences where they needed to correct the user because the user was touching a different part of the model than the one being explained. Two main correction methods were identified: physically guiding the user's hand to the correct location, and providing verbal instructions only, allowing the user to move their hands independently.

Eleven participants reported using direct hand guidance as a correction method. All of them emphasized that they always obtained the user's permission before touching their hand. P3 explained, *"We guide their hand because continuing with verbal explanations would only cause more confusion."* P20 stated, *"As a blind person myself, I don't mind having my hand guided. It helps me reach the* *correct part of the model faster and with less stress."* In contrast, P22 responded, *"I absolutely dislike having my hand touched. Instead of physical guidance, clear verbal instructions that help me reach the correct part on my own are better for independent exploration."* These contrasting views indicate that preferences for hand guidance vary among users. Additionally, P1 noted, *"If the goal is for users to explore independently, some form of voice-based guidance is preferable,"* suggesting that physical hand guidance, while effective, may be limited in its applicability.

Furthermore, eleven participants reported guiding users verbally to the correct location. P1, P4–P6, and P8 gave directional cues based on the user's current hand position, such as, *"Move a little to the right from where you are now."* Notably, P7 stated, *"Before correcting, I first describe the part the user is currently touching. Then I redirect them to the part I intended to explain,"* showing a flexible verbal strategy that incorporates acknowledgment of the user's current position before redirection. However, P6 added that *"Even when I say 'right' or 'left,' users often move their hands too much, making accurate guidance difficult."* They found that providing more specific instructions such as *"move 3 cm to the right"* was more effective in promoting understanding.

3.4.5 Communication Style and Personalization: Adapting to Visual Function, Learner Characteristics, and Group Settings. All participants reported that they flexibly adjusted both their language and explanatory strategies based on users' visual capabilities and prior visual experiences. In particular, P21 stated, *"Learners with acquired blindness tend to rely on visual memory to grasp concepts, so referencing visual features and colors is effective. On the other hand, congenitally blind learners, while generally adept at tactile exploration, require metaphors for explaining color and scale."*

Six science communicators (P1–P6) also revealed that when explaining scientific topics to BLV individuals, they frequently use easy-to-imagine metaphors. P4 pointed out that comparative expressions such as *"about the size of two people" or "a distance equivalent to two train stations"* are effective in conveying size. However, P3 commented, *"When presenting an exhibit on viruses, we received feedback that using metaphors still made it difficult for visitors to form a clear mental image."* This indicates that, for certain scientific topics, words alone may not be sufficient.

Eight participants (P9–P16) emphasized the need to tailor communication based not only on visual ability but also on factors such as age, prior knowledge, tactile experience, language proficiency, and personality. For instance, P14 stated, *"While kindergarten children generally require more structured guidance and concise explanations, middle and high school students tend to engage in deeper conceptual questioning and abstract reasoning."*

The mode of communication varied depending on the number of BLV users and facilitators. One-on-one instruction, often used in mobility training or early tactile literacy development, allowed communicators to tailor their guidance to each individual's learning style. In contrast, settings such as museum tours or classroom lessons typically involved small groups (3–6 participants), where providing clear explanations to individuals with different backgrounds and perspectives in the same space was more challenging. P5 noted, *"I led a tour for multiple participants, and instructions intended for one person were misunderstood by another, causing confusion."* This highlights the need for personalized guidance even in group-based learning environments.

3.5 Design Requirements

Based on the results of interviews from Section 3.4, we have identified five key design requirements (D1–D5) that are critical for supporting independent tactile exploration by BLV users and deepening their understanding of scientific concepts.

- **D1 Use of Multiple Tactile Contents and Disassemble Models:** By using multiple tactile models for a single theme, BLV users can independently grasp the interrelationships among its various components. Furthermore, incorporating disassemble models provides users with the opportunity to explore internal structures, thereby making abstract concepts more tangible.
- **D2 Choice and Switching of Explanation Methods:** The system should offer users the option to choose between engaging in self-directed exploration with spontaneous questioning and receiving story-based guide, and to switch freely between them during use. In both modes, the system should first provide an overview before delving into detailed explanations.
- **D3 Minimal Intervention and Respect for Autonomy:** To preserve learner autonomy, tactile exploration should involve minimal explanation and intervention. If a learner is confused, subtle prompts such as "Where are you getting lost?" can be used to encourage their sense of control.
- **D4 Option for Guided Assistance:** With minimal intervention, some users may experience uncertainty about which areas to explore. Therefore, the system should provide an optional "directional guidance mode" that can be activated when needed, enabling a balanced approach between self-directed exploration and corrective support.
- **D5 Adaptive Communication Tailored to Individual Characteristics:** Communication should be adapted to each user's characteristics, such as visual abilities, background, and experience. In group instruction settings, special care must be taken to ensure that no participant is left behind.

These design requirements aim to support BLV users in engaging in self-directed tactile exploration, deepening their understanding and interest in scientific themes, and ultimately generating new questions. While our long-term goal is to develop a interactive "Touch and Talk" system that utilizes a camera to track user behavior and provide adaptive feedback, the present study focuses primarily on exploring dynamic timing and interactive communication methods to highlight the benefits of adaptive, dialogue-driven touch interactions, beyond the purely technical aspects such as recognition.

4 STUDY 2: TACTILE EXPLORATION USING A TECHNOLOGY PROBE ON SCIENTIFIC THEMES

To investigate the specific design of an adaptive and interactive "Touch and Talk" system, we conducted a second study with 10 blind and low vision (BLV) participants using a low-fidelity technology probe. While informed by the design requirements identified in Study 1 (Section 3.5), this probe prioritized core features—such as dynamic explanation modes and directional guidance—that were most relevant for probing interaction patterns and investigating user difficulties. Through this exploratory study, we examined how BLV users integrate information across multiple tactile sources, and how such an interactive system can be designed to foster self-directed, meaningful engagement with complex scientific concepts, deepen interest and understanding, and stimulate the generation of new questions.

4.1 Technology Probe Design

4.1.1 Multimodal Content: Scientific Themes and Tactile Models. Drawing on feedback from expert interviews, we selected two complex scientific themes that encompass multiple perspectives, making them difficult to convey through a single tactile model. We chose "Earthquake and Tsunami Mechanisms" and "The Journey of Hayabusa2". Each theme was represented through three interrelated tactile models. The models were designed not only to convey content through touch but also to emphasize distinct learning goals: the Earthquake materials aimed to present multiple conceptual perspectives (e.g., internal structure vs. causality), while the Hayabusa2 materials focused on illustrating relationships between different objects, including enlarged and disassemble parts.

"Earthquake and Tsunami Mechanisms" Theme (hereafter referred to as Earthquake): As shown in Figure 1, this set of tactile models was designed to introduce different views of the phenomenon, including both structural composition and causal mechanisms:

- A **3D-printed tactile globe** with protrusions representing tectonic plates, one-quarter of which was disassemble to allow tactile exploration of its internal layers, created using publicly available design data[4]. (Figure 1 (b))
- A **tactile diagram representing tsunami mechanisms**, constructed with laser-cut overlapping elements to clearly distinguish between tectonic plates and tsunami flows. Different textures represent the Eurasian and Pacific plates[5]. (Figure 1 (c))
- A contrasting **tactile diagram of tsunami sedimentation**: the left side features laser-engraved textures based on an image of real sediments, capturing naturally ambiguous boundaries[6]; the right side uses laser-cut overlapping layers with exaggerated boundaries and varied texture densities to enhance clarity. (Figure 1 (d))

"The Journey of Hayabusa2" Theme (hereafter referred to as Hayabusa2): As shown in Figure 2, this theme aimed to support exploration of object interrelationships, such as between spacecraft components and their journey to the asteroid:

- A **3D printed model of the asteroid Ryugu**, the celestial body visited by Hayabusa2, created using publicly available data[7]. (Figure 2 (b))
- A **3D printed model of the Hayabusa2 spacecraft**, created using publicly available design data[8]. (Figure 2 (c))
- A **3D model of Hayabusa2's re-entry capsule**, the component that returned to Earth. The model is designed to be disassembled in two stages—first, the heat shield can be removed, and then the capsule separates into three parts to reveal the internal sample container. Braille labels and tactile arrows were added to guide the disassembly and reassembly process. (Figure 2 (d))

4.1.2 Adaptive Interaction: Wizard-of-oz Interface. To explore how a future interactive "Touch and Talk" system might support self-directed tactile exploration, we implemented a low-fidelity, voice-based interface (Figure 7) using a Wizard-of-Oz approach. Our long-term goal is to develop a fully autonomous system that leverages camera-based gesture recognition and speech-based interaction to adaptively present information in response to users' movements and verbal input.

In this study, we simulated this vision through a controlled setup that combined pre-recorded audio descriptions (generated using OpenAI's text-to-speech API[9]) and real-time responses (generated using OpenAI's GPT-4 and text-to-speech API). The system was operated on-site by a author acting as the "wizard." The author listened to participants' verbal input, observed their hand movements, and selected appropriate responses. During the experiment, participants were not informed about the Wizard-of-Oz method and were debriefed afterward. This Wizard-of-Oz setup allowed us to simulate a responsive conversational system while avoiding the technical limitations of real-time gesture and speech recognition.

To limit unintended human intervention, we employed a strict can't-answer routine: whenever a participant's action or request was ambiguous, the wizard played the fallback prompt "Sorry, I can't answer your question. Please try another way." If the same ambiguity occurred three times, a second prompt — "Please discuss with the researchers for system improvement." —terminated the session.

To support diverse exploration styles and learning needs, the system followed a three-phase adaptive interaction flow:

(1) **Initial Overview**: The system first introduced all three tactile models for the selected scientific theme, offering a brief explanation of each to help participants choose their starting point.
(2) **Two Exploration Modes**: Participants could choose between two explanation modes and switch between them at any time:
    (a) **Question-Based Interaction**: Participants explored the models freely and asked questions verbally. If a matching response existed in the pre-recorded audio, it was played immediately. Otherwise, the question was forwarded to

[4]https://www.thingiverse.com/thing:1750333

[5]The original tactile graphic was adapted from the tactile book Earthquakes and Tsunamis Through Touch. Because raised-line tactile graphics can be difficult for some BLV individuals to perceive, especially when distinguishing larger areas, we enhanced the tactile clarity using a 2.5D representation created with a laser cutter. Each layer was elevated by 2.5 mm, allowing users to more easily distinguish between elements such as tectonic plates and tsunami flows.

[6]The actual sediment specimens are those exhibited at the science museum. Copyright law permits making these exhibits accessible to visitors with visual impairments and sharing them individually for personal use.

[7]https://www.hayabusa2.jaxa.jp/en/galleries/3D/

[8]https://www.thingiverse.com/thing:3712838

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

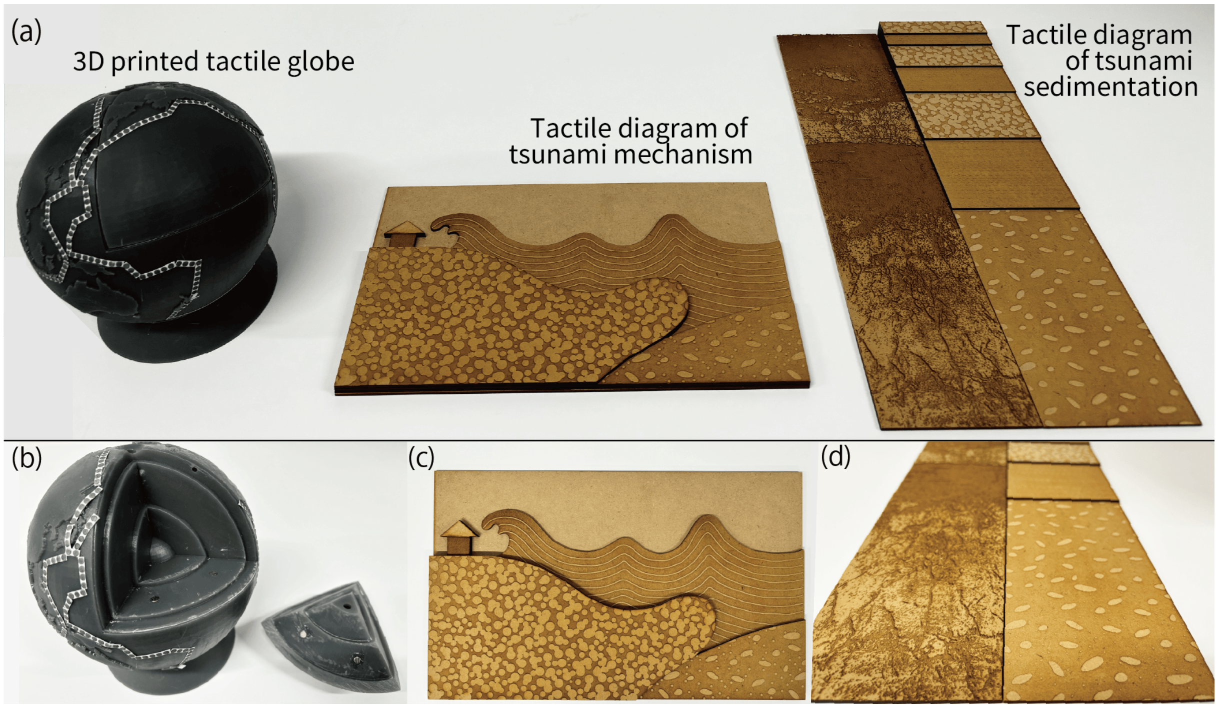

Figure 1: Three models used for the theme “Earthquake and Tsunami Mechanisms.” (a) From left to right: a 3D printed tactile globe, a tactile diagram representing tsunami mechanisms, and a contrasting tactile diagram depicting tsunami sedimentation. (b) The globe can be disassembled; its main body consists of multiple layers. (c) The tactile diagram is overlaid with different textures to facilitate understanding through touch. (d) Both tactile diagrams represent tsunami sedimentation; the left version directly replicates the real object’s features, while the right version incorporates varied textures to emphasize sediment boundaries. The textures and reliefs used in (c) and (d) are solely intended to enhance tactile readability and carry no inherent meaning.

the GPT-4 API. If GPT-4 could not provide a satisfactory answer, the system responded: “Sorry, I cannot answer that well. Please discuss with the researcher for improvement.”

(b) **Storytelling Narration**: The system delivered a sequential, pre-scripted explanation that began with a general overview and gradually introduced details while linking across different models. The narration paused at each step, allowing participants to proceed at their own pace by saying “next.”

The transcript of pre-recorded audio of each theme was detailed in Appendix B.

(3) **(Optional) Directional Guidance**: If participants had difficulty identifying a specific tactile feature, they could activate a guidance mode. In this mode, the system gave directional cues such as “I will start to guide you through your right index finger,” “Move left,” or “A bit more to the left,” helping guide their touch to the correct location.

4.2 Participants

Ten BLV individuals (5 male, 5 female; average age = 47.83, SD = 13.09) were recruited via the organization’s mailing list. Three participants self-identified as having low vision, while seven reported being totally blind. Detailed participant information is provided in Table 1. Eight participants had prior experience receiving explanations through 2D tactile graphics. Seven participants reported having received explanations using 3D tactile models including those depicting yoga body movements, rockets, human anatomy models, and architectural structures. Three participants had experience with explanations that simultaneously employed multiple tactile contents, such as rockets, maps, and human anatomy models. All participants stated that explanations are essential during tactile exploration and expressed a desire to conduct exploration independently, without assistance. Regarding braille literacy, five participants reported that they cannot read braille at all, two stated that they can decode braille slowly, one mentioned that they learned braille in the past but cannot recall it now, and two indicated that they can read braille fluently.

4.3 Procedure

Each session lasted approximately 90 minutes, with breaks offered as needed. The session was conducted individually in a quiet, private room. The recruitment and study procedure were approved by an Institutional Review Board.

4.3.1 Preparation. After signing the consent form and providing demographic information, participants were introduced to the study

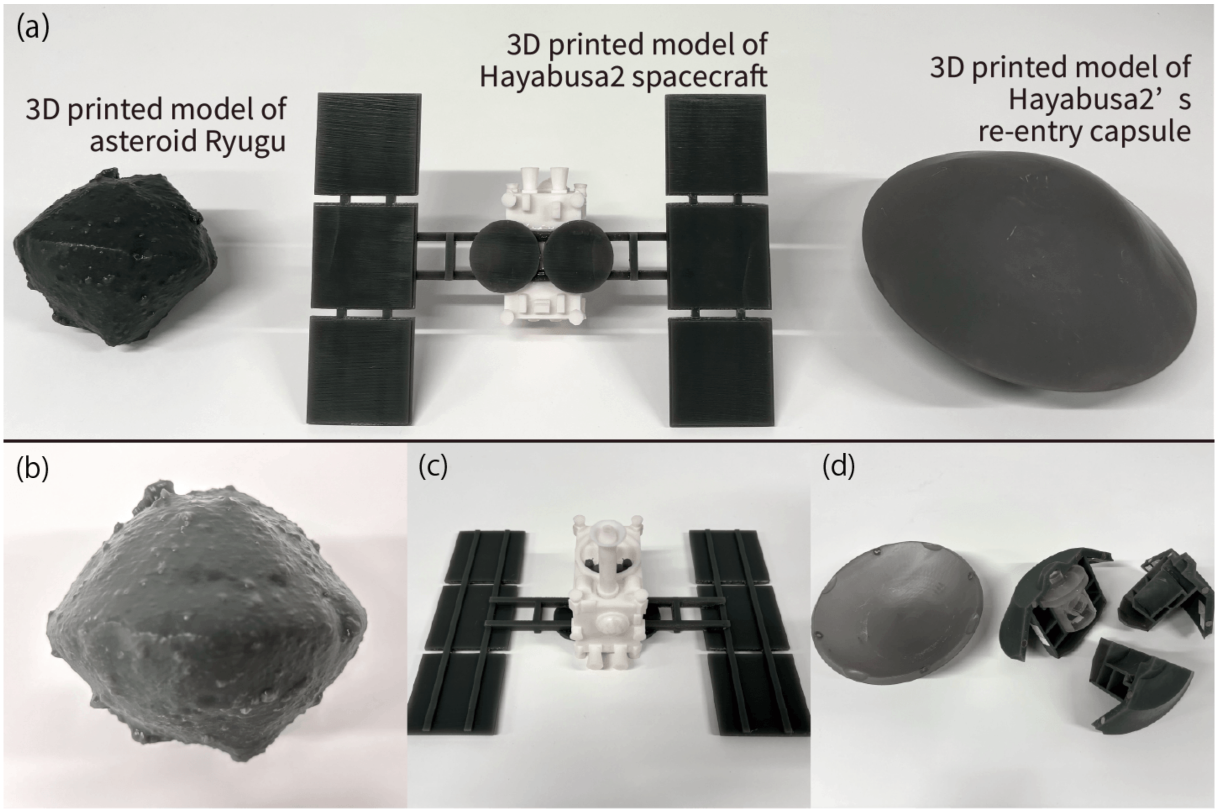

Figure 2: Three models used for the theme "The Journey of Hayabusa2." (a) From left to right: a 3D-printed Asteroid Ryugu, a 3D-printed Hayabusa2 spacecraft, and a 3D-printed re-entry capsule of Hayabusa2. (b) Asteroid Ryugu features textured surfaces that give it a rough, cratered feel when touched. (c) The opposite side of the Hayabusa2 spacecraft. (d) Hayabusa2's re-entry capsule can be disassembled in two stages. First, the capsule splits into two halves (left side of (d)); then it can be further dismantled (right side of (d)). Once disassembled, the chambers A, B, and C which store the parachute and sand can be explored through touch.

goals and given a short demonstration of how to interact with the voice-based system. This briefing session lasted approximately 20 minutes.

4.3.2 Main Study. Each participant explored one of two scientific themes. P1, P3, P5, P7, and P9 explored the Earthquake theme, while P2, P4, P6, P8, and P10 explored the Hayabusa2 theme. For each theme, three tactile models were presented simultaneously, and participants were free to choose which to explore first. Before the start of the session, participants were asked to rate their prior understanding of and interest in the theme using a 7-point Likert scale, from 1 (not at all) to 7 (extremely). To promote purposeful exploration, participants were asked to keep three tactile reasoning tasks in mind as they interacted with the system and models: (1) Describe the overall concept or structure of the tactile models. (2) Explain how the different models relate to anothers. (3) Reflect on whether your understanding of the theme deepened and whether the exploration led you to form new questions.

During the touch exploration session shown in Figure 3, participants interacted with the system to receive guidance and ask questions. If the system could not adequately respond, such as when an answer was unclear or unavailable, the interaction was paused, and the author engaged the participant in a brief discussion. These moments helped clarify participants' needs and informed potential improvements to the system design. The session concluded when the participant had explored all three models, completed the three tasks, and indicated that no further support was needed. This session took around 50 minutes.

4.3.3 Post-Exploration Interview. After completing the tactile exploration tasks, participants took part in a semi-structured interview lasting approximately 20 minutes. The interview covered two main areas: (1) Task reflection and perception: Participants were invited to reflect on the three tactile reasoning tasks and rate their understanding and interest in the theme using a 7-point Likert scale (1 = not at all, 7 = extremely). (2) Feedback on theme-specific support: Participants were prompted to share what aspects of the system were helpful and what could be improved to enhance their understanding and interest in the theme. They were also instructed to compare the experience with that of being guided by a human facilitator.

ID	Field of vision	Visual impairment status	Age	Gender	Experience in 3D models	Experience in 2D models	Experience in multi models	Braille literacy
P1	Low vision (0.05 in both eyes; with a central scotoma)	10 years ago	35	Male	No	No	No	Cannot read at all
P2	Totally blind	Congenital	64	Female	Yes	Yes	No	Cannot read at all
P3	Totally blind (with light perception)	12 years ago	52	Female	Yes	Yes	Yes	Cannot read at all
P4	Totally blind	Onset at age 15	28	Male	Yes	Yes	Yes	Can read and use it fluently
P5	Totally blind	13 years ago	21	Male	Yes	Yes	Yes	Can read and use it fluently
P6	Totally blind (right eye can perceive light and shadows)	Right eye since 1 month old; left eye since age 3	52	Female	No	Yes	No	Can read one character at a time
P7	Low vision (0.02 in left eye)	6 years ago	56	Male	No	No	No	Cannot read at all
P8	Low vision (central field of 1, poor color perception)	5 years ago	52	Female	Yes	Yes	No	Cannot read at all
P9	Totally blind (capable of recognizing silhouettes)	16 years ago	52	Female	Yes	Yes	No	Can read one character at a time
P10	Legally blind (right eye is blind; left eye has light perception)	Congenital	57	Male	Yes	Yes	Yes	Cannot remember

Table 1: Demographic information for the 10 BLV participants in Study 2.

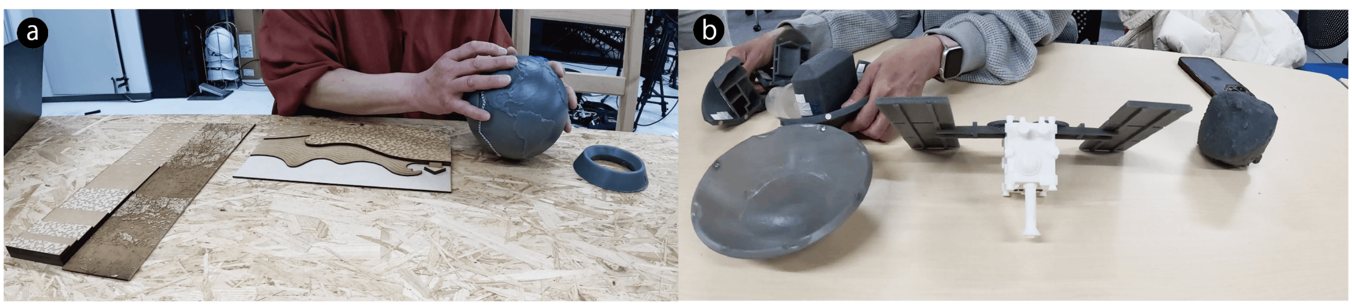

Figure 3: Participants exploring the three tactile contents. (a) A participant in the earthquake theme freely lifted and touched the 3D printed globe model, actively exploring its features. (b) A participant in the Hayabusa2 theme was able to freely disassemble the model.

4.4 Data Collection and Analysis

All sessions were audio and video recorded for later analysis. System logs from the Wizard-of-Oz interface were also collected to capture real-time interactions during tactile exploration. During the tactile exploration, authors identified usability challenges by noting events that disrupted the experience and moments that led to discussions about potential improvements. Audio recordings were transcribed using Notta and manually verified for accuracy. Thematic analysis [6] was conducted on both post-exploration interviews and in-session notes and discussions. Two authors independently coded the transcripts and collaboratively grouped the codes into themes through an iterative comparison process. The analysis followed the procedure described in Section 3.3.

4.5 Results

This section reports findings organized into two categories: (1) interaction patterns and challenges, and (2) the impact on understanding and engagement. The first focuses on how participants engaged with the system, highlighting both effective elements and areas for improvement. The second presents changes in self-reported understanding and interest, along with reflections on how the system

ID	Theme	Understanding Before	Understanding After	Interest Before	Interest After
P1	Earth-quake	2	3	5	5
P2	Hayabusa2	2	5	4	6
P3	Earth-quake	2	3	6	6
P4	Hayabusa2	1	3	5	6
P5	Earth-quake	4	7	7	7
P6	Hayabusa2	2	4	4	7
P7	Earth-quake	2	5	3	4
P8	Hayabusa2	1	5	2	6
P9	Earth-quake	4	6	6	7
P10	Hayabusa2	4	5	7	7

Table 2: Before and after the tactile-exploration experiment, 10 BLV participants rated their understanding and interest in each theme on a 1–7 Likert scale.

can better supported learning. Together, these findings inform the design of an accessible and engaging "Touch and Talk" system.

4.5.1 Interaction Patterns and Challenges. **Mode Selection and Adaptability.** Nine participants (all except P8) chose the storytelling narration and reported that being able to select their preferred mode reduced their cognitive load. P1 stated, *"The storytelling approach deepens my understanding,"* while P2 commented, *"It reduced stress, especially with difficult topics."* During storytelling narration, all participants used voice commands such as "Next" and "Please proceed," and six (P1, P3–P7) provided spontaneous verbal feedback phrases such as "Ah, I see," "Thank you," "Understood," "I didn't know that," and "That's interesting", demonstrating real-time engagement and comprehension.

P8 was the only participant to explore in question base interaction, though they did not explicitly select it. Instead, P8 interrupted the system's initial overview with questions such as *"What is this?"* and continued tactually exploring via further questions. As a result, their inquiries became scattered, leading to a misunderstanding: *"I thought Ryugu was not an independent asteroid but a component attached to Hayabusa2. Only after repeated questioning did I realize my mistake."* P8's behavior reveals system limitations. Participants may start interacting before fully understanding available modes indicating a need for more adaptive and responsive mode switching. While the misunderstanding was ultimately resolved through continued interaction, concerns remain about whether users could correct similar errors independently without precise system support.

Guidance Features and Spatial Language. All participants encountered situations where they did not know which part of the model to touch based on the system's instructions, prompting them to activate the directional guidance mode. For example, when the system said "move to the right," four participants (P2, P4, P6, P7) moved their fingers much farther right than intended. When the system then instructed "move to the left" to correct this, their fingers again deviated significantly. These four participants reported feeling considerable stress when guided through finger movements on the 3D model.

It also proved difficult to guide participants to a specific target on the 3D model using only the index finger. Of the five participants who explored the Hayabusa model, four (all except P10) failed to find the correct component based on the system's directions. When told to "move backward," participants often drifted away from the model and asked in confusion, *"What does 'backward' mean on this model?"* (P4). Similarly, in the re-entry capsule model, P1 and P3 could not correctly identify the internal layers of the disassembled sphere because directing them to use "one finger" was insufficient for small, adjacent regions with overlapping boundaries.

Notably, P9 did not follow the system's instruction to "guide with your right index finger." Instead, P9 held the disassembled portion of the model in their right hand and freely explored the globe's main body with their right middle finger, while their left hand stabilized and rotated the globe. As the model rotated, the system was unable to provide accurate guidance. Only when P9's thumb accidentally landed on the target area and the system responded "That's it" was the correct location confirmed. Figure 4 illustrates P9's actual tactile exploration. Because we employed a Wizard-of-Oz approach, the authors were able to monitor the situation and manually intervene when necessary.

Six participants (P2, P5–P7, P9, P10) suggested that a camera-based system capable of tracking finger position would support independent exploration. P5 commented, *"If my finger covers Japan on the globe and the camera captures the whole Earth, it could show that the hidden area is Japan, letting me identify it without guidance."* P10 added, *"Most of the time I want to explore freely, but I need help when I get stuck. However, I want that help to be something I can handle on my own,"* emphasizing the need for both autonomous tactile exploration and camera-based finger tracking in future systems. These findings suggest the necessity of providing spatial instructions anchored to clear reference points and consistent model orientations. Instructions should dynamically adapt to the user's hand position and support multi-finger guidance when required.

Timing and System Responsiveness During Tactile Exploration. Seven participants (P1–P4, P7, P8, P10) frequently encountered challenges related to the timing of audio guidance during tactile exploration. Many reported that the system's spoken instructions did not align with the rhythm of their hand movements, disrupting concentration and creating hesitation. In particular, delays caused by transcription and GPT-4 response time led to what 5 participants described as "timing stress." As P8 explained, *"I didn't know when the system would talk, so I hesitated to speak"*

Conversely, P5, P6 and P9 attempted to adapt their hand movements to match the system's pacing. For example, P5 slowed down their exploration speed to match the system. These behaviors reflected a strong desire to maintain agency and regulate the interaction flow according to their own rhythm, rather than being constrained by the system's timing.

5 participants (P1–P4, P7) also noted that, unlike human guides who can intuitively adjust their explanations using gaze or gestures, the system lacked contextual responsiveness. Instructions such as "touch here" or "move to the right" were often confusing

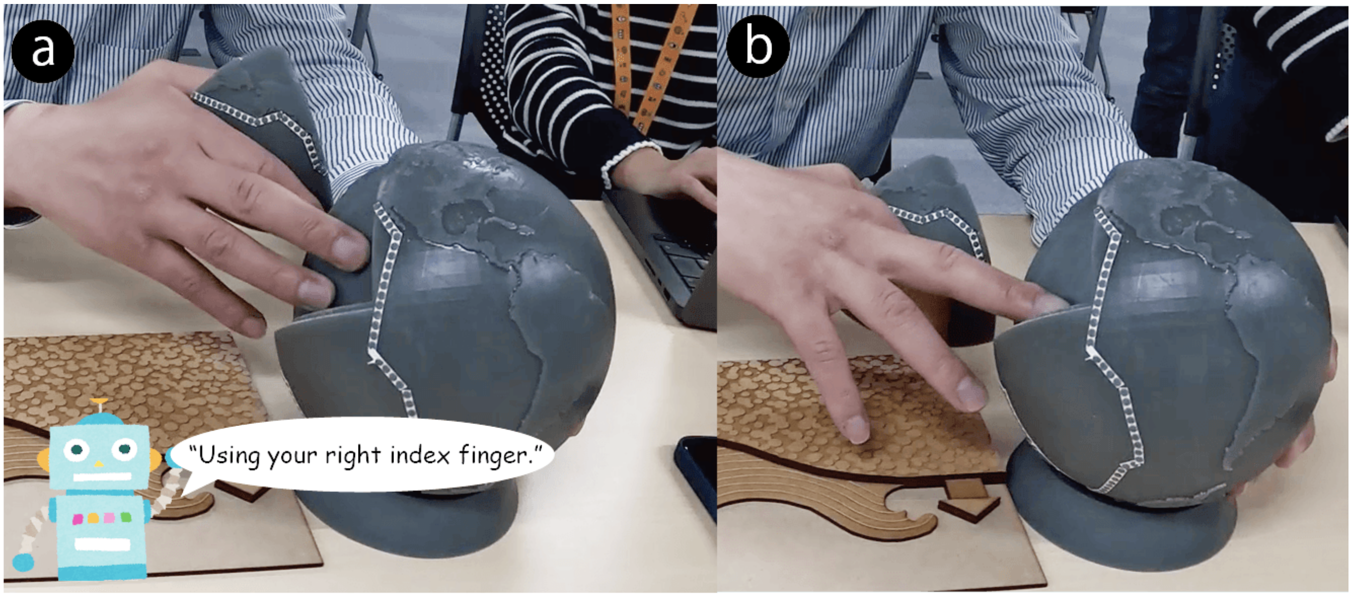

Figure 4: P9's actual tactile exploration. P9 ignored the system's guided mode instructions and conducted a self directed exploration. (a) P9 holds the disassembled segment in the right hand while exploring the globe model with the middle finger. (b) P9 touches the mantle layer of the globe with the right middle finger.

when delivered too early, too late, or without clear spatial reference. These comments highlighted the limitations of audio-only guidance without real-time contextual awareness and underscored the importance of delivering instructions with precise timing and clarity. P7 appreciated the system's consistent and detailed explanations, they emphasized that its inability to adapt to their actions in real time disrupted the flow of exploration and limited sustained engagement.

To improve interaction flow, future systems should synchronize guidance timing with users' tactile behavior. This could involve detecting touch activity or hand movement to better align system responses with user pace. Audio instructions should be concise, context-aware, and delivered at appropriate moments to support uninterrupted, self-directed exploration.

Personalization and Adaptive Explanation. Participants exhibited diverse interaction patterns based on their visual abilities and familiarity with the topic. Three low-vision participants (P1, P7, P8) could generally recognize the total number and layout of the models independently but needed audio support to inspect finer details. As P1 noted, *"I can tell something is on the table, but even when I get closer, I can't see the finer parts."* In contrast, the totally blind participants (P2–P6, P9, P10) required additional structural guidance and spatial cues at the start to understand aspects like model count and arrangement.

Five participants (P2, P6, P8–P10) also valued the ability to ask follow-up questions via the GPT-based interface. P9 said, *"I use AI regularly. If I can ask questions at any time during tactile exploration, I'd prefer the system over asking a person."* This feature allowed for some personalization of question timing and content, helping users explore unfamiliar concepts or clarify specific model details. Moreover, six participants (P1, P3, P6, P8–P10) desired quiz-style prompts or periodic questions from the system. P6 remarked, *"If I'm just listening to an explanation, I can't tell if I truly understand. I might get bored. If the system checked in with me along the way, my comprehension would deepen."* Such prompts can foster self-reflection and correct misunderstandings, and several participants suggested they could also help catch errors when exploring multiple models.

All participants requested more detailed explanations. For instance, P4 said, *"I don't understand technical terms, but I want to hear the correct term first, then ask for an explanation in simpler language."* These suggestions indicate that even in a guided structure, letting users control the pace promotes active, responsive interaction. Additionally, six participants (P1, P2, P7–P10) expressed a desire to adjust the difficulty of explanations based on their prior knowledge during the interaction. This underscores the importance of dynamically adapting language and tailoring content to meet individual user needs.

Explaining scale proved challenging as well. Although the system supplied precise numerical data (e.g., "X kilometers"), five participants (P3, P4, P8–P10) found it hard to form meaningful mental images. P8 even misheard "kilometers" as "meters" and believed the asteroid was physically part of the Hayabusa2 spacecraft. To aid conceptual understanding, P3 suggested using comparative expressions like "about three soccer balls in size" or "several Tokyo Domes worth."

These results highlight the necessity for a more personalized tactile exploration system. Future systems should include an onboarding process to assess users' topic familiarity and let them set their preferred level of explanation complexity. Moreover, explanations should adjust dynamically and continuously in response to users' behaviors, questions, and knowledge levels throughout the interaction. Scale should be communicated through familiar, concrete analogies. Additionally, the system should integrate tactile input with real-time, context-aware dialogue to support a truly multimodal understanding.

Confidence and Clarity for Operating Disassemblable Models. All participants responded positively to the opportunity to interact with disassemblable models, showing interest in exploring how internal components fit together. However, all participants regardless of visual condition faced difficulties when attempting to disassemble the models. The instruction "this model can be disassembled" was perceived as vague and insufficient. Three participants (P3, P6 and P8) expressed uncertainty about how much force they could safely apply or which direction the parts should move. As P6 remarked, *"I was afraid I might break it and didn't know how much force I could apply."*

As shown in Figure 5, the Hayabusa2 re-entry capsule was equipped with embossed braille and arrow markings on its halved interior components to indicate the correct reassembly positions. However, these proved ineffective for all users. Five participants (P1–P3, P7 and P8) could not read braille. However, P4 commented, *"There's a risk I might mistake the braille bumps for a decorative feature."* Figure 6 illustrates P4 locating and touching the braille. However, without an initial system prompt, they might mistake it for decoration highlighting that guidance methods relying solely on braille or visual cues do not serve all BLV users.

All participants indicated that they expected more explicit, step-by-step guidance when interacting with removable components. They anticipated verbal instructions that would clearly describe where to place their hands, how to move the parts (e.g., pull forward, twist, slide), and what kind of resistance to expect. Without such instructions, they were hesitant to fully explore the models. These findings suggest that to support safe, confident, and independent interaction with decomposable tactile models, systems must provide multimodal guidance. Verbal instructions should describe precise actions, supported by non-visual feedback mechanisms. Additionally, models should incorporate intuitive tactile and visual cues to accommodate both blind and low vision users.

Understanding Model Relationships Participants reported that interacting with multiple tactile models helped them better understand the themes. Four participants (P1, P5, P8 and P10) reported that interacting with multiple models increased their curiosity and prompted them to seek additional information. P1 reflected that realizing their own desire to explore further was, in itself, a meaningful outcome.

However, participants differed in their ability to integrate multiple models into a coherent mental representation. P4, who explored the Hayabusa2 theme, initially misinterpreted the relationships among the models because Ryugu was introduced first in the narrative, they assumed it was a component of the Hayabusa2 spacecraft. However, after manipulating the disassembled models with guided assistance, P4 corrected this misunderstanding and developed a more accurate understanding of each individual component.

P5, who explored the earthquake theme, was able to understand the entire sequence from plate movement through tsunami generation to sediment layer formation by examining each model in the order presented by the system's explanation. P5 noted that this step-by-step interaction helped clarify the conceptual links between the models.

However, P1 and P3 were only able to partially understand the relationships. While they grasped the connection between the tactile globe and the tactile diagram representing tsunami mechanisms, the link between these phenomena and the tactile diagram of tsunami sedimentation remained unclear. They suggested that improvements are needed in how the sequence and logical relationships between tactile models are communicated. Morever, P1 and P10 expressed confusion when explanations lacked clarity in scale and oreitation relationships. For example, whether a model was being described in cross-section or plan view—highlighting the need for better framing of spatial information.

The findings suggest that to support integrated understanding, future systems should clearly sequence model presentation, reinforce conceptual relationships, and provide consistent cues about spatial perspective and scale. Narration should explicitly explain how models connect, helping users form a coherent and meaningful mental map.

4.5.2 Impact on Understanding and Engagement. Before conducting the tactile exploration, participants were asked to rate their level of understanding and interest on a 7-point scale both before and after experiments on two different themes. The results are shown in Table **??**.

Earthquake Theme: Participants showed improved understanding, with average scores increasing from 2.8 to 4.8. Tactile exploration supported comprehension of complex concepts such as tectonic plates, tsunami mechanisms, and sedimentation processes. P7 said, *"I learned that there are many plates around Japan. That's why so many earthquakes occur."* P9 reflected, *"On TV, they only say an earthquake occurred... But through tactile exploration, I was truly able to understand that it's due to the collision of tectonic plates."* These responses highlight how tactile interaction helped ground abstract phenomena in more concrete, relatable experiences.

Interest also increased slightly, with average scores rising from 5.4 to 5.8. P5, who rated interest as 7 both before and after, commented, *"If I could have selected a score higher than 7, I would have."* P1, connecting the topic to personal experience, shared, *"I can't see tsunamis, but through tactile exploration, I was able to learn how terrifying a tsunami really is."* These findings suggest that linking tactile models to personal knowledge and lived experiences can enrich emotional engagement and sustained interest.

Hayabusa2 Theme: Participants began with limited prior knowledge of the Hayabusa2 mission. Average understanding scores increased significantly from 2.0 to 4.4, with P4 and P8 initially rating their knowledge as "1" (no knowledge) and most reporting only brief familiarity through television. The disassemble re-entry capsule model was a key highlight. Participants commented that being able to physically take apart the capsule helped them understand its internal structure. P4 remarked, *"So the collected sand is stored in this part,"* while tracing the sample container. P6 noted, *"I was surprised by how many craters there were,"* and P10 added, *"I was shocked to discover its actual shape. I had imagined asteroids to be smooth and round."*

Interest scores increased from 4.4 to 6.4. Notably, P10 asked spontaneous follow-up questions such as, *"Where did the re-entry capsule land with its parachute?"* and *"Is that place accessible to people?"* after the session ended. They shared, *"Touching multiple*

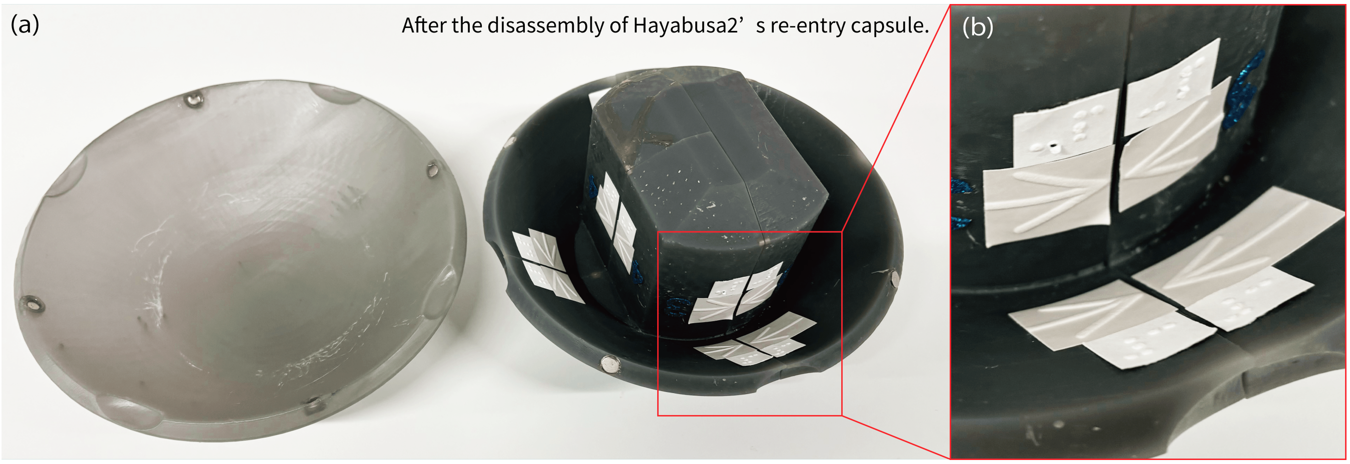

Figure 5: Figure of the Hayabusa2 re-entry capsule disassembly. (a) The capsule is split in half. The model on the right can be further separated. Embossed arrows and braille labels have been added to guide BLV users in reassembling the parts, the arrows indicate the correct orientation for joining. (b) Close-up of the embossed arrows and braille markings.

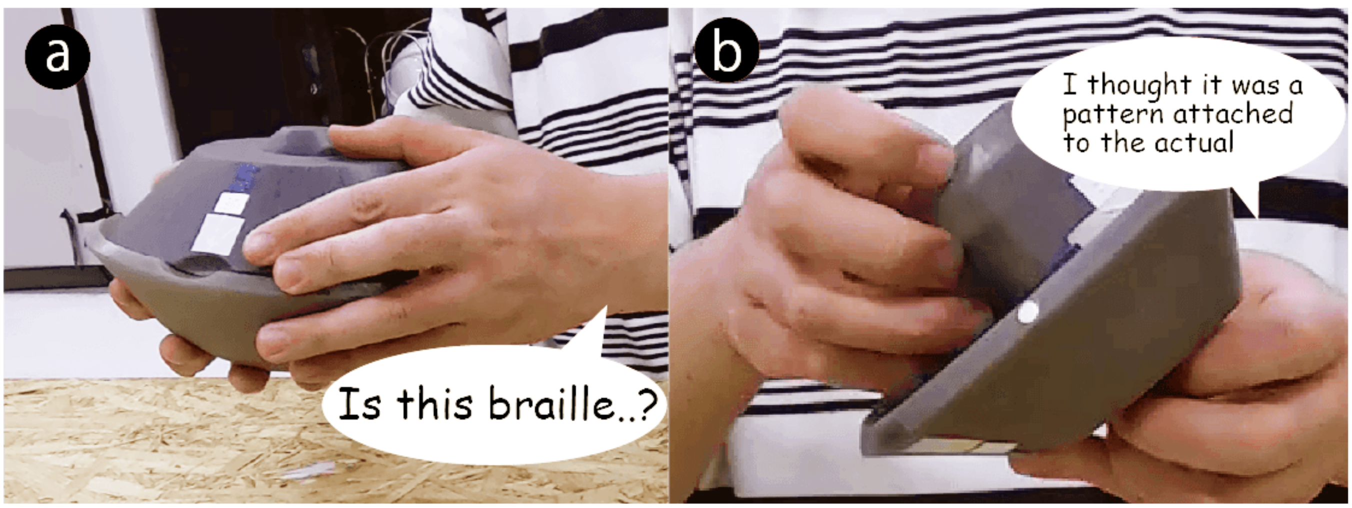

Figure 6: P4 locating and exploring the Braille on the model. (a) P4 vocalizes "Is this braille?" while touching the raised dots. (b) P4 then reports, "I thought this was a decorative pattern on the capsule, not braille," revealing that when Braille or embossing is applied to a model, clear notification is essential. It also underscores that braille-based guidance cannot accommodate users who cannot read braille.

models helped me understand the connections, and now I want to learn even more. I'm really happy about that." These findings highlight that unfamiliar themes can particularly benefit from rich, multi-model tactile exploration to foster curiosity, emotional connection, and active inquiry.

Across both themes, tactile exploration led to measurable increases in understanding and interest. In familiar topics earthquakes, the models helped reinforce and personalize existing knowledge. In less familiar themes such as Hayabusa2, they served as an entry point for curiosity and independent learning. The results underscore the value of multi-model tactile systems in deepening both conceptual grasp and motivation to explore further. At the same time, careful consideration must be given to how models are ordered, how relationships are explained, and how scale and perspective are conveyed to ensure that connections are clear and meaningful.

5 DISCUSSION

5.1 Reflections on the Study's Design Strategy and Key Findings

This study explored how to design tactile exploration systems that support conceptual learning, curiosity, and autonomy for BLV users. Building on findings from our Study 1, which examined (1) tactile content for conveying abstract concepts, (2) instructional strategies for tactile exploration, and (3) strategies for communication and adaptations based on learner needs, we developed "Touch and Talk" system, a set of low-fidelity technology probes to investigate how BLV users engage with multiple tactile models and audio-based explanations.

Our system design focused on two key themes: (1) providing multiple tactile models to represent complex scientific themes, and (2) enabling users to dynamically control how information is delivered through question base interaction and storytelling narration. These probes were tested in study 2, controlled sessions with 10 BLV participants, revealing not only how different users made sense of the tactile content, but also how their needs diverged in relation to explanation complexity, spatial guidance, and model manipulation.

Through Study 2, both the strengths and limitations of our approach were revealed. Participants demonstrated an ability to form an integrated understanding when multiple tactile models were clearly ordered and their relationships were mentally linked. However, confusion arose when the system did not clearly convey information regarding spatial orientation, model positioning, terminology, and scaling.

The feature allowing users to toggle between storytelling and question-based modes was generally well received; however, discrepancies between the timing of the system's explanations and the users' hand movements during tactile exploration led to reduced engagement for some. Similarly, while all participants agreed that using detachable models enhanced comprehension, the lack of specific guidance for disassembling and reassembling the models resulted in hesitations and frustration. These findings underscore the importance of designing the Touch and Talk system with adaptive, clear, and multimodal support—especially when conveying abstract or unfamiliar content. The next section outlines six design recommendations for developing future Touch and Talk systems that are more comprehensive, effective, and user-driven.

5.2 Design Implications for the Future "Touch and Talk" System

Based on the findings and observations in this study, we identify six key design implications for future tactile exploration systems aimed at supporting BLV users in engaging with complex scientific content. They are arranged from broad to specific and aligned with the user journey, starting from pre-exploration setup through to end-of-exploration reflection.

- **Enable Dynamic Personalization Through Onboarding and Real-Time Adaptation.** Our findings reaffirm the importance of personalized guidance for BLV users, echoing prior work [60], and identify three key needs for science learning: terminology preservation, user-driven customization, and dynamic adaptation. Participants preferred hearing accurate technical terms, with the option to request simplified explanations. Choosing explanation detail and pace helped deepen understanding and supported learner autonomy. A customizable onboarding step—where users indicate visual condition and topic familiarity (e.g., beginner to advanced)—enables the system to tailor content accordingly, consistent with recent findings [59, 61]. Study 2 further discovered the value of real-time and dynamic adaptation based on user behavior and comprehension, enhancing both learning outcomes and learner agency. We suggest that future tactile systems integrate onboarding with real-time, conversational adaptation. Personalization that is continuous and context-aware can enable more engaging and effective science communication for BLV learners.
- **Use Analogies and Descriptive Language to Convey Scale and Model Features.** Many participants struggled to visualize abstract numerical dimensions (e.g., "X kilometers"). To support mental modeling, the system should complement precise measurements with relevant analogies (e.g., "about as large as a soccer field"). This approach helps anchor abstract concepts within a familiar frame of reference. Furthermore, by using vivid and descriptive language that emphasizes the relationship between space and tactility, abstract scientific concepts can be made more concrete and easier to understand. Phrases such as "a surface with uneven textures resembling a crater" help users associate verbal explanations with tactile input. The need for easily comprehensible language is consistent with [59].
- **Real-time Position Awareness Through Hand Tracking.** Accurate spatial guidance remains a central challenge in tactile exploration, and purely vision-based hand tracking suffers when users' fingers occlude the model. To mitigate this limitation, an autonomous system should fuse complementary sensing strategies. One avenue is to maintain a virtual proxy of the model anchored by an AR marker that stays visible to an external camera [68]; the proxy preserves model pose even when parts are hidden behind the user's hands. A second, occlusion-robust option is to reposition the camera onto the user. For example, finger-mounted or ring-mounted cameras that simultaneously capture fingertip location and points of contact [22]. For surface details too small for reliable computer vision, we envisage embedding low-cost pressure, capacitive, or Hall-effect sensors directly into printed components. In parallel, adding CV-friendly landmarks—distinct colours, fiducial labels, or magnetic markers can simplify part recognition; colour-coded blocks have already proven effective for assembly guidance [52]. By combining these modalities, the system could deliver context-sensitive feedback such as "Your thumb and index finger are touching the edge of the solar panel." This approach helps reduce confusion, improve the accuracy of guidance, and support users in confidently exploring complex 3D models.
- **Adaptive Exploration Assistance Based on Tactile Behavior.** To maintain flow and reduce "timing stress," the system must dynamically adjust its pace and the granularity of its guidance based on tactile input. For example, if a user stops interacting for an extended period, the system might provide follow-up information or ask, "Would you like to learn more about this section?" This approach fosters a sense of agency and promotes interactive engagement rather than passive instruction. The need for user-driven tactile exploration support is consistent with [59].

- **Multimodal Feedback for Manipulating Disassembled Models.** The use of disassembled models is recommended in [60], but participants expressed concerns that handling such models might involve applying too much force or damaging components, leading to hesitation. To address this issue, the system should provide clear, step-by-step audio instructions that specify precise actions (e.g., "Hold the base with your left hand and pull the top upward with your right hand"), while anticipating users' concerns regarding movement and resistance. These instructions should be reinforced with non-visual feedback such as tactile click sounds or vibrations to confirm successful disassembly or assembly. Additionally, labels should incorporate clear tactile features (e.g., texture, raised shapes) and high-contrast visual elements to support both blind and low vision users, rather than relying solely on Braille.
- **Promote Reflection Through Quizzes and Recaps.** Several participants expressed a desire to verify their understanding during or after exploration. Prompting users with reflective questions (e.g., "Where was the sample capsule located?") or short quizzes can reinforce learning, clarify relationships between models, and encourage deeper curiosity—particularly in open-ended, self-guided environments.

6 LIMITATIONS AND FUTURE WORK

While this study yielded valuable insights into the design of tactile-exploration systems for BLV users, several limitations and corresponding research directions remain.

Individual vs. Group Use. Our investigation centred on single-user exploration. Although Study 1 revealed communication needs in group settings, Study 2 did not examine shared interaction dynamics. Future work should analyse how multiple BLV users coordinate turn-taking, engage in collaborative inquiry, and cope with differences in pace and understanding.

Sample Size and Participant Diversity. Study 2 involved ten participants—appropriate for an exploratory technology-probe study, yet insufficient for broad generalisation. Replication with a larger and more demographically diverse cohort is required to confirm robustness and uncover edge-case behaviours.

Wizard-of-Oz Prototype. The Wizard-of-Oz protocol allowed us to investigate the prospective capabilities of the "Touch and Talk system". However, it differs from full automation in two critical respects. First, because a human "wizard" interpreted ambiguous user actions, potential breakdowns that an autonomous system would expose may have been obscured, even though we employed the strict can't-answer routine described in Section 4.1.2. Second, the wizard's response latency and precision exceeded what current sensing and dialogue modules can deliver, particularly for fine-grained part recognition. Consequently, future work should evaluate real-time prototypes that integrate multimodal sensing—for example, finger-mounted cameras or embedded pressure sensors—with adaptive dialogue to test reliability and mode switching without human intervention.

Roadmap Toward an Autonomous System. To bridge our high-level design guidelines and low-level implementation challenges, we propose a three-stage roadmap that also accommodates vision-based capabilities. In Step 1, replaces the wizard with webcam-based finger-tap and coarse-gesture detection to validate timing and conversational flow. Step 2 introduces part recognition and spatial mapping by combining color-coded or sensor-embedded components with emerging vision models such as GPT-4V, enabling richer audio feedback and precise spatial references even under hand occlusion. Step 3 logs user behavior to drive adaptive prompts, refine feedback granularity, and iteratively converge on reliability that matches or surpasses the Wizard-of-Oz baseline. Integrating large-scale vision models in this pipeline may also support cross-model comparison and higher spatial accuracy, features that current technical constraints prevented us from evaluating.

Narrative Structure and Model Design. Participants explored three tactile models per theme, but we did not systematically vary narrative sequencing, abstraction level, or physical layout. Future research should examine how these design choices influence conceptual understanding, cognitive load, and long-term knowledge retention—especially for unfamiliar STEM topics.

Authorship and Support for Scientific Tactile Content Creation. Future work will focus on an authoring framework that brings together the professionals who already serve BLV learners. We will begin with a qualitative study of their current workflows, challenges, and envisioned use cases. Insights from this study will inform the design of authoring tools for rapid 3D model prototyping, semi-automatic audio annotation, and shared version control. Integrating these tools into our three-stage roadmap (baseline dialogue, part recognition and spatial mapping, adaptive feedback) should enable Touch and Talk to scale beyond the laboratory and sustain a rigorously curated, accessibility-focused library of tactile learning materials.

These limitations outline concrete opportunities for extending our work: scaling to larger and group-based studies, replacing the wizard with robust multimodal sensing, leveraging state-of-the-art vision models, and refining tactile narratives across diverse learning contexts.

7 CONCLUSION

This study investigated the impact of interactive dialogue and temporal dynamics on tactile exploration among BLV users, particularly within the context of science communication. First, we conducted semi-structured interviews with 22 tactile guidance experts to elucidate strategies for tactile explanations and interventions in various settings. In addition, we employed a voice-based "Touch and Talk" system in a Wizard-of-Oz experiment with 10 BLV participants to examine tactile exploration within scientific themes.

The experimental results revealed that participants emphasized the effectiveness of an explanation approach that allows users to switch between storytelling narration and question base interaction on their own. Furthermore, the findings confirmed that using multiple disassemblable models within a single theme facilitates an easier understanding of the relationships among the models. Moreover, adaptive and context-sensitive communication strategies not only enhanced the comprehension of abstract and spatially complex scientific topics but also contributed to maintaining interest and stimulating new questions.

These results suggest that flexible and adaptive system designs are essential for BLV users to autonomously explore scientific phenomena and develop a more comprehensive understanding. Future system development should build on the insights obtained in this study by incorporating personalized guidance tailored to users' visual capabilities and learning characteristics, interactive feedback, dynamic adjustments of timing and content, as well as camera-based tracking functionality.

Acknowledgments

We gratefully acknowledge all participants who generously contributed their time and insights to this study. We also extend our sincere thanks to the team at Sony Semiconductor Solutions for their invaluable assistance in coordinating and supporting the experimental sessions.

References

[1] Nazatul Naquiah Abd Hamid and Alistair D.N. Edwards. 2013. Facilitating route learning using interactive audio-tactile maps for blind and visually impaired people. In CHI '13 Extended Abstracts on Human Factors in Computing Systems (Paris, France) (CHI EA '13). Association for Computing Machinery, New York, NY, USA, 37–42. doi:10.1145/2468356.2468364

[2] Saki Asakawa, João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. 2018. The Present and Future of Museum Accessibility for People with Visual Impairments. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (Galway, Ireland) (ASSETS '18). Association for Computing Machinery, New York, NY, USA, 382–384. doi:10.1145/3234695.3240997

[3] Atheer Awad, Aliya Yao, Sarah J Trenfield, Alvaro Goyanes, Simon Gaisford, and Abdul W Basit. 2020. 3D printed tablets (printlets) with braille and moon patterns for visually impaired patients. Pharmaceutics 12, 2 (2020), 172.

[4] Catherine M. Baker, Lauren R. Milne, Jeffrey Scofield, Cynthia L. Bennett, and Richard E. Ladner. 2014. Tactile graphics with a voice: using QR codes to access text in tactile graphics. In Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility (Rochester, New York, USA) (ASSETS '14). Association for Computing Machinery, New York, NY, USA, 75–82. doi:10.1145/2661334.2661366

[5] Hugh Beyer and Karen Holtzblatt. 1999. Contextual design. interactions 6, 1 (1999), 32–42.

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

[7] Anke Brock, Philippe Truillet, Bernard Oriola, Delphine Picard, and Christophe Jouffrais. 2012. Design and user satisfaction of interactive maps for visually impaired people. In Computers Helping People with Special Needs: 13th International Conference, ICCHP 2012, Linz, Austria, July 11-13, 2012, Proceedings, Part II 13. Springer, 544–551.

[8] Anke M Brock, Philippe Truillet, Bernard Oriola, Delphine Picard, and Christophe Jouffrais. 2015. Interactivity improves usability of geographic maps for visually impaired people. Human–Computer Interaction 30, 2 (2015), 156–194.

[9] Craig Brown and Amy Hurst. 2012. VizTouch: automatically generated tactile visualizations of coordinate spaces. In Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction. 131–138.

[10] Harold Burton, Abraham Z Snyder, Thomas E Conturo, Erbil Akbudak, John M Ollinger, and Marcus E Raichle. 2002. Adaptive changes in early and late blind: a fMRI study of Braille reading. Journal of neurophysiology 87, 1 (2002), 589–607.

[11] Matthew Butler, Leona Holloway, Kim Marriott, and Cagatay Goncu. 2017. Understanding the graphical challenges faced by vision-impaired students in Australian universities. Higher Education Research & Development 36, 1 (2017), 59–72.

[12] Luis Cavazos Quero, Jorge Iranzo Bartolomé, and Jundong Cho. 2021. Accessible visual artworks for blind and visually impaired people: comparing a multimodal approach with tactile graphics. Electronics 10, 3 (2021), 297.

[13] Kathy Charmaz. 2006. Constructing grounded theory: A practical guide through qualitative analysis. Sage.

[14] Elyse D. Z. Chase, Alexa Fay Siu, Abena Boadi-Agyemang, Gene S-H Kim, Eric J Gonzalez, and Sean Follmer. 2020. PantoGuide: A Haptic and Audio Guidance System To Support Tactile Graphics Exploration. In Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility (Virtual Event, Greece) (ASSETS '20). Association for Computing Machinery, New York, NY, USA, Article 77, 4 pages. doi:10.1145/3373625.3418023

[15] Helen Chatterjee. 2020. Touch in museums: Policy and practice in object handling. Routledge.

[16] Helen Chatterjee, Sonjel Vreeland, and Guy Noble. 2009. Museopathy: Exploring the healing potential of handling museum objects. Museum & Society 7, 3 (2009), 164–177.

[17] James Coughlan, Huiying Shen, and Brandon Biggs. 2020. Towards Accessible Audio Labeling of 3D Objects. Journal on technology and persons with disabilities : ... Annual International Technology and Persons with Disabilities Conference 8 (01 2020), 210–222.

[18] Josh Urban Davis, Te-Yen Wu, Bo Shi, Hanyi Lu, Athina Panotopoulou, Emily Whiting, and Xing-Dong Yang. 2020. TangibleCircuits: An Interactive 3D Printed Circuit Education Tool for People with Visual Impairments. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376513

[19] Fabio D'Agnano, Caterina Balletti, Francesco Guerra, Paolo Vernier, et al. 2015. Tooteko: A case study of augmented reality for an accessible cultural heritage. Digitization, 3D printing and sensors for an audio-tactile experience. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 40, 5W4 (2015), 207–213.

[20] Polly Edman. 1992. Tactile graphics. American Foundation for the Blind.

[21] Alistair DN Edwards. 1989. Soundtrack: An auditory interface for blind users. Human–Computer Interaction 4, 1 (1989), 45–66.

[22] Leah Findlater, Lee Stearns, Ruofei Du, Uran Oh, David Ross, Rama Chellappa, and Jon Froehlich. 2015. Supporting Everyday Activities for Persons with Visual Impairments Through Computer Vision-Augmented Touch. In Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (Lisbon, Portugal) (ASSETS '15). Association for Computing Machinery, New York, NY, USA, 383–384. doi:10.1145/2700648.2811381

[23] Giovanni Fusco and Valerie S. Morash. 2015. The Tactile Graphics Helper: Providing Audio Clarification for Tactile Graphics Using Machine Vision. In Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (Lisbon, Portugal) (ASSETS '15). Association for Computing Machinery, New York, NY, USA, 97–106. doi:10.1145/2700648.2809868

[24] John A Gardner. 2002. Access by blind students and professionals to mainstream math and science. In International conference on computers for handicapped persons. Springer, 502–507.

[25] Giuseppe Ghiani, Barbara Leporini, and Fabio Paternò. 2009. Vibrotactile feedback to aid blind users of mobile guides. Journal of Visual Languages & Computing 20, 5 (2009), 305–317.

[26] Uttara Ghodke, Lena Yusim, Sowmya Somanath, and Peter Coppin. 2019. The Cross-Sensory Globe: Participatory Design of a 3D Audio-Tactile Globe Prototype for Blind and Low-Vision Users to Learn Geography. In Proceedings of the 2019 on Designing Interactive Systems Conference (San Diego, CA, USA) (DIS '19). Association for Computing Machinery, New York, NY, USA, 399–412. doi:10.1145/3322276.3323686

[27] Stéphanie Giraud, Anke M Brock, Marc J-M Macé, and Christophe Jouffrais. 2017. Map learning with a 3D printed interactive small-scale model: Improvement of space and text memorization in visually impaired students. Frontiers in psychology 8 (2017), 930.

[28] Daniel Goldreich and Ingrid M Kanics. 2003. Tactile acuity is enhanced in blindness. Journal of Neuroscience 23, 8 (2003), 3439–3445.

[29] Reginald G Golledge. 1993. Geography and the disabled: a survey with special reference to vision impaired and blind populations. Transactions of the institute of British geographers (1993), 63–85.

[30] Timo Götzelmann. 2016. LucentMaps: 3D Printed Audiovisual Tactile Maps for Blind and Visually Impaired People. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (Reno, Nevada, USA) (ASSETS '16). Association for Computing Machinery, New York, NY, USA, 81–90. doi:10.1145/2982142.2982163

[31] T. Götzelmann. 2018. Visually Augmented Audio-Tactile Graphics for Visually Impaired People. ACM Trans. Access. Comput. 11, 2, Article 8 (June 2018), 31 pages. doi:10.1145/3186894

[32] Alfons O Hamm, Almut I Weike, Harald T Schupp, Thomas Treig, Alexander Dressel, and Christof Kessler. 2003. Affective blindsight: intact fear conditioning to a visual cue in a cortically blind patient. Brain 126, 2 (2003), 267–275.

[33] Liang He, Cheng Xu, Ding Xu, and Ryan Brill. 2015. PneuHaptic: delivering haptic cues with a pneumatic armband. In Proceedings of the 2015 ACM international symposium on wearable computers. 47–48.

[34] Morton A Heller and Edouard Gentaz. 2013. Psychology of touch and blindness. Psychology press.

[35] Marion A Hersh and Michael A Johnson. 2008. Assistive technology for visually impaired and blind people. Vol. 1. Springer.

[36] Leona Holloway, Matthew Butler, Alex Waddell, and Kim Marriott. 2025. 3D Printing for Accessible Education: A Case Study in Assistive Technology Adoption. In Proceedings of the 2025 CHI Conference on Human Factors in Computing

Systems (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 273, 19 pages. doi:10.1145/3706598.3713689

[37] Leona Holloway, Kim Marriott, and Matthew Butler. 2018. Accessible Maps for the Blind: Comparing 3D Printed Models with Tactile Graphics. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3173772

[38] Jorge Iranzo Bartolome, Luis Cavazos Quero, Sunhee Kim, Myung-Yong Um, and Jundong Cho. 2019. Exploring Art with a Voice Controlled Multimodal Guide for Blind People. In Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction (Tempe, Arizona, USA) (TEI '19). Association for Computing Machinery, New York, NY, USA, 383–390. doi:10.1145/3294109.3300994

[39] Lynette A Jones and Nadine B Sarter. 2008. Tactile displays: Guidance for their design and application. Human factors 50, 1 (2008), 90–111.

[40] Shaun K. Kane and Jeffrey P. Bigham. 2014. Tracking @stemxcomet: teaching programming to blind students via 3D printing, crisis management, and twitter. In Proceedings of the 45th ACM Technical Symposium on Computer Science Education (Atlanta, Georgia, USA) (SIGCSE '14). Association for Computing Machinery, New York, NY, USA, 247–252. doi:10.1145/2538862.2538975

[41] Shaun K. Kane, Jeffrey P. Bigham, and Jacob O. Wobbrock. 2008. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility (Halifax, Nova Scotia, Canada) (Assets '08). Association for Computing Machinery, New York, NY, USA, 73–80. doi:10.1145/1414471.1414487

[42] Shaun K Kane, Jacob O Wobbrock, and Richard E Ladner. 2011. Usable gestures for blind people: understanding preference and performance. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 413–422.

[43] Akif Khan and Shah Khusro. 2021. An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities. Universal Access in the Information Society 20, 2 (2021), 265–298.

[44] Barbara Landau, Lila R Gleitman, and Barbara Landau. 2009. Language and experience: Evidence from the blind child. Vol. 8. Harvard University Press.

[45] Steven Landau and Lesley Wells. 2003. Merging tactile sensory input and audio data by means of the Talking Tactile Tablet. In Proceedings of EuroHaptics, Vol. 3. 414–418.

[46] Franklin Mingzhe Li, Lotus Zhang, Maryam Bandukda, Abigale Stangl, Kristen Shinohara, Leah Findlater, and Patrick Carrington. 2023. Understanding Visual Arts Experiences of Blind People. 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 60, 21 pages. doi:10.1145/3544548.3580941

[47] Eric MacDonald and Ryan Wicker. 2016. Multiprocess 3D printing for increasing component functionality. Science 353, 6307 (2016), aaf2093.

[48] Nasif Mahmud, RK Saha, RB Zafar, MBH Bhuian, and Shahzad S Sarwar. 2014. Vibration and voice operated navigation system for visually impaired person. In 2014 international conference on informatics, electronics & vision (ICIEV). IEEE, 1–5.

[49] Samantha McDonald, Joshua Dutterer, Ali Abdolrahmani, Shaun K. Kane, and Amy Hurst. 2014. Tactile aids for visually impaired graphical design education. In Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility (Rochester, New York, USA) (ASSETS '14). Association for Computing Machinery, New York, NY, USA, 275–276. doi:10.1145/2661334.2661392

[50] Jennifer M Morris, Peter L Dumble, and M Ramsay Wigan. 1979. Accessibility indicators for transport planning. Transportation Research Part A: General 13, 2 (1979), 91–109.

[51] Ruth G Nagassa, Matthew Butler, Leona Holloway, Cagatay Goncu, and Kim Marriott. 2023. 3D Building Plans: Supporting Navigation by People who are Blind or have Low Vision in Multi-Storey Buildings. 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 539, 19 pages. doi:10.1145/3544548.3581389

[52] Ruth Galan Nagassa, Andre Ky Pham, Matthew Butler, Leona Holloway, Kalin Stefanov, Skye de Vent, and Kim Marriott. 2025. Enhancing Tactile Learning: A Co-Designed System for Supporting Speech Interaction with Multi-Part 3D Printed Models by Students who are Blind. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 294, 18 pages. doi:10.1145/3706598.3713706

[53] National Federation of the Blind. 2009. Braille Literacy Report. Online. Retrieved from https://nfb.org/images/nfb/documents/pdf/braille_literacy_report_web.pdf.

[54] Mark Paterson. 2020. The senses of touch: Haptics, affects and technologies. Routledge.

[55] Andreas Reichinger, Anton Fuhrmann, Stefan Maierhofer, and Werner Purgathofer. 2016. Gesture-Based Interactive Audio Guide on Tactile Reliefs. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (Reno, Nevada, USA) (ASSETS '16). Association for Computing Machinery, New York, NY, USA, 91–100. doi:10.1145/2982142.2982176

[56] Andreas Reichinger, Stefan Maierhofer, and Werner Purgathofer. 2011. High-quality tactile paintings. J. Comput. Cult. Herit. 4, 2, Article 5 (Nov. 2011), 13 pages. doi:10.1145/2037820.2037822

[57] Andreas Reichinger, Moritz Neumüller, Florian Rist, Stefan Maierhofer, and Werner Purgathofer. 2012. Computer-aided design of tactile models: taxonomy and case studies. In Computers Helping People with Special Needs: 13th International Conference, ICCHP 2012, Linz, Austria, July 11-13, 2012, Proceedings, Part II 13. Springer, 497–504.

[58] Samuel Reinders, Swamy Ananthanarayan, Matthew Butler, and Kim Marriott. 2023. Designing conversational multimodal 3d printed models with people who are blind. In Proceedings of the 2023 ACM Designing Interactive Systems Conference. 2172–2188.

[59] Samuel Reinders, Swamy Ananthanarayan, Matthew Butler, and Kim Marriott. 2023. Designing Conversational Multimodal 3D Printed Models with People who are Blind. In Proceedings of the 2023 ACM Designing Interactive Systems Conference (Pittsburgh, PA, USA) (DIS '23). Association for Computing Machinery, New York, NY, USA, 2172–2188. doi:10.1145/3563657.3595989

[60] Samuel Reinders, Matthew Butler, and Kim Marriott. 2020. "Hey Model!" – Natural User Interactions and Agency in Accessible Interactive 3D Models. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376145

[61] Samuel Reinders, Matthew Butler, and Kim Marriott. 2025. " It Brought the Model to Life": Exploring the Embodiment of Multimodal I3Ms for People who are Blind or have Low Vision. arXiv preprint arXiv:2502.14163 (2025).

[62] Brigitte Röder, Mark Wallace, Amir Amedi, Noa Raz, Haim Azulay, Rafael Malach, and Ehud Zohary. 2010. Cortical activity during tactile exploration of objects in blind and sighted humans. Restorative neurology and neuroscience 28, 2 (2010), 143–156.

[63] Elen Sargsyan, Bernard Oriola, Marc J-M Macé, Marcos Serrano, and Christophe Jouffrais. 2023. 3D Printed Interactive Multi-Storey Model for People with Visual Impairments. 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 540, 15 pages. doi:10.1145/3544548.3581304

[64] William Schiff and Emerson Foulke. 1982. Tactual perception: a sourcebook. Cambridge University Press.

[65] Lei Shi, Holly Lawson, Zhuohao Zhang, and Shiri Azenkot. 2019. Designing Interactive 3D Printed Models with Teachers of the Visually Impaired. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3290605.3300427

[66] Lei Shi, Idan Zelzer, Catherine Feng, and Shiri Azenkot. 2016. Tickers and Talker: An Accessible Labeling Toolkit for 3D Printed Models. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 4896–4907. doi:10.1145/2858036.2858507

[67] Lei Shi, Yuhang Zhao, and Shiri Azenkot. 2017. Designing Interactions for 3D Printed Models with Blind People. In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (Baltimore, Maryland, USA) (ASSETS '17). Association for Computing Machinery, New York, NY, USA, 200–209. doi:10.1145/3132525.3132549

[68] Lei Shi, Yuhang Zhao, and Shiri Azenkot. 2017. Markit and Talkit: A Low-Barrier Toolkit to Augment 3D Printed Models with Audio Annotations. In Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (Québec City, QC, Canada) (UIST '17). Association for Computing Machinery, New York, NY, USA, 493–506. doi:10.1145/3126594.3126650

[69] Jennie Small, Simon Darcy, and Tanya Packer. 2012. The embodied tourist experiences of people with vision impairment: Management implications beyond the visual gaze. Tourism Management 33, 4 (2012), 941–950.

[70] Kengo Tanaka, Tatsuki Fushimi, Ayaka Tsutsui, and Yoichi Ochiai. 2023. Text to Haptics: Method and Case Studies of Designing Tactile Graphics for Inclusive Tactile Picture Books by Digital Fabrication and Generative AI. In ACM SIGGRAPH 2023 Labs (Los Angeles, CA, USA) (SIGGRAPH '23). Association for Computing Machinery, New York, NY, USA, Article 10, 2 pages. doi:10.1145/3588029.3595471

[71] Andrew F Tatham. 1991. The design of tactile maps: theoretical and practical considerations. Proceedings of international cartographic association: mapping the nations (1991), 157–166.

[72] Anne Theurel, Arnaud Witt, Philippe Claudet, Yvette Hatwell, and Edouard Gentaz. 2013. Tactile picture recognition by early blind children: the effect of illustration technique. Journal of experimental psychology: Applied 19, 3 (2013), 233.

[73] Catherine Thinus-Blanc and Florence Gaunet. 1997. Representation of space in blind persons: vision as a spatial sense? Psychological bulletin 121, 1 (1997), 20.

[74] Touch Graphics Inc. 2025. Touch Graphics. https://www.touchgraphics.com/. Accessed: 2025-04-15.

[75] Ayaka Tsutsui, Tatsuki Fushimi, Takahito Murakami, Ryosei Kojima, Kengo Tanaka, and Yoichi Ochiai. 2024. HIFU Embossment of Acrylic Sheets. 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 431, 23 pages. doi:10.1145/3613904.3642890

[76] Ayaka Tsutsui, Tatsuki Fushimi, Kengo Tanaka, Takahito Murakami, and Yoichi Ochiai. 2023. Ultrasonic Embossment of Acrylic Sheets with Transparency Control. In ACM SIGGRAPH 2023 Labs (Los Angeles, CA, USA) (SIGGRAPH '23). Association for Computing Machinery, New York, NY, USA, Article 11, 2 pages. doi:10.1145/3588029.3595475

[77] Steven Wall and Stephen Brewster. 2006. Feeling what you hear: tactile feedback for navigation of audio graphs. In Proceedings of the SIGCHI conference on Human Factors in computing systems. 1123–1132.

[78] Xiyue Wang, Seita Kayukawa, Hironobu Takagi, and Chieko Asakawa. 2023. TouchPilot: Designing a Guidance System that Assists Blind People in Learning Complex 3D Structures. 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 5, 18 pages. doi:10.1145/3597638.3608426

[79] Xiyue Wang, Seita Kayukawa, Hironobu Takagi, and Chieko Asakawa. 2024. BentoMuseum: 3D and Layered Interactive Museum Map for Blind Visitors. Commun. ACM 67, 11 (Oct. 2024), 93–102. doi:10.1145/3617678

[80] Zheshen Wang, Baoxin Li, Terri Hedgpeth, and Teresa Haven. 2009. Instant tactile-audio map: enabling access to digital maps for people with visual impairment. In Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility (Pittsburgh, Pennsylvania, USA) (Assets '09). Association for Computing Machinery, New York, NY, USA, 43–50. doi:10.1145/1639642.1639652

[81] DJ Weiss, Andy Nelson, HS Gibson, W Temperley, Stephen Peedell, Allie Lieber, Matt Hancher, Eduardo Poyart, Simão Belchior, Nancy Fullman, et al. 2018. A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature 553, 7688 (2018), 333–336.

[82] Chang-Chun YIN, Bo ZHANG, Yun-He LIU, and Jing CAI. 2015. 2.5-D forward modeling of the time-domain airborne EM system in areas with topographic relief. Chinese Journal of Geophysics 58, 4 (2015), 1411–1424.

[83] Qian-Yi Zhou and Ulrich Neumann. 2010. 2.5 d dual contouring: A robust approach to creating building models from aerial lidar point clouds. In Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part III 11. Springer, 115–128.

A Participant Demographics and Interview Questions for Study 1

Table 3 lists the participants who took part in Study 1, and Table 4 presents the questions asked during the interviews.

B The Transcript of Pre-recorded Audio used in the Wizard-of-oz Interface

B.1 Earthquake and Tsunami Mechanisms

The storytelling narration is based on the tactile graphic book Earthquakes and Tsunamis Through Touch and the accompanying texts from a science museum's earthquake-themed exhibition. It follows the steps outlined below, and its content has also been adapted into responses for the question-based interaction mode.

Overview. There are three tactile objects in front of you. From left to right: A 3D model of Earth's continents. A 3D model of tectonic plates and geological strata. A 2.5D tactile diagram. The left half shows how tsunamis are generated and reach land, and the right half shows tsunami sediment layers. Please explore them by touch.

1. We will now explain what tectonic plates are—these are one of the causes of tsunamis. Focusing on the plates off the coast of Japan in the Pacific Ocean, we'll describe how plate movement causes earthquakes. Then, we explain how the movement and distortion of plates directly lead to tsunamis. Finally, we describe the geological structure and tsunami sediment collected in March 2013 from Arahama in Wakabayashi Ward, Sendai City.

2. Please touch the globe on the left. Some geographical features and islands are omitted in the model. Japan is shown as a long, narrow shape slightly above center. First, find Japan on the 3D Earth model.

3. The raised lines on the model represent tectonic plate boundaries. This model shows the plates near Japan.

4. The rigid rock layer covering the Earth's surface is called the "plate." Earth's surface is made up of about ten large and small plates. These plates are not fixed, but slowly move horizontally at speeds of several to several tens of centimeters per year.

5. As plates move, stress builds up between them. When this stress exceeds a limit, sudden movement occurs—this movement is what causes earthquakes.

6. When an earthquake occurs on land, we can see its effects immediately: ground cracks or buildings collapse. But when an earthquake occurs underwater, it's hard to know what's happening. To understand what's happening on the seafloor, please explore the tactile diagram of tsunami generation and arrival.

7. Please touch the tactile diagram in the middle. It shows a cross-section of the seafloor on the Pacific side of Japan, where the Eurasian Plate and Pacific Plate meet.

8. On the left is the Eurasian Plate, and on the right is the Pacific Plate. The Eurasian Plate on the left overlaps the Pacific Plate on the right. Please touch it.

9. The wavy part at the top represents the Pacific Ocean surface. Please touch it.

10. The land on the Eurasian Plate includes homes. Please feel them.

11. As the Pacific Plate subducts, it drags down the Eurasian Plate with it. Eventually, the Eurasian Plate resists the pull and rebounds. When the stress surpasses a limit, it springs upward.

12. This rebound motion of the Eurasian Plate pushes the ocean above, raising the sea level and generating a tsunami. The tsunami then travels toward land. This example using the Eurasian and Pacific Plates near Japan explains how tsunami waves are caused by plate friction.

13. Now we'll explain tsunami sediment. But first, return to the 3D Earth model you touched earlier. This model can be partially disassembled—parts are connected by magnets. Disassemble it and feel the inside.

14. Inside, you'll find the plates and the Earth's internal structure. The outermost layer is the tectonic plate, directly involved in earthquakes. Please touch it. The plate thickness extends about 100 km from the surface downward.

15. Beneath the plate is a layer of molten rock called the mantle. This model includes both upper and lower mantle sections. Please touch them.

16. The upper mantle lies below the plate and its movement causes earthquakes, volcanic eruptions, and plate tectonics.

17. Inside the mantle lies a layer of magma. Please touch it.

18. At the very center is the inner core. Please touch it.

19. Now, let's move to the tactile diagram on the right. It shows tsunami sediment layers located below the Earth's surface. The topmost area of the diagram represents the ground surface. On the left is a tactile diagram based on an actual photograph of tsunami sediment. On the right is a simplified version with added textures to clarify the layer structure.

ID	Age	Gender	Occupation	Experience (years)	Details
P1	27	Female	Science-museum communicator	1	
P2	27	Male	Same as above	2	
P3	34	Male	Same as above	3.5	
P4	29	Female	Same as above	1	
P5	30	Female	Same as above	5	
P6	36	Male	Same as above	4.5	
P7	28	Male	Same as above	2.5	
P8	57	Male	Tactile-museum staff	34	
P9	67	Male	TVI (Science)	25	
P10	62	Male	TVI (Acupuncture)	34	Low vision
P11	54	Female	TVI (Self-reliance instruction)	30	
P12	46	Female	TVI (Math)	2	
P13	46	Female	TVI (English)	14	
P14	40s	Female	TVI (English)	14	
P15	50s	Female	TVI (Massage, acupuncture, moxibustion)	9	
P16	47	Female	TVI (Anatomy)	7	
P17	52	Female	Staff at a facility for BLV individuals	30	
P18	50	Male	Same as above	12	
P19	63	Female	Same as above	10	
P20	53	Male	Same as above	24	Blind
P21	29	Female	Same as above	4	
P22	48	Male	Accessibility researcher	16	Blind

Table 3: Details of the 22 participants with diverse backgrounds who took part in the Study 1 interviews.

(1) Tactile Content Used for Conveying Concepts	
Q1	Which types of content (e.g., 2D or 3D) have you used to convey a concept?
Q2	Have you ever used multiple content types together, for example combining 2D and 3D materials, when conveying a concept?
Q3	Do you have experience creating tactile content? What kinds of tactile materials would you like to see for BLV users in the future?
Q4	How could technology help BLV individuals learn concepts independently through tactile content, and what kinds of technology might be used?
(2) Instructional Approaches and Correction During Tactile Exploration	
Q5	What approaches do you use when explaining or guiding BLV users?
Q6	When a BLV user touches an area different from the one you are explaining, how do you correct them?
Q7	Do you face challenges in deciding the right timing for explanations? What is required to deliver guidance at the most appropriate moment?
(3) Strategies for Communication and Adaptations Based on Learner Needs	
Q8	When delivering explanations, what is the typical group size?
Q9	When using multiple tactile contents, what factors do you focus on and how do you adapt your explanations?
Q10	How do you tailor explanations for BLV individuals with different residual vision?
Q11	What challenges do you face when explaining to BLV users?

Table 4: Questions posed to the 22 participants in Study 1, organized into three thematic categories.

20. Experts in geology study these tsunami sediment layers left by past events to estimate the timing and extent of past trench-type earthquakes and use this data to forecast future events.

21. At the bottom is marine sediment. Please touch it. This includes volcanic ash and biological remains deposited by natural processes.

22. One layer above is wetland sediment. Please touch it.

23. The next layer is tsunami sediment from the Jogan Earthquake of 869 CE. Please touch it.

24. Then another wetland sediment layer. Wetlands are areas that are regularly or seasonally submerged in water or have saturated soil. Please touch it.

25. Next is tsunami sediment from the Kyotoku Earthquake of 1454 CE. Please touch it.

26. The following layer is topsoil, also called "cultivated soil," where agriculture is conducted. Please touch it.

27. The topmost tsunami layer represents sediment from the 2011 Great East Japan Earthquake and Tsunami. Please touch it.

B.2 The Journey of Hayabusa2

The storytelling narration is based on the accompanying texts from a science museum's space-themed exhibition. It follows the steps outlined below, and its content has also been adapted into responses for the question-based interaction mode.

Overview: In front of you, there are three tactile models. From left to right: A 3D model of the asteroid Ryugu. A 3D model of the asteroid explorer Hayabusa2. A 3D model of Hayabusa2's re-entry capsule. Please feel them with your hands.

1. Using these three models, we will explain the story of how Hayabusa2 traveled to asteroid Ryugu, collected surface and subsurface samples, and returned them to Earth.

2. Please touch the model in the center. It is the Hayabusa2 asteroid explorer. This spacecraft was equipped with instruments that enabled it to create an artificial crater on an asteroid and retrieve subsurface materials.

3. Please touch the model at the left. It is a model of the asteroid Ryugu. This is the asteroid that Hayabusa2 explored. It has a shape like two spinning tops stacked together. The actual asteroid is about 900 meters in diameter.

4. This shape is believed to have formed due to a history of rapid rotation. Ryugu also has a low density and is covered with many rocks. These features suggest it was formed from the destruction and reaccumulation of a parent celestial body.

5. Next, we will explain how Hayabusa2 performed a touchdown on Ryugu. Please touch the Hayabusa2 model again.

6. Feel the flat panels on both sides. These are the solar panels.

7. Let's talk about digging. Please touch the cylindrical object located underneath.

8. This is the impactor device. Because Ryugu's gravity is too weak for tools like shovels or drills to be effective, Hayabusa2 used a 2 kg copper projectile, fired at 2 km per second, to create an artificial crater on the surface.

9. If you'd like to feel a crater, please touch the asteroid Ryugu model.

10. The indented areas are craters. There are many of them.

11. Let's move on to sample collection. Please touch the long part underneath.

12. This is the sampler horn. To collect material, the sampler horn was placed onto Ryugu's surface. A small projectile was then fired from inside the device to stir up dust, which was collected. The mouth of the horn is designed with hooks to help trap the material more effectively.

13. Lastly, we'll explain how the collected dust was stored and brought back to Earth. Please touch the semicircular part above the sampler horn.

14. This semicircle is the re-entry capsule. The dust that was stirred up was stored in this capsule and returned to Earth. When re-entering Earth's atmosphere, the capsule experiences temperatures close to 10,000℃, so its outer shell is designed to protect the asteroid material from heat.

15. This is the re-entry capsule. The dust stirred up by Hayabusa2 was stored inside and brought back to Earth. During atmospheric re-entry, temperatures reach nearly 10,000℃, and the capsule's shell protects the contents from this intense heat.

16. Now, please touch the model on the right. It is the model of the re-entry capsule, and it can be disassembled. Please touch the outermost layer of the capsule—the heat shield.

17. The heat shield protects the contents from the extreme heat during atmospheric re-entry.

18. Try separating the heat shield into two parts.

19. The other half of the heat shield can also be disassembled. Please give it a try.

20. Inside the separated halves, parachutes are stored on both outer sides.

21. The compartmented area in the middle contains electronic instruments that measured speed and other data.

22. The object in the center is called the sample container. Please touch it.

23. Inside the sample container is a catcher used to store asteroid material. Please touch it.

24. The interior is divided into three chambers—A, B, and C. Chamber A held surface dust, while Chamber C held subsurface dust. Please touch them.

C The Wizard-of-oz Interface

Figure 7 illustrates the interface used in Study 2.

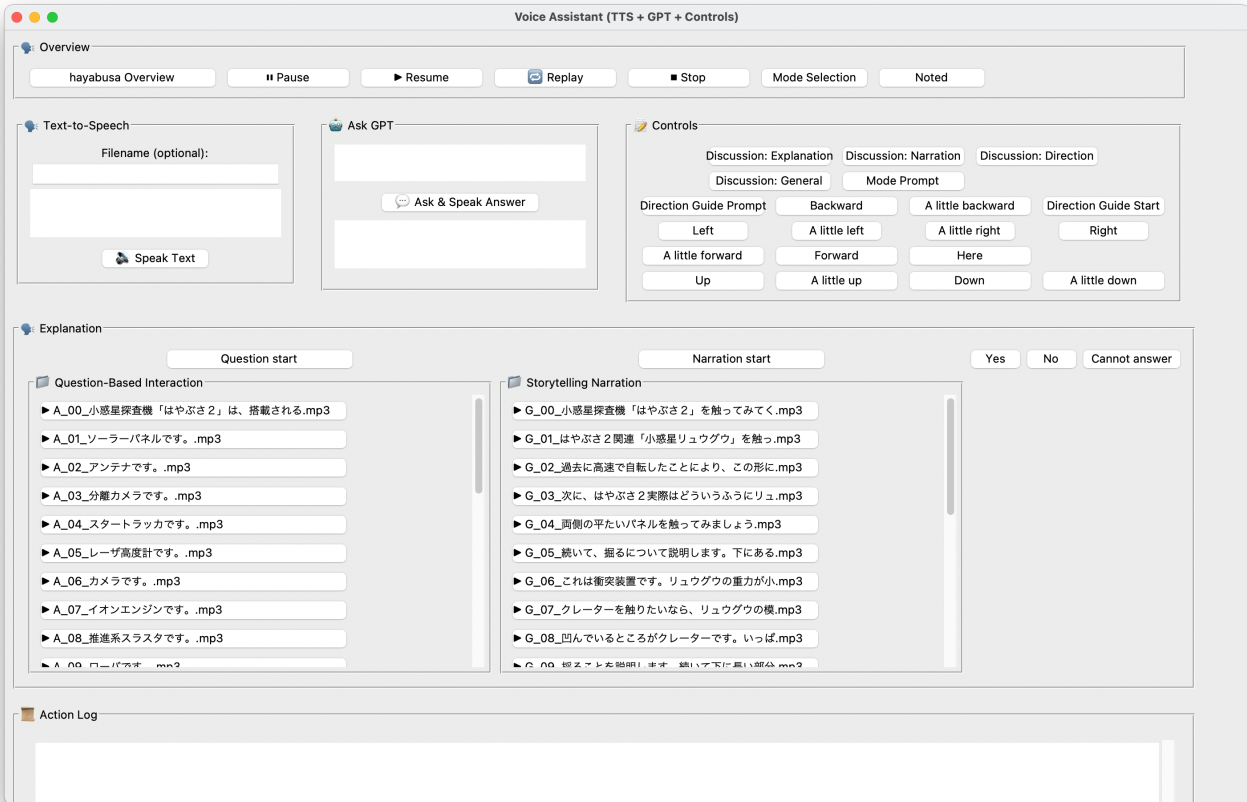

Figure 7: Screenshot of the Wizard-of-Oz interface used in Study 2. The interface allowed the researcher to trigger different text-to-speech audio responses to simulate a voice-based "Touch and Talk" system. It supported pausing and resuming audio, playing pre-recorded storytelling and question-based explanations, delivering GPT-generated answers, and issuing directional guidance.