# Designing Conversational Multimodal 3D Printed Models with People who are Blind

Samuel Reinders
samuel.reinders@monash.edu
Monash University
Melbourne, Victoria, Australia

Swamy Ananthanarayan
swamy.ananthanarayan@monash.edu
Monash University
Melbourne, Victoria, Australia

Matthew Butler
matthew.butler@monash.edu
Monash University
Melbourne, Victoria, Australia

Kim Marriott
kim.marriott@monash.edu
Monash University
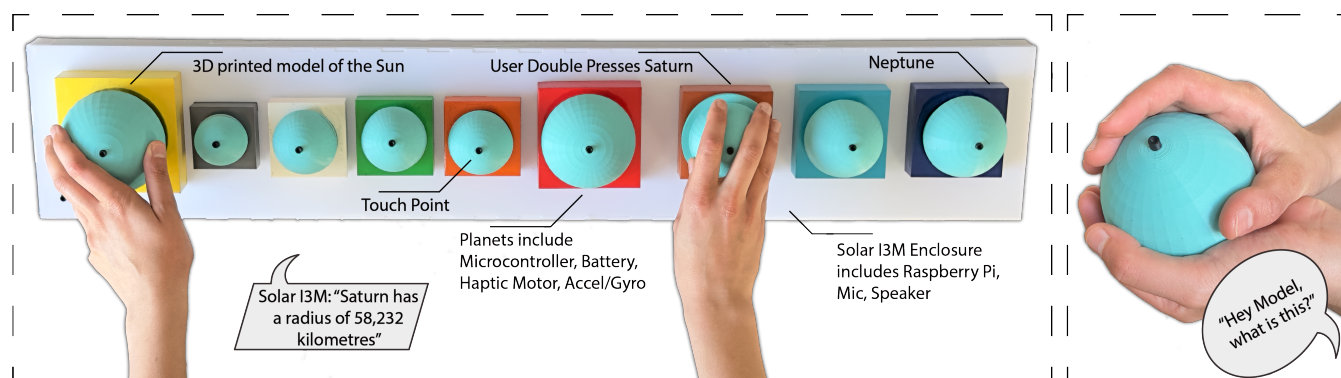Melbourne, Victoria, Australia

Figure 1: Solar I3M is designed to be touched, held, or picked up. The eight planets and the Sun sit upon an acrylic enclosure. Blind users can perform touch gestures on the planets to extract audio labels, or they can engage in conversational dialogue. Solar I3M also uses haptic vibratory feedback to add additional context during interactions.

## ABSTRACT

3D printed models have been used to improve access to graphical information by people who are blind, offering benefits over conventional accessible graphics. Here we investigate an interactive 3D printed model (I3M) that combines a conversational interface with haptic vibration and touch to provide more natural and accessible experiences. Specifically, we co-designed a multimodal model of the Solar System with nine blind people and evaluated the prototype with another seven blind participants. We discuss our journey from a design perspective, focusing on touch, conversational and multimodal interactions. Based on our experience, we suggest design recommendations that consider blind users' desire for independence and control, customisation, comfort and use of prior experience.

## CCS CONCEPTS

• **Human-centered computing → Accessibility**.

## KEYWORDS

Blind; Accessibility; 3D Printed Models; Conversational Interface; Multimodal Interaction;

## 1 INTRODUCTION

People who are blind experience difficulty accessing graphical information, particularly educational content [13, 48]. 3D printed models are becoming increasingly used as a technology for providing accessible graphics, with widespread interest amongst blind users in how they can be applied in everyday life. These models have benefits over traditional tactile graphics, including better representation of three-dimensional concepts, most notably in science education.

However, much like tactile graphics, 3D printed models often include braille labelling and legends, which can be problematic.

Braille labels can occupy considerable space, limiting the amount of information that can be conveyed directly on the model. Separate legends can be provided, but these add to cognitive load as blind users may have to constantly shift between the model and the legend [31]. Moreover, the use of braille is not ideal due to limited braille literacy among blind users [46]. As a result, researchers have developed interactive 3D printed models (I3Ms) with pre-recorded audio labels [21, 23, 24, 27, 31, 52, 59]. Although these I3Ms are an improvement, they offer limited passive interactivity, consisting only of basic contextual information retrieval.

Recent research has identified that blind users want rich interactions with I3Ms in ways similar to their personal technology, using modalities such as touch and conversational interfaces [53]. Our current work presents the next step in this nascent area, with the co-design of an I3M of the Solar System – Solar I3M. Building on earlier works that elicited I3M interaction modalities and strategies directly from blind end-users [53, 60], our work explores the creation of a fully functional multimodal 3D model that includes touch interaction, haptic vibratory feedback and a conversational interface (Figure 1). Solar I3M was developed with people who are blind across multiple co-design sessions, and shares a tight integration of modalities. Our contributions include:

- A co-design focused account of Solar I3M and its support for richer multimodal interactions that enable more natural interaction paradigms. To the best of our knowledge, this is the first I3M for blind users that includes a conversational interface (with demonstrative pronouns) associated with a physical tangible model.
- A usability evaluation of Solar I3M that revealed that blind users share a desire for richer multimodal interactions that are physically embodied and the ability to customise these according to their personal preferences.
- Design recommendations based on our co-design process and evaluation to help facilitate future I3M research and production. Particularly important was identifying how key areas of functionality – touch, conversational and multimodal interaction – could be designed to better support blind users' desire for: use of *prior experience*, interactions that uphold their *independence and control* and respect their sense of *confidence and comfort*, and allow *customisation and personalisation*.

The enthusiastic evaluation of Solar I3M by blind users demonstrates the potential of these kinds of models in an educational context. They could also be used in museums and other cultural institutions. We therefore believe that our research and Solar I3M will be of interest to practitioners creating accessible education products and accessible exhibits at museums and other cultural institutions.

## 2 RELATED WORK

### 2.1 Accessible Graphics & 3D Printed Models

Accessing graphical information can prove difficult for people who are blind, impacting education opportunities [13] and making independent travel difficult [57]. These can lead to reductions in confidence and quality of life [37]. Tactile graphics are used by blind people to assist in classroom learning [4, 54] and orientation and mobility training [6, 55] and are commonly provided with braille labels and legends. However, they can only convey limited amounts of contextual information, chiefly due to the amount of space that braille requires [9, 31, 59]. Additionally, conventional production methods can only convey height/depth to a limited degree [31], restricting the types of graphics that can be produced.

3D printed models have increasingly become an alternative, allowing a broader range of graphics to be produced when compared to conventional tactile graphics, including more complex three-dimensional concepts. The cost and effort involved in producing 3D printed models have fallen more in line with that of tactile graphics, resulting in 3D printing being studied broadly across many accessible graphics application areas: mapping and navigation [28, 30, 31, 33]; special education [12]; books [38, 64]; mathematics [10, 34]; graphic design [43]; science [29, 67] and programming curricula [36]. While 3D printed models offer benefits over tactile graphics, braille labelling remains an issue: the low-fidelity of 3D printed braille can limit the readability of braille labels [10, 59, 65], there is limited space for labels and updates and additions to braille require model reprinting [31]. Additionally, as with tactile graphics, reliance on braille labels is not ideal for the majority of blind people who have limited braille literacy [46].

### 2.2 Interactive 3D Printed Models (I3Ms)

To help overcome labelling limitations, 3D printed models are increasingly being combined with low-cost electronics and smart devices to produce interactive 3D printed models (I3Ms). I3Ms have been created and applied across many blind-specific contexts: mapping and navigation [27, 31, 62]; art [32, 35]; and education [23, 53, 58]. Many I3Ms include button or touch-triggered audio labels that when activated describe different details of the model [21, 23, 24, 27, 31, 52, 59]. Audio labels can be stored as either pre-recorded files or synthesised from text in real-time. Audio labels are particularly useful for blind users who are not braille readers, and the label information can be updated easily.

The design space of I3Ms is still very much in its infancy, with much of the current research in the field focused on simple touch interactions and extraction of basic information using passive audio labelling. Recent works using more user-driven design approaches have begun to uncover a desire amongst blind users for richer, more interactive experiences [53, 58, 60]. One study used a Wizard-of-Oz (WoZ) method and identified that blind users want to interact with I3Ms across three modalities – gestures, speech, and buttons [60]. Further work found that, in addition to passive auditory labelling, I3Ms should allow blind users to ask questions using more conversational language [58], while our earlier work found that blind users want to interact in ways similar to their personal technology [53], including the use and combination of touch gestures, conversational dialogue, and haptic vibratory feedback in more multimodal interactions.

### 2.3 Conversational Interfaces

People who are blind find voice interaction convenient [5], and adoption rates of devices offering conversational interfaces – e.g. Siri (Apple), Alexa (Amazon) and Google Assistant – are high

amongst this group [50]. Recent work into the accessibility of conversational interfaces integrated into devices like smartphones and smart speakers has found that they often assume a strict human-to-human conversational model that can limit the interactions of blind users [2, 8]. [20] found that the slow speed of conversational output might conflict with the speech comprehension rates of many blind users, with other research identifying a lack of customisation options relating to the voice used by conversational interfaces, and response length [2, 8].

Little work has been undertaken exploring how conversational interfaces can be integrated into I3Ms. Such models could allow blind users to engage and ask questions using natural language, allowing for more natural interaction paradigms and richer interactive experiences. [17] combined a tactile graphic with a conversational interface focused on indoor navigation, but voice interactions were performed through a connected smartphone rather than the model itself. [18, 35] created voice-controlled guides to help blind users navigate art pieces, however, their implementations supported simple and rudimentary commands rather than more expansive conversational dialogue. Other work has identified, but not comprehensively designed or tested, a set of I3M conversational interactions. [58] suggested allowing blind users to extract information using conversational language, while our previous work showcased a simple WoZ implementation of wake words that could be used to invoke an I3M conversational interface [53].

## 2.4 Multimodal Interfaces

The provision of multiple modalities can increase the adaptability of a user interface [51], the resolution of information that the interface can communicate, and allow more natural interaction [7]. Adaptability is particularly important for blind users, affording choice on how they interact. For example, they might choose to interact with an interface that offers both voice and touch input differently based on their ability, environment or even level of comfort [2]. Richer resolution of information is also important for blind users, as the combination of accessible modalities – e.g. tactile and auditory output – may more closely match the capacity of vision and help overcome the *"bandwidth problem"* [22]. Multiple modalities can also improve the confidence and independence of blind users [19]. Modalities can also be combined to create more *"natural user modalities"*, as described by the *"Put-That-There"* paradigm [7]. This work combined gestural and speech inputs using demonstrative pronouns, and is foundational to the gestural-based interactions seen in virtual and augmented reality [42].

Multimodal accessible aids have been created for blind users combining tactile graphics or touch screens with auditory output [1, 19, 49], haptic vibrations [26, 47, 49], olfactory and gustatory perception [11], and visual feedback for low-vision users that still possess levels of residual vision [3]. I3Ms are multimodal and largely combine physical models with auditory output, albeit in a rather limited fashion [17, 21, 23, 27, 31, 66]. However, our previous work found that blind users desired to also engage with haptic vibratory output, desiring seamless combinations of touch, haptic vibratory feedback, and conversational dialogue [53].

## 2.5 Motivation

With much of the work into I3Ms being preliminary in nature, often dealing with hypothetical or limited prototypes, we lack guidelines on how I3Ms should be designed. In particular, the design of I3Ms that integrate multiple modalities and make use of conversational interfaces is virtually unexplored despite the potential of such interfaces to support richer and more natural interaction.

## 3 SOLAR I3M: A CO-DESIGN JOURNEY

### 3.1 Overview

To address this gap we designed and evaluated a fully functional multimodal I3M that supports touch gesture interaction, haptic vibratory feedback, audio labels, and a conversational interface. We chose to build a model of the Solar System. This not only reflects the high demand by blind students for more accessible STEM materials [13] and the rich opportunities for multimodal interaction revealed in our previous work [53], but also affords us the opportunity to contrast our design findings with earlier work where we used a partial-WoZ I3M implementation.

Importantly, we co-designed Solar I3M with multiple blind co-designers. The use of more participant-oriented methods, like co-design, is imperative when designing for and working with people with disability, as the researchers may not possess lived experience of the relevant disability or impairment. Co-design has been used in the blind accessibility space [23, 25, 45], and such methods have been recommended to address the low involvement of blind users across the design and evaluation of accessible graphic research [14, 39]. This is particularly important as the degree of how fit-for-purpose accessible aids/tech are can influence abandonment rates [44].

We held two design workshops, each with different groups of blind co-designers, and two one-on-one design sessions with a single blind expert co-designer (Table 1). Recruited through our lab's contact pool, our S1 & S2 co-designers had no prior co-design experience, while our expert co-designer, involved in S3 & S4, had prior design experience from unrelated research projects.

**Table 1: Outline of the four co-design sessions**

| Session | Users | Duration | Session Focus |
|---------|-------|----------|---------------|
| S1 | 4 | 90mins | Touch Gestures |
| S2 | 4 | 90mins | Audio Labelling Conversational Interface Haptic Vibration |
| S3 | 1 | 90mins | Touch Gestures Audio Labelling Conversational Interface |
| S4 | 1 | 90mins | Audio Labelling Conversational Interface Haptic Vibration |

Solar I3M's design was inspired by prior work eliciting hypothetical I3M interaction modalities from blind users using WoZ experiments [53, 60]. What makes Solar I3M unique, in addition to being the first I3M with a working conversational interface integrated directly into the model, is that conversational dialogue,

audio, haptic vibration, and touch are tightly integrated to deliver a rich multimodal experience. The following illustrates a typical interaction (Figure 2):

(1) A blind user picks up 3D printed models of Jupiter and Saturn, gathering details like shape and size tactually.
(2) They proceed to press the touch points on each planet; Solar I3M identifies them via an audio label using an attached speaker – *"Activated Jupiter"*, *"Activated Saturn"*.
(3) Continuing, the user engages in conversation, asking Solar I3M a question – *"Hey Model, which of these is bigger?"*.
(4) Solar I3M provides the spoken response – *"Jupiter has a radius of 69,911 kilometres, while Saturn has a radius of 58,232 kilometres"* – emitting haptic vibrations from each planet model as they are referred to.
(5) The user continues their interaction, asking *"Hey Model, how many stars are in the Milky Way?"*.
(6) Solar I3M replies – *"I don't know the answer to: 'How many stars are in the Milky Way', but would you like me to search for it?"* – to which the user responds – *"Yes, search for it"*.
(7) Solar I3M speaks – *"According to NASA.gov, there are over 100 billion stars in the Milky Way"* – and the user decides to conclude their current interaction.

Our co-design journey is broken down around three areas – **touch, conversational and multimodal interaction** – highlighting how our blind co-designers and existing research informed Solar I3M's design.

## 3.2 Co-Designing Touch Interactions

Like traditional tactile graphics, I3Ms are designed to be held and touched, allowing blind end-users to undertake information gathering through tactile exploration. What makes I3Ms unique is that in addition to touch being used to explore an object, touch can be used as an input to trigger I3M functionality. Many I3Ms make use of touch to trigger audio labels that playback audio labels [21, 23, 24, 31, 59]. However, unlike smartphones, where the enclosure is gripped and gestures performed on a dedicated touch surface, I3Ms are held and tactually explored uninterrupted, with touch points sharing the same faces of the object being actively explored. This presents a design challenge, with research showing that blind participants can become frustrated during I3M tactile exploration if interrupted by the accidental activation of system output [31, 53].

*3.2.1 Creating Easily Identifiable Touch Points.* Using electrically conductive printing filament, it is possible to create 3D prints where specific features of areas of the print are conductive. When connected to capacitive touch sensors, these areas can act as tappable surfaces or touch points. During S1 we explored designing easily discoverable touch areas. Our co-designers were handed three 3D printed planet models (Figure 3): (1) a planet with a tappable area integrated flush into the surface of its northern pole, (2) another with a small indented touch surface at the top of the planet, and (3) a planet with a discrete raised touch point that protruded out of the northern pole. The first two designs were influenced by an existing recommendation that touch points shouldn't distort the surface of a print [31]. We asked our co-designers which designs were most distinctive. Three co-designers gave immediate preference towards

the protruded design, suggesting that the shape, size, and feel of the touch point made it easier to locate amongst other details of the printed model. The protruded design deviates from [31]'s guideline, suggestive that this recommendation might be specific to domains requiring a greater degree of touch exploration, e.g. I3Ms of maps that have more salient tactual details.

*3.2.2 Choosing Touch Gestures & Their Uses.* Touch gestures are well understood due to the proliferation of touch input devices, e.g. smartphones, with high rates of use amongst blind users. [31] implemented Single, Double and Long touches in their I3M to playback audio labels, while [53] observed participants designing their own interactions based on smartphone analogous touch gestures. We drew from these, implementing touch gestures used in end-users' personal technology as follows:

- **Single Tap/Press**: Identifies the object being interacted with using an audio label, e.g. *"this is Jupiter"*;
- **Double Tap/Press**: Rotates a selection of audio label facts about the object being interacted with – colour, size, order, distance/location, composition, and number of moons, e.g. *"Mars is reddish brown, with no clouds, and has white ice caps"*;
- **Long Tap/Press**: Engages Solar I3M's conversational interface.

### Table 2: Touch gestures supported by Solar I3M

| Gesture | Timing | Output |
|---|---|---|
| Activate/Single Press | 500ms | First press activates a touch point, subsequent press identifies touch point |
| Double Press | 500ms | Rotates through a series of touch point facts |
| Long Press | 2s | Engages Solar I3M's conversational interface |

*3.2.3 Ensuring That Touch Interactions Are Deliberate.* Initially, Solar I3M's touch gestures were influenced by existing standards [1], for consistency with personal technology use. A Single Tap was designed to trigger after a touch event of 250ms had been performed on a touch point, a Double Tap when two touch events were detected within 250ms, and a Long Tap for touch events lasting longer than 500ms.

During S1 we asked our four co-designers to repeatedly perform each touch gesture on the touch point of one of Solar I3M's printed planets. All co-designers experienced problems reliably performing the touch gestures, including Double Taps being detected as successive Single Taps. One co-designer became visibly irritated during this activity and indicated that the touch gestures made them feel frustrated, as they found the timing windows too brief/precise. With the detection timing windows of gestures adjustable on the spot, we took turns with our co-designers adjusting the timing of each touch gesture until a setting was found that our co-designers could reliably and consistently perform. The timing of each gesture was increased (Table 2), suggesting that blind users may require

---

[1]https://developer.apple.com/design/human-interface-guidelines/inputs/touchscreen-gestures/
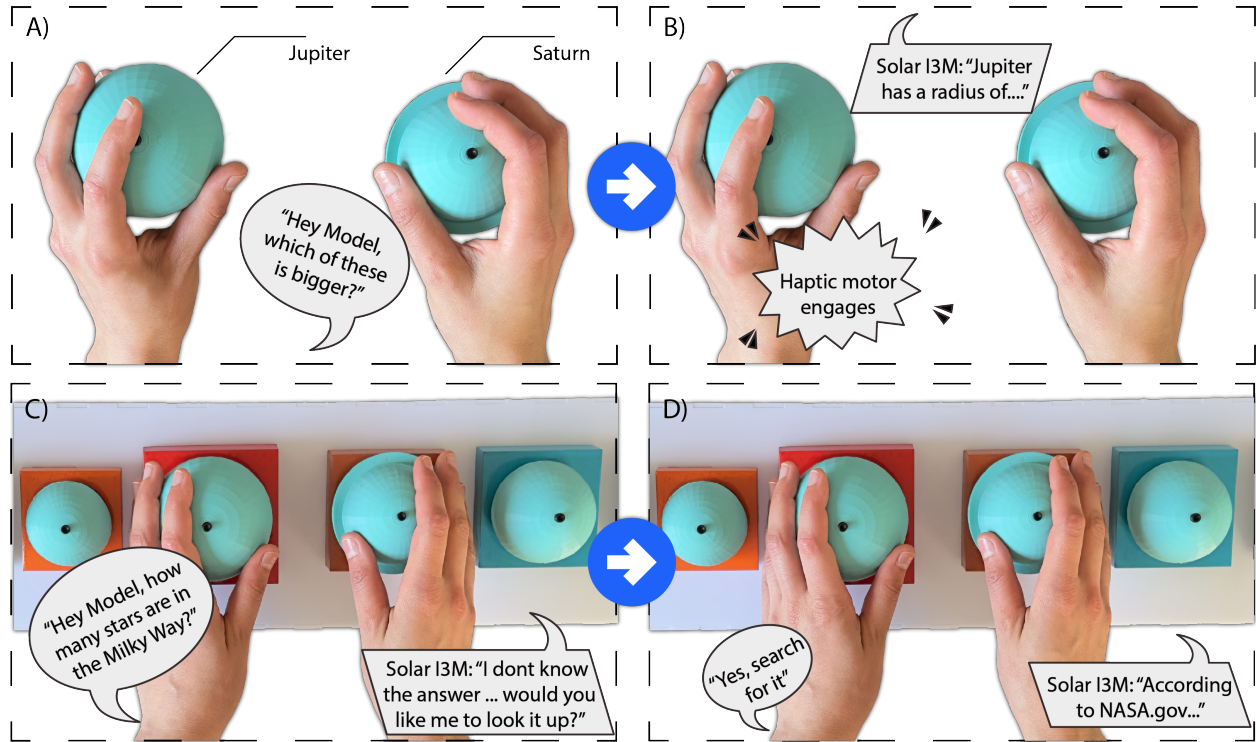
**Figure 2: Solar I3M allows blind users to engage in rich multimodal experiences. A) Users can engage in conversational dialogue using wake words and demonstrative pronouns; B) Solar I3M combines auditory responses simultaneously with haptic vibratory feedback; C & D) Users can perform searches to extend their knowledge.**
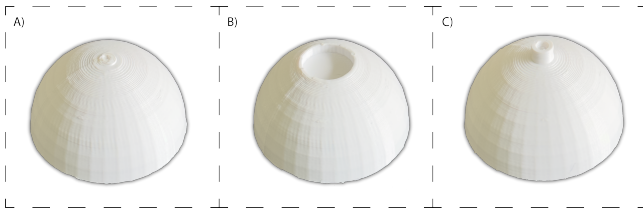


**Figure 3: Touch point designs from S1 design session. A) tappable area integrated flush; B) indented touch surface; C) raised touch point.**

additional time and concentration when performing touch gestures on I3Ms compared to personal technology with dedicated touch surfaces.

We returned to touch gestures in S3, and while our co-designer was able to reliably perform all of the touch gestures, they suggested improvements:

- **Touch gesture names should reflect how they are performed**. Touch gestures required longer presses compared to the quicker taps they performed on their smartphone. Our co-designer suggested that referring to touch gestures as *Presses* rather than *Taps* would be more illustrative (Table 2);

- **Touch points should require activation**. While feeling in control during tactile exploration, our co-designer suggested that an initial *Single Press* should be required to *activate* a touch point before subsequent touch gestures could be performed. They detailed that this would increase their likelihood of engaging in tactile exploration as it would reduce inadvertent interruptions from touch gesture-driven audio label playback. A parallel was made to the screen reader built into the co-designers smartphone which requires a gesture to activate the currently selected interface element.

*3.2.4 Design Outcomes.* After S1 & S3, we had a clear outline of how our blind co-designers wanted to undertake touch interactions. **Independence during tactile exploration** was important, allowing users to freely explore Solar I3M without the interruption of accidental audio label playback. Maintaining a **sense of control during touch interactions** was also seen as crucial, supporting users in clearly identifying which areas of Solar I3M touch interactions can be performed, and ensuring that touch gestures work reliably and with confidence.

## 3.3 Co-Designing Conversational Interactions

Conversational interfaces allow blind users to navigate and complete tasks in a non-visual manner. Despite their widespread use

by people who are blind, recent work investigating the accessibility of mainstream conversational interfaces has found that their design can limit blind users' interactions, as dialogue is based on human-to-human conversational models [2, 8].

Little work has been undertaken exploring how conversational interfaces can be integrated into I3Ms. [17] paired a conversational interface with an I3M to assist blind users with indoor navigation. However, their conversational interface was activated through a connected smartphone rather than the I3M itself. [58] suggested that blind students should be able to ask I3Ms questions about modelled content, while [53] found that blind users wanted to talk directly to I3Ms using natural language to fill gaps in their knowledge using wake words.

We built Solar I3M's conversational interface using Dialogflow[2] and its Python API[3], and drew from prior work into the accessibility of mainstream conversational interfaces.

*3.3.1 Customising Conversational Output.* During S2 & S3, we ran a series of activities to understand how our co-designers wanted to customise responses, and made adjustments on-the-fly using Speech Synthesis Markup Language (SSML).

During S2, we played back a series of audio labels generated at speeds comparable to standard human conversation. We asked our co-designers if there were any scenarios where they would like to change the speed of auditory output. Our co-designers suggested that using Solar I3M to revise content, or in a public space, that it should speak faster. One co-designer indicated that they would like Solar I3M to speak twice as fast. We modified the SSML of the audio label responses, generating versions at 150% and 200% speed. Our co-designers liked the faster-spoken responses, indicating that use might vary according to environment. The speed of Solar I3M's auditory responses can be changed using co-designer suggested voice commands – *"Can you speak faster [or] slower?"* and *"Can you speak at [100/150/200%] speed?"* (Table 3).

The synthesised voice that Solar I3M's conversational interface employs can also be changed. One of our co-designers indicated that the initial voice had an unpleasant quality, and sounded robotic. To explore this further, we opened the Google Text-to-Speech web page and began generating speech with different vocoders[4]. All four of our co-designers found speech generated using the WaveNet vocoders, which use neural networks to synthesise more human-like speech, more pleasant. One co-designer suggested that they would like to select different types of voices, similar to how Siri offers a selection of different voices. After the design session concluded, we implemented two WaveNet voices – masculine and feminine – to complement the existing Standard vocoder synthesised voice. The voice Solar I3M uses to generate auditory responses can be changed using the co-designer suggested voice command – *"Can you speak with a [feminine/masculine/robotic] voice?"* (Table 3).

The length and detail of Solar I3M's auditory responses can also be modified, influenced by [2]'s finding that blind users can find responses verbose or insufficiently detailed. During S3, we asked our one-on-one co-designer whether the length of a typical response – e.g. *"Mars has a radius of 3389 kilometres"* – could be

increased meaningfully. This co-designer felt that additional detail should only be added if it provided greater context and meaning, e.g. how the radius of Mars compared to Earth. We provided a more detailed response – *"Mars has a radius of 3389 kilometres, making it around half the size of Earth"* – which our co-designer found useful, especially for objects or concepts they were learning about for the first time. After the session, we modified all auditory responses to have two variants – *low* and *high* verbosity – adjusted using the voice command – *"Can you give me [longer/shorter] answers?"* (Table 3).

**Table 3: Users can modify Solar I3M's auditory responses based on voice, speed and length**

| Voice | Vocoder | Speed | Verbosity/Length |
|---------|----------|---------------|------------------|
| Feminine | WaveNet | 100, 150, 200 | Shorter, Longer |
| Masculine | WaveNet | 100, 150, 200 | Shorter, Longer |
| Robotic | Standard | 100, 150, 200 | Shorter, Longer |

*3.3.2 Ensuring That Users Are In Control During Interactions.* When designing Solar I3M's conversational interface, we wanted to ensure that blind users could undertake conversational interactions that upheld and respected their sense of control. During S2, we ran a series of activities to understand how blind users wanted to invoke and engage with Solar I3M's conversational interface.

We began by asking our co-designers what wake words they wanted to invoke the conversational interface with. We provided an example: *"Hey Model"*. One co-designer thought that *"Hey Model"* was memorable. Two of the other co-designers initially suggested that Solar I3M could use a more descriptive wake word, e.g. *"Solar"* or *"Solar System"*, but later indicated that using a more generalisable phrase, like *"Hey Model"*, meant they wouldn't have to remember different wake words between I3Ms. We used Picovoice Porcupine[5] to generate the wake word *"Hey Model"*. Our co-designers were comfortable taking turns saying this phrase.

We asked our co-designers if it was important to offer a manual method of activating Solar I3M's conversational interface. One co-designer drew connections to their smartphone, highlighting how they primarily activated their conversational interface by holding down a button, as they found it more reliable. This aligns with [2]'s finding that manual forms of activation are important for blind users, as they can become quickly frustrated when wake words aren't properly detected. We outlined our *Long Press* gesture (Section 3.2.2), with one co-designer directly relating this to how they would hold down the *Side Button* on their iPhone.

To further investigate the design of Solar I3M's conversational interface, during S4 we gave our co-designer ample time to initiate conversational interactions. While this co-designer enjoyed their time with Solar I3M, they offered the following suggestions:

- **A voice command to repeat auditory outputs**. Our co-designer found that longer auditory responses, particularly those returned externally from Google Knowledge Graph (Section 3.3.3), were difficult to fully comprehend. They proposed a voice command that would repeat Solar I3M's most

---

[2]https://cloud.google.com/dialogflow
[3]https://github.com/googleapis/python-dialogflow
[4]https://cloud.google.com/text-to-speech/docs/voices

[5]https://picovoice.ai/platform/porcupine/

recent auditory response, suggesting the phrase – *"Can you repeat that?"*;

- **Longer recording windows**. Our co-designer suggested that the microphone timeout window of Solar I3M was too short and cut them off whilst they were still asking their question. We modified the default microphone timeout of the Dialogflow Python API, adjusting it so that Solar I3M would continually record a user's request until no further microphone input was detected;

- **Appropriate feedback when recording**. Initially, when Solar I3M's conversational interface was invoked, the microphone would automatically begin recording a user's speech. Our co-designer felt that additional feedback was required for two reasons: (1) to take the guesswork out of when they could begin their request, and (2) it made them feel more secure knowing when the I3M was listening to them. The phrases – *"Recording"* and *"I am listening"* – were suggested to help improve their sense of control over conversational interactions;

*3.3.3 Comprehensive Conversational Interactions.* Solar I3M's conversational interface uses Dialogflow to provide natural language understanding. Operating in the cloud, this instance has been trained on a corpus of information relating to the Solar System. This covers the same information and knowledge accessible through touch gesture-driven audio labels (Section 3.4.1). Dialogflow also allows users to ask questions, allowing richer interactions where Solar I3M can fulfil a wider variety of requests, and help users in filling gaps in their knowledge. If Solar I3M's conversational interface cannot directly answer a question, it will offer to perform a search by asking – *"I don't know the answer, but would you like me to search for it?"*. Users can authorise by responding – *"Yes, search for it"* – activating Google Knowledge Graph [6] to fetch a response from sources including Wikipedia, Wikidata and CIA World Factbook.

During the S4 one-on-one session, we wanted to understand whether this feature would be meaningful, and how its design could support blind users. We prompted our co-designer to ask Solar I3M a series of questions that it had not been trained to answer. After Solar I3M successfully returned responses from Google Knowledge Graph, our co-designer indicated that being able to perform searches would help to increase their confidence that their interactions would be fulfilled, and offered the following suggestions:

- **Repeat the user's question before searching**. Initially, Solar I3M would prompt a user to perform a search by asking – *"I don't know the answer, would you like me to search for it?"*. Our co-designer felt it was important for this prompt to repeat their question, allowing them to confirm that the conversational interface had correctly interpreted and understood their request. The format – *"I don't know the answer to [question]..."* – was suggested;

- **Appropriate references should be provided**. Our co-designer felt that presenting a source for all external responses would be integral, particularly if used in education contexts, allowing them to make their own judgement call on

the trustworthiness of responses. The format – ***"According to [source], [searched result]..."*** – was discussed;

- **Asking for more information**. After asking a question and receiving an answer, our co-designer suggested that users should be able to ask for additional information as required. The voice command – *"Can you tell me more?"* – was suggested as a way of rotating through a planet's facts, and the conversational interface prompt – *"According to [source], [answer]. **Would you like to know more?**"* – was discussed.

*3.3.4 Design Outcomes.* Our S2, S3 & S4 co-designers clearly expressed the ways in which they wanted to engage in conversational interactions. This centred on **customisable conversational outputs**, allowing users to modify and control how Solar I3M's auditory responses are generated based on preference, ability and environment. Additionally, ensuring **a sense of control during conversational interactions** was important, allowing users to choose how they initiate dialogue – using wake words or touch gestures – and ensuring appropriate feedback is provided so that interactions can be effectively understood and completed.

## 3.4 Co-Designing Multimodal Interactions

As previously discussed, multimodal interaction offers several advantages to blind users: richer resolution of information [22], more natural interaction [7], and increased adaptability [51].

*3.4.1 Allowing Users To Choose How They Interact.* Between co-design sessions, we were particularly interested in making Solar I3M more adaptable, providing additional choices to users for them to determine how they could undertake interactions. To increase adaptability, we took the corpus that Solar I3M's conversational interface was trained on and made sure that all information accessible through the conversational interface could also be extracted through touch gesture-driven audio labels. Users are afforded choice in how they interact – e.g. a user may not feel comfortable engaging in conversational dialogue in public spaces [2], or may simply be tired, and may prefer to use a touch gesture to access information played back using an audio label (excluding search results discussed in Section 3.3.3).

*3.4.2 Facilitating More Natural Multimodal Interactions.* As early as the 1980s, the *Put-That-There* paradigm [7] described how speech and gestural inputs can be combined to converge on more *"natural user modalities"*. Their system allowed the use of demonstrative pronouns – e.g. *"that"* – to refer to the objects in speech commands when pointing fingers at shapes on a screen. Demonstrative pronouns are foundational to gestural-based systems in virtual and augmented reality [42]. We were interested in how demonstrative pronouns could be used in conversational dialogue during touch exploration, or when one or more planet models were picked up and held – e.g. *"what is this?"* – or compared. Demonstrative pronouns could enable blind end-users to engage more naturally in interactions that combine different modalities. Mainstream conversational interfaces, such as Siri and Alexa, do not support demonstrative pronouns, but [40] modified the Google Assistant SDK to support *"this"* and *"these"* pronouns in a WoZ application.

---

[6]https://developers.google.com/knowledge-graph

During S2, we explained the concept of demonstrative pronouns, and had our co-designers complete a hypothetical interaction: holding and comparing two planets, and asking – *"which of **these** is bigger?"*. While not yet supported by Solar I3M, we asked our co-designers what they thought of the interaction, and whether it was natural. Two co-designers suggested that it would allow them to perform interactions with objects they had yet to identify. One continued mentioning that they would feel less stressed about having to remember the name of each object. Our co-designers suggested additional pronouns including – *"this, it* and *that"*.

After the session concluded, we embedded accelerometers/gyroscopes into each planet model to allow tracking of which planets were being physically manipulated or held. Solar I3M's conversational interface was modified to support the use of demonstrative pronouns in voice commands and conversational dialogue. To explore this further, during S3 we asked our co-designer to complete a series of comparison tasks using the demonstrative pronouns. They felt that referring to planets using pronouns made them more willing to engage in conversational interactions, as it made them more natural and human-like. They also suggested that demonstrative pronouns should have some persistence between interactions – e.g. the ability to pick up Mars and ask *"how large is **this**?"*, before placing Mars back in its stand, picking up another planet and asking – *"is **it** larger than **this** planet?"*.

*3.4.3 Augmenting Interactions with Haptic Vibratory Output.* The majority of previous work into multimodal accessible aids has focused on combinations of tactile graphics/printed models with auditory output [23, 27, 31, 66]. [53] found a desire for combinations with haptic vibratory feedback, however, this was untested. Haptic vibratory feedback has been used in touch display systems [26, 49], but by their very nature, they are less tangible than physical I3Ms. During S2, S3 & S4, we explored how our co-designers wanted to engage in multimodal interactions, and how Solar I3M could utilise combinations of auditory and haptic inputs/outputs. As the components of Solar I3M were designed to be picked up and held, we were particularly interested in combinations of haptic vibratory feedback and auditory output.

During S2, we handed our co-designers a printed planet model with an embedded haptic motor, demonstrating basic haptic vibrations by turning the motor on and off. We asked our co-designers if they could think of any situations where they would like haptic vibrations to be used. One co-designer expressed excitement and suggested that vibrations could be used to help identify specific objects of interest – e.g. a particular planet. We designed for this during the session, so that whenever an auditory response made reference to a particular planet, a one-second haptic vibration would be emitted from that planet model. We asked our co-designers to find Jupiter, and upon receiving the auditory response – *"Jupiter is the fifth planet from the Sun"* – with a haptic vibration emitting from Jupiter, all co-designers agreed that:

- When engaging in tactile exploration, haptic vibrations combined with auditory responses would provide the information necessary to find objects or to gain an approximate location of where they could be found;
- The use of the haptic vibration helped to better convey the location of the planets in a non-visual form.

Progressing, one co-designer asked if they could perform their own interaction. They picked up Neptune and Venus, performed a *Long Press* on Neptune, and asked – *"Do Neptune and Venus have any moons?"*. Receiving the auditory response – *"Neptune has 14 moons, including..., while Venus has 0 moons."* – this co-designer expressed confusion regarding the fact that haptic vibrations had triggered from both planet models at the same time, and suggested that haptic vibrations should be carefully timed and synced. We made this change during the session, and when our co-designer performed this interaction again, they agreed that the haptic vibrations added more meaning to their interactions, and helped to better relate information regarding any tangible objects being held.

During S4, we revisited haptic vibratory feedback. Our co-designer suggested that haptic vibrations could provide feedback to confirm any touch gestures performed. They indicated that it would be useful to know when performing a touch gesture if they had executed it correctly, connecting this to the use of their smartphone screen reader that emits small haptic vibrations in response to gestural inputs. During the session, we created three unique haptic vibrations of varying duration, each mapped to a specific touch gesture (Table 4). Resuming, our co-designer then performed various touch gestures on Mars, and indicated that the haptic vibrations made their interactions more engaging and that confirming their inputs helped give them a greater sense of control.

Table 4: Each touch gesture triggers a unique haptic vibration

| Touch Gesture | Vibration Duration |
|---|---|
| Activate/Single Press | 150ms |
| Double Press | 150ms (x2) |
| Long Press | 500ms |

*3.4.4 Design Outcomes.* Across S2, S3 & S4, our co-designers provided clear design feedback on how they wanted to engage in richer multimodal interaction. This included **using haptics to augment other modalities**, and facilitating non-visual experiences that holistically offer more meaning. **Natural interaction paradigms** were also desired, allowing users to engage in conversational dialogue using more human-like language, such as the use of demonstrative pronouns.

## 4 EVALUATION

All design outcomes observed throughout our co-design journey (Section 3) were implemented into Solar I3M. We conducted a formal user study with blind participants in order to evaluate the design and usability of Solar I3M. Our evaluation consisted of a training and customisation period followed by both structured tasks and unstructured exploration in order to better understand the practical implications of our end-user-led design decisions.

### 4.1 Participants

Seven blind participants were recruited from our lab's participant contact pool. As a result of the low-incidence of blindness and associated difficulty in recruiting appropriate participants [14], many studies in the blind accessibility space involve between 6-12 participants [30, 45, 53, 60, 62]. Our work falls within this range. For user

studies that follow co-design or participatory processes and involve blind users, it is not uncommon for evaluations to be conducted with even fewer participants [23, 25]. None of the participants recruited for the user study were involved in our co-design journey.

**Table 5: Participant demographic information**

| Participant | P1 | P2 | P3 | P4 | P5 | P6 | P7 |
|---|---|---|---|---|---|---|---|
| **Level of Vision:** | | | | | | | |
| Blind | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Low-Vision | - | - | - | - | - | - | - |
| **Recent Accessible Formats Used:** | | | | | | | |
| Braille | ✓ | ✓ | - | ✓ | ✓ | ✓ | ✓ |
| Audio | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Raised Line | ✓ | - | - | - | ✓ | ✓ | ✓ |
| Models | - | - | ✓ | ✓ | ✓ | ✓ | - |
| **Familiarity (1: Not Familiar - 4: Very Familiar):** | | | | | | | |
| Tactile Graphics | 2 | 2 | 1 | 3 | 3 | 4 | 2 |
| Tactile Models | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| Interactive Models | 1 | 1 | 2 | 1 | 2 | 2 | 1 |
| **Recent Conversational Interfaces Used:** | | | | | | | |
| Alexa | - | - | - | ✓ | ✓ | - | - |
| Bixby | - | - | - | - | ✓ | - | - |
| Cortana | - | - | - | - | ✓ | ✓ | - |
| Google Assistant | - | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Siri | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **Accessed on Devices:** | | | | | | | |
| Phone | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Tablet | - | - | - | - | ✓ | - | ✓ |
| Speaker/Display | - | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Computer | - | - | - | - | ✓ | - | - |
| Television | - | - | - | - | ✓ | - | - |

All participants self-reported as blind (summarised in Table 5), with some reporting minimal low-levels of light perception. Age was spread evenly, from mid-20s to mid-70s. All participants regularly made use of text in the form of audio, while six knew how to read braille. Regular use of tactile models was not as common, however, six participants reported some or substantial past exposure to tactile graphics. In terms of technology use, all participants regularly used at least one conversational interface, with many reporting the use of several of these through smartphones and speakers. Common uses of conversational interfaces included checking the weather, dictating messages, and making phone calls.

## 4.2 Procedure

Each user study session lasted approximately 90 minutes, and began with a researcher providing a brief overview of the project, before guiding participants through the following components:

(1) **Training**. Participants were first asked to explore Solar I3M tactually with all system outputs turned off in order to build an initial level of comfort and confidence with the model. They were then taught to perform all of the touch gestures Solar I3M supported, as well as how to engage the conversational interface and supported voice commands. Participants were also guided through a series of sample interactions, and encouraged to perform their own interactions to build experience and comfort. This part of the session lasted 20-25 minutes.

(2) **Customisation**. Participants were then shown how they could customise Solar I3M's conversational outputs. Lasting 20-minutes, this included choosing a voice type *[feminine, masculine, robotic]*, voice speed *[100%, 150%, 200%]* and response length *[longer, shorter]*. For each customisable setting, participants were shown all available options, asked to choose their preferred setting, and informed that changes could be made at any time throughout the remainder of the user study.

(3) **Activity**. Once participants had built an understanding of how to use and customise Solar I3M, they completed one of two tasks: a set of five short structured activities, or an unstructured activity. Lasting a total of 10-15 minutes, the structured activities involved simple information gathering (e.g. discovering the largest planet, the temperature of Mars, etc). During the unstructured activity, participants were given up to 10 minutes to freely explore and interact with Solar I3M to discover and learn something previously unknown about the Solar System. The length of the study precluded participants from doing both, and separating them allowed us to compare participant behaviour.

## 4.3 Data Collection

We used two primary data collection methods, including a modified System Usability Scale (SUS) questionnaire, and a 30-minute semi-structured interview. Minor adjustments were made to the SUS questions, in consultation with our S3/S4 co-designer to ensure they would be meaningful in a blind context. All statements used in the SUS can be found in Appendix: Table 6. The semi-structured interview focused on participants' experiences with Solar I3M across key aspects of design and functionality, including touch, conversational, and multimodal interactions. Questions also explored whether Solar I3M's design respected participants' sense of agency, independence, and expectations around use. All participant sessions were video recorded for transcription.

## 4.4 Analysis

The goal of the user study was to assess the design and usability of Solar I3M from a qualitative perspective. We conducted a thematic analysis of: comments made by participants during the training, customisation and activity components; researcher observations of participant interactions throughout; and participant responses from the semi-structured interview. Based on our observations from the co-design sessions, we derived an initial set of codes and themes. Three members of the research team then independently coded one of the evaluation sessions, before meeting to extend the initial codes. All transcribed participant sessions were subsequently coded, with any conflicts between classifications reconciled in a subsequent meeting. Themes were then further refined and consolidated.

## 5   RESULTS

As alluded to earlier, our participants were split between those completing structured tasks (3 participants) and those engaging in unstructured exploration (4 participants). There were very minor differences in the types of interactions and strategies employed by the two groups. All participants started with a general idea of what they wanted to accomplish or discover, and chose a series of interactions that they perceived would be the most efficient means of accomplishing their goal.

We first present results from the SUS questionnaire followed by participant outcomes with respect to touch, conversation, and multimodal interaction, as captured throughout all components of the evaluation. Each of these topics is presented from the point of view of our final themes, which include how **prior experience** with personal technology influenced participant expectations, how participants desired interactions that upheld their **independence and control** and respected their **confidence and comfort**, and the desire for **customisation and personalisation** of the model. We found the themes to be highly interlinked; for example, being able to customise the voice of the Solar I3M made participants comfortable, and this in turn positively affected their confidence in using the model.
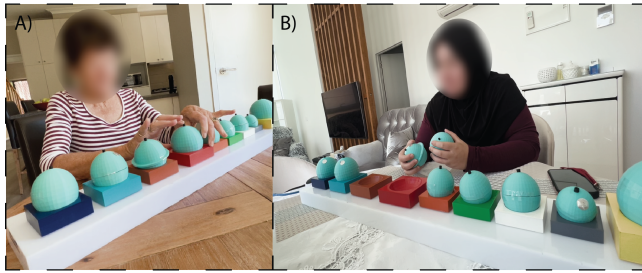


**Figure 4: Participants interacting with Solar I3M, A) performing a Double Press touch gesture on Jupiter, and B) holding and tactually comparing Jupiter and Saturn while engaging in conversation with Solar I3M.**

### 5.1   SUS Results

Participants generally had overall positive experiences with Solar I3M, with an average SUS score of 84.29 (Table 6). All participants indicated that they would be interested in using I3Ms that represented other content (SUS1), and the consensus was that the use of similar I3Ms would improve participant's uunderstanding of the related concept (SUS5). Additionally, all felt comfortable when interacting with Solar I3M (SUS9). While participants generally found that it was easy to interact with Solar I3M (SUS3), some felt that they would be more comfortable interacting with I3Ms alongside a support person (SUS4). Upon further enquiry, P3 indicated that further experience and training might mitigate this need and could build up their comfort and familiarity, aligning with SUS10.

### 5.2   Touch Interaction

All participants were able to successfully engage with and perform touch interactions during their time with Solar I3M (Figure

4a). Touch interaction generally transitioned from tactile exploration, assisting with immediate information gathering, to more comprehensive use of touch gestures to trigger audio label playback and invoke Solar I3M's conversational interface. As participants gained confidence in performing touch gestures, participants often picked up one or more of the planets to partake in more complex interactions.

*5.2.1   Prior Experience.* Participants drew heavily upon previous experiences with personal technology when performing touch interactions. This was seen across both *tactile exploration* and when performing *touch gestures*. All participants engaged in tactile exploration techniques when asked to explore Solar I3M. Participants more familiar with tactile graphics took part in more directed information gathering, while one participant with very limited experience – *P3* – engaged in more limited exploration.

Solar I3M's touch gestures – *Single Press, Double Press* and *Long Press* – were largely understood by all participants. Participants drew strongly from experiences with personal technology, expecting the touch gestures to be designed largely analogous to use on smartphones and tablets. This led to some initial confusion as to how they were to be performed. *P4* expected Single Press to act like the Single Tap gesture on their smartphone, describing *"[the Single Press] was not intuitive for me... it is usually a quick tap... not a tap and hold"*. There was a clear understanding amongst participants about what each touch gesture was used for, with two participants directly connecting the use of the *Long Press* to how they invoke the conversational interface on their smartphone (*P4 & P6*).

*5.2.2   Independence & Control.* Participants agreed that the physical design of Solar I3M's touch points – raised and protruding from the surface – helped them to undertake independent tactile exploration and to easily identify what parts of the model they were able to perform touch interactions on, which directly aligned with our co-design finding. *P4* suggested that touch points that were flush with the surface of a model could result in them *"accidentally hitting something when [they weren't] ready"*, while *P5* felt that the design of the touch points helped in orienting model components, *"[the touch point] aids in placing the planet properly and making sure it is upright"*. One participant, *P4*, advised that the design of the touch points could be further refined so that only the top would be sensitive to touch (pressable), finding that there were a few instances where they were able to accidentally trigger touch gestures by brushing the sides of the touch points.

The need for the Activate Press touch gesture was understood among participants. Three participants stated that the gesture made their touch interactions more deliberate (*P3, P6 & P7*), with *P6* reporting that it gave them the control to *"feel as though [they] wouldn't accidentally activate different planets ... unless it was explicitly what [they] wanted"*.

*5.2.3   Confidence & Comfort.* Continued use influenced the confidence of participants and their willingness to perform touch interactions. While it was common for participants to experience slight confusion or hesitation when first introduced to the touch gestures, participants became more confident in their interactions over the course of the session. *P3* spoke of how it took them time to *"get used to how long or hard [they] had to press the touch points"*, while

**Table 6: System Usability Scale responses**

| Str. Disagree | Disagree | Neutral | Agree | Str. Agree | Mean |
|---|---|---|---|---|---|
| SUS1: I would like to use I3Ms like this to access other content/concepts | | | | | 4.86 |
| - | - | - | 1 | 6 | |
| SUS3: The I3M was easy to use and interact with | | | | | 4.00 |
| - | 1 | 1 | 2 | 3 | |
| SUS5: I think using I3Ms like this would help me understand the relevant concept better | | | | | 4.71 |
| - | - | - | 2 | 5 | |
| SUS7: I imagine that most people would learn to use I3Ms like this very quickly | | | | | 4.71 |
| - | - | - | 2 | 5 | |
| SUS9: I felt very confident and comfortable using the I3M | | | | | 4.71 |
| - | - | - | 2 | 5 | |
| SUS2: The I3M was unnecessarily complex/difficult to use | | | | | 1.71 |
| 3 | 3 | 1 | - | - | |
| SUS4: I think I would need the use of a support person to be able to use I3Ms like this | | | | | 2.29 |
| 1 | 4 | 1 | 1 | - | |
| SUS6: I think the information you can get with the I3M is too limited | | | | | 1.86 |
| 4 | 1 | 1 | 1 | - | |
| SUS8: I think the I3M was cumbersome to use | | | | | 1.71 |
| 4 | 2 | - | 1 | - | |
| SUS10: I needed to learn a lot of things before I felt comfortable using the I3M | | | | | 1.71 |
| 3 | 3 | 1 | - | - | |

*P5* suggested that a practice mode similar to the VoiceOver screen reader on their smartphone would increase their confidence if they were to encounter similar I3Ms in day-to-day life. One participant, after experiencing difficulty reliably performing the Single Press gesture early on during their session, was even confident enough to state towards the end of the session that *"with use and with regularity ... the presses could take over as [my] preferred way [to interact]"* (*P7*).

*5.2.4 Customisation & Personalisation.* During the user study sessions a number of participants emphasised a desire to customise how touch interactions could be performed. Despite being able to reliably perform the touch gestures, *P6* suggested that the sensitivity of the touch gestures should be customisable and that this would be particularly useful for users who might have different levels of dexterity. Recalling their smartphone usage, *"on Android you can change the tap sensitivity ... with certain mobility issues, [touch gestures] would be trickier [to perform]"* (*P6*). *P4* desired control over the timing windows of touch gestures in order to make their interactions quicker and more efficient – e.g. being able to make *Single Press* faster and more analogous to the use of the Single Tap on their smartphone.

## 5.3 Conversational Interactions

Participants were largely comfortable engaging with the conversational interface from the moment it was introduced. Throughout the course of the user study, participants began gravitating more towards this form of interaction, which was seen as faster and more efficient.

*5.3.1 Prior Experience.* The expectations that participants had when undertaking conversational interactions were largely influenced by their experiences with the conversational interfaces found on their personal technology. This extended to how participants *expected to be able to engage* with Solar I3M's conversational interface.

The use of a *wake word* – *"Hey Model"* – was well understood amongst participants. This specific phrase was seen as being analogous to phrases used on their smart devices, *"well this is sort of going back to Apple and Siri... [Hey Siri]"* (*P7*). Using the *Long Press* touch gesture as a method of physically waking the conversational interface was also well understood. *P4* directly spoke of how a similar touch gesture was used on their smartphone, *"[this] is how you manually wake up voice assistants, there is continuity there"*.

*5.3.2 Independence & Control.* The conversational interface was seen as being the most efficient method of interaction among participants. Participants felt that engaging in a conversational interaction would be more straightforward, giving them the greatest level of control. The *"Hey Model"* wake word was also considered to be the fastest way of invoking the conversational interface, with *P7* mentioning how it was quicker and more direct. *P4* opted to use the wake word because they felt it gave them more control, finding the *Long Press* gesture to be *"arduous ... [it] was too long"*.

Despite not actively using it, all participants agreed that it was imperative to have a voice command that could repeat auditory responses – *"can you repeat that?"*. This directly validated our co-design finding. *P4* even suggested this command before it was covered in the user study, *"there is nothing as frustrating as if I can't have it repeat what is said... then I have to cycle through things... or just ignore it"*.

Participants felt that being able to seek responses from Google Knowledge Graph was particularly empowering, allowing them to undertake more extensive interactions. One participant, *P3*, spoke of how when using a conversational interface on their personal technology, they would often experience frustration when it responded with the message – *"I don't have any information on that"*. The majority of participants felt in control when Solar I3M entered its search flow – *"I don't know the answer to [question], but would you like me to search for it?"*. *P6* described how it made their interactions with Solar I3M *"[feel] more like a conversation rather than just using my voice as a keyboard"*.

A major area of friction during conversational interactions included the support of limited-phrase variations in voice commands. This often resulted in participants uttering voice commands that went undetected by Solar I3M, causing confusion. This occurred when Solar I3M would prompt participants if it should enter its search flow. Expecting the response – *"Yes, search for it"* – some participants would answer – *"Yes"* – which would go undetected. *P2* discussed how voice commands should support more phrase-agnostic expressions, saying that *"there is a lack of intuitiveness when you [say] 'yes go ahead, look it up', I want the process to be simpler... why not just yes"*.

*5.3.3 Confidence & Comfort.* Participants were comfortable asking questions using the conversational interface. This was no surprise to the researchers, as all seven participants indicated regular use of conversational interfaces on their personal technology (Table 5). *P5* openly indicated their comfort engaging in conversational interactions based on their prior experience, *"honestly because I am accustomed to using smart devices like Alexa, it comes naturally to me to ask questions"*. *P4* outlined that they would not be comfortable engaging in conversational interaction in public spaces, in part to not inconvenience others and due to privacy, describing *"in this [environment] where it is private, I will speak to it, but in more public environments, I prefer physical touch interactions"*. *P4* suggested that a headphone option might help somewhat mitigate this, but still felt uncomfortable engaging using speech and conversation open in public.

*P7* described how when the I3M offered to perform a search, having it admit that it couldn't answer them directly – *"I don't know the answer to [query/question]..."* – increased their confidence and trust, describing it as *"being genuine... it enhanced my confidence, because it was aware of what it knew and what it [didn't]"*. *P6* elaborated further, *"It helps build trust, rather than the teacher just making something up, it admits that it is fallible... it makes it seem more genuine"*. All seven participants emphasised that providing references was critical to their comfort, allowing them to come to their own interpretations as to how trustworthy returned information was. This directly aligned with one of our co-design findings.

*5.3.4 Customisation & Personalisation.* The ability to modify conversational output was seen as critical amongst participants, allowing them to reduce friction and consume information in a way that suited them. All participants were able to successfully use voice commands to choose their preferred voice type, speed, and response verbosity. Individual preferences were crucial, *"what suits one person may not suit another... but what I have chosen was good for*

*me"* (*P3*). Participants also described how their preferences could change based on multiple variables, including:

- **Their own *ability*:** Participants said that differing levels of hearing and comprehension could greatly influence how they customised conversational output. *P4 & P2* indicated that due to mild hearing impairment, they found human voices easier to understand compared to highly synthesised voices, while *P7* outlined that a memory impairment meant that they wouldn't confidently be able to follow at speeds greater than 100%.
- **Their own *personal preference*:** Participants indicared that they chose voices based on pleasantness, with more human-like WaveNet voices having more satisfying inflection and intonation. *P6* preferred WaveNet voices, but suggested that they might choose the eSpeak voice on their screen reader of its familiarity.
- **The *environment* where the I3M is being used:** *P6* said that in loud environments – e.g. a busy train – they might decrease the speed of conversational output so that they can better interpret speech, and also suggested that in a quiet environment, they might increase the speed and use a more synthetic-sounding robotic voice.
- **The *content* the I3M represents:** *P1, P2, P3 & P5* discussed how if the I3M represented content that they were particularly interested in learning about, they might prefer slower responses that were more easily consumable, with *P5* indicating that *"if you want to enjoy something and think about it deeply, then the speed needs to be at the pace of your thought process"*. *P6* said that for more technical and difficult-to-digest content – e.g. university materials — they would prefer the more verbose output.
- **The *purpose* the I3M is being used for:** *P1, P4, P5 & P7* indicated that if they were using an I3M to revise information they had already learnt, they might increase the speed of the conversational output, or reduce the response verbosity. *P4* spoke of how with critical tasks they might make output slower and more verbose in order to focus, *"if is it important for my life, like a new environment I want to physically go to, I would change it"*.

Participants were eager to make additional voice configuration suggestions. In particular, *P6* requested a voice type that was gender-neutral, in part because they felt their preference for the feminine voice might have been based on gender stereotypes, describing *"there might be some bias there... like the female secretary type thing... I don't know"*.

## 5.4 Multimodal Interactions

Throughout the user study, participants seamlessly performed interactions involving combinations of auditory and haptic input/output. It was not uncommon for participants to pick planets up, perform touch gestures on them whilst being held, and then engage in conversational interactions. Participants found the use of haptic vibratory feedback and demonstrative pronouns appealing, creating more natural experiences that helped to bind their interactions

together. Amongst some participants, there was a desire for interactions and experiences with I3Ms that would extend further into the space of more physically embodied models.

*5.4.1 Prior Experience.* Participants weren't explicit about whether any of their previous experiences with personal technology influenced their interaction strategy or willingness to combine and switch between modalities during interactions. It is, however, our belief that participants may have unknowingly been influenced by their use of personal technology, specifically smartphone use. When asked about their strategy, *P6* described their use of different modalities as *"...just [doing] what came to mind"*. We feel that the touch reading experience participants had also likely played a role in their willingness to engage and combine tactile information gathering and more directed touch – e.g. picking planets up – during their interactions.

*5.4.2 Independence & Control.* Participants expressed that elements of Solar I3M's multimodal design helped them to stay in control and independent during interactions. This was seen by how *adaptable* Solar I3M was, and how haptic vibratory feedback was used to *augment interactions* by providing added feedback and meaning. Participants provided direct suggestions on how vibratory feedback could further augment interactions to enhance their sense of control.

Participants found it empowering being able to choose their preferred interaction modalities, and how they could extract and access the same information using their choice of touch gesture input or conversational interaction. Participants began describing different situations where they might shift between the two, including based on the environment or the task. *P7* spoke of how they might choose to use touch gesture input to trigger audio label playback because they perceived that it was quicker than engaging in conversation, while *P5* said that if they were using an I3M to introduce themselves to a subject, they wanted to *"immerse [themselves] the way they would with a tactile image"*. *P4* indicated that the environment could be a determining factor, discussing how *"in this [environment] where it is private, I will speak to it, but in more public environments, I prefer physical touch interactions"*. Participants emphasised that it was vital that they weren't made to feel that they were missing out on information based on the interaction modality they chose to use – e.g. if touch gesture audio label playback gave shorter responses compared to conversational output.

There was agreement amongst all participants that the use of haptic vibratory feedback to augment touch and conversational interactions helped to increase a sense of control during interactions. Aligning directly with our co-design finding, participants found that haptic vibratory feedback helped to confirm their touch gesture inputs, *P3* described *"you knew that you'd done the right thing ... Single Press, Double Press, Long Press ... [haptics] confirmed it... "*. *P5* was more direct about how it improved their sense of control, speaking *"if I didn't have the feedback, I would keep [repeatedly] pressing"*.

*5.4.3 Confidence & Comfort.* During interactions, participants were seen comfortably changing between and combining different modalities, perceiving them as creating a **more natural experience**.

Participants enjoyed referring to planets using demonstrative pronouns during interactions. One participant, *P5*, was particularly fascinated by the concept, and repeatedly performed comparisons while holding a planet in each hand – asking questions including *"how large are **these**?"* and *"which of **these** is closer to the Sun?"* (Figure 4b). When asked, *P5* suggested that using pronouns made them more willing to pick planets up as part of their interactions, and that it made them more confident to explore and undertake more complex interactions. This aligned with the experiences of our co-designers.

Following on, *P5* was so comfortable during their interactions with Solar I3M that they began suggesting feedback that would allow them to partake in richer natural interactions with a more embodied Solar I3M. Asking if auditory output could be emitted from a speaker embedded directly within the planet they were holding, *P5* described that *"it would change the experience, right now they feel like spheres with no life other than when they vibrate, information just comes out of a speaker"*. Further elaborating, they suggested that a personality or character could be imbued within Solar I3M, detailing *"... or a guide... Einstein? ... [or] I am the Sun, be careful when you pick me up, I might be too hot!"*.

*5.4.4 Customisation & Personalisation.* While all participants found the use and combination of haptic vibratory feedback during touch and conversational interactions valuable, there was a desire amongst some participants to have more control over how it was used. One participant spoke of how due to sensory considerations, some users may want to customise any haptic vibratory feedback, specifically changing the intensity and length of the vibrations or even turning them off altogether (*P4*). Despite classifying themself as a *"very pro-haptics person"*, they further described how during one interaction, they wanted to interact focusing solely on Solar I3M's auditory responses, uninterrupted by any haptic vibratory feedback. *P4*'s response perhaps suggests that when it comes to modalities used to augment I3M outputs, some blind users may need better support to choose how multimodal or unimodal the output will be.

## 6 DISCUSSION

The role of personal technology played a significant factor during Solar I3M's design and evaluation. Specifically, our blind co-designers and participants desired interactions that drew from built-up knowledge from their *prior experience* with personal technology – e.g. smartphones. This applied to the touch gestures users could perform on Solar I3M and how the conversational interface operated, helping to increase the *confidence and comfort* of our users. The use of touch gestures is well understood amongst blind users due to the proliferation of touch input devices, and prior work has identified that blind users desire I3Ms that support consistent touch-based experiences [60, 61]. However, the degree to which prior experience influences conversational interfaces is less clear. [53] observed participants using smart speaker-analogous wake words with a WoZ I3M, but found that built-up experience with conversational interfaces seemingly didn't influence the likelihood of users engaging in interactions with conversational dialogue. Our observations deviated from this, with participants finding the conversational interface the most efficient way of interacting with Solar I3M, drawing a direct connection to their use of smart devices. One possible

reason for this might be that since this work in 2020, blind users may have become more reliant on and comfortable with conversational interfaces. It is clear that going forward I3Ms should be designed to be consistent with interaction experiences available on personal technology, allowing blind users to transfer knowledge.

The use of Solar I3M also contrasted with our previous work involving a WoZ I3M of the Solar System, where participants employed an interaction hierarchy: from tactile exploration to touch gesture audio labels and natural language questioning. Our findings suggest that the choice of interaction strategy may be more nuanced, with other variables impacting how blind users choose to interact with an I3M – including *task, environment, and familiarity with tactile graphics and conversational interface* – e.g. one participant was adamant that they wouldn't engage in any conversational interactions in public.

Our co-designers and participants provided design feedback on how Solar I3M could better support their *independence and control*. The need for independent experiences has been identified in prior work, allowing blind users to use I3Ms without the need of support workers [58], and to come to their own interpretations [53]. [31, 53] both observed blind users becoming frustrated when interrupted by I3M audio label playback during tactile exploration. Our findings aligned with these works, providing direct feedback on the physical design of touch points and activation gestures. But as Solar I3M is also the first I3M that integrates a fully-functional conversational interface inside the model itself, we also observed that the desire for independence extends beyond touch to conversational dialogue, and gathered direct design feedback that supports the independence of blind users when using Solar I3M to ask questions and perform searches.

The possibilities of multimodal interactions informed Solar I3M's design. Whereas most I3Ms combine printed models with auditory output in a limited fashion [23, 24, 27, 31], Solar I3M involvedtightly integrated conversational dialogue, audio, haptic vibration, and touch during interactions. Our co-designers directly suggested ways in which haptic vibratory feedback could be used to augment auditory responses with additional context/meaning, and confirm the successful activation of their touch gestures. We also explored the use of demonstrative pronouns, markedly extending previous work in accessibility research which used conversational interfaces in service of voice-driven menu navigation and extraction of audio labelling [18, 35]. Demonstrative pronouns were integral to facilitating blind participants' engagement in more natural dialogue during interactions, leading one participant to desire I3Ms with more physically embodied characters.

The ability to *customise and personalise* Solar I3M was strongly desired amongst our participants during their interactions. This extended across all aspects of Solar I3M's design: from touch gestures, to the conversational interface, and to associated auditory output and haptic vibratory feedback. Our participants indicated that being afforded this level of control would allow them to determine their own pace during interactions, and that preferences could change based on multiple variables, including ability and environment. Prior work has already identified that customisation of conversational interfaces is needed to help remove restrictions on blind users' interactions [2, 8, 20]. Our findings align with this work, highlighting the importance of customisation, but extending

to a new context – when conversational interfaces are integrated inside tangible I3Ms. We also observed that our users expected to be able to customise other interactions, including how touch gestures were performed, and the timing, sensitivity and intensity of haptic vibratory feedback. To our knowledge, this level of I3M customisation has not been previously considered, and highlights a need to support blind users in customising interactions to cater to very personal preferences and abilities.

## 6.1 I3M Design Recommendations

We developed a set of design recommendations based on our observations and the results from Solar I3M's design journey and evaluation. These recommendations coalesce across our identified themes, and focus on the design of I3Ms that support blind users' desire for – use of **prior experience**, interactions that uphold their **independence and control**, sense of **confidence and comfort**, and allow **customisation and personalisation**. The following design recommendations are being put forward to the accessibility research community to help guide the implementation of future I3Ms:

- *Interruption-free tactile exploration:* I3Ms should be designed to respect the *independence and control* of blind users, especially when any tactile exploration takes place to discover/interpret information. Touch points should be physically designed to be easily identifiable, and may protrude or recess [31] the surface of the model based on the degree of salient details on the model. Additional features, including an Activate Press touch gesture, on/off buttons [60], or voice commands, should be used to prevent unintended touch or conversational interactions from occurring, reducing friction for blind users during initial exploration or subsequent interactions.

- *Leverage prior interaction experience with personal technology:* Blind users expect to be able to interact with I3Ms in ways similar to their smartphones and smart speakers, aligning with [53]. I3Ms should be designed to take advantage of this *prior experience*, supporting interactions that are similar to those used with personal technology. Users should be able to perform gestures – including Single, Double and Long Taps or Presses – during touch interactions to extract basic information. Users should also be able to use wake words, touch gestures, and easy-to-remember voice commands – similar to Siri and Google Assistant – when engaging in conversational interactions to extract richer information that fills gaps in knowledge.

- *Support customisation and personalisation:* I3M interactions should be designed to be customised, allowing blind users to determine their own pace and preference when performing interactions. Based on personal preference, task or environment, *customisation and personalisation* should extend to how users both perform inputs and receive output during their interactions, influencing their *confidence and comfort*. Touch gestures should have adjustable timing and sensitivity, while haptic vibratory effects should have adjustable intensity and timing, or be able to be turned off to

support any users with sensory considerations. Conversational output should have customisable high-quality voice types, voice speeds, and response lengths [2, 8], allowing blind users to determine the pace and form in which auditory responses are consumed.

- **Support more natural dialogue:** I3Ms should allow blind users to leverage conversational interfaces to fill gaps in their knowledge, empowering their desire for *independence and control* during interactions. Users should be able to use demonstrative pronouns [7], including *"these, this, it, that"*, to refer to any I3M model components being touched or held. Demonstrative pronouns help support more intuitive dialogue, which is particularly important for blind users, allowing them to engage conversationally with objects that are yet to be verbally identified, versus a sighted user who might identify an object using visual information (e.g. colour). Additionally, users should be able to authorise I3Ms to perform searches and be able to respond to follow-up queries using straightforward language – e.g. command-agnostic responses *"yes", "okay", "sure"* over command-specific phrases like *"yes, search for it"* – increasing their *confidence* and *comfort* when engaging conversationally.

- **Tightly coupled haptic feedback:** Haptic vibratory feedback should augment speech and touch to create more intuitive and natural interactions. It can be used to confirm the successful activation of touch gestures, increasing the *independence* and *control* of blind users as they validate that their interactions are being fulfilled. Haptic vibratory feedback localised within model components allows blind users to better relate conversational responses to specific components that are being held, compared to sighted users who are able to use visual information to help achieve this. In general, tightly coupled haptic feedback supports richer interactions – e.g. comparing multiple components held simultaneously, and could be used with an I3M of a map in order to more effectively direct a user to a destination using combinations of audio and haptic feedback.

## 7 LIMITATIONS & FUTURE WORK

Solar I3M's design and evaluation resulted in the creation of our design recommendations. However, a limitation is that our recommendations were distilled from the design of a single I3M. A major focus of our future work will be to examine and expand these using different models – e.g. an I3M map used for orientation and mobility training. Additionally, despite seeing a large degree of consistency across our participant group, we would like to run evaluations with a larger participant base. We would also like to run a similar study with low-vision participants, as we suspect the tactile and conversational nature of Solar I3M will be useful in low-vision contexts. These will add value and strengthen the generalisability of our recommendations.

We also aim to explore the design of more physically embodied I3Ms. Some of our participants desired interactions that were more personal and *alive*. This included wanting to address conversational commands directly to physical components being held (e.g. *"Mars*

*how big are you?"*), having auditory responses delivered from speakers inside individual components rather than an enclosure/stand, and deeper use of haptic vibratory feedback. This would be particularly interesting in that the embodiment of interactive interfaces and agents traditionally relies on the perception of many visual attributes [56]. I3Ms could have modes further along the human-to-human conversational model, introducing additional conversational protocols [16] and a more human-like social presence [41].

We are also interested in investigating the role of trust between blind users and I3Ms. Trust has proven to be a significant factor in other under-served populations when deploying new systems (e.g. the elderly in healthcare [15, 63]). Developing a trust relationship is particularly important for people who are blind, as the usefulness of an accessible graphic, aid, or tool depends entirely on whether the user is willing to accept and engage with the information it provides. With Solar I3M, the stakes were relatively low; if incorrect information was provided, users could merely ignore it or cease using Solar I3M. However, an opportunity exists to explore the concept of trust further, particularly with I3Ms in contexts where trust and confidence are more significant – e.g. with orientation and mobility training, in which erroneous information could have safety implications.

## ACKNOWLEDGMENTS

## 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)*. ACM, New York, NY, USA, 37–42. https://doi.org/10.1145/2468356.2468364

[2] Ali Abdolrahmani, Ravi Kuber, and Stacy M. Branham. 2018. "Siri Talks at You": An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) *(ASSETS '18)*. ACM, New York, NY, USA, 249–258. https://doi.org/10.1145/3234695.3236344

[3] Jérémy Albouys-Perrois, Jérémy Laviole, Carine Briant, and Anke M. Brock. 2018. Towards a Multisensory Augmented Reality Map for Blind and Low Vision People: A Participatory Design Approach. 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, 1–14. https://doi.org/10.1145/3173574.3174203

[4] Frances K. Aldrich and Linda Sheppard. 2001. Tactile graphics in school education: perspectives from pupils. *British Journal of Visual Impairment* 19, 2 (2001), 69–73. https://doi.org/10.1177/026461960101900204 arXiv:https://doi.org/10.1177/026461960101900204

[5] Shiri Azenkot and Nicole B. Lee. 2013. Exploring the Use of Speech Input by Blind People on Mobile Devices. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (Bellevue, Washington) *(ASSETS '13)*. ACM, New York, NY, USA, Article 11, 8 pages. https://doi.org/10.1145/2513383.2513440

[6] Mark Blades, Simon Ungar, and Christopher Spencer. 1999. Map Use by Adults with Visual Impairments. *The Professional Geographer* 51, 4 (1999), 539–553. https://doi.org/10.1111/0033-0124.00191 arXiv:https://doi.org/10.1111/0033-0124.00191

[7] Richard A. Bolt. 1980. "Put-That-There": Voice and Gesture at the Graphics Interface. In *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques* (Seattle, Washington, USA) *(SIGGRAPH '80)*. Association for Computing Machinery, New York, NY, USA, 262–270. https://doi.org/10.1145/800250.807503

[8] Stacy M. Branham and Antony Rishin Mukkath Roy. 2019. Reading Between the Guidelines: How Commercial Voice Assistant Guidelines Hinder Accessibility for Blind Users. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association

for Computing Machinery, New York, NY, USA, 446–458. https://doi.org/10.1145/3308561.3353797

[9] Anke M. Brock, Philippe Truillet, Bernard Oriola, Delphine Picard, and Christophe Jouffrais. 2012. Design and User Satisfaction of Interactive Maps for Visually Impaired People. *Computers Helping People with Special Needs* 7383 (2012), 544–551. https://doi.org/10.1007/978-3-642-31534-3_80

[10] 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* (Kingston, Ontario, Canada) *(TEI '12)*. ACM, New York, NY, USA, 131–138. https://doi.org/10.1145/2148131.2148160

[11] Emeline Brule, Gilles Bailly, Anke Brock, Frederic Valentin, Grégoire Denis, and Christophe Jouffrais. 2016. MapSense: Multi-Sensory Interactive Maps for Children Living with Visual Impairments. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 445–457. https://doi.org/10.1145/2858036.2858375

[12] Erin Buehler, Niara Comrie, Megan Hofmann, Samantha McDonald, and Amy Hurst. 2016. Investigating the Implications of 3D Printing in Special Education. *ACM Trans. Access. Comput.* 8, 3, Article 11 (March 2016), 28 pages. https://doi.org/10.1145/2870640

[13] 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. https://doi.org/10.1080/07294360.2016.1177001 arXiv:https://doi.org/10.1080/07294360.2016.1177001

[14] Matthew Butler, Leona M Holloway, Samuel Reinders, Cagatay Goncu, and Kim Marriott. 2021. Technology Developments in Touch-Based Accessible Graphics: A Systematic Review of Research 2010-2020. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 278, 15 pages. https://doi.org/10.1145/3411764.3445207

[15] Felix Carros, Johanna Meurer, Diana Löffler, David Unbehaun, Sarah Matthies, Inga Koch, Rainer Wieching, Dave Randall, Marc Hassenzahl, and Volker Wulf. 2020. Exploring Human-Robot Interaction with the Elderly: Results from a Ten-Week Case Study in a Care Home. 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, 1–12. https://doi.org/10.1145/3313831.3376402

[16] Justine Cassell. 2000. More than just another pretty face: Embodied conversational interface agents. *Commun. ACM* 43, 4 (2000), 70–78.

[17] Luis Cavazos Quero, Jorge Iranzo Bartolomé, Dongmyeong Lee, Yerin Lee, Sangwon Lee, and Jundong Cho. 2019. Jido: A Conversational Tactile Map for Blind People. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 682–684. https://doi.org/10.1145/3308561.3354600

[18] Luis Cavazos Quero, Jorge Iranzo Bartolomé, Seonggu Lee, En Han, Sunhee Kim, and Jundong Cho. 2018. An Interactive Multimodal Guide to Improve Art Accessibility for Blind People. 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, 346–348. https://doi.org/10.1145/3234695.3241033

[19] 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). https://doi.org/10.3390/electronics10030297

[20] Dasom Choi, Daehyun Kwak, Minji Cho, and Sangsu Lee. 2020. "Nobody Speaks That Fast!" An Empirical Study of Speech Rate in Conversational Agents for People with Vision 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. https://doi.org/10.1145/3313831.3376569

[21] 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. https://doi.org/10.1145/3313831.3376513

[22] Alistair D. N. Edwards, Nazatul Naquiah Abd Hamid, and Helen Petrie. 2015. Exploring Map Orientation with Interactive Audio-Tactile Maps. In *Human-Computer Interaction – INTERACT 2015*, Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler (Eds.). Springer International Publishing, Cham, 72–79.

[23] 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. https://doi.org/10.1145/3322276.3323686

[24] 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, 930 (2017). https://doi.org/10.3389/fpsyg.2017.00930

[25] Tom Giraud, Ines Di Loreto, and Matthieu Tixier. 2020. The Making of Accessibility to Rural Place for Blind People: The Relational Design of an Interactive Map. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1419–1431. https://doi.org/10.1145/3357236.3395527

[26] Cagatay Goncu and Kim Marriott. 2011. GraVVITAS: Generic Multi-touch Presentation of Accessible Graphics. In *Human-Computer Interaction – INTERACT 2011*, Pedro Campos, Nicholas Graham, Joaquim Jorge, Nuno Nunes, Philippe Palanque, and Marco Winckler (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 30–48.

[27] Timo Götzelmann, Lisa Branz, Claudia Heidenreich, and Markus Otto. 2017. A Personal Computer-based Approach for 3D Printing Accessible to Blind People. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments* (Island of Rhodes, Greece) *(PETRA '17)*. ACM, New York, NY, USA, 1–4. https://doi.org/10.1145/3056540.3064954

[28] Jaume Gual, Marina Puyuelo, Joaquim Lloverás, and Lola Merino. 2012. Visual Impairment and urban orientation. Pilot study with tactile maps produced through 3D Printing. *Psyecology* 3, 2 (2012), 239–250.

[29] Eric Hasper, Rogier A. Windhorst, Terri Hedgpeth, Leanne Van Tuyl, Ashleigh Gonzales, Britta Martinez, Hongyu Yu, Zoltan Farkas, and Debra P. Baluch. 2015. Methods for Creating and Evaluating 3D Tactile Images to Teach STEM Courses to the Visually Impaired. *Journal of College Science Teaching* 44, 6 (2015), 92–99. http://www.jstor.org/stable/43632001

[30] Leona Holloway, Swamy Ananthanarayan, Matthew Butler, Madhuka Thisuri De Silva, Kirsten Ellis, Cagatay Goncu, Kate Stephens, and Kim Marriott. 2022. Animations at Your Fingertips: Using a Refreshable Tactile Display to Convey Motion Graphics for People Who Are Blind or Have Low Vision. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 32, 16 pages. https://doi.org/10.1145/3517428.3544797

[31] Leona Holloway, Matthew Butler, and Kim Marriott. 2018. Accessible maps for the blind: Comparing 3D printed models with tactile graphics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, QC, Canada) *(CHI '18)*. ACM. https://doi.org/10.1145/3173574.3173772

[32] Leona Holloway, Kim Marriott, Matthew Butler, and Alan Borning. 2019. Making Sense of Art: Access for Gallery Visitors with Vision Impairments. 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, 1–12. https://doi.org/10.1145/3290605.3300250

[33] Leona Holloway, Kim Marriott, Matthew Butler, and Samuel Reinders. 2019. 3D Printed Maps and Icons for Inclusion: Testing in the Wild by People Who Are Blind or Have Low Vision. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 183–195. https://doi.org/10.1145/3308561.3353790

[34] Michele Hu. 2015. Exploring New Paradigms for Accessible 3D Printed Graphs. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers &#38; Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. ACM, New York, NY, USA, 365–366. https://doi.org/10.1145/2700648.2811330

[35] 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. https://doi.org/10.1145/3294109.3300994

[36] 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*. ACM, 247–252.

[37] Jill Keeffe. 2005. Psychosocial Impact of Vision Impairment. *International Congress Series* 1282 (2005), 167–173.

[38] Jeeeun Kim and Tom Yeh. 2015. Toward 3D-printed movable tactile pictures for children with visual impairments. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2815–2824.

[39] N. W. Kim, S. C. Joyner, A. Riegelluth, and Y. Kim. 2021. Accessible Visualization: Design Space, Opportunities, and Challenges. *Computer Graphics Forum* 40, 3 (2021), 173–188. https://doi.org/10.1111/cgf.14298

[40] Jaewook Lee, Sebastian S. Rodriguez, Raahul Natarrajan, Jacqueline Chen, Harsh Deep, and Alex Kirlik. 2021. What's This? A Voice and Touch Multimodal Approach for Ambiguity Resolution in Voice Assistants. In *Proceedings of the 2021 International Conference on Multimodal Interaction* (Montréal, QC, Canada) *(ICMI '21)*. Association for Computing Machinery, New York, NY, USA, 512–520. https://doi.org/10.1145/3462244.3479902

[41] Michal Luria, Samantha Reig, Xiang Zhi Tan, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. 2019. Re-Embodiment and Co-Embodiment: Exploration of Social Presence for Robots and Conversational Agents. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 633–644. https://doi.org/10.1145/3322276.3322340

[42] Sven Mayer, Jens Reinhardt, Robin Schweigert, Brighten Jelke, Valentin Schwind, Katrin Wolf, and Niels Henze. 2020. Improving Humans' Ability to Interpret Deictic Gestures in Virtual Reality. 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, 1–14. https://doi.org/10.1145/3313831.3376340

[43] 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)*. ACM, New York, NY, USA, 275–276. https://doi.org/10.1145/2661334.2661392

[44] Betsy Phillips M.S. and Hongxin Zhao Ph.D. 1993. Predictors of Assistive Technology Abandonment. *Assistive Technology* 5, 1 (1993), 36–45. https://doi.org/10.1080/10400435.1993.10132205 arXiv:https://doi.org/10.1080/10400435.1993.10132205 PMID: 10171664.

[45] 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. https://doi.org/10.1145/3544548.3581389

[46] National Federation of the Blind. 2009. The Braille Literacy Crisis in America: Facing the Truth, Reversing the Trend, Empowering the Blind. (2009). Available from https://nfb.org/images/nfb/documents/pdf/braille_literacy_report_web.pdf.

[47] Helen Petrie, Christoph Schlieder, Paul Blenkhorn, Gareth Evans, Alasdair King, Anne-Marie O'Neill, George T. Ioannidis, Blaithin Gallagher, David Crombie, Rolf Mager, and Maurizio Alafaci. 2002. TeDUB: A System for Presenting and Exploring Technical Drawings for Blind People. In *Computers Helping People with Special Needs*, Klaus Miesenberger, Joachim Klaus, and Wolfgang Zagler (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 537–539.

[48] Mahika Phutane, Julie Wright, Brenda Veronica Castro, Lei Shi, Simone R. Stern, Holly M. Lawson, and Shiri Azenkot. 2021. Tactile Materials in Practice: Understanding the Experiences of Teachers of the Visually Impaired. *ACM Transactions on Accessible Computing (TACCESS)* (2021). https://doi.org/10.1145/3508364

[49] Benjamin Poppinga, Charlotte Magnusson, Martin Pielot, and Kirsten Rassmus-Gröhn. 2011. TouchOver Map: Audio-tactile Exploration of Interactive Maps. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (Stockholm, Sweden) *(MobileHCI '11)*. ACM, New York, NY, USA, 545–550. https://doi.org/10.1145/2037373.2037458

[50] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. "Accessibility Came by Accident": Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 459, 13 pages. https://doi.org/10.1145/3173574.3174033

[51] Leah M. Reeves, Jennifer Lai, James A. Larson, Sharon Oviatt, T. S. Balaji, Stéphanie Buisine, Penny Collings, Phil Cohen, Ben Kraal, Jean-Claude Martin, Michael McTear, TV Raman, Kay M. Stanney, Hui Su, and Qian Ying Wang. 2004. Guidelines for Multimodal User Interface Design. *Commun. ACM* 47, 1 (Jan. 2004), 57–59. https://doi.org/10.1145/962081.962106

[52] 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)*. ACM, New York, NY, USA, 91–100. https://doi.org/10.1145/2982142.2982176

[53] 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. https://doi.org/10.1145/3313831.3376145

[54] L. Penny Rosenblum and Tina S. Herzberg. 2015. Braille and Tactile Graphics: Youths with Visual Impairments Share Their Experiences. *Journal of Visual Impairment & Blindness* 109, 3 (2015), 173–184. https://doi.org/10.1177/0145482X1510900302 arXiv:https://doi.org/10.1177/0145482X1510900302

[55] Jonathan Rowell and Simon Ungar. 2005. Feeling our way: tactile map user requirements – a survey. In *In Proceedings of 22nd International Cartographic Conference. International Cartographic Association*.

[56] Ameneh Shamekhi, Q. Vera Liao, Dakuo Wang, Rachel K. E. Bellamy, and Thomas Erickson. 2018. Face Value? Exploring the Effects of Embodiment for a Group Facilitation Agent. 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. https://doi.org/10.1145/3173574.3173965

[57] Rebecca Sheffield. 2016. International Approaches to Rehabilitation Programs for Adults who are Blind or Visually Impaired: Delivery Models, Services, Challenges and Trends.

[58] 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)*. ACM, New York, NY, USA, Article 197, 14 pages. https://doi.org/10.1145/3290605.3300427

[59] 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 34rd Annual ACM Conference on Human Factors in Computing Systems (CHI'16)*. https://doi.org/10.1145/2858036.2858507

[60] 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)*. ACM, New York, NY, USA, 200–209. https://doi.org/10.1145/3132525.3132549

[61] 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* (Qu&#233;bec City, QC, Canada) *(UIST '17)*. ACM, New York, NY, USA, 493–506. https://doi.org/10.1145/3126594.3126650

[62] Lei Shi, Yuhang Zhao, Ricardo Gonzalez Penuela, Elizabeth Kupferstein, and Shiri Azenkot. 2020. Molder: An Accessible Design Tool for Tactile Maps. 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, 1–14. https://doi.org/10.1145/3313831.3376431

[63] Brendan Spillane, Emer Gilmartin, Christian Saam, and Vincent Wade. 2019. Issues Relating to Trust in Care Agents for the Elderly. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (Dublin, Ireland) *(CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 20, 3 pages. https://doi.org/10.1145/3342775.3342808

[64] Abigale Stangl, Chia-Lo Hsu, and Tom Yeh. 2015. Transcribing Across the Senses: Community Efforts to Create 3D Printable Accessible Tactile Pictures for Young Children with Visual Impairments. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers &#38; Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. ACM, New York, NY, USA, 127–137. https://doi.org/10.1145/2700648.2809854

[65] Brandon T. Taylor, Anind K. Dey, Dan P. Siewiorek, and Asim Smailagic. 2015. TactileMaps.net: A web interface for generating customized 3D-printable tactile maps. In *Proc. ACM SIGACCESS Conference on Computers & Accessibility*. ACM, 427–428. https://doi.org/10.1145/2700648.2811336

[66] Lauren Thevin, Christophe Jouffrais, Nicolas Rodier, Nicolas Palard, Martin Hachet, and Anke M. Brock. 2019. Creating Accessible Interactive Audio-Tactile Drawings Using Spatial Augmented Reality. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) *(ISS '19)*. Association for Computing Machinery, New York, NY, USA, 17–28. https://doi.org/10.1145/3343055.3359711

[67] Henry B Wedler, Sarah R Cohen, Rebecca L Davis, Jason G Harrison, Matthew R Siebert, Dan Willenbring, Christian S Hamann, Jared T Shaw, and Dean J Tantillo. 2012. Applied computational chemistry for the blind and visually impaired. *Journal of Chemical Education* 89, 11 (2012), 1400–1404.