



PDF Download
3706598.3713706.pdf
09 February 2026
Total Citations: 2
Total Downloads: 1223

Latest updates: <https://dl.acm.org/doi/10.1145/3706598.3713706>

RESEARCH-ARTICLE

Enhancing Tactile Learning: A Co-Designed System for Supporting Speech Interaction with Multi-Part 3D Printed Models by Students who are Blind

RUTH GALAN NAGASSA, Monash University, Melbourne, VIC, Australia

ANDRE KY PHAM, Monash University, Melbourne, VIC, Australia

MATTHEW BUTLER, Monash University, Melbourne, VIC, Australia

LEONA M HOLLOWAY, Monash University, Melbourne, VIC, Australia

KALIN STEFANOV, Monash University, Melbourne, VIC, Australia

SKYE DE VENT, Monash University, Melbourne, VIC, Australia

[View all](#)

Open Access Support provided by:

[Monash University](#)

Published: 26 April 2025

[Citation in BibTeX format](#)

CHI 2025: CHI Conference on Human Factors in Computing Systems
April 26 - May 1, 2025
Yokohama, Japan

Conference Sponsors:
[SIGCHI](#)

Enhancing Tactile Learning: A Co-Designed System for Supporting Speech Interaction with Multi-Part 3D Printed Models by Students who are Blind

Ruth Galan Nagassa
Monash University
Melbourne, VIC, Australia
ruth.nagassa@monash.edu

Andre Ky Pham
Monash University
Melbourne, Victoria
Australia
andrekypham@gmail.com

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

Leona Holloway
Monash University
Melbourne, VIC, Australia
leona.holloway@monash.edu

Kalin Stefanov
Monash University
Melbourne, Australia
kalin.stefanov@monash.edu

Skye de Vent
Monash University
Melbourne, Australia
skyedevent@gmail.com

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

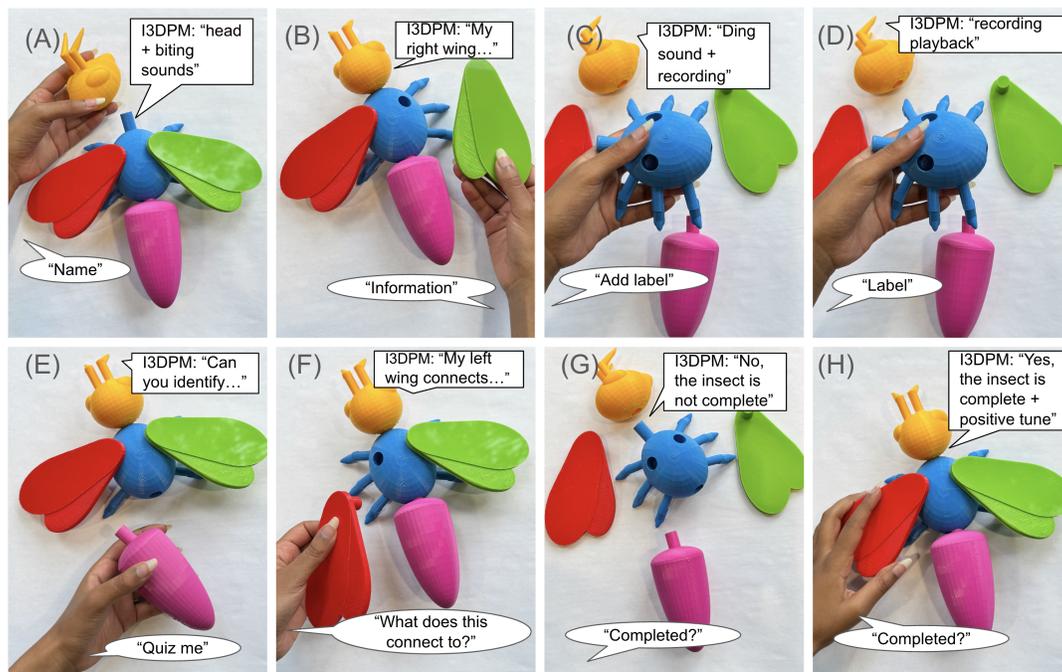


Figure 1: The iPad application supports interaction with an interactive 3D printed model (I3DPM) of an insect through speech, enabling blind users to explore the insect model's anatomy (A-B), create annotations (C-D), test their knowledge (E), and gain support for assembly tasks (F-H).

Abstract

3D printed models (3DPMs) are increasingly used to support the education of students who are blind or have low vision (BLV). As 3DPMs are more widely-adopted, educators are using more complex

multi-part models. However, with this increased complexity comes additional challenges for their use, such as supporting audio labels of multiple parts as well as guiding the assembly and disassembly of the model. This work explores the co-design and evaluation of a system that supports the use of multi-part 3DPMs by BLV students. Working with BLV adults and children, as well as educators, an iPad application was developed to support interaction with an insect model, including speech interaction and support for assembly. Evaluation showed that the system was strongly enjoyed by students and educators were enthusiastic as they believed it would



increase classroom engagement and inclusion, and its support for voice annotation could be used for assessment.

CCS Concepts

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

Keywords

Accessibility, 3D printed models, computer vision, blind, low vision, education

ACM Reference Format:

Ruth Galan Nagassa, Andre Ky Pham, Matthew Butler, Leona Holloway, Kalin Stefanov, Skye de Vent, and Kim Marriott. 2025. Enhancing Tactile Learning: A Co-Designed System for Supporting Speech Interaction with Multi-Part 3D Printed Models by Students who are Blind. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan*. ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/3706598.3713706>

1 Introduction

Students who are blind or have low vision (BLV) encounter numerous challenges accessing the school curriculum. Access to visual graphics, such as images and diagrams, is particularly problematic. This severely impacts BLV student engagement and learning outcomes in disciplines such as science, technology, engineering and mathematics (STEM) which make heavy use of graphics [36, 49].

3D printed models (3DPMs) have proven to be a promising approach to overcome this barrier and are increasingly used to teach concepts in STEM. This includes astronomy [2], programming [35, 38], chemistry [65] and biology [5]. As 3DPMs become more widely used in education, there has been growing interest in multi-part 3DPMs [28, 59]. Such models facilitate the exploration of relationships between different model components [59], foster greater engagement [50, 54, 56] and provide BLV learners with opportunities to practice hands-on motor skills and independently build their own models [22]. However, to date facilitating meaningful interactions with multi-part models remains under-explored.

Multi-part 3DPMs raise two difficult interaction challenges. The first, is how to provide guidance for BLV students when trying to understand the relationships between parts and the way in which they fit together. Previous work into laser cut multi-part models has shown that vision plays a major role in the physical manipulation of model parts, particularly for alignment and assembly, and that BLV participants experience greater challenges during assembly and disassembly [15].

The second challenge is how to best provide the BLV student with detailed information about individual components. Due to formatting requirements, only short braille labels are suitable for 3DPMs and these are often accompanied by a legend or key which adds to the cognitive load [30]. In response to this, there has been a surge in research exploring options for audio labelling [10]. However, most research into labelling has considered single part models and audio labelling of multi-part 3DPMs remains challenging because of the number of components and their typically smaller size.

Here, we describe a novel computer vision-based app that addresses these two challenges (Figure 1). Prior research into computer

vision-based audio labelling has considered single-part 3DPMs [18]. Here, we consider the challenges of labelling multi-part models and providing guidance during assembly. Importantly, the app was co-designed with BLV adults (one of whom is a co-author) and educators of BLV students, to ensure that it meets the needs of BLV students and their educators. In particular, the app runs on tablet devices (iPads) already used in the classroom and utilises only ‘off-the-shelf’ software for object and voice recognition. Our specific contributions are:

- The first computer vision-based system with a voice-based interface for BLV students to interact with multi-part 3D printed models that provides information about individual components and can guide assembly;
- An evaluation of the system with BLV students and their educators demonstrating the potential educational value of the system and its ability to foster classroom inclusion;
- Design considerations based on our co-design process and evaluation that relate to supporting the use of multi-part 3DPMs within educational settings, meeting the needs and preference of educators for use in the classroom, and practicality for implementation and ongoing use.

Our research provides strong evidence that multi-part 3DPMs utilising computer vision-based audio interaction can enable BLV students to overcome the barriers they currently face in accessing graphical information, particularly in STEM disciplines. Our findings reveal strong educator interest in integrating these into their teaching practice, as well as key functional, educational and practical requirements essential to the successful design and implementation of future computer vision-supported educational tools.

2 Related Work

2.1 3D Printed Models in Education

3D printing has emerged as a fast, inexpensive and highly individual production method for tactile materials [9, 44]. Today, 3DPMs are increasingly used as an alternative format for the accessible representation of graphics in areas including orientation and mobility training [31, 45], arts and culture [12], accessible design [41] and particularly in the context of education [37, 65]. Given the advantages of representing graphics in 3D to touch readers [29], 3DPMs have been used to teach children’s books [37], maths [6], astronomy [3] and science [66, 71]. With a range of 3D models available on open source platforms [7] and the increasing presence of 3D printers in community maker spaces, schools and libraries, now more than ever 3DPMs are accessible to the community. As 3D printing continues to become more accessible, its presence in education can be expected to significantly expand.

2.2 Interactive 3D Printed Models

Interactive 3D printed models (I3DPMs) that provide audio labels have gained considerable research attention. Specifically, within the context of special education, it has encouraged STEM engagement and the development of educational aids for accessible curriculum content [8, 9]. This includes I3DPMs to teach tactual shape perception and spatial awareness sub-concepts [32], as well as geography [24] and hardware computing [14, 20]. Researchers have explored

different approaches to adding interactive audio elements to 3DPMs for a more informative and engaging user experience. These approaches have involved the use of embedded electronics [25, 42], overlays on touchscreens [20, 69], acoustic sensing [58, 61, 67], NFC or QR technology [11, 46] and PenFriend [24, 57].

Interestingly, research into I3DPMs have predominantly focused on single-part models. A notable exception is the work by Reinders and colleagues who developed an interactive solar system model with removable planets [55, 56]. Each planet was embedded with electronics and a touch point, and allowed users to pick up each planet from the base. A conversational agent interface permitted natural language interrogation of the I3DPM. This was valuable to confirm and deepen BLV user understanding [56]. One of the notable features of the system was its support for model assembly, in which it confirmed the correct or incorrect placement of models [56]. However, this approach employs bespoke software and hardware, and models must be large enough to compartmentalise electronics. These requirements limit its application within the education context where a range of 3D models are required, often consisting of small parts. For this reason, we identified that a computer vision-based approach may be a promising alternative.

2.3 Computer Vision-Based Systems for I3DPMs

Computer vision technology has emerged as a prominent tool in accessibility research for BLV people [68]. As such, computer vision-based systems are increasingly being used to augment 3DPMs with interactive audio elements.

Extensive work by Shi and colleagues [59, 60, 62–64] has established the design foundation for computer vision-based systems with 3DPMs. Their work began with the use of acoustic signals produced when the user tapped a marker on the 3DPM which was then recognised by a smartphone that would play the corresponding label information [58, 61]. The authors then pivoted to a computer vision-based system. Through the addition of a cube-shaped physical marker on the 3DPM, Magic Touch [60] could detect the users hand gesture and playback audio labels associated with specific locations on the 3DPM. A toolkit consisting of applications Markit and Talkit [64] was then developed to support the creation of and interaction with 3DPMs. Similar to MagicTouch, the system required the addition of a physical cube-shaped tracker on the 3DPM, as well as the application of a red sticker on the user's finger to support recognition. This toolkit initially relied on a RGB camera and a microphone to later utilising the built-in RGB camera of an iPad [62]. Further exploration of the system with educators of students who are BLV, emphasised that 3DPMs should have distinct tactile features between model components, provide auditory and visual feedback, and consider pedagogical methods [59]. However, the physical cube interfered with the exploration of the model and the need for a sticker on the user's finger meant the system was not well suited to real-world use in the classroom.

Biggs, Coughlan and colleagues [4, 16–18] have also explored computer vision-based systems with 3DPMs. The authors developed CamIO [18]. The system required the 3D model to be mounted on a board with barcodes, and a writing and reading stylus that allowed users to add and retrieve audio labels [18]. Later work, then expanded CamIO to 3D maps [4]. The system utilised a laptop and

webcam to capture the map for testing. This was later expanded to run on an iPhone [16] in which users could point to hotspots using a stylus and retrieve corresponding audio labels. Based on this work, the authors developed Point-and-Tap [47]. The system tracked the user's fingers, however it has yet to be expanded to 3D models.

Wang and colleagues [70] also explored computer vision technology for 3DPMs found in a science museum. TouchPilot [70] provided step-by-step guidance for the exploration of 3DPMs by users who are BLV, which was evaluated to be more beneficial than a pinpointing system. The setup of the system required the use of a checker board pattern to mark the world coordinate and did not support real-time re-calibration. Thus this required the 3DPMs to be placed on a stand and firmly mounted onto a base. The authors highlighted the need for future work to support interactions with movable models and employing the expertise of educators to investigate the use of such systems in classrooms. Thus we see that previous research into vision-based systems has focused on fixed models [4, 16–18, 70] or is impractical for use in the classroom. To the best of our knowledge, research into such systems has yet to focus on models with multiple components.

2.4 Our approach

The computer vision-based system described differs from previous research by:

- Supporting audio labelling of multi-part 3DPMs;
- Guiding the assembly of multi-part 3DPMs;
- Adopting a method that utilises colour and shape information similar to Gupta [26] and Narcisi [47], avoiding physical trackers that obscure the tactile features of 3DPMs [64].

Importantly, it differs in its insights into the design of such systems for use in education:

- Expanding on related work that has supported annotation of 3DPMs by sighted users [59, 62, 64], to allow BLV student-made annotations of audio labels which have been reported as a desired feature of learning materials for BLV students [59] and assistive software broadly [27];
- Utilising iPads and 'off-the-shelf' software to prioritise technological simplicity for implementation by educators.

3 Methodology

We employed a co-design approach, as recommended for assistive technology research [43]. As well as BLV participants it is important that educators partake in both the co-design phase and evaluation phase given the extensive evidence that suggests usability of assistive technology can be a significant barrier to its use in classrooms [23]. Our research had two distinct phases as illustrated in Figure 2: Co-design and Evaluation.

3.1 Co-Design

The co-design phase comprised three cycles of app design, informal evaluation and refinement with different stakeholders. Following each co-design session, the researchers implemented design ideas for a new iteration of the vision app. In Part 1, a basic vision app prototype was developed to act as a prompt for design exploration.

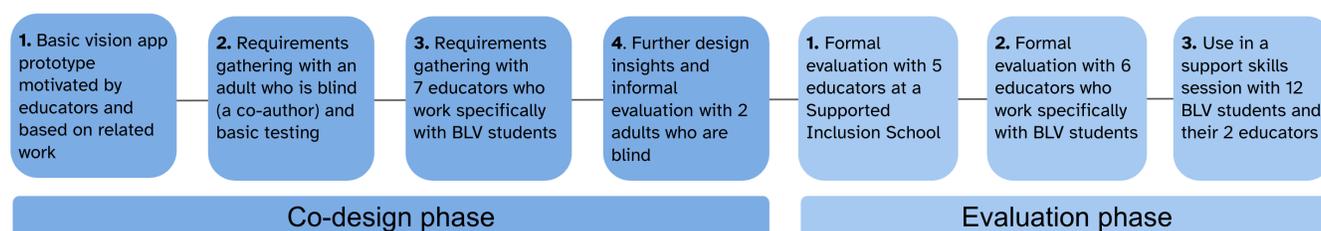


Figure 2: The study procedure is divided into two distinct phases: Co-design and Evaluation.

This is described in Section 4.1. Part 2 focused on requirements gathering with a BLV adult (a co-author) using the vision app technology prompt, as well as some basic testing of the prompt itself. Part 3 undertook requirements gathering and informal evaluation with 7 educators of BLV students. After further work on the technology, Part 4 focused on a more thorough evaluation with 2 BLV adults, as well as obtaining further design insights.

3.2 Evaluation

A final iteration of the vision app was developed as a result of the co-design phase. The app then undertook three rounds of evaluation with different stakeholders. Part 1 involved formal evaluation with 5 educators at a Supported Inclusion School. Part 2 undertook the same evaluation, however with 6 educators who work specifically with BLV children. The final stage of evaluation was through use in a support skills session with 12 BLV students and their two educators.

3.3 Data Collection

This research was approved by the home institution Human Research Ethics Committee. Due to guidelines of the research ethics approval, it was not possible to include students as direct participants in this study or to record their interactions with the system. Instead, similar to other work [59], we used the data provided by educators who participated in the student sessions and observed their use of the system. This procedure is detailed further in Section 5.2.

4 Co-Design Phase

4.1 Part 1: Initial App Design

4.1.1 Motivation. Seven of our 10 co-designers were educators from a specialised service of a state government department that actively produce and distribute 3DPMs to educators and students who are BLV to support access to the curriculum. Initial discussions with these co-designers motivated the concept of this study and the initial prototype vision app. The educators, some who are BLV and sighted, all have experience working with BLV students. Note, educators from this service were also involved in Co-design Phase Part 3 (Section 4.3), and Evaluation Phase Part 2 (Section 5.1) and Part 3 (Section 5.2). The small pool of educators within the specialised service necessitated their continued involvement across the study phases.

Initial feedback from the service was that iPads were widely available and distributed to students. As such, an important design criteria was that the developed vision app should be designed for

iPad technology, rather than use a bespoke combination of other hardware and software. These early discussions influenced the development of a simple technology ‘prompt’ that would be used to stimulate discussion and inform the design and development of a more fully realised system described in Section 4.5.

4.1.2 3D Model. We chose to focus on one 3DPM from a catalogue of the specialised service that is used to teach concepts in STEM. We selected “Beesly The Educational Insect” by Laurk¹. This model has components with distinct structural and functional relationships and requires assembly by the user, both characteristic features of multi-part 3DPMs used in STEM education. For example, it shares commonalities with molecular models in organic chemistry that support exploration of relationships between model sub-parts [48], and the disassembly and reassembly of anatomical models, such as a dissected frog model with removable organs [56]. The choice of model was supported by the educators. We 3D printed the model, printing each part in a different colour to support recognition as detailed in Section 4.1.4. The forewings and hindwings of the Beesly model differed in texture as a result of the printing process.

4.1.3 Design and Implementation. A simple iOS vision-based application was developed for BLV students to interact with the Beesly model. The app ran on an iPad and was equipped with vision and a voice user interface, designed for object recognition using the iPad’s camera, recognised several simple verbal commands, and provided simple text-to-speech feedback.

The app leveraged established software components: Swift for the programming language; UIKit for user interface development; AVFoundation for multimedia processing; the Speech framework for speech-to-text conversion; Create ML for training the machine learning model; and the Vision framework for hand recognition. The system architecture is outlined in Figure 3.

4.1.4 Object Detection. Object detection, or rather detection of the Beesly 3D model, relied on a computer vision model. There were three major steps involved in the development of the computer vision model: (1) image dataset collection and annotation; (2) computer vision model training and evaluation; and (3) integration in the application.

The computer vision model was based on the YOLOv2 architecture [53] and was trained in CreateML as an object detector and classifier. High-resolution photos of the 3DPM were taken on a white background. Each of the five model parts were annotated: head, thorax, abdomen, left wing and right wing. We began by

¹<https://www.thingiverse.com/thing:2446409>

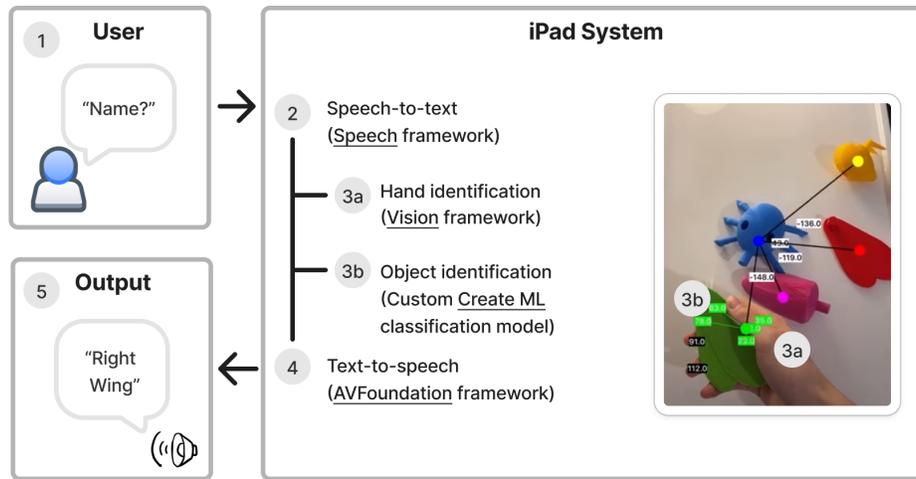


Figure 3: System architecture.

taking and annotating 747 images, of which 199 became validation images. The remaining 548 were augmented to produce 3 versions of each image. This resulted in 1644 training and 199 validation images. All 1644 training and 199 validation images are publicly available². The training data was explicitly captured in different lighting conditions and with hands covering portions of the 3D parts. The training data also had an exposure and orientation augmentation added for even more variation.

The colour of the 3D printed parts had significant impact on the performance of the computer vision model. Distinct colours increased the model’s success in differentiating and detecting different parts. The YOLOv2 architecture, and more broadly convolutional neural networks, learn to extract features such as edges, textures, and colours to be used for detection and classification. Directly correlating certain objects with specific colours (in this case, the insect 3D printed parts) allowed the computer vision model to learn associations between colours and objects to better classify and detect them.

When evaluated in a controlled environment, the computer vision model reached 87% accuracy. This data is presented in Appendix A.

4.1.5 Hand Detection. Hand detection was performed using Apple’s Vision framework. A `VNDetectHumanHandPoseRequest` is created that provides the normalised positions of the anatomical joints on all detected hands. We denormalised both the model’s bounding boxes and the hand joint positions to compare their distances and relationships to each other. We used their proximity to detect when the user is holding an insect part, and which one. We specifically use the joints `thumbTip`, `indexTip`, `middleTip`, `ringTip`, `indexMCP`, `middleMCP`, `ringMCP`. Hands-to-object distance threshold was obtained from extensive system testing. Using this threshold and additional logic we know the most probable object currently held. Object release is implemented as the change in that state.

4.1.6 2D Logic. All decisions were based on carefully crafted and tested 2D logic specific to the Beesly 3D model. The system was designed to detect when the user held a particular 3D model part. It relied on proximity detection between the user’s hands and the position of the detected 3D model. This required the model part of interest to be in the camera’s field of view. This appeared to be a challenge at times during the study (Section 4.4, 5.1 and 5.2). To address this, we introduced a base that indicated the camera’s field of view (Section 4.4.3). To manage this during the study, a researcher would notify the user if the model part of interest was out of the camera’s field of view.

While accurate, the system was not infallible. For instance, the app’s current version could recognise the wrong 3D part in cases where the user’s hand was close to multiple models. This was also the case when the user held a model part above another in such a manner that both models are visible to the camera. This emerged in Section 4.4. To address this, users were instructed to hold one model part at a time and the app was capable of multi-hand detection, notifying the user if they were holding more than one part of the 3DPM. To manage this during the study, a researcher encouraged users not to position their held 3D model over another.

Additionally, the system recognised model assembly completion and incompleteness, while partial completions were not detected. It could also recognise when the left wing and right wing were swapped and notify the user that the model was incorrectly assembled. This was achieved by: determining the position of all parts; calculating 4 angles formed by the positions of each triplet of neighbouring parts, i.e. {head, torso, left wing}, {head, torso, right wing}, {left wing, torso, abdomen}, and {right wing, torso, abdomen}; determining whether the 4 angles were within known ranges; determining whether the sum of the 4 angles was in the expected range; and checking that the detected bounding boxes intersect, i.e. model parts are indeed close to each other and in the expected configuration.

4.1.7 Set up and User Interactions. The set up of the system consisted of: an iPad Pro 11-inch (4th generation); iPad stand; Beesly

²https://drive.google.com/file/d/1zQzEH5_lNexEeOij5LSJx8F-PmKhCosa

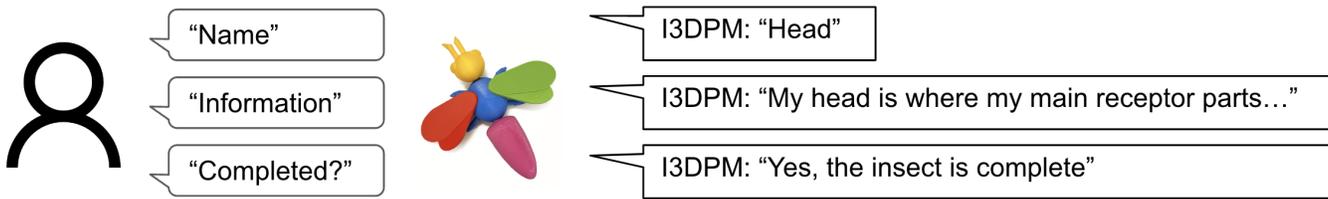


Figure 4: The initial app design allowed users who are BLV to gain basic details about the 3DPM, using the commands ‘name’, ‘information’ and ‘completed.’

3D model; and an A3 piece of paper in which the 3D model could be placed upon to support recognition. The iPad was mounted above the 3D model.

Based on initial conversations with the educators, the app was designed to detect when the user was holding a part of the 3DPM and would respond to a range of voice commands supporting engagement with the part of interest and the 3D model as a whole. This consisted of three basic commands: ‘Name’ (for the name of a model part); ‘Information’ (for more detailed information of a model part); and ‘Completed’ (to check if the model had been correctly assembled) (Figure. 4). Additionally, the app was capable of “let go” detection in which it identified when the user let go of a specific part and stopped any ongoing audio description of the part. The following illustrates a typical interaction:

- A BLV user tactually explores the 3DPM and picks up a part of interest with one hand and uses a speech command – ‘Name.’
- The system identifies the model part and provides an audio label – ‘Head.’
- Continuing, the user may ask for more – ‘Information.’
- The system responds with further details about the model part – ‘My head is where my main receptor parts are located.’ The user may place the part back down to stop the audio at any point.
- The user continues their interaction, assembles the five-part 3DPM and asks whether the parts are correctly assembled – ‘Completed.’
- The system identifies the position of the 3DPM parts and responds – ‘Yes, the insect is complete’ or ‘No, the insect is incomplete’ – depending on the accuracy of assembly.

4.2 Part 2: Co-design Session with Co-author

4.2.1 Co-designer and Procedure. We conducted an in-person session with a co-designer and co-author – a disability advocate and activist who is blind. The one-hour session involved: (1) an exploration of the Beesly 3D model alone to better understand touch reading of the 3D model; (2) a discussion of their understanding of the 3D model, challenges and ideation of how an app can support interaction with the 3D model; (3) an exploration of the app with the 3D model; (4) an evaluation of the app involving discussion of their initial thoughts about the app, its role in supporting understanding with the 3D model and the perceived usefulness; and lastly (5) co-design activities which explored the ideal experience in interacting with the 3D model through the app. Here, we ideated

design elements that would better support understanding of the relationship between various model parts and the assembly process.

4.2.2 Results. The co-designer provided insights into the usefulness of such a system, how the system could be used to tackle the inherent complexities of multi-part 3DPMs, as well as provide key design and functionality insights.

Usefulness and Broader Application: Initial impressions were very positive. Indeed, the co-designer expressed that they would have appreciated educational tools like this when they were in school:

“I love it, it’s really cool, I would have loved this, I still love it, [I’m] just a bit older.”

Importantly, they could see the application of the app beyond the Beesly 3D model:

“Could you extend this to different things as well.. I’m just thinking musical instruments or models of body parts, viruses.. skeletal system.”

An interesting use case was for the app to support learning of maths, as they believed that “maths is such a visual thing at school”.

Assembly Complexity: Discussions began with the assembly process of the 3D model, noting that parts can be mistakenly misplaced, suggesting a need for the design to either prevent or utilise such errors to support learning. As explained, it was challenging to identify different model parts:

“Conceptually it’s hard to work out what things are.”

Presenting the 3DPM in its assembled form provided context for model assembly and assisted in the identification of model parts:

“Obviously they are wings, but that’s because I saw it when it’s together, it might be a challenge to figure which goes where.”

Additionally, the reliance on memory alone was challenging:

“I’m struggling a bit to put it back like to remember where it goes.”

This co-design session confirmed the importance of a dedicated feature to provide feedback on assembly. Similar to TouchPilot [70], a guidance system and assembly clues were suggested:

“Getting it to be able to say where your hands are on the model and get it to be able to recognise that. And there need to be clues about how you put it together.”

Supporting specific interaction with the 3D model: The app was praised for providing an informative experience with the 3D

model. Further, the co-designer suggested using the app to test student learning by presenting questions to users:

“You could get it to ask questions to students.”

The importance of making the learning experience fun and engaging was highlighted, with suggestions for sound effects and personalised feedback to provoke curiosity and encourage exploration. When asked about how to best support understanding and engagement with the 3D model, the co-designer suggested the use of descriptive sounds:

“Like digesting when you’re holding the abdomen..
and like crunching when you got the head.”

Sounds were also discussed in the context of feedback mechanisms to support active learning and problem-solving. Ideas involved indicating when a piece is correctly assembled or offering hints about parts’ locations:

“That’d be cool to have sound effects to say you’re
way off, getting closer or like a ding ”

4.2.3 Iteration in Design. The co-design session discussed provided invaluable insights into how the app could be developed to better meet the needs of its users and enhance interaction with the Beesly 3D model:

- *Descriptive sound effects:* we introduced realistic sound effects corresponding to different parts of the model, such as sounds of digestion for the abdomen and biting sounds for the head, to make the learning process more immersive and fun.
- *Supporting cognitive load and guiding assembly:* we introduced a new command where the user may ask “What does this connect to?” to diminish the reliance on memory and guide assembly upon user enquiry.
- *Positive feedback upon assembly:* we introduced a positive short tune upon correct assembly by the user.
- *Assessment of student learning:* we introduced command “Quiz me” an interactive quiz mode that tests the user’s knowledge and understanding of the model and its assembly.
- *Functionality improvements:* the session identified bugs that required fixing and the need for improved reliability in the app’s voice and gesture recognition to ensure smoother user interactions.

4.3 Part 3: Co-design Session with Educators who work specifically with Students who are BLV

4.3.1 Co-designers and Procedure. We conducted a one-hour virtual session with 7 educators from the specialised service described in Section 4.1. Staff from this service were also involved in Evaluation Phase Part 2 (Section 5.1). Two of these educators were BLV and five were sighted.

The session began with a demonstration of the app by a researcher. At this stage, development of design elements from Part 1 were still underway and as such the app only provided the three basic commands described in Section 4.1. Following the demonstration, the educators were asked about their initial thoughts and perceived usefulness of the app. In order to gain an understanding of current teaching practices, we asked the educators how they

would normally use this type of 3D model with students. Further, we inquired whether there were additional app features or functions they thought could better enhance student interactions with 3DPMs. Following this session, a researcher continued discussions about the app design with one of the educators at the service. The main points from the session and ongoing correspondence are reported collectively.

4.3.2 Results. As with our first co-designer, all educators (n=7) perceived the app to be useful in teaching with 3DPMs and could see many applications for the app to support learning a range of concepts, particularly those in science such as the anatomy of the heart, metamorphosis cycle (biological life cycles), geography, maps, puzzles and molecular models:

“This would be great to use with Molymod³ (commercially available kits for chemical modelling). Having it trained to identify the individual atoms (each colour represents a particular atom) as well as tell you when you’ve made a particular molecule would be very cool.”

Additionally, the educators reported that the app would foster greater independence and allow students to learn at their own pace. There was an interest in using the app to support learning across different year levels with the same 3DPM but with different levels of information.

The potential of the app to support collaborative learning in the classroom was also recognised:

“Beesly could work as a station in ‘rotations’ which is a common thing in a primary classroom, and be used by all students.”

Further, the educators raised that teaching objectives with such 3DPMs was very specific to each student, as well as the teaching style of the educator. The educators recommended an annotation feature to support note-taking and revision by students to support independent learning experience.

4.3.3 Iteration in Design. The co-design session with 7 educators provided a valuable ‘sense check’ of app features including the label retrieval method (in which the user holds a part of interest and asks for more information) and the audio “stop” function (in which users placed parts down). The session yielded the following additional key design elements:

- *Clear instructions:* alongside a verbal description, we introduced a list of instructions on how to use the app in both large print and braille. This included a brief description and a list of the commands. This allowed users to refer back to these details at their convenience and support autonomous use of the app.
- *Annotation of audio labels:* we introduced an annotation feature in which BLV users can create and apply their own audio labels. We introduced the command “Add label” to record an annotation while the user holds a 3D model part of interest. The user could then ask “Label” to playback their recording.

³<https://molymod.com/>

4.4 Part 4: Co-design Session with Adults who are Blind

4.4.1 Co-designers and Procedure. We conducted an in-person session with two totally blind adults aged 67 (P1) and 30 (P2). The procedure detailed in Section 4.2.1 was repeated to elicit further design insights and informally evaluate the further developed prototype with BLV users. The interviews were conducted in-person, each lasting approximately 60 minutes.

4.4.2 Results. Differing Attitudes Towards Technology:

While P1 viewed technology as an opportunity to enhance and personalise learning, P2 championed traditional, less technology-reliant methods for their simplicity and ability to teach fundamental skills effectively. An examination of these contrasting perspectives, sheds light on the critical balance between innovation and inclusion in the classroom, and the ongoing value of tactile and traditional learning tools. For narrative clarity, the perspectives of P1 and P2 are (largely) presented separately.

Embracing Technology as a Learning Tool [P1]: The usefulness of the app was clear as explained by P1, particularly in the case of independent and collaborative learning for students:

“I love it. I think it’s a really fantastic idea. It makes the child capable of independent learning that’s for sure. I think it potentially is good for being used in groups. So, you know, the vision impaired child can be working in a group and they can all be doing the same thing.” [P1]

P1 responded positively to the descriptive sound effects that were introduced as a result of the previous design sessions, and indicated that they were a notable feature of the app:

“[laughs] I love the sound.” [P1]

Indeed, the descriptive sound effects for each model part were useful in conveying its identity and function:

“So it’s suggested by the noises it’s making that it’s a fly.” [P1]

P1 was interested in a longer recording time for the annotation feature to support additional learning and highlighted an opportunity to use the feature to support assessments. Continuing their discussion of the educational benefit of the system, P1 emphasised learning through mistakes, specifically, in cases of incorrect assembly of model parts:

“Is it possible for the person to put the head and tail on the wrong way? Because that way then you could get more feedback from it about which is the correct way around.” [P1]

A researcher observed that P1 struggled at times to determine the camera’s field of view. Interestingly, after becoming familiar with the system, P1 chose to place all model parts aside, then proceeded to bring a part of interest into the camera’s field of view, one-by-one. This was an effective workaround to a flaw in the current version of the app whereby it would potentially recognise the wrong part. However, this would need to be addressed in the next version. Overall, the excitement and curiosity expressed by P1 suggested a potential of the app to engage and motivate students, making learning with 3D models like Beesly more enjoyable:

“Oh, it’s flying. It is flying.” [P1]

Prioritising Traditional Learning Methods over Technology [P2]: Despite acknowledging the potential appeal of the 3D model (“He is kinda cute”), P2 disagreed with the use of technology for student learning and thus did not see the usefulness of the app. P2 expressed a strong preference for traditional learning methods, such as braille and tactile graphics, over the use of technology like iPads. They highlighted concerns about the potential for technology to make BLV students stand out in a mainstream classroom, potentially leading to feelings of exclusion:

“I don’t want to be the only blind student in a mainstream class with all this bulky technology.” [P2]

On a similar note, P2 emphasised the importance of discreteness and the ability to quickly set up and use educational tools in a classroom setting. They argued that using technology could be cumbersome and time-consuming, potentially causing delays in learning and again drawing unwanted attention to the student:

“Whereas if you had access to a braille book, for example, like, when I did maths, all of my stuff was in braille and like, in a book and on paper and that made me feel like that’s more normal.” [P2]

Importance of Inclusiveness [P1, P2]: In fact both P1 and P2 agreed on the importance of inclusive learning materials in the classroom. The point of difference was whether the app would be inclusive or not. P1 was far more positive:

“It to me seems that it’s not something that’s specifically for the vision impaired person. So it makes then the whole interaction with peers more enticing. I’ve learned from talking with educators that if you’ve got interactive things for your vision impaired student, all the others want to know how it works.” [P1]

The contrasting views of P1 and P2 highlight the importance of considering the unique needs and preferences of users in the design and implementation of technological educational tools. It suggests that while technology has the potential to improve learning, it must be carefully integrated in a way that does not exacerbate feelings of exclusion.

4.4.3 Iteration in Design. This co-design session yielded the introduction of several additional design elements and considerations for the app’s development:

- **Physical indication of camera field of view:** we introduced a base that indicated the camera’s field of view following researcher observation. The base was white to support model recognition.
- **Practicality of technology set up:** we ensured the system involved as minimal set up requirements as possible. This included an iPad, iPad stand and base.
- **Inclusiveness:** it is essential that any assistive technology introduced in the classroom is discrete and does not exacerbate feelings of isolation for students who are BLV. For this reason, we did not introduce any additional hardware such as microphones, speakers or camera’s.
- **Robustness:** performance optimisation and better recognition of additional variations of commands.

4.5 Final App Design

As a result of the co-design phase a range of functions and features were introduced to the system described in Section 4.1.3. These related to identification of model parts, assembly guidance and support for learning (Figure 5).

Speech commands such as “Name” and “Information” allowed users to identify and explore the insect model’s anatomy. Assembly tasks were supported through straightforward voice prompts such as “What does this connect to?” and “Completed?”, making the educational process both engaging and accessible. The app further supported student learning with commands including “Add label” to support student annotation of audio labels and “Label” for playback. Finally, the addition of a “Quiz me” command provided an opportunity for students to test their knowledge by working through questions related to the insect. This question-and-response quiz mode supported both auditory and visual answers. For instance one question being “Can you identify my left wing?” detected when the user identified and held the correct part.

The final set up of the system consisted of a base in which the 3DPM could be placed atop while simultaneously indicating the camera’s field of view (Figure 6).

We welcome researchers, developers and interested parties to build on this work and thus the app source code has been shared on GitHub⁴, and the beta version of the app is available on TestFlight⁵.

5 Evaluation Phase

The evaluation phase consisted of a three-part evaluation of the final app design with: (1) educators at a supported inclusion school; (2) educators from the specialised service for students who are BLV described in Section 4.1; and (3) students who are BLV. It was important to evaluate the system with both the students and the educators that would advocate for its use in the classroom. The same procedure was conducted with all educators in Parts 1 and 2 of the Evaluation Phase, thus we collectively present these findings.

5.1 Part 1 & 2: Evaluation with Educators

5.1.1 Participants. Participant Group 1 consisted of 5 educators at a Supported Inclusion School, a mainstream school designed to cater for a higher than usual proportion of students with disability. The newly developed secondary school had one student in year 7 with low vision at the time. All 5 educators worked with this student. All educators in Participant Group 1 were sighted.

Participant Group 2 consisted of 6 educators from the specialised service described in Section 4.1. Two of the educators were blind and 4 were sighted.

5.1.2 Procedure. Three researchers conducted the evaluation sessions with the 11 educators. All interviews with educators were recorded for later review and analysis. Each session ran for approximately 70 minutes, consisting of three parts.

First, an introduction and demonstration of how the app works was presented.

Second, the educators tried the app for themselves in a free-exploration, with suggested interactions based on the demonstration. This was followed by a questionnaire that asked the educators to rate statements on a 5-point Likert scale. The statements enquired about interest, usability, efficacy as a tool for inclusive teaching, and the perceived usability and engagement of the app for BLV students, as well as all students in their class. Only 8 of the 11 educators completed the questionnaire. All 11 educators shared their thoughts and initial impressions with the group.

The third part was a discussion on how to extend the system to support other 3D models. In order to stimulate discussion, we selected an additional three 3DPMs from the service’s catalogue, each of which consisted of multiple components. The session concluded with an open discussion about the usefulness of the system, its efficacy in the educational context, its use cases and finally practical challenges and considerations.

The session was recorded and transcribed. An inductive thematic analysis approach was adopted in which two researchers independently coded the transcripts. Next, the researchers reviewed, cross-checked and revised these codes. These codes were collectively reviewed and collated to generate a set of themes. Note that [W] denotes responses from Participant Group 1 and [S] denotes responses from Participant Group 2.

Usefulness & Broader Application. All 8 educators who completed the questionnaire strongly agreed (n=6) or agreed (n=2) that they’d like to use the app to teach with the Beesly model. All educators (n=8) strongly agreed that they were interested in using 3DPMs to support student learning.

The educators recognised the app’s potential to engage students deeply and facilitate learning across various subjects with 3DPMs. Educators noted its engaging nature and the positive reinforcement provided by sound effects. They were positive about the app’s versatility, engagement factor and the potential for broad application across educational settings. In particular, educators acknowledged the educational value of multi-part 3D models as tools to break down complex concepts in a way that would appeal to students:

“If you’re talking about insects in lower primary, like they would love to be able to take it apart.” [W]

Further, manipulation involving assembly and disassembly of the 3DPMs could support understanding of relationships between different model parts:

“She [BLV student] could disassemble it and assemble it, disassemble and assemble it and she’d get more confident where the parts went.” [W]

The educators highlighted the versatility of the app in teaching a range of concepts. There were many ideas for use cases including - pre-mapping spatial tests, music, sport, art, science, humanities, geography, history, food tech and maths:

“I’m thinking there’s applications right across the board.” [W]

Usability & Engagement. Of the 8 educators who completed the questionnaire, the majority strongly disagreed (n=3) or disagreed (n=4) in response to the statement ‘I found the app unnecessarily complex,’ while one response was neutral (n=1). All educators strongly agreed (n=5) or agreed (n=3) that the app was easy to use.

⁴<https://github.com/monash-assistive-tech/model-interact-ios>

⁵<https://testflight.apple.com/join/LD8nTYlk>

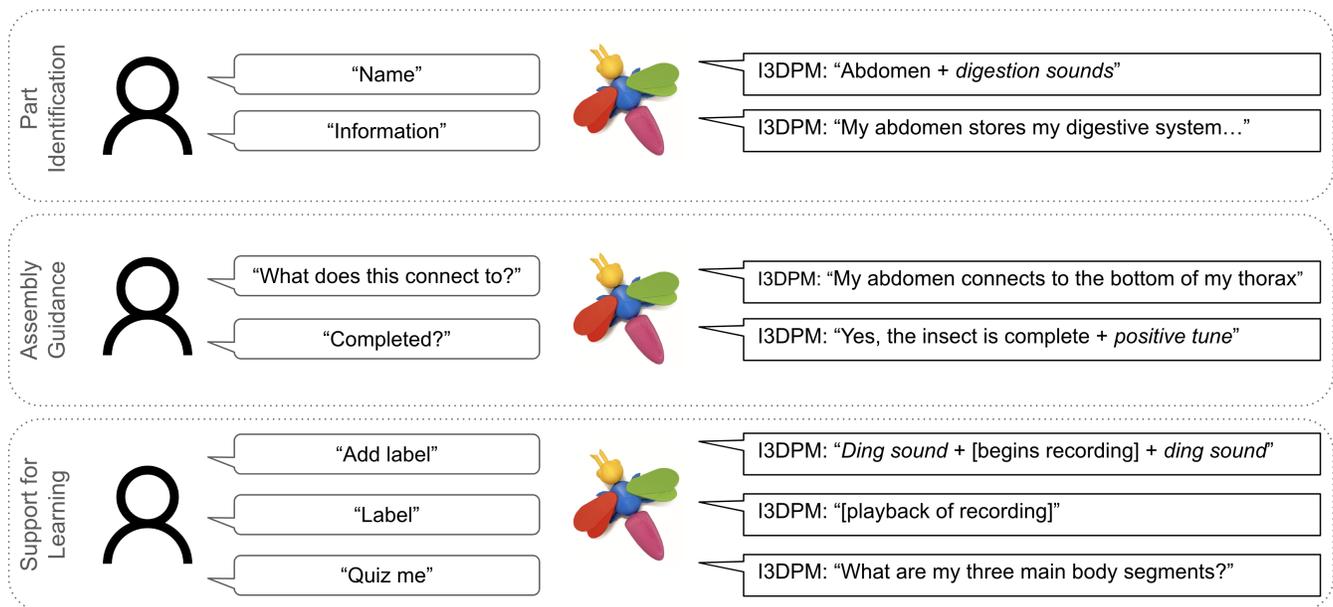


Figure 5: The final app design allowed students who are BLV to engage meaningfully with the Beesly 3DPM.

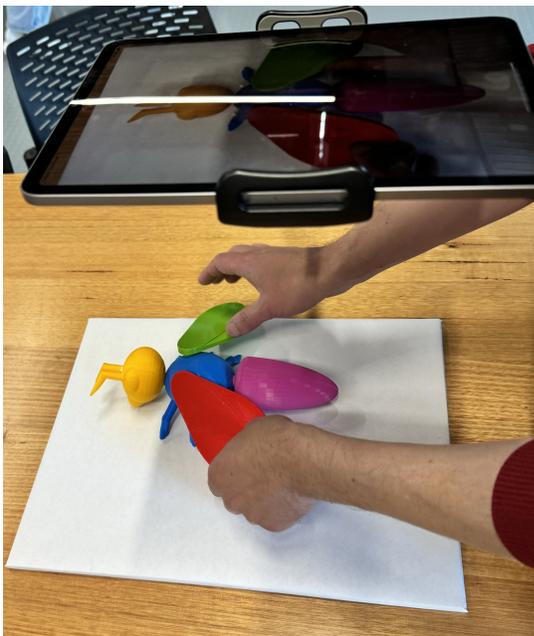


Figure 6: The set up of the final system consisting of an (1) iPad, (2) iPad stand, (3) Beesly 3DPM and (4) Base that physically indicated the camera's field of view. A user interacts with the system, exploring the 3DPM while engaging with the iPad application.

Interestingly, results from the supported inclusion school were varied regarding their confidence in using the app and their perception of required support (Figure 7). In contrast, educators from the

specialised service were more assured (Figure 8). This may be explained by their experience and expertise in assistive technologies.

BLV educators encountered some challenges when using the system, particularly in identifying the camera's field of view. In order to address this, one of the educators suggested that a physical stand or tactile indicators be introduced to the base:

"Would it be useful to have like a stand that you can line up with the camera to put the piece that you want to focus on and get your hands out of the way." [S]

In response to the perceived usability by students who are BLV, all 8 educators strongly agreed (n=4) or agreed (n=4) that they'd imagine that most students who are BLV would learn to use the app very quickly. Indeed, educators explained that they'd imagine students who are BLV would find it easy to use:

"All the kids will just pick it up." [S]

All 8 educators also strongly agreed (n=5) or agreed (n=3) they would imagine that most students who are BLV would find this app engaging. The educators could see how the app would appeal to their students who are BLV:

"Whereas sometimes when you employ technology, it makes it more interesting for the kids. They'd love these on the iPad or they'd like the sound effects." [S]

It was clear that the commands were usable and effective, however there was a desire for a more conversational enquiry. When asked about a conversational agent like Siri or Alexa as an additional app feature, the educators agreed that it would further support understanding with the 3D models. As explained by an educator, this would support further inquiry by students:

"Yes. You can ask for more information... if the kid really wants to know more, they can say, tell me more." [W]

Fostering Student Autonomy & Confidence. Educators highlighted how the app could foster student confidence and autonomy through self-discovery and learning by doing. The app was also perceived to foster greater autonomy for students:

“That way she’s [BLV student] exploring it for herself as well and it’s creating more independence on her part.” [W]

As well as confidence:

“It might help her [BLV student] with her speaking and getting into different groups and things.” [W]

Guiding Assembly. An educator explained how the app could provide valuable feedback for the students progress in assembly tasks, to further support the arrangement of model parts, and give further feedback in cases of incorrect assembly:

“Is it completed? Is it correct?.. she [BLV student] might put things that are in the wrong order... if you could do that on the app, that would be awesome.” [W]

Suggestions included the use of music or sounds to indicate patterns of arrangement:

“Putting like the right combinations, like the music when we’re clapping patterns and stuff, she’s [BLV student] got to put the correct one. Or it could like play a pattern for her.” [W]

As described by an educator who is BLV, it would be valuable to include descriptive details of points of connection between different model parts:

“Just describe, place the flat end of the whatever piece into the slot on the right hand side of the whatever. Just be quite descriptive about things.” [S]

Annotation, Revision & Alternative Assessment Methods. The usefulness of the annotation feature was recognised by all educators. Educators were interested in expanding the annotation feature to support note-taking for student revision:

“So if she was going to do a test that she could do revision and listen to what she had learnt before the test and then come in and do it potentially.” [W]

Further, educators explained that they often faced challenges designing assessments for students who are BLV. Educators highlighted an opportunity for the app to address this and deepen educator understanding of student learning:

“It’s another opportunity to understand the thinking processes of students and how their approach to things are, which means you can also start to tailor more things...” [W]

Efficacy as an Inclusive Learning Tool. In response to the perceived usability by ALL students in the class, just as when asked about BLV student response, all 8 educators strongly agreed (n=5) or agreed (n=3) that they’d imagine that all students in the class would learn to use this app very quickly. Again, all educators strongly agreed (n=5) or agreed (n=3) they would find this app engaging.

All 8 educators strongly agreed (n=5) or agreed (n=3) this app would be an effective tool for inclusive teaching in the classroom:

“I think we can all see the use of it whole classroom-wide as well, which is positive I think. That makes it very inclusive too.” [W]

As highlighted by P2 during the co-design phase, it is important that educational tools introduced in the classroom do not perpetuate the feeling of isolation or exclusion amongst their peers:

“... what I love is that it’s just so inclusive of all students. I think that’s one thing that she [BLV student] struggles with is the fact that she has different materials and has different equipment than some of the other kids.” [W]

The educators also acknowledged the apps broader inclusion for students who have difficulties with language and words:

“So I’m thinking about students with dyslexia or who struggle to write.” [W]

The value for sighted students in their class was clear to educators:

“I’m just thinking this could also be good for everyone, not just because some students are visual learners. So by having it very visual in front of you and explicitly telling you where things go and telling you it’s not correct will help a lot more.” [W]

Practicality of Implementation: Planning Time, Technical Support & Scalability. Planning time was a significant practical consideration for educators. To address this, educators were interested in the alignment of the app with the curriculum, with relevant activities and assessments that could be tailored by the educator:

“Even if it’s like functions, then you can add in things that are content specific. This is a quiz, you can now make your own quiz.” [W]

Additionally, educators were strongly in favour of the app being delivered as an educational package, with 3D models and pre-set learning activities to simplify planning time:

“If it came with like a box of just all these pre-fills and things for starting subjects and stuff that would be useful. Because we don’t have to think about what we need to get, if it just comes with everything here.” [W]

Technical support was an additional practical consideration, with some mentioning that they’d need time for familiarisation with the app:

“I just need technical help. I think you’d need time to play.” [W]

The scalability and sustainability of the system was an initial concern:

“I would want to know how we could make it sustainable in terms of us inputting our own content because we potentially could be developing and working with partnerships to create our own 3D printed things. How then would we label and all that stuff to make it something that we can keep going at the school?” [W]

A suggestion was to incorporate the app as a library resource with its own catalogue of 3D models:

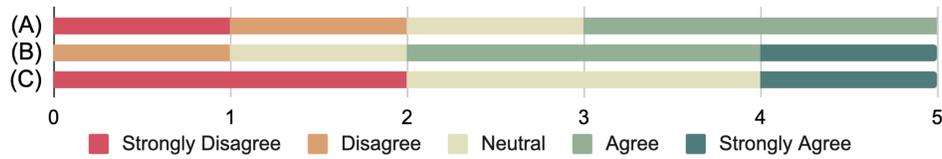


Figure 7: Educator responses from Participant Group 1 (supported inclusion school) to the following statements: (A) I think that I would need the support of a technical person to be able to use this app; (B) I felt very confident using the app; (C) I needed to learn a lot of things before I could get going with this app.

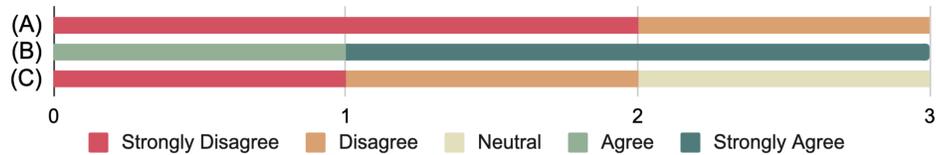


Figure 8: Educator responses from Participant Group 2 (specialised service) to the following statements: (A) I think that I would need the support of a technical person to be able to use this app; (B) I felt very confident using the app; (C) I needed to learn a lot of things before I could get going with this app.

“I am thinking what if we had a catalogue, our own little catalogue of things that we had created that lived as part of our library resources that anyone could borrow... loan it out to each class.” [W]

5.2 Part 3: Evaluation with Students who are Blind or have Low Vision

5.2.1 Participants. We conducted an evaluation session with 12 students who are BLV who attended a support skills session at the specialised service. The students in attendance spanned Grades 4 to 6 and their level of vision varied from: no usable vision; no or very limited usable vision; and some usable vision. As the research ethics approval did not allow students as direct participants (as discussed in Section 3.3), no further student demographic information could be collected. Similar to previous work [59], we analysed and report data provided by the educators who participated in the sessions and observed the student behaviour. Two educators from the service co-facilitated the session alongside two researchers.

5.2.2 Procedure. The 50 minute session involved three 15 minute rotations. Each group consisted of four students. The students were given a quick introduction to the 3D model and a demonstration of the app. Each student was then invited to try the app for themselves. Students were presented with a list of the speech commands in braille and large print.

One educator facilitated the session by guiding the students and providing necessary support during their interactions; while the other educator primarily observed and took notes on the students’ experience with the app. The note-taking educator was given a list of questions regarding their initial impression on the students’ experience with the app, the usability, engagement and feedback for improvement. They were encouraged to note down any other points of interest regarding the students’ interaction with the app. While useful tools like an adapted System Usability Scale (SUS) for children exist [51], it was impractical for us to use this given

time constraints and class arrangement. Directly after the session, two researchers interviewed the educator who had taken notes during the student session. The semi-structured interview ran for 30 minutes.

Usefulness & Broader Application. Upon observation and reflection of the student evaluation session, the educator reported that the app was indeed useful.

“Being able to point to each one and get a bit more information. Like diagrams where you have the arrows pointing in. I guess that’s kind of what they’re replacing to a certain extent. They sort of fill in the picture a bit more.” [S]

The app was useful in providing a more informative experience for students:

“It’s kind of adding that extra level of information.” [S]

The broader application of the app with variations of 3DPMs was noted by students:

“One of them was talking about having multiple, you know, if you had other animals in a set, that would be quite good. You know, as in different examples of the same things, I think would be a good way to sort of extend just this one.” [S]

Usability & Engagement. All students were able to use the app, all were eager to interact with the system and the educator felt the app extended engagement with the 3DPM. The age group of grade 4-6 was well suited to the self-discovery approach adopted during the student sessions:

“They have enough capability to kind of put things together and interact with it pretty naturally.” [S]

While students were able to easily interact with the model and overall system, as highlighted in Parts 1 & 2, some students struggled to identify points of connection between the different model parts:

“It was amazing how many of them tried to put these into the holes, which makes sense, these being the antennae. And finding these holes seem to be tricky for a few of them.” [S]

As in the Co-Design Phase, some students faced some difficulty ensuring the 3DPMs were in the camera’s field of view:

“Where to hold it was tricky, especially, even for the ones who had some vision it was still, they picked it up quicker but it was still tricky to kind of position.” [S]

After student observations, the educator suggested that the app should provide immediate and clear feedback on misplacement of 3DPMs in relation to the camera’s field of view. It was also suggested to have a stand to place model parts of interest.

Despite the initial learning curve of how the app works, once the students became more familiar, they were able to deepen their engagement relatively quickly given the short sessions:

“Once it started going, seemingly once they got it once, it seemed to kind of work really well.” [S]

The descriptive sound effects and positive tune upon correct assembly were received very positively by the students:

“I just think the sound effects worked, it just kind of connected to the real, you know, the real thing.” [S]

In addition to these sound effects, creating and sharing annotations as a group was particularly enjoyed by students:

“I’ve got a few laughs [noted].” [S]

Despite this, there was room for improvement when it came to the voice of the system. Both the students and educator were in favour of a more relatable, less-robotic voice. The educator also suggested incorporating further elements of playfulness by incorporating entertaining questions and responses:

“If there were the occasional entertaining question you could ask where it gives a slightly funny answer.” [S]

Fostering Student Autonomy & Confidence. Importantly, following the student session, the educator was confident that the app could support independent learning. The app provided an opportunity for students to learn at their own pace. Indeed, student self-discovery was an important aspect of interaction:

“There’s something nice about them being like, all right, this is a model of an insect. You’re going to do some exploring and find out what the parts are. This is how you explore. Off you go.” [S]

Guiding Assembly. An indication of the student’s progress of model assembly was reiterated, as well as definitions and further details upon enquiry such as through hints:

“There are so many pieces, like a definition would be helpful for all of those when you’re interacting with it. And then if you wanted some like a hint, like a hint mode almost.” [S]

Annotation, Revision & Alternative Assessment Methods. The annotation feature was a huge hit amongst the students. Not only did the students enjoy recording and replaying their own annotations, but they also enjoyed hearing that of their peers:

“They liked recording themselves, that was my other observation. The novelty was high. They seem to enjoy hearing, you know, when someone said something and they all enjoyed that.” [S]

Additionally, the opportunity to expand the annotation feature into an alternative assessment method was emphasised:

“You could almost use it as an alternative assessment tool where they add their own labels, like they could do an audio kind of thing rather than having to draw a diagram, for example.” [S]

Efficacy as an inclusive learning tool. The app’s value as an inclusive learning tool in the classroom was again reiterated:

“Say it’s a, you know, blind kid with three other peers who are all sighted. They can work on it together and it’s not like the blind kid has to have the person who’s there the whole time to give them the information.” [S]

The educator envisioned using the app for group activities in which different roles are allocated to students such as 3D model explorer, timekeeper and scribe:

“But if you’re doing like a classroom activity, you probably need to like have defined roles for a group to interact with it.” [S]

6 Discussion

6.1 Educational Value

The co-design and evaluation phases provided strong evidence of the system’s educational value and potential to address current educational challenges facing BLV students and their educators, especially in STEM. In particular:

- **Information Access:** Educators reported that the app not only complemented the 3DPM with an “extra level of information” but also helped “fill in the picture a bit more.” Furthermore audio-labelled multi-part 3DPMs allow BLV students to breakdown complex concepts. This is consistent with previous work that highlights the value of multi-part models to break down concepts and explore relationships [42, 59].
- **Independent Exploration:** Assembly of multi-part 3DPMs can be challenging for BLV users [13, 15]. Educators were excited by the use of audio feedback to reduce this challenge and so support independent exploration. Many educators noted that system offered an opportunity for BLV students to learn with 3DPMs independently. The educators reported that the system could foster student confidence and autonomy through self-discovery and learning by doing.
- **Engagement:** It was clear that students found the system engaging, in particular, sound effects and voice annotation increased the sense of fun. Educators were also interested in incorporating further elements of playfulness by incorporating entertaining questions and responses.
- **Inclusion:** Educators felt that the inclusive and engaging nature of the system would encourage collaboration between sighted and BLV students in the classroom and reduce any

stigma associated with the use of assistive technology by BLV students.

6.2 Design Considerations for Future Computer Vision-Supported Educational Tools

Our research has revealed a number of design considerations to guide the design of future computer vision-supported educational tools for learning with multi-part 3DPMs. These relate to supporting the use of multi-part 3DPMs, meeting the needs and preference of educators for use in the classroom, and practicality for implementation and ongoing use.

6.2.1 Supporting Multi-Part 3DPMs. There are a number of core design considerations that relate to the basic functionality of a vision-system supporting the use of multi-part 3DPMs.

- **Identification of Individual Components and Obtaining Detailed Information:** Users must be able to easily and intuitively access information about the individual parts of a multi-part 3DPM. This includes identifying information, such as the name of the part, along with descriptions that support the educational objectives and understanding of the relationship between parts. This should be a combination of speech inquiry and touch gestures, and allow flexible information access, such as repeating information upon request and to continue from where it was paused.
- **Offering Assembly Guidance:** Given the complexity of working with multi-part 3DPMs, systems must be designed to offer guidance for assembly. Consistent with previous work [15], it was challenging for BLV users to determine the alignment and connections between model parts, as well as the order of assembly. To address this, educators desired intuitive audio cues to indicate progress of assembly including music or non-speech sounds to indicate patterns of arrangement. Moreover, guidance should consist of descriptions and detail regarding points of connection between the different parts of the 3DPM. Additionally, systems should provide validation of correct placement of parts, as well as clear guidance when there are issues regarding the correctness of assembly. Greater explanation was desired to support times when the model was incorrectly assembled. Systems may utilise assembly errors as learning opportunities, whereby feedback may explain the reason for incorrect assembly allowing users to make and learn from mistakes [21].
- **Field of View and Physical Workspaces:** Given the vision-based system relies on the model parts being within the camera's field of view, systems should provide immediate and clear feedback on their misplacement. The lack of such feedback can cause confusion and frustration during interaction. Determining the camera's field of view is a known challenge of vision-based systems, with musical tones, volume and speech directions having been suggested to address this [33]. Further, in accompaniment to such features, our study highlights an interest among BLV educators for a defined physical workspace, with a physical stand or tactile indicators, to overcome this challenge. Such features may also assist organisation of model parts and assembly tasks.

- **Environment and Object Recognition Performance:** There should be careful consideration as to how to best design such systems for real-world environments such as classrooms. Lighting conditions and the presence of other objects are known challenges of vision-based systems [39, 40]. Additionally, classrooms are often noisy which can interfere with voice recognition.

6.2.2 Application in the Education Context. There are additional design considerations when considering systems to support the use of multi-part 3DPMs specifically in the educational context. This includes educator needs, teaching objectives and how best to support student understanding and engagement.

- **Annotation and Note Taking:** Along with providing audio labels and descriptions of model parts, the ability for annotation by BLV students is crucial. Educators desire the ability to annotate to support student note taking, revision and assessment. Annotations can provide insights into a student's comprehension and areas of difficulty, as well as allow for more personalised learning experiences. This aligns with previous work [27], where annotation was a highly desired feature. Interestingly, educators were excited about the possibility of using voice annotations as an assessment tool.
- **Engaging Audio:** Having audio, both voice and sound effects, that is relatable and engaging is important in an educational setting, especially with children. Students reported an interest in a more relatable voice and language, aligning with previous findings [1, 55] where customisable voice types and language levels were offered based on user preferences. Each part of the model should also have associated sounds that depict its function or nature. Similarly, distinct non-speech sounds, tones and musical instrument notes have been shown to be valuable in distinguishing between different elements of accessible diagrams [72]. Such features not only further reinforces the nature of the model part, but can provoke feelings of excitement from students.
- **Implementation in the Classroom:** Educators expressed interest in implementing such systems as a stationed activity in the classroom, in which groups of students could rotate and participate in group learning activities. Educators envisioned group activities in which different roles were allocated to students such as 3D model explorer, timekeeper and scribe. Future designs should support multi-hand interactions so as to support simultaneous use by multiple students.
- **Educational Package:** Educators were in favour of the system being delivered as an educational package. Given support for educators is a barrier to adoption [23], systems should align with the curriculum so as to become a natural part of daily teaching activities, incorporating pre-set content, learning activities and customisation.

6.2.3 Considerations for Practicality, Adoption and Ongoing Use. The extensive involvement of educators offered a number of insights into the practicalities of classroom adoption and ongoing use:

- **Compatibility with Devices Used in Practice:** Computer vision-supported educational tools should be compatible with devices commonly used in practice. Alongside previous work, our results reiterate the prevalent use and preference for iPads by BLV people [19] and educators of students who are BLV [62]. Compatibility with such devices lowers the barrier to adoption by educators.
- **Simplicity of Set Up and Technology Support:** Systems should prioritise technology simplicity, require minimal technical know-how and have easy setup. Given that these are reported barriers to assistive technology adoption [23], their consideration should be at the forefront of system design. This will ensure integration of the system into existing educational practices is as easy as possible. Additionally, systems should be readily supported within school technology support infrastructures.

6.3 Future Directions

The Co-Design and Evaluation sessions uncovered key future directions for the development of systems for multi-part 3DPMs in education.

- **Scalability to a Range of 3DPMs:** Both educators and students were in favour of introducing a wide range of 3DPMs to the system. Future designs should provide an opportunity for educators to be able to readily extend the system to handle a new 3DPM and provide vision recognition, part identification and assembly assistance. In order to facilitate this using the system, a semi-automated data collection and annotation process would be highly beneficial. A possible implementation would be to: (1) record a video of one or more objects from different angles, distances, and grasps; (2) employ a large state-of-the-art model (e.g., SAM2 [52]) to track the object(s) in the video - this will require a simple and fast manual initialisation of the object(s) locations in the first frame of the video; and (3) use the tracking results to annotate all frames in the video. We are in the process of developing and testing this functionality to scale the system to other 3DPMs used in education.
- **Natural Dialogue and Conversational Enquiry:** Educators and students desired a more natural dialogue that allowed them to more independently gain further information upon enquiry. Participants were particularly interested in a conversational interface that provided more information when interacting with a specific part of the 3DPMs e.g. student may ask, “what are the pointy things on the head?” A desire for supporting this conversational enquiry is consistent with previous research into 3DPMs [55] and suggests future exploration of large language model based conversational agents.
- **Tactile Differentiation of Model Parts:** The Beesly model employed different textures for the forewing and hindwing. Although it was not a deliberate design choice, it was noted by some participants. Previous work has utilised texture to support recognition [34] and tactually distinguish between model parts [69]. Indeed, previous work into similar systems has highlighted educator interest in tactile differentiation to

make different model parts more distinct from one another [59]. Hence, the use of texture would be a valuable design choice for future multi-part 3DPMs.

6.4 Limitations

Our research should be viewed as an initial exploration of computer-vision based voice interfaces for multi-part 3DPMs. As such it has a number of limitations.

6.4.1 Technical Limitations. One of the app’s limitations is in its speech-to-text reliability. This is a core component of the application, being the basis for recognising commands and auditory quiz responses. The application relies on Apple’s Speech framework for accurately transcribing text in real-time. While this framework is a state-of-the-art model for transcribing audio, its intention for generalised speech recognition combined with its context-sensitive approach creates limitations in the app’s ability to recognise commands and quiz responses. Typically, specialised models are used in these contexts, however our focus on ‘off-the-shelf’ technologies imposed a limitation that must be investigated.

There are also limitations in the model’s accuracy in being able to detect interactions between the Beesly 3DPM’s components and the user’s hands. While accurate, the model is not infallible to incorrectly detecting which component the user is holding. Furthermore, we used a controlled environment to achieve the best possible performance with the current implementation. This included a white background and the user holding one model part at a time.

Now that we have established the usefulness of the app and that computer vision is a viable approach to supporting interaction with multi-part 3DPMs in this paper, future work will focus on: improving the robustness of the system; implementing depth-sensing capabilities using the iPad LiDAR Scanner to address limitations in the 2D logic approach (Section 4.1.6); and quantitative evaluation of the model’s accuracy in a classroom environment.

Lastly, aligned with educator interest in iPad technology (discussed in Section 4.1.1) the app was designed for iOS systems. The proposed approach is relatively general and uses established computer vision techniques and workflows, hence we do not expect any issues with deploying similar functionality on Android. However, efficiency will need to be evaluated.

6.4.2 Methodology Limitations. There are also limitations to acknowledge regarding aspects of the co-design and evaluation sessions. Although multiple 3DPMs were used as prompts in the evaluation sessions (Section 5.1.2), the system only functioned with the ‘Beesly’ model. This one 3D model does not embody all the relevant features of multi-part 3DPMs used in STEM education. In order to understand other design implications, a range of 3DPMs should be incorporated into the system and evaluated.

While the system was learned quickly by students, and evidence was obtained regarding its efficacy, further insight could be derived from evaluating its extended use in a classroom setting and its impact on student learning. In addition, we did not conduct a controlled evaluation comparing the efficacy of the system with a baseline in which the app was not used to support interaction with the model. Both of these evaluations would allow for further data collection and quantitative analysis which was not present in

this work and potentially provide further evidence for the system's efficacy.

Lastly, in order to align with the guidelines of the research ethics approval, we could not collect data directly from the students themselves. Future work will involve students as direct participants in the evaluation of the system.

7 Conclusion

This work presents the co-design, iterative design process and evaluation of an iPad based application that supports the use of multi-part 3DPMs used in STEM education. The co-designed app presented a meaningful way of interacting with multi-part 3DPMs, and our findings illuminate strong educator interest in integrating these into their teaching practice. By supporting detailed interaction with multi-part 3DPMs, our results suggest that apps like this have the potential to make STEM education more accessible, engaging and inclusive. Evaluation showed that the system was strongly enjoyed by students and educators were enthusiastic for its implementation in the classroom.

The design considerations outline the key functional, educational and practical requirements essential to the successful design and implementation of future computer vision-supported educational tools. Our next major focus will be on scalability, ensuring BLV users and accessible format producers are able to train the system on a 3D object of their choice. Further evaluation of the app by students who are BLV will also be undertaken to evaluate the effect of the system on student learning outcomes.

Acknowledgments

We extend our gratitude to our co-designers and participants for their invaluable input, expertise and time.

References

- [1] 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). Association for Computing Machinery, New York, NY, USA, 249–258. doi:10.1145/3234695.3236344
- [2] Kimberly Arcand, April Jubett, M. Watzke, Sara Price, Kelly Williamson, and Peter Edmonds. 2019. Touching the stars: improving NASA 3D printed data sets with blind and visually impaired audiences. *Journal of Science Communication* 18 (07 2019). doi:10.22323/2.18040201
- [3] Kimberly Kowal Arcand, April Jubett, Megan Watzke, Sara Price, Kelly T. S. Williamson, and Peter Edmonds. 2019. Touching the stars: improving NASA 3D printed data sets with blind and visually impaired audiences. arXiv:1906.06457 [astro-ph.IM]
- [4] Brandon Biggs, James M. Coughlan, and Peter Coppin. 2021. Design and evaluation of an interactive 3D map. *Rehabilitation Engineering and Assistive Technology Society of North America* (2021).
- [5] Abena Boadi-Agyemang, Elizabeth Jeanne Carter, Alexa F Siu, Aaron Steinfeld, and Melisa Orta Martinez. 2023. Understanding Experiences, Attitudes and Perspectives towards Designing Interactive Creative Tools for Teachers of Visually Impaired Students. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) (ASSETS '23). Association for Computing Machinery, Article 62, 5 pages. doi:10.1145/3597638.3614512
- [6] 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). Association for Computing Machinery, New York, NY, USA, 131–138. doi:10.1145/2148131.2148160
- [7] Erin Buehler, Stacy Branham, Abdullah Ali, Jeremy J. Chang, Megan Kelly Hofmann, Amy Hurst, and Shaun K. Kane. 2015. Sharing is Caring: Assistive Technology Designs on Thingiverse. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 525–534. doi:10.1145/2702123.2702525
- [8] 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 (mar 2016), 28 pages. doi:10.1145/2870640
- [9] Erin Buehler, Shaun K. Kane, and Amy Hurst. 2014. ABC and 3D: Opportunities and Obstacles to 3D Printing in Special Education Environments. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility* (Rochester, New York, USA) (ASSETS '14). Association for Computing Machinery, New York, NY, USA, 107–114. doi:10.1145/2661334.2661365
- [10] Matthew Butler, Leona M Holloway, Samuel Reinders, Catagay 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, Article 278, 15 pages. doi:10.1145/3411764.3445207
- [11] Matthew Butler, Erica J Tandori, Vince Dziekan, Kirsten Ellis, Jenna Hall, Leona M Holloway, Ruth G Nagassa, and Kim Marriott. 2023. A Gallery In My Hand: A Multi-Exhibition Investigation of Accessible and Inclusive Gallery Experiences for Blind and Low Vision Visitors. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) (ASSETS '23). Association for Computing Machinery, Article 9, 15 pages. doi:10.1145/3597638.3608391
- [12] Virginio Cantoni, Luca Lombardi, Alessandra Setti, Stanislav Gyoshev, Dimitar Karastoyanov, and Nikolay Stoimenov. 2018. Art Masterpieces Accessibility for Blind and Visually Impaired People. In *Computers Helping People with Special Needs*, Klaus Miesenberger and Georgios Kouroupetroglou (Eds.). Springer International Publishing, Cham, 267–274.
- [13] Ruei-Che Chang, Chih-An Tsao, Fang-Ying Liao, Seraphina Yong, Tom Yeh, and Bing-Yu Chen. 2021. Daedalus in the Dark: Designing for Non-Visual Accessible Construction of Laser-Cut Architecture. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) (UIST '21). Association for Computing Machinery, New York, NY, USA, 344–358. doi:10.1145/3472749.3474754
- [14] Ruei-Che Chang, Wen-Ping Wang, Chi-Huan Chiang, Te-Yen Wu, Zheer Xu, Justin Luo, Bing-Yu Chen, and Xing-Dong Yang. 2021. AccessibleCircuits: Adaptive Add-On Circuit Components for People with Blindness or Low Vision. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 670, 14 pages. doi:10.1145/3411764.3445690
- [15] Ruei-Che Chang, Seraphina Yong, Fang-Ying Liao, Chih-An Tsao, and Bing-Yu Chen. 2023. Understanding (Non-)Visual Needs for the Design of Laser-Cut Models. 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 639, 20 pages. doi:10.1145/3544548.3580684
- [16] James M. Coughlan, Brandon Biggs, and Huiying Shen. 2022. Non-visual Access to an Interactive 3D Map. In *Computers Helping People with Special Needs*, Klaus Miesenberger, Georgios Kouroupetroglou, Katerina Mavrou, Roberto Manduchi, Mario Covarrubias Rodriguez, and Petr Penáz (Eds.). Springer International Publishing, Cham, 253–260.
- [17] James M. Coughlan, Brandon Biggs, and Huiying Shen. 2022. Point and Listen: Bringing a 3D Map to Life with Audio-Based AR. *Frameless* (2022).
- [18] J. M. Coughlan, H. Shen, and B. Biggs. 2020. Towards Accessible Audio Labeling of 3D Objects. *J Technol Pers Disabil* 8 (2020), 210–222.
- [19] Michael Crossland, Rui Silva, and Antonio Macedo. 2014. Smartphone, tablet computer and e-reader use by people with vision impairment. *Ophthalmic and Physiological Optics* 34 (07 2014). doi:10.1111/opo.12136
- [20] Josh Urban Davis, Te-Yen Wu, Bo Shi, Hanyi Lu, Athina Panotopoulou, Emily Whiting, and Xing-Dong Yang. 2020. TangibleCircuits: An Interactive 3D Printed Circuit Education Tool for People with Visual Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376513
- [21] Danyang Fan, Gene S-H Kim, Olivia Tomassetti, Shloke Nirav Patel, Sile O'Modhrain, Victor R Lee, and Sean Follmer. 2024. Tangible Stats: An Embodied and Multimodal Platform for Teaching Data and Statistics to Blind and Low Vision Students. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems* (CHI EA '24). Association for Computing Machinery, New York, NY, USA, Article 310, 9 pages. doi:10.1145/3613905.3650793
- [22] Zheng Fang, Yanhua Fu, and Tianyou Chai. 2009. A low-cost modular robot for research and education of control systems, mechatronics and robotics. In *2009 4th IEEE Conference on Industrial Electronics and Applications*. 2828–2833. doi:10.1109/ICIEA.2009.5138725
- [23] Melissa Fanshawe, Georgina Barton, Monique Mandarakas, Melissa Cain, and Nicole Todd. 2023. Enablers and barriers to equitable participation for students with blindness or low vision in Australian mainstream secondary schools. *International journal of inclusive education* (2023), 1–17.

- [24] Uttara Ghodke, Lena Yusim, Sowmya Somanath, and Peter Coppin. 2019. The Cross-Sensory Globe: Participatory Design of a 3D Audio-Tactile Globe Prototype for Blind and Low-Vision Users to Learn Geography. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) (DIS '19). Association for Computing Machinery, New York, NY, USA, 399–412. doi:10.1145/3322276.3323686
- [25] Stéphanie Giraud, Anke M. Brock, Marc J.-M. Macé, and Christophe Jouffrais. 2017. Map Learning with a 3D Printed Interactive Small-Scale Model: Improvement of Space and Text Memorization in Visually Impaired Students. *Frontiers in Psychology* 8 (2017). doi:10.3389/fpsyg.2017.00930
- [26] Ankit Gupta, Dieter Fox, Brian Curless, and Michael Cohen. 2012. DuploTrack: A Real-Time System for Authoring and Guiding Duplo Block Assembly. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA) (UIST '12). Association for Computing Machinery, New York, NY, USA, 389–402. doi:10.1145/2380116.2380167
- [27] Jaylin Herskovitz, Andi Xu, Rahaf Alharbi, and Anhong Guo. 2023. Hacking, Switching, Combining: Understanding and Supporting DIY Assistive Technology Design by Blind People. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 57, 17 pages. doi:10.1145/3544548.3581249
- [28] Leona Holloway. 2024. *3D Printing for Access to Graphics by Touch Readers*. Thesis. Faculty of Information Technology. doi:10.26180/26300848.v1
- [29] Leona Holloway, Kim Marriott, and Matthew Butler. 2018. Accessible Maps for the Blind: Comparing 3D Printed Models with Tactile Graphics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal, QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3173772
- [30] 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. doi:10.1145/3308561.3353790
- [31] 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. doi:10.1145/3308561.3353790
- [32] Rabia Jafri, Asmaa Mohammed Aljuhani, and Syed Abid Ali. 2017. A tangible user interface-based application utilizing 3D-printed manipulatives for teaching tactual shape perception and spatial awareness sub-concepts to visually impaired children. *International Journal of Child-Computer Interaction* 11 (2017), 3–11. doi:10.1016/j.ijcci.2016.12.001
- [33] Chandrika Jayant, Hanjie Ji, Samuel White, and Jeffrey P. Bigham. 2011. Supporting blind photography. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility* (Dundee, Scotland, UK) (ASSETS '11). Association for Computing Machinery, New York, NY, USA, 203–210. doi:10.1145/2049536.2049573
- [34] Jacqueline Johnstone, Madhuka Nadeeshani, Hanmin Chen, Mohith Vemula, Erica J Tandori, Kate Stephens, Hashimi Senaratne, Kirsten Ellis, and Swamy Ananthanarayan. 2024. Designing Accessible Adaptations for an Electronic Toolkit with Blind and Low Vision Users. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) (ASSETS '24). Association for Computing Machinery, New York, NY, USA, Article 43, 15 pages. doi:10.1145/3663548.3675652
- [35] Shaun K. Kane and Jeffrey P. Bigham. 2014. Tracking @stemxcomet: teaching programming to blind students via 3D printing, crisis management, and twitter. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education* (Atlanta, Georgia, USA) (SIGCSE '14). Association for Computing Machinery, New York, NY, USA, 247–252. doi:10.1145/2538862.2538975
- [36] Özgül Keleş, Naim Uzun, and Gülistan Yalçın. 2023. Barrier-Free Science Education for Visually Impaired Students: An Activity for Life Cycle in Plants. *Science Insights Education Frontiers* 14 (01 2023), 1945–1964. doi:10.15354/sief.23.or082
- [37] Jeeun 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* (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 2815–2824. doi:10.1145/2702123.2702144
- [38] Varsha Koushik and Shaun K. Kane. 2017. Tangibles + Programming + Audio Stories = Fun. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) (ASSETS '17). Association for Computing Machinery, New York, NY, USA, 341–342. doi:10.1145/3132525.3134769
- [39] Kyungjun Lee, Jonggi Hong, Simone Pimento, Ebrima Jarjue, and Hernisa Kacorri. 2019. Revisiting Blind Photography in the Context of Teachable Object Recognizers. 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, 83–95. doi:10.1145/3308561.3353799
- [40] Kyungjun Lee and Hernisa Kacorri. 2019. Hands Holding Clues for Object Recognition in Teachable Machines. 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. doi:10.1145/3290605.3300566
- [41] Myungjoong Lee, Seung-Won Kim, and Hyunki In. 2024. Archi-TangiBlock: A Modular Block Based Tangible Media as a CAD Tool for Visually Impaired People. *IEEE Access* 12 (2024), 46582–46595. doi:10.1109/ACCESS.2024.3382311
- [42] Barbara Leporini, Valentina Rossetti, Francesco Furfari, Susanna Pelagatti, and Andrea Quarta. 2020. Design Guidelines for an Interactive 3D Model as a Supporting Tool for Exploring a Cultural Site by Visually Impaired and Sighted People. *ACM Trans. Access. Comput.* 13, 3, Article 9 (aug 2020), 39 pages. doi:10.1145/3399679
- [43] Charlotte Magnusson, Per-Olof Hedvall, and Héctor Caltenco. 2018. *Co-designing together with Persons with Visual Impairments*. Springer International Publishing, Cham, 411–434. doi:10.1007/978-3-319-54446-5_14
- [44] Jonathan Muth, Andre Klunker, and Christina Völlmecke. 2023. Putting 3D printing to good use—Additive Manufacturing and the Sustainable Development Goals. *Frontiers in Sustainability* 4 (2023). doi:10.3389/frsus.2023.1196228
- [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. doi:10.1145/3544548.3581389
- [46] Ruth G Nagassa, Matthew Butler, Leona M Holloway, Erica J Tandori, and Kim Marriott. 2024. Push, Scan, Tap: Embracing Commodity Technology for the provision of Accessible Information in Cultural Institutions. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '24). Association for Computing Machinery, New York, NY, USA, Article 525, 6 pages. doi:10.1145/3613905.3637122
- [47] Andrea Narcisi, Huiying Shen, Dragan Ahmetovic, Sergio Mascetti, and James M. Coughlan. 2024. Accessible Point-and-Tap Interaction for Acquiring Detailed Information About Tactile Graphics and 3D Models. In *Computers Helping People with Special Needs*, Klaus Miesenberger, Petr Peňáz, and Makoto Kobayashi (Eds.). Springer Nature Switzerland, Cham, 252–259.
- [48] Jon T. Njardarson. 2023. Introductory Organic Chemistry (First-Semester) for Blind and Visually Impaired Students: Practical Lessons and Experiences. *Journal of Chemical Education* 100, 10 (2023), 3960–3967. doi:10.1021/acs.jchemed.3c00616
- [49] Jill Opie, Joanne Deppeler, and Jane Southcott. 2017. 'You have to be like everyone else': Support for students with vision impairment in mainstream secondary schools. *Support for Learning* 32, 3 (2017), 267–287. doi:10.1111/1467-9604.12169 arXiv:https://nasejournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9604.12169
- [50] Emilia Christie Picelli Sanches, Juliana Bueno, and Maria Lucia Leite Ribeiro Okimoto. 2021. Designing 3D Printed Audio-Tactile Graphics: Recommendations from Prior Research. In *Universal Access in Human-Computer Interaction. Design Methods and User Experience*, Margherita Antona and Constantine Stephanidis (Eds.). Springer International Publishing, Cham, 461–472.
- [51] Cynthia Putnam, Melisa Puthenmadom, Marjorie Ann Cuerdo, Wanshu Wang, and Nathaniel Paul. 2020. Adaptation of the System Usability Scale for User Testing with Children. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '20). Association for Computing Machinery, New York, NY, USA, 1–7. doi:10.1145/3334480.3382840
- [52] Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, Eric Mintun, Junting Pan, Kalyan Vasudev Alwala, Nicolas Carion, Chao-Yuan Wu, Ross Girshick, Piotr Dollár, and Christoph Feichtenhofer. 2024. SAM 2: Segment Anything in Images and Videos. arXiv:2408.00714 [cs.CV] https://arxiv.org/abs/2408.00714
- [53] Joseph Redmon and Ali Farhadi. 2016. YOLO9000: Better, Faster, Stronger. arXiv:1612.08242 [cs.CV] https://arxiv.org/abs/1612.08242
- [54] Samuel Reinders. 2021. Accessible Interactive 3D Models for Blind and Low-Vision People. *SIGACCESS Access. Comput.* 129, Article 6 (mar 2021), 7 pages. doi:10.1145/3458055.3458061
- [55] Samuel Reinders, Swamy Ananthanarayan, Matthew Butler, and Kim Marriott. 2023. Designing Conversational Multimodal 3D Printed Models with People Who Are Blind. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) (DIS '23). Association for Computing Machinery, New York, NY, USA, 2172–2188. doi:10.1145/3563657.3595989
- [56] Samuel Reinders, Matthew Butler, and Kim Marriott. 2020. "Hey Model!" – Natural User Interactions and Agency in Accessible Interactive 3D Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York,

- NY, USA, 1–13. doi:10.1145/3313831.3376145
- [57] Elen Sargsyan, Bernard Oriola, Marc J-M Macé, Marcos Serrano, and Christophe Jouffrais. 2023. 3D Printed Interactive Multi-Storey Model for People with Visual Impairments. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 540, 15 pages. doi:10.1145/3544548.3581304
- [58] Lei Shi. 2015. Talkabel: A Labeling Method for 3D Printed Models. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (Lisbon, Portugal) (ASSETS '15). Association for Computing Machinery, New York, NY, USA, 361–362. doi:10.1145/2700648.2811327
- [59] Lei Shi, Holly Lawson, Zhuohao Zhang, and Shiri Azenkot. 2019. Designing Interactive 3D Printed Models with Teachers of the Visually Impaired. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3290605.3300427
- [60] Lei Shi, Ross McLachlan, Yuhang Zhao, and Shiri Azenkot. 2016. Magic Touch: Interacting with 3D Printed Graphics. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (Reno, Nevada, USA) (ASSETS '16). Association for Computing Machinery, New York, NY, USA, 329–330. doi:10.1145/2982142.2982153
- [61] Lei Shi, Idan Zelzer, Catherine Feng, and Shiri Azenkot. 2016. Tickers and Talker: An Accessible Labeling Toolkit for 3D Printed Models. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 4896–4907. doi:10.1145/2858036.2858507
- [62] Lei Shi, Zhuohao Zhang, and Shiri Azenkot. 2018. A Demo of Talkit++: Interacting with 3D Printed Models Using an IOS Device. 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, 429–431. doi:10.1145/3234695.3241004
- [63] Lei Shi, Yuhang Zhao, and Shiri Azenkot. 2017. Designing Interactions for 3D Printed Models with Blind People. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) (ASSETS '17). Association for Computing Machinery, New York, NY, USA, 200–209. doi:10.1145/3132525.3132549
- [64] Lei Shi, Yuhang Zhao, and Shiri Azenkot. 2017. Markit and Talkit: A Low-Barrier Toolkit to Augment 3D Printed Models with Audio Annotations. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) (UIST '17). Association for Computing Machinery, New York, NY, USA, 493–506. doi:10.1145/3126594.3126650
- [65] Ishu Singhal and B. S. Balaji. 2020. Creating Atom Representations Using Open-Source, Stackable 3D Printed Interlocking Pieces with Tactile Features to Support Chemical Equation Writing for Sighted and Visually Impaired Students. *Journal of Chemical Education* 97, 1 (2020), 118–124. doi:10.1021/acs.jchemed.9b00255 arXiv:https://doi.org/10.1021/acs.jchemed.9b00255
- [66] Ishu Singhal and B. S. Balaji. 2022. Open-Source, Tactile 3D Printed Interlockable Tiles Incorporating Valency, Bonding, and Hybridization for Molecular Representation for Sighted and Visually Impaired Students. *Journal of Chemical Education* 99, 4 (2022), 1708–1714. doi:10.1021/acs.jchemed.1c01278 arXiv:https://doi.org/10.1021/acs.jchemed.1c01278
- [67] Carlos Tejada, Osamu Fujimoto, Zhiyuan Li, and Daniel Ashbrook. 2018. Blow-hole: Blowing-Activated Tags for Interactive 3D-Printed Models. In *Proceedings of the 44th Graphics Interface Conference* (Toronto, Canada) (GI '18). Canadian Human-Computer Communications Society, Waterloo, CAN, 131–137. doi:10.20380/GI2018.18
- [68] Yong-Joon Thoo, Maximiliano Jeanneret Medina, Jon E. Froehlich, Nicolas Ruffieux, and Denis Lalanne. 2023. A Large-Scale Mixed-Methods Analysis of Blind and Low-vision Research in ACM and IEEE. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) (ASSETS '23). Association for Computing Machinery, New York, NY, USA, Article 20, 20 pages. doi:10.1145/3597638.3608412
- [69] Xiyue Wang, Seita Kayukawa, Hironobu Takagi, and Chieko Asakawa. 2022. BentoMuseum: 3D and Layered Interactive Museum Map for Blind Visitors. 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 35, 14 pages. doi:10.1145/3517428.3544811
- [70] Xiyue Wang, Seita Kayukawa, Hironobu Takagi, and Chieko Asakawa. 2023. TouchPilot: Designing a Guidance System that Assists Blind People in Learning Complex 3D Structures. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) (ASSETS '23). Association for Computing Machinery, New York, NY, USA, Article 5, 18 pages. doi:10.1145/3597638.3608426
- [71] 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. doi:10.1021/ed3000364 arXiv:https://doi.org/10.1021/ed3000364

- [72] Yichun Zhao, Miguel A Nacenta, Mahadeo A. Sukhai, and Sowmya Somanath. 2024. TADA: Making Node-link Diagrams Accessible to Blind and Low-Vision People. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 45, 20 pages. doi:10.1145/3613904.3642222

A Computer Vision Accuracy

Quantitative analysis of the computer vision model on the validation dataset is as follows.

Class	Count	Correct	FP	FN	Precision	Recall
Wing (red)	85	70	9	6	89%	92%
Wing (green)	100	80	16	4	83%	95%
Thorax (blue)	74	62	5	7	93%	90%
Head (yellow)	82	72	6	4	92%	95%
Abdomen (magenta)	84	74	7	3	91%	96%

Table 1: Performance metrics for the classification of 3D printed parts of the Beesly model. FP = False Positive, FN = False Negative.