

Visualizing the Invisible: Co-creating an AI-Assisted Tactile Picture Book With a Congenitally Blind Student

Chu-Yu Cheng, Southern Taiwan University of Science and Technology, Taiwan
Kuan-Ju Ho, Southern Taiwan University of Science and Technology, Taiwan

The European Conference on Education 2026
Official Conference Proceedings

Abstract

This study explores how congenitally blind students perceive and imagine biological forms through co-creating an AI-assisted tactile picture book. Our research laboratory specializes in developing educational materials for visually impaired students. After completing several projects helping students understand biological characteristics, we became curious: how do students without visual experience conceptualize living creatures? We used AI as a creative tool and aesthetic design as a communication channel to externalize a blind student's imaginative world. The visually impaired (VI) student, along with design professionals, collaborated using Google Labs Mixboard to co-create a picture book about urban insects. The 15-year-old congenitally blind student developed characters, invented storylines, and orally described scenes. Inspired by Chinese martial arts novelist Louis Cha Leung-yung, he created a ten-legged cockroach protagonist with martial arts-style actions and settings. Prompts derived from his descriptions were input into Mixboard for image generation, with designers refining the outputs. The final 24-page book was produced using swell paper printing technology, creating tactile graphics accompanied by Chinese Braille. This collaborative project reveals that VI students' biological understanding often stems from verbal descriptions and tactile toys, resulting in imaginative yet scientifically inaccurate conceptions. The methodology demonstrates how AI-assisted co-creation can bridge imagination and reality while providing valuable insights into alternative cognitive frameworks and celebrating the boundless creativity of students with visual impairments. Its contribution is an account of a working method and a set of transferable principles, not a measure of learning effectiveness.

Keywords: practice-based design research, congenital blindness, tactile picture book, generative image models, authorship

iafor

The International Academic Forum
www.iafor.org

Introduction

Generative AI is now used in classrooms, including in special education (Holmes & Tuomi, 2022; Khlaif et al., 2026). A language model can turn what a blind author imagines into a picture, and digital fabrication can turn that picture into an object he can touch. This was not possible before (Tanaka & Ochiai, 2023). However, most special education still treats blind students as people who receive information that others prepare for them (Khlaif et al., 2026).

This study began during earlier work on tactile teaching materials. A 15-year-old congenitally blind student described the urban insects (cockroaches) in his home: the scratching sound, the damp smell, the segmented body under his fingers. He had never seen one, and his account came entirely from hearing, touch, and what others had told him. Asked how many legs a cockroach has, he said ten, attributing the figure to a plush toy he had handled for years.

That is not a defect in his account. Congenitally blind people build mental images from what their other senses give them, not from remembered sight, and those images hold together on their own terms (Cattaneo & Vecchi, 2011; Cattaneo et al., 2008). This study asks how such a student builds the shape of an animal he has never seen, whether generative tools can carry that shape forward instead of correcting it, and what a working process must do to keep him in the author's role.

Materials for blind students are built to deliver content. Braille, tactile graphics, and multimodal books all make it reachable (Lahav & Mioduser, 2004; Valente et al., 2024). Blind learners have been brought into design decisions, though rarely in schools. Three of 21 teachers of the visually impaired involved students in making the materials they would later use (Phutane et al., 2022). Participation at that stage shapes how this subject is presented. The subject is still chosen by teachers, publishers, and curriculum designers.

A route in the other direction runs through three technologies. Language processing can work from what the learner says and keep his sensory vocabulary instead of replacing it with a visual one. A text-to-image model turns that description into a picture. Swell-paper printing returns the picture to his hands. This study traces that sequence through a single sustained case: the co-creation of a 24-page tactile picture book with the student described above. The design brief was set by his sensory account rather than by an existing curriculum framework, and this paper reports what that reversal produced.

Literature Review

Sensory Perception and Learning Characteristics of VI Students

Blind and visually impaired learners come to rely more on hearing and touch than on any remaining sight (Cattaneo & Vecchi, 2011; Cattaneo et al., 2008). Touch recognizes an object by sampling it part by part. This is slower than vision, but what it builds is stable. Teaching methods adapted to non-visual channels help blind students learn about shapes and raise their motivation to learn (Wu, 2023).

The point that matters here is about images rather than about access to information. Congenitally blind people form and handle mental images built from what their senses give them, not rebuilt from remembered sight (Cattaneo et al., 2008). An image made this way is not a rough copy of a visual original. It is a picture with its own logic. Valente et al. (2024)

tested how well children with visual impairments understand a multimodal book, taking comprehension as the outcome. That question is about reception: how well a reader takes in content prepared by others. What happens when the blind reader is also the source of that content is a different question, and it is the one this study takes up.

AI in Education and the Direction of Sensory Translation

Generative AI and natural language processing are already used in teaching, as adaptive tutors and as tools for making content (Holmes & Tuomi, 2022). In special education, AI has been used for visual impairment mainly in one direction: it turns visual material into non-visual output for a blind user (Khlaif et al., 2026). Automatic image description, scene narration, and AI-made tactile graphics from existing diagrams all run this way, as does the text-to-touch pipeline of Tanaka and Ochiai (2023).

In all of these the blind user is at the receiving end. The content starts with a sighted author, a curriculum, or a camera. AI translates it. The blind user takes in the result. This makes access the goal, and the design question becomes how much of the visual original survives the conversion.

The same technology can run the other way. Because it works from what the learner says, it can keep his sensory vocabulary instead of replacing it with a visual one. Paired with image generation, this makes his description the starting point of the translation rather than its destination, and each round of his comments changes what the system produces next.

Tactile Design and Swell-Paper Production

A tactile image only works if the hand can follow it. That means clear outlines, a clear order to explore in, and shapes the reader can recognize (Wu et al., 2022). Thermal swell paper is the most common way to make such images. It is cheap and needs only heat unit by printing (Phutane et al., 2022). Good design also needs varied line weights, enough space between parts, Braille labels, and good to go with the touch (Wu et al., 2022).

Tactile picture books sit inside this literature. Valente et al. (2024) test how well children with visual impairments understand a multimodal book. Both look at books made for blind readers. Sighted designers and teachers chose the subjects, wrote the stories, and made the design decisions. Researchers have answered this by bringing blind learners into design, but little of it has reached classrooms. Phutane et al. (2022) interviewed 21 teachers of the visually impaired and ran a three-week diary study with eight of them. Three had involved students in making tactile materials. Where they did, the making itself became the lesson, and students used the results with more confidence. Participatory design is well established in assistive technology research and largely missing from education, and earlier systems either automated the making or treated the teacher as the only user. The same interviews show why substitutes are common. Teachers wanted three-dimensional models and real objects, could not get them, and made replacements from craft supplies and cheap household goods.

The 4F Reflection Model

The 4F model sets out four stages of reflection: recording what happened, noting how it felt, working out what it means, and planning what to do next. It grows out of the experiential learning tradition running from Dewey through Kolb (2015) and Schön (1983), and cycles of

this kind give participants a sequence for turning experience into words. Here it works as a turn-taking structure rather than a measure. At each round the author had to say what he noticed, how he felt, what he concluded, and what he wanted changed, so the next image answered him rather than a designer.

Research Gap

Each of these literatures stops at the same point. Research on blind cognition shows that congenitally blind learners build coherent images from non-visual input (Cattaneo et al., 2008), but does not ask what happens when those images are made available to others. Work on AI in visual impairment education has developed the visual-to-tactile pipeline in detail (Khlaif et al., 2026; Tanaka & Ochiai, 2023), but its direction is fixed: the source is visual, the blind user is the destination, and success means staying faithful to the original. Studies of tactile materials have brought blind learners into the making, yet the teacher still chooses the subject (Phutane et al., 2022). The gap is about position, not technology, since the tools themselves are available and documented (Lahav & Mioduser, 2004).

Methodology

Research Positioning

This is practice-based design research. What it produces is an object, a 24-page tactile picture book, and the knowledge it offers is the kind such research offers: an account of how something was made, the decisions taken along the way, and principles that carry over to similar work. It does not offer general measures of learning. The single participant is not a sample. He is the author of the book, and the study reports what became possible once he was given that role.

Participant and Ethical Framework

One 15-year-old student, congenitally blind with no visual experience from birth, participated as the author. His guardians and school provided written consent, and his own assent was obtained at the outset and confirmed at each stage. He directed creative decisions, approved all outputs, and could decline to proceed at any point. Interviews were conducted in his home, where he could move and handle objects with confidence. Identifying details other than age and visual status have been withheld.

Two features of this arrangement were adopted deliberately. The student was named as author of the resulting book rather than described as a participant in it. Where his account of an animal conflicted with the biological record, the conflict was documented rather than resolved, and no attempt was made during production to bring his descriptions into line with the taught form.

Tools

Google Labs Mixboard, an interface that combines language understanding with generative image editing, turned the author's spoken descriptions into images. He directed all revisions. Researchers worked the interface and turned his instructions into prompts. A 4F protocol shaped three formal interviews, covering sensory facts, feelings and creative ideas, and the building of the story. Final images were made into tactile graphics on a TactPlus thermal swell-paper printer, which raises treated paper under heat into readable relief. This is a raised-line method, not additive fabrication. No three-dimensional printing was used at any stage.

Procedure

Stage 1 (Months 1–2): Interview and Exploration

Three formal interviews followed the 4F structure. The author described household insects, the sensory experiences attached to them, and his emotional responses, and the team documented his language, metaphors, and comparisons. Preliminary character concepts emerged here.

Stage 2 (Months 2–5): Co-creation

His descriptions were transcribed into prompts and submitted to Mixboard by an University student. He reviewed each generated image through a researcher's verbal description and issued revisions. Twenty-four cycles were completed across the two characters, a cockroach and a gecko, and the narrative developed alongside the images with his approval at each part.

Stage 3 (Months 5–8): Production and Author Verification

Final images were converted to tactile form using the thermal swell-paper printer, and the 24-page book was completed with Chinese and English text alongside Chinese Braille. The author then read the finished pages by touch. He checked mainly the details of setting that carry the sensory core of each part, among them the textures for humidity, the stickiness of the trap, and the wall cracks, and he confirmed that the printed surfaces matched the scenes he had described. He expressed clear satisfaction with the finished object and with the story as his own. Reading by other visually impaired students is planned as the next stage of this work and had not been conducted at the time of writing.

Data

The corpus comprises roughly fifteen hours of verbatim interview recordings, field notes on the author's reactions and creative decisions, an archive of more than 240 generated images across all iteration cycles, recordings of his responses to images and printed pages, process photographs, and the book itself.

Findings

The Author's Sensory World

Asked about the cockroaches in his home, the author did not describe their appearance. He described the scratching in the walls, the damp smell near the kitchen, and one remembered occasion of touching one: smooth segments, legs in sequence. His knowledge came from three sources of unequal weight. Most came from what family members had told him. A plush cockroach he had handled many times gave less, but what it gave was more specific. The direct encounter gave least of all, being brief and offering no chance to count or compare.

His language for movement was precise in ways visual description does not capture. The spoke of cockroaches rushing like water through cracks, and of the body trembling when touched. These temporal and textural terms carried into the images he later commissioned.

The Sources of Structural Knowledge

Across the interview stages the author's accounts of animal form returned repeatedly to objects he had handled. Three cases are presented in descending order of evidential strength, since the objects differ in whether they could be examined.

The Frog: Object Available

He described frogs as having forelimbs much longer than their hind limbs, which reverses the real proportion. The toy he named still exists and we examined it. It comes from the IKEA SAGOSKATT range, a limited yearly collection in which children's competition drawings are made into plush toys (IKEA Museum, n.d.). The published selection criteria cover originality, colour, expressiveness, and whether a design can be mass-produced. Anatomical accuracy is not one of them (IKEA, 2021). The toy sits in a crouch. Its forelimbs reach outward as slender continuous lines ending in separate digits. Its hind limbs fold against the body as a compact mass with less defined feet.

Figure 1

The IKEA SAGOSKATT Frog



Its proportions therefore come from a child's drawing that was made up without correction, which places the distortion before the toy rather than in it. Its posture also sets the forelimbs along a longer path than the hind limbs, so a hand following the shape measures the path it travels rather than the length of the limb.

Neither feature matters for a sighted child, who sees frogs in ponds, in books, and on film, and reads the toy as a stylized version. For this author, touch was the only way to know the toy, and the toy was his only reported source for the proportions of a frog. What a sighted audience reads as charm, and corrects without effort, reached him as the only account of the shape he had.

Two Further Cases, Objects Unavailable

Asked how many legs a cockroach has, he answered ten, attributing this to a plush cockroach his family bought when he was around ten. The toy no longer survives, so the attribution is reported as his account rather than a verified property of the object. Two readings remain open: the toy may have been made with ten legs, or it may have had six alongside antennae, cerci, or seams indistinguishable from legs under the fingers, in which case the count arose in the reading. Asked separately whether a praying mantis has six legs, he again answered ten, again

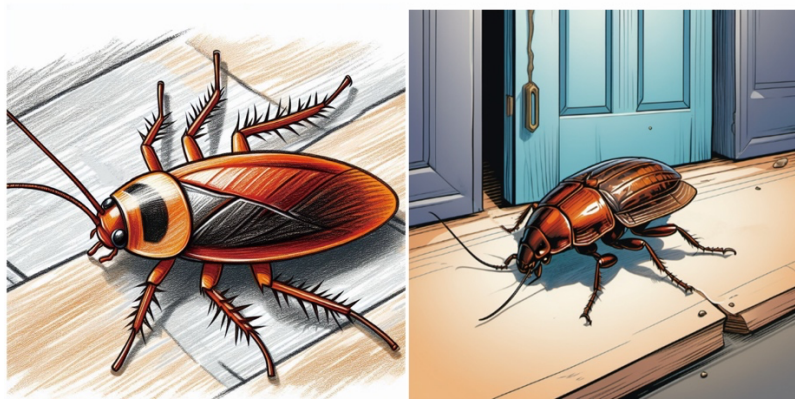
citing a plush toy; none survives and the team holds no record of one. Insect anatomy is taught in his elementary curriculum and he had received that teaching. The mantis case is of interest for its scope rather than its object. Ten legs was not a fact he held about cockroaches but the body plan he held for insects, applied to a second species when he was asked.

Directing the Image

The author's descriptions had to become pictures, and the first round showed what that risked. He had asked for a cockroach with ten legs. The model returned two images, each showing a conventional six-legged insect (Figure 2). The training data holds the standard body plan, and it supplied that plan instead of the description it was given. This was not a failure of prompting. It was the model imposing what it already held on an author who departed from it, and it marks the point where his version could most easily have been lost.

Figure 2

First-Round Outputs for the Protagonist, Showing the Six-Legged Form Produced by the Model Before Revision



The images were described to him aloud. He rejected the leg count himself, without prompting from the research team, and required ten. The revision was made, and the ten-legged form held through the twenty-four cycles that followed. The elder cockroach who appears later in the narrative was given ten legs on the same basis.

Later revisions followed the same pattern. Early antennae were too delicate, and he corrected the researchers directly: they needed to be stronger, with more feeling in them. A version was rejected because the legs were resting rather than poised to run. His instructions concerned posture and readiness rather than appearance, which is what a description assembled from sound and touch would be expected to priorities.

World-Building Without Vision

Across the narrative the author specified details he cannot perceive. A wall calendar reads 1968; a second, in a later scene, reads 30 April 1981. Two Giant electric bicycles stand outside a lobby, one blue carrying the plate 1856 and one white carrying 93156. A yellow storage cabinet beside a sofa is priced at 1,111 yuan. Five violets sit by a window. The building manager reads a newspaper, on the left.

None of this reaches him through hearing or touch. Years on calendars, digits on plates, prices on labels, the position of a figure in a frame are all registered by sight, and he supplied them

without hesitation. Some of this is the conventional furniture of verbal description. He has spent years with audio recordings of Jin Yong's novels, where rooms contain calendars, calendars carry years, vehicles carry plates, and goods carry prices. He knows what a described world contains and populated his own accordingly. The values are arbitrary. Their presence is not.

Building the Narrative

Before co-creation began, the research team taught the author the classical four-part Chinese narrative structure of *qi*, *cheng*, *zhuan* and *he*: opening, development, turn, and resolution. He then wrote to that structure. It came from the team, not from him, and we report it as a scaffold rather than as something he found for himself. What went inside it, and in what order, was his.

The 4F protocol served a different function, organizing the interviews rather than the story. The facts stage recorded what he had experienced: scratching in walls, water pipes, the damp smell near the kitchen. The feelings stage drew out his response, and he described the insects as trapped, searching for food, and frightened. In the findings stage this sensory knowledge met a larger idea, drawn from the martial arts fiction he listens to, of a protagonist facing impossible odds.

A cockroach wandering beneath a bedroom cabinet enters a newly laid trap out of curiosity and wakes to find others stuck around him. Relatives arrive, and an elder proposes sending the smallest out for help. The small one recruits geckos, who agree while privately anticipating a meal. They carry the trap to a puddle and free the colony, then turn on them, and fifty-eight are eaten. One escapes, crosses a pine wood, and reaches an apartment lobby, where it finds an abandoned nest behind a sofa. Eating crumbs there, it decides to open a school of cockroach survival, teaching others to evade geckos, avoid traps, and recognize human footsteps.

In the future stage we asked how he would explain the story to another blind student without the book in hand. He said he would describe what he had felt, the trapped feeling, then hope, then the shock of betrayal, because that was the real story. His written ending makes the same move. The protagonist is content at the close not because it survived, but because it has finally worked out how to live.

Discussion

AI as Translation Medium

The generative model worked neither as a teacher nor as a corrector but as a translator. It took the author's own sensory vocabulary as input and did not ask him for visual terms he did not have. Given descriptions full of tactile metaphor and timing, it produced images he could then revise. The tool moved from delivering information to carrying expression outward.

This depended on an arrangement that had to be kept up deliberately, and the first round showed why. Asked for the author's cockroach, the model produced a six-legged one, and the designers working the interface held the same expectation. Every point where his description differed from it was a point where the difference could have disappeared without anyone deciding to remove it. What kept the ten legs was not the tool and not the operators. It was the rule that only the author issued revisions, and the fact that he issued this one.

The Position of the Doubters

Every challenge to the ten-legged form came from sighted people, and each drew on a channel the author cannot access: textbook illustrations, diagrams, remembered images. For those raising it, the objection that a cockroach has six legs is visually acquired. For the author it had only ever been a verbal report, and it did not displace what his hands had established.

The sighted responses divided. When design students at our institution first saw the generated character, one said it looked like an alien, then that it had the quality of a strange and marvellous creature, and that the author's world was unlike the one their own visual education had furnished. The strangeness registered first as error and then as interest. At a public presentation, an audience member asked whether the work had any meaning at all.

The answer is the position of this paper. As biology instruction the work has no value and could mislead. As creative work it opens a world that visual convention does not constrain. Both responses mark the same line from opposite sides, between judging a blind author's imagery against sighted norms and reading it on his terms. The ten-legged cockroach shows that this study held the second position. A finished character with six legs would have meant that the priors of the model and its sighted operators had replaced the author's account.

What This Study Yields

The study yields an object, a documented process, and four principles that carry over to similar work. The first concerns input. The author's words should reach the model unedited, because turning them into visual vocabulary is where authorship is first lost. The second concerns control: only the author issues revisions, and collaborators work the interface without judging it. Third, places where his account differs from the biological record belong in the record of the process, not among its corrections, since they are what the process exists to carry. The fourth is checking. An object the author cannot read for himself has been made for him rather than by him.

Limitations

This is a single case. The study reports what one working process made possible. It does not measure learning, retention, or transfer. The narrative structure was not the author's own discovery, since the four-part form was taught to him at the start and only the content inside it was his. The book has two characters, so nothing here speaks to scale. No other visually impaired student has read it yet, so nothing can be said about how it works for its audience. Of the three cases of knowledge about shape, only one involves an object we could examine. The interfaces were built for sighted operators, which is why researchers rather than the author worked them, and questions of ownership over co-created work remain open.

Conclusion

This study set out to see what would follow from putting a congenitally blind student in the author's role and using generative tools to carry his account outward rather than correct it. The result is a 24-page tactile picture book whose main character has ten legs, made through a documented sequence that runs from spoken description to prompt, to generated image, to raised-line print, and checked by the author's own hands.

Three contributions follow. The paper documents a working method for authorship without vision and says where control over the work is gained and lost within it. It records cases in which an author's knowledge of animal shape came from handled objects never meant to teach, one of which can be traced back to how it was made, which suggests that the shortage of purpose-made tactile materials has effects reaching further than access. And it describes a kind of world-building that draws on learned narrative convention as well as on sensory record, visible in the calendars, plates, and price tags he specified for scenes he cannot see.

The ten legs carry the argument. The model itself gave six on the first attempt, drawing on the anatomy in its training data, and a finished cockroach with six legs would have meant that this expectation, shared by the sighted collaborators, had quietly replaced the author's account. The ten legs did not survive by default. The author heard the images described, rejected the count, and asked for them back. That he did so, and that they held through every later cycle, is the clearest sign available that the author's role, once given, was used.

Acknowledgements

We thank the student author, whose work this book is, and his family and school for their support throughout. This work was supported by the National Science and Technology Council, Taiwan, under grant NSTC 114-2410-H-218-007-MY2 and NSTC 114-2813-C-218-005-H.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

Generative AI was used in this project in two distinct capacities, which are declared separately. As a research instrument, Google Labs Mixboard was used to generate and revise the images that constitute the artefact under study. This use is the subject of the research and is described in full in the Methodology. As a writing aid, the authors used Claude (Anthropic) to improve the language, clarity, and structural organization of the English text. No text was generated for substantive content, and no analysis, finding, or interpretation originated with the tool. The authors reviewed and edited all output and take full responsibility for the content of this publication.

References

- Cattaneo, Z., & Vecchi, T. (2011). *Blind vision: The neuroscience of visual impairment*. MIT Press.
- Cattaneo, Z., Vecchi, T., Cornoldi, C., Mammarella, I., Bonino, D., Ricciardi, E., & Pietrini, P. (2008). Imagery and spatial processes in blindness and visual impairment. *Neuroscience and biobehavioral reviews*, 32(8), 1346–1360. <https://doi.org/10.1016/j.neubiorev.2008.05.002>
- Coulson, D., & Harvey, M. (2013). Scaffolding student reflection for experience-based learning: a framework. *Teaching in Higher Education*, 18(4), 401–413. <https://doi.org/10.1080/13562517.2012.752726>
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542–570. <https://doi.org/10.1111/ejed.12533>
- IKEA. (2021, March 9). *Five kids drawings come to life for SAGOSKATT collection*. IKEA Global Newsroom. <https://www.ikea.com/global/en/newsroom/collaborations/five-kids-drawings-selected-to-become-the-soft-toy-limitededition-saguskatt-collection-210309/>
- IKEA Museum. (n.d.). *SAGOSKATT soft toys from 2015*. Retrieved July 30, 2026, from <https://ikeamuseum.com/en/explore/product-stories/saguskatt-soft-toy-2016/>
- Khlaif, Z. N., Alshakhshir, R., Hamamra, B., & Joma, A. (2026). Assistive generative AI for visually impaired learners: Personalization and inclusion in higher education. *International Journal of Artificial Intelligence in Education*, 36(1), 100008. <https://doi.org/10.1016/j.ijaied.2026.100008>
- Kolb, D. A. (2015). *Experiential learning: Experience as the source of learning and development* (2nd ed.). Pearson Education.
- Lahav, O., & Mioduser, D. (2004). Exploration of Unknown Spaces by People Who are Blind Using a Multi-Sensory Virtual Environment. *Journal of Special Education Technology*, 19(3), 15–23. <https://doi.org/10.1177/016264340401900302>
- Phutane, M., Wright, J., Castro, B. V., Shi, L., Stern, S. R., Lawson, H. M., & Azenkot, S. (2022). Tactile materials in practice: Understanding the experiences of teachers of the visually impaired. *ACM Transactions on Accessible Computing*, 15(3), 1-34. <https://doi.org/10.1145/3508364>
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Tanaka, K., & Ochiai, Y. (2023). From text to touch: Methods and case studies for designing tactile graphics via digital fabrication and AI. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 1–13). ACM.

- Valente, D., Chennaz, L., Archambault, D., Négrerie, S., Blain, S., Galiano, A. R., & Gentaz, E. (2024). Comprehension of a multimodal book by children with visual impairments. *British Journal of Visual Impairment*, 42(1), 276–286.
- Wu, C. F., Wu, H. P., Tu, Y. H., Yeh, I. T., & Chang, C. T. (2022). A study on the design procedure of three-dimensional printable tactile graphics for individuals with visual impairments. *Journal of Visual Impairment & Blindness*, 116(4), 507–516.
<https://doi.org/10.1177/0145482X221122754>
- Wu, H. P. (2023). Improving learning of figurative concepts in individuals with blindness: Adopting teaching strategies to enhance learning motivation. *Journal of Visual Impairment & Blindness*, 117(4), 292–302.
<https://doi.org/10.1177/0145482X231193021>