

# Teaching Tiktaalik:

## Comparing Interactive and Narrated Formats for a Museum Exhibit

Nathan Lachenmyer, Sadiya Akasha, Jia Meng, and Sray Martin

Sitara Systems

Findings from February 2025 · Report updated 2025 March 27

### Abstract

We compared two ways of teaching museum visitors about *Tiktaalik*, an early fossil animal that sits close to the origin of life on land: a short narrated video and a self-guided interactive prototype. Both carried the same information. Ten adults each experienced one of the two formats and talked us through what they learned. Going in, we expected the interactive version to teach more, on the assumption that active exploration beats passive watching. The opposite happened. Video viewers came away with a firmer grasp of why *Tiktaalik* matters, while prototype users learned the animal's anatomy in finer detail but more often missed the larger point. The two formats did not differ in quality so much as in what they made easy to understand. This report documents the study and recommends a hybrid design that keeps the video's clear storyline while adding the prototype's moments of hands-on detail. These are patterns from a small, exploratory study, not statistical claims.

## 1 Background and Purpose

Museums increasingly ask visitors to touch, click, and explore rather than simply watch. This reflects a well-founded tradition: constructivist accounts of museum learning hold that visitors build understanding through active engagement [1, 2], and interactive exhibits demonstrably capture and hold attention [3]. The open question is not whether interactivity engages — it does — but whether that engagement reliably converts into conceptual understanding, or whether interactive features can leave a visitor busy without deepening what they grasp [4]. This study puts that narrower question to a modest test.

The exhibit was developed for the Yale Peabody Museum of Natural History, an institution with a long research tradition in vertebrate paleontology and in the fish-to-land transition in particular [5]; the museum displays a cast of the *Tiktaalik* fossil. *Tiktaalik* is a roughly 375-million-year-old animal whose fossils show a creature part fish and part land animal [6]. It belongs to the lobe-finned lineage long studied as the bridge between fishes and the first four-limbed vertebrates [5], and it is often described as a *transitional species*: one of the stages that mark a major transition in vertebrate history, the move from water onto land [7]. The finished exhibit — including the touch-

screen kiosk evaluated here — is installed in the museum’s “At Water’s Edge” gallery (Figure 1). Reconstructing that history — and, in particular, how such an animal actually moved — rests on inference from the fossil record: the evidence-based phylogenetic reasoning that underpins modern vertebrate paleontology [8], and, increasingly, the biomechanical analysis of what an extinct animal’s joints and limbs could and could not do [9]. That story of movement, from swimming to walking, is the “big picture” we hoped visitors would take away.



**Figure 1.** The touchscreen kiosk (foreground) installed in the Yale Peabody Museum’s “At Water’s Edge” gallery, beside the tetrapod skeleton and fossils it interprets. The evaluation reported here tested a remote prototype of this kiosk’s content; the installed exhibit is shown for context.

We built two versions of the same lesson. One is a short narrated video animation. The other is an interactive prototype that lets the visitor click through the same material at their own pace. Our working hypothesis was that the interactive prototype would produce better learning, since it asks the visitor to do something rather than sit back. As the findings below show, the video proved the stronger teacher of the central idea, and the prototype the stronger teacher of detail. The interesting result is not that one format “won,” but that each shaped a different kind of understanding.

## 2 Method

This was a *formative, comparative* evaluation: its purpose was to inform the design of the exhibit by comparing two prototypes, not to deliver a summative verdict or a statistically powered test. The methods below reflect that aim, and follow a growing move to adapt user-experience evaluation methods from the technology sector to the assessment of museum exhibits [10].

### 2.1 Participants

We recruited ten adults through UserInterviews.com, screening for a mix that resembles a general museum audience: a range of ages, household types, and income levels, all of whom reported visiting museums at least occasionally. Table 1 summarizes the panel. Five participants were assigned to the video and five to the prototype.

| ID  | Age | Parent? | Museum Visitor? | Educational Media         |
|-----|-----|---------|-----------------|---------------------------|
| P1  | 29  | ✗       | ✓               | YouTube, Reddit, AI tools |
| P2  | 31  | ✗       | ✓               | Red Note, YouTube         |
| P3  | 32  | ✗       | ✓               | Reddit, Instagram         |
| P4  | 37  | ✓       | ✓               | Various                   |
| P5  | 38  | ✓       | ✓               | TikTok                    |
| P6  | 39  | ✓       | ✓               | Various                   |
| P7  | 43  | ✓       | ✓               | None                      |
| P8  | 43  | ✓       | ✓               | YouTube, PBS Kids         |
| P9  | 50  | ✓       | ✓               | YouTube                   |
| P10 | 61  | ✓       | ✓               | Instagram                 |

**Table 1.** Screening survey summary of the ten participants. Participant names and locations are withheld; each participant is identified only by a code (P1–P10).

### 2.2 Materials

The video version is a 1:12-minute narrated animation. The interactive version is a click-through prototype, built in Figma, that presents the same content but lets the visitor set their own pace and choose what to look at; representative screens from the prototype appear in Figure 3 (Appendix A). Links to both stimuli are given in Appendix A.

### 2.3 Procedure

We ran ten remote sessions over Zoom in February 2025, about 30 minutes each. After a brief introduction and consent to record, each participant either watched the video or explored the prototype while sharing their screen. We asked them to think aloud, and we followed up with questions about what they found engaging, what they understood, and what they would change. Because each person saw only one format, the comparison is between groups rather than within

a single viewer. The full interview guide is reproduced in Appendix B.

## 2.4 Measuring what people learned

To compare learning across formats, we scored each participant on a seven-level knowledge hierarchy — a technique we adopted directly from Deborah Perry’s *What Makes Learning Fun?* [11] — running from no interest in *Tiktaalik* up to grasping how fossils let us reconstruct how extinct animals moved (Table 2). The individual levels are specific to this exhibit’s content; the method of placing visitors on a content-specific knowledge hierarchy is Perry’s. For each participant we estimated a starting level from their opening remarks and an ending level from how they described *Tiktaalik* afterward; the difference is their *growth*, a change-from-baseline measure in the spirit of the contextual model of museum learning, in which learning is read as movement from each visitor’s own starting point rather than against a fixed standard [2, 12]. Each participant’s session was scored independently by at least two members of the research team, and differences were reconciled through discussion. Because the hierarchy is an ordinal scale — the step from one level to the next is not guaranteed to be an equal interval — we treat the group averages below as a heuristic summary and report the range of scores alongside them, not as precise measurements.

The technique also has deeper roots in educational psychology, in Gagné’s theory of *learning hierarchies*: the principle that grasping a higher-level concept depends on first mastering the lower-level concepts beneath it [13, 14]. As Section 3.3 shows, our data bear the marks of exactly such a prerequisite structure.

| Level | What the visitor understands                                                                 |
|-------|----------------------------------------------------------------------------------------------|
| 0     | I don’t know anything about <i>Tiktaalik</i> and I don’t care.                               |
| 1     | I don’t know anything about <i>Tiktaalik</i> , but I am curious and would like to know more. |
| 2     | I understand that <i>Tiktaalik</i> existed.                                                  |
| 3     | I understand how <i>Tiktaalik</i> moved from water to land.                                  |
| 4     | I understand which anatomical features let <i>Tiktaalik</i> move that way.                   |
| 5     | I understand that <i>Tiktaalik</i> represents a turning point in how vertebrates moved.      |
| 6     | I understand that the fossil record lets us reconstruct how ancient creatures moved.         |

**Table 2.** The seven-level knowledge hierarchy, a technique adopted from Perry [11], used to score understanding before and after each session.

## 3 Findings

### 3.1 How much people learned

Both formats taught something: every participant ended higher on the hierarchy than they began, and each described learning something new. But the video group gained more ground. Video viewers grew by an average of 4.2 levels (range 3–5) and prototype users by 3.0 (range 2–4); Figure 2 and Table 3 show the individual trajectories. Several video viewers reached the top of

the hierarchy, articulating not just that *Tiktaalik* existed but why it matters — one described it as “a link between past and present.” Prototype users made real gains too, but tended to stop a step or two lower, often rich in anatomical detail: one came away describing the animal as “amphibious” without connecting that to the broader evolutionary story.

| ID  | Format    | Start | End | Growth | What they took away                                                                              |
|-----|-----------|-------|-----|--------|--------------------------------------------------------------------------------------------------|
| P8  | Video     | 1     | 6   | 5      | Grasped <i>Tiktaalik</i> 's role in evolution; valued seeing the fossil become a living creature |
| P4  | Video     | 1     | 6   | 5      | Recognized the transitional nature; understood how it moved between environments                 |
| P10 | Video     | 1     | 5   | 4      | Called it “a link between past and present”; understood the movement and its significance        |
| P5  | Video     | 1     | 4   | 3      | Struck by its ability to live in both worlds; focused on adaptations and bone structure          |
| P9  | Video     | 1     | 5   | 4      | Called it an “intermediary in the evolutionary process”; noted land-animal-like adaptations      |
| P1  | Prototype | 1     | 5   | 4      | Saw it as the “first to really migrate from water to land,” but did not expand on why            |
| P3  | Prototype | 1     | 4   | 3      | Focused on anatomy such as the legs; called it “amphibious”; missed the deeper context           |
| P2  | Prototype | 1     | 4   | 3      | Understood how anatomy shapes movement and breathing; stayed on mechanical detail                |
| P6  | Prototype | 1     | 4   | 3      | Described its land-and-water ability and flat-headed form; limited evolutionary context          |
| P7  | Prototype | 0     | 2   | 2      | Recalled basic anatomical features but could not articulate the significance                     |

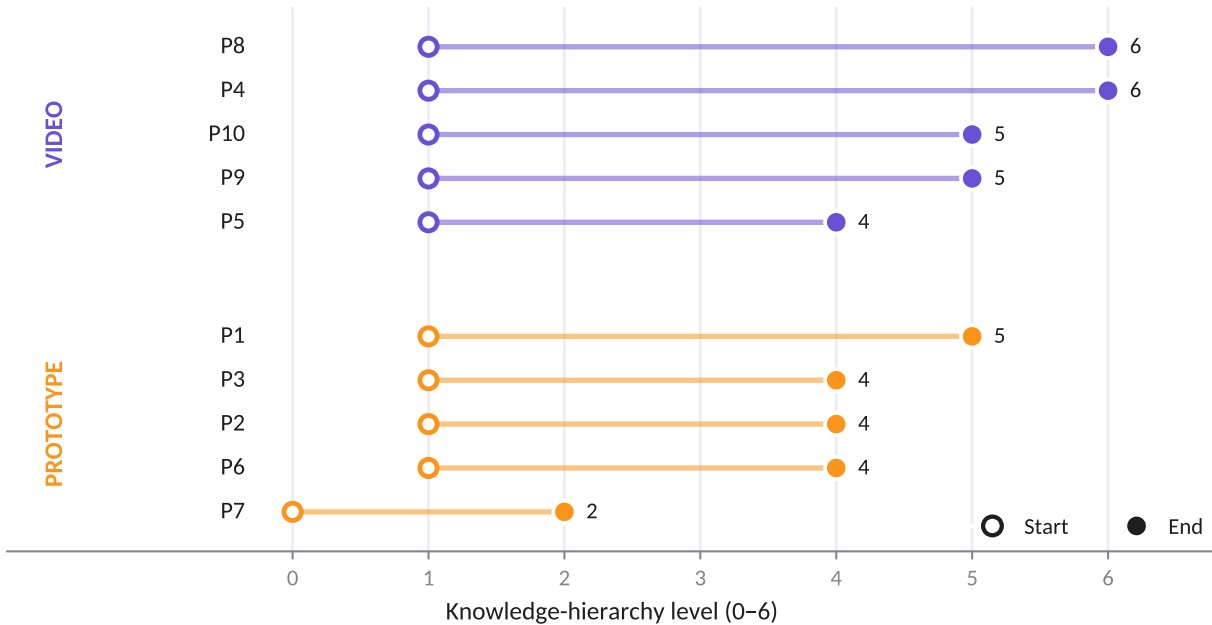
**Table 3.** Estimated understanding before and after each session. Average growth: video 4.2 levels (range 3–5), prototype 3.0 (range 2–4).

### 3.2 What each format did well

The two formats had distinct strengths, and they line up with how each presents the material.

The video was better at conveying movement and the overall story. All five video viewers named the animal's movement, its transition between water and land, as a key takeaway, compared with two of five prototype users. Four of five video viewers also drew a clear link between the fossil and the living animal, whereas several prototype users were unsure how the two related. Viewers credited the combination of narration, imagery, and motion for making the idea land, though some wished for more on-screen text alongside the audio.

The prototype was better at detail and control. Three of five prototype users valued being able to explore at their own pace, and at least two video viewers wished aloud for the same control over playback. Four of five prototype users named specific anatomical features in what they took away, a finer-grained vocabulary than most video viewers used, although one feature, the spiracle (a breathing opening on the head), was a recurring point of confusion. In short, the prototype taught



**Figure 2.** Each participant’s estimated understanding before (hollow marker) and after (filled marker) the session, on the seven-level knowledge hierarchy. Video viewers (purple) grew more on average and reached higher final levels (group end mean 5.2) than prototype users (orange; group end mean 3.8).

the parts well; the video taught the point well.

### 3.3 A pattern in how understanding formed

The scores suggest more than a difference in amount; they hint at a difference in route. Video viewers often appeared to “skip” the anatomical level of the hierarchy (level 4) yet still arrived at the broader significance above it. Prototype users who did not first grasp how *Tiktaalik* moved from water to land (level 3) tended to stall at that anatomical level, accumulating detail without connecting it to the larger story.

This is close to what Gagné’s theory of learning hierarchies would predict [13]: grasping *Tiktaalik*’s significance (level 5) appears to depend on first understanding the water-to-land move (level 3), so prototype users who missed that prerequisite could not reach the higher concept, however much anatomical detail they accumulated. The video route is the more interesting wrinkle — viewers often reached level 5 while apparently skipping level 4 — which suggests that anatomical detail deepens the story but is not itself a strict prerequisite for grasping it. The practical reading is that the hierarchy is prerequisite-structured around the ideas that carry the “big picture,” and that a format which secures those load-bearing levels will teach the significance even if it lingers less on the parts.

Read together, the linear video seems to carry the “big picture” more reliably, while the interactive prototype conveys specific features well but does not always tie those features back to why they matter. This is the study’s central finding, and it runs against our starting assumption that hands-

on interaction would teach more. Interaction taught *differently*, not more.

### 3.4 Accessibility and usability

Each format also raised its own practical concerns. With the video, some older viewers found the pace too fast to absorb the information, several wanted subtitles, and several wanted playback controls to pause or rewatch; a few also wanted to see the creature from more than one angle. With the prototype, every user reached the end unaided, but some mistook highlighted labels for clickable hotspots, some expected to rotate the 3D model and could not, and several wanted a clearer link between the fossil and the animated creature. Notably, the two formats' shortcomings are close to complements: what frustrated video viewers (no control, no captions) is what the prototype offers, and what confused prototype users (a weak fossil-to-creature link) is what the video conveys well.

## 4 How This Compares to Prior Work

The instinct to make exhibits interactive is well grounded, and it is worth saying so plainly before we complicate it. Constructivist theories of museum learning hold that visitors build understanding by doing, not only by watching [1, 2], and interactivity has a strong empirical claim on visitors' attention: Sandifer found that technological novelty and open-endedness are among the best predictors of how long visitors stay with an exhibit [3]. Our own prototype bears this out — every user reached the end unaided, and prototype users left with a richer vocabulary for *Tiktaalik*'s anatomy than video viewers did. This study is therefore not a verdict against interactivity, and the widespread conviction among exhibit designers that interaction supports learning is not misplaced.

What our result qualifies is the next step — the assumption that engagement converts automatically into conceptual understanding. Here the research is more cautious than the design instinct. Allen framed the tension as the gap between exhibits that entertain and exhibits that teach, and argued for designs that are immediately apprehendable rather than merely absorbing [4]. In HCI, Hornecker and Stifter found that visitors at interactive installations often engage *operationally* — working out what to click or manipulate — rather than reflecting on the content [15], and recent work continues to probe this interaction–engagement–learning relationship directly [16]. That is a close description of our prototype users, who gathered detail but frequently stalled before its evolutionary meaning: engagement was necessary but, on its own, not sufficient for the “big picture.”

Why, then, did the passive-seeming video carry the larger idea? Multimedia-learning theory predicts it. Mayer's modality and coherence principles hold that pairing spoken narration with synchronized visuals helps a viewer build a single coherent mental model, while added interactive control can introduce extraneous cognitive load that competes with understanding [17, 18]. This also points to what may be the true active ingredient. Our video did not differ from the prototype only in being linear; it was also narrated, and Mayer's principles are specifically about spoken narration paired with visuals. The video's advantage may therefore owe as much to having a guid-

ing voice as to its linearity — a distinction our two-condition design cannot separate, and one a follow-up should test by holding narration constant. If narration is indeed the driver, that is good news for the hybrid, because a guiding voice can be carried into an interactive experience.

Read against this backdrop, our finding is neither an anomaly nor a rebuke of interactive design; it is a boundary condition. Interactivity excels at engagement and at conveying detail; a well-made, narrated linear sequence excels at synthesis. The honest reading of the literature is not that interactivity fails to teach — it plainly teaches some things well — but that the format best suited to a goal depends on the goal. That is why we read the two as complementary, and it is precisely what the hybrid design below sets out to exploit.

## 5 Recommendations

The results point toward a single conclusion: neither format is complete on its own, and the strengths of one answer the weaknesses of the other. We recommend a hybrid experience that keeps the video’s clear narrative spine and adds the prototype’s moments of hands-on detail.

### 5.1 Primary recommendation: a guided experience with interactive pauses

Build a narrated, video-led experience that pauses at a few points to let the visitor explore. At each pause, the visitor could rotate the model, highlight a feature, or read a short note, then return to the storyline. This preserves the narrative that carried the “big picture” for video viewers while giving the control and detail that prototype users valued. Throughout, keep the anatomy tied explicitly to the story: state plainly that *Tiktaalik* is a transitional species, and connect each feature to the movement it made possible, so detail never floats free of its significance.

### 5.2 Strengthen the fossil-to-creature link

Both groups wanted the relationship between the fossil and the living animal made clearer. Consider an overlay that shows the fossil and the reconstructed creature in the same frame, and an intermediate skeleton layer that makes visible how paleontologists infer movement from bones. Done well, this also teaches the highest level of the hierarchy: that the fossil record is what lets us reconstruct how extinct animals moved.

### 5.3 Design for a range of visitors

Support different paces and needs directly: offer speed controls, provide captions or subtitles as a text alternative to narration, and allow both guided and self-directed paths through the same content. These measures address the specific accessibility gaps each format showed and widen the audience the exhibit can reach.

## 5.4 Format-specific refinements

If the video and prototype continue as separate pieces in the near term, two smaller sets of changes apply. The **video** should gain basic playback controls, richer visuals (more angles, overlays), and captions with speed control. The **prototype** should make the fossil-to-creature connection explicit and let users rotate both the fossil and the animated model.

## 6 Limitations

This was a small, exploratory study, and its results should be read that way. Ten participants, five per format, is enough to reveal patterns worth acting on — a sample well suited to formative, qualitative usability work, where a handful of users surfaces most of what a design needs to fix [19, 20] — but not enough to measure the size of those patterns with confidence; the reported averages summarize this group rather than predict a population. Each format was represented by a single stimulus (one video, one prototype), and the two differed in more than interactivity: the video was also narrated, shorter, and more heavily produced. As noted above, guided narration is therefore a plausible driver of its advantage, not linearity alone, so a cleaner test would hold narration constant across formats. The sessions were also remote and one-on-one, with participants watching on a shared screen; the real kiosk is encountered standing, amid ambient noise and social groups, often mid-visit (Figure 1). Because the contextual model of museum learning locates much of what visitors take away in exactly this physical and social setting [2], remote testing strips out variables the model treats as central, and in-gallery behavior may differ. Finally, understanding was inferred from what people said aloud rather than from a formal test — well suited to surfacing how people think, less so to precise measurement. We therefore read these findings as strong directional guidance for design, and as hypotheses worth confirming with a larger, in-context study.

## Acknowledgments

This work was carried out for the Yale Peabody Museum of Natural History. We thank our scientific content advisors — Jacques Gauthier, Armita Manafzadeh (then a postdoctoral researcher at Yale), and Bhart-Anjan (Anjan) Bhullar — whose expertise in vertebrate paleontology, the biomechanics of locomotion, and major evolutionary transitions shaped the material that visitors encountered in both formats. We are also grateful to our primary contacts at the museum — Kailen Rogers, Assistant Director of Exhibitions; Susan Butts, Director of Collections & Research; and Collin Moret — who facilitated the project throughout. The user testing, analysis, and any errors in this report are our own.

## References

- [1] George E. Hein. *Learning in the Museum*. Routledge, London, 1998.
- [2] John H. Falk and Lynn D. Dierking. *Learning from Museums: Visitor Experiences and the*

*Making of Meaning*. AltaMira Press, Walnut Creek, CA, 2000.

- [3] Cody Sandifer. Technological novelty and open-endedness: Two characteristics of interactive exhibits that contribute to the holding of visitor attention in a science museum. *Journal of Research in Science Teaching*, 40(2):121–137, 2003.
- [4] Sue Allen. Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education*, 88(S1):S17–S33, 2004.
- [5] Keith S. Thomson. The biology of the lobe-finned fishes. *Biological Reviews*, 44(1):91–154, 1969.
- [6] Edward B. Daeschler, Neil H. Shubin, and Farish A. Jenkins. A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. *Nature*, 440:757–763, 2006.
- [7] Bhart-Anjan S. Bhullar, Zachary S. Morris, Elizabeth M. Sefton, Atalay Tok, Masayoshi Tokita, Bumjin Namkoong, Jasmin Camacho, David A. Burnham, and Arhat Abzhanov. A molecular mechanism for the origin of a key evolutionary innovation, the bird beak and palate, revealed by an integrative approach to major transitions in vertebrate history. *Evolution*, 69(7):1665–1677, 2015.
- [8] Jacques Gauthier, Arnold G. Kluge, and Timothy Rowe. Amniote phylogeny and the importance of fossils. *Cladistics*, 4(2):105–209, 1988.
- [9] Armita R. Manafzadeh, Robert E. Kambic, and Stephen M. Gatesy. A new role for joint mobility in reconstructing vertebrate locomotor evolution. *Proceedings of the National Academy of Sciences*, 118(7):e2023513118, 2021.
- [10] Ellie King, M. Paul Smith, Paul F. Wilson, Janet F. Stott, and Mark A. Williams. Evaluating museum exhibits: Quantifying visitor experience and museum impact with user experience methodologies. *Curator: The Museum Journal*, 68(1):163–200, 2024.
- [11] Deborah L. Perry. *What Makes Learning Fun? Principles for the Design of Intrinsically Motivating Museum Exhibits*. AltaMira Press, Lanham, MD, 2012.
- [12] John H. Falk and Martin Storksdieck. Using the Contextual Model of Learning to understand visitor learning from a science center exhibition. *Science Education*, 89(5):744–778, 2005.
- [13] Robert M. Gagné. Learning hierarchies. *Educational Psychologist*, 6(1):1–9, 1968.
- [14] Robert M. Gagné. *The Conditions of Learning*. Holt, Rinehart and Winston, New York, 1965.
- [15] Eva Hornecker and Matthias Stifter. Learning from interactive museum installations about interaction design for public settings. In *Proceedings of the 18th Australia Conference on Computer-Human Interaction (OZCHI '06)*, pages 135–142. ACM, 2006.
- [16] Adina Friedman. The relationship between interaction paradigm, engagement, and learning. In *Proceedings of the Nineteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '25)*, pages 1–4. ACM, 2025.

- [17] Richard E. Mayer. *Multimedia Learning*. Cambridge University Press, New York, 2nd edition, 2009.
- [18] Roxana Moreno and Richard E. Mayer. Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91(2):358–368, 1999.
- [19] Jakob Nielsen. Why you only need to test with 5 users. Nielsen Norman Group, 2000. <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>.
- [20] Jakob Nielsen and Thomas K. Landauer. A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, pages 206–213. ACM, 1993.

## Appendix A: The Interactive Prototype

Figure 3 shows representative screens from the interactive prototype (the “kiosk”). Visitors moved through a fixed set of sections — moving around, breathing, and looking landward — each pairing a short caption with a rotatable 3D fossil model whose features were labelled as tappable highlights. The spiracle labels visible in panels (d) and (e) correspond to the recurring point of visitor confusion noted among the prototype’s usability findings in Section 3.

**Stimuli.** Video animation (1:12 min): [https://drive.google.com/file/d/1FyCm4ApE2RW6hWxz\\_nkt8OnWzZgpKqfz/view?usp=drive\\_link](https://drive.google.com/file/d/1FyCm4ApE2RW6hWxz_nkt8OnWzZgpKqfz/view?usp=drive_link)

Interactive prototype (Figma): <https://www.figma.com/proto/Q8VzBxmSkojynd1bL5mwPY/AV103.9-Locomotion?page-id=1270%3A39&node-id=2990-958&viewport=316%2C359%2C0.04&t=q1UHTezFASOuWhsC-1&scaling=contain&content-scaling=fixed&starting-point-node-id=2990%3A958>

Per-participant scores appear in Table 3 and Figure 2; anonymized raw data are available from the authors on request.

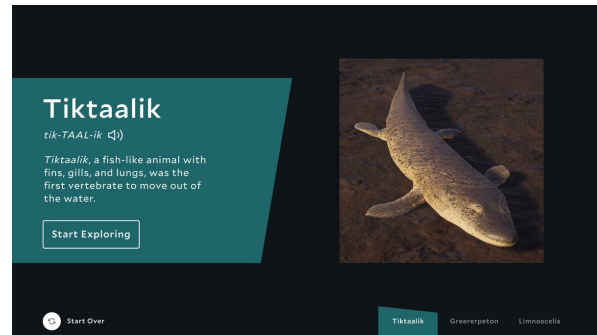
## Appendix B: Interview Guide

Reproduced from the study’s interviewer guide (updated 2025 January 21). Interviewer names have been replaced with “[Interviewer].” Bracketed notes marked “prototype” or “video” applied to that format only.

**Part 1 — Introduction.** [Introduce ourselves.] “Hi! I’m [Interviewer], and I work at Sitara Systems. This is my colleague [Interviewer]. What’s your name? Thanks for coming in today — we’re always trying to improve our experiences, and your honest feedback is an important part of that. Before we start, do I have your permission to record the session, for our own internal use? You’re free to take a break or leave at any time.” The participant is then told the session has three parts: a few background questions; a walk-through of “some things we’re working on” (think aloud, ask questions, don’t worry about getting stuck); and follow-up questions about what they saw (candid feedback is welcome; there are no right or wrong answers).



(a) Title screen



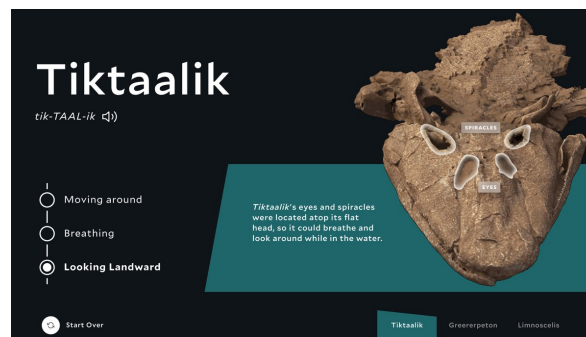
(b) Species introduction



(c) "Moving around": front limbs



(d) "Breathing": lungs and spiracles



(e) "Looking landward": eyes and spiracles

**Figure 3.** Representative screens from the interactive prototype (kiosk) tested against the linear video. Panels (c)–(e) show the labelled anatomical highlights — front limb, lungs, spiracles, eyes — that let visitors explore specific features at their own pace.

[For the prototype: “If you get stuck, I’ll try not to answer or tell you what to do — I want to see what you would do on your own — but I’ll help if you’re completely stuck. This is a prototype, so some links or buttons may not work; you can still click anywhere, and I’ll let you know when something isn’t working.”]

**Part 2 — Context questions.** Open-ended questions about the participant’s background and habits. [Demographic items — age, education, and whether they have children — were captured in the screening survey.]

- How often do you visit museums, and what kinds?
- Do you usually visit alone, or with friends or family?
- Do you use social apps like TikTok or Instagram for educational purposes (for yourself or your children)?

**Part 3 — Testing the experience.** “Do you mind sharing your screen? It helps to see how you interact with the content. Feel free to think out loud and ask questions as you go.” Follow-up prompts used during the session included:

- What is this? What is it for?
- What did you think of that? What happened there?
- What goes through your mind as you look at this? What would you do next, and why?
- Is there another way to do that? How would you change this to make it better for you?
- What additional information would have helped? Was that what you expected?

[For the video: “Do you want to watch it again? Feel free to use the playback bar to skip around.”]

**Part 4 — Debrief.** *Usability.* Were there technical or navigation difficulties (prototype only)? Did you find the experience intuitive — why or why not?

*Comprehension and retention.* Can you summarize what you learned? Did you learn something new — if so, what? What do you remember most about *Tiktaalik*? Could you summarize what it is and why it is important? How would you describe it to a friend or your family — or, if you were a teacher, to your students?

*Experience and engagement.* How did you feel while using the kiosk or video (engaged, bored, curious)? Pick three words to describe how it made you feel. Did anything stand out or surprise you? Was anything confusing or unclear?

*Format preferences.* What did you like or dislike? If you had three wishes to make it better, what would they be, and why? Is this a good format for conveying this kind of information — what other formats would you suggest? Would it be effective in a museum setting (and if not, under what circumstances would you use it)?