

Hands That Read: The Book Body as Interface in Interactive Picture Books (2010-2025).

A theoretical framework for material and hybrid meaning-making in post-digital children's literacy

Milan Mašát

Department of Czech Language and Literature, Faculty of Education,
Palacky University Olomouc, Olomouc, Czech Republic

Contact: milan.masat@upol.cz

ABSTRACT

This article offers a theoretical account of interactive picture books as epistemic objects (i.e., objects that invite questions and discovery rather than only deliver information), proposing the book body as interface that stages meaning through material and hybrid operations. Spanning 2010-2025, from engineered print to page-registered AR, we synthesize visual-culture, multimodality, and material engagement theory to show how flaps, windows, pull-tabs, volvelles, leporello/accordion folds, transparencies, and AR overlays implement a mechanic of discovery (cue → action → reveal → re-framing). We formalize three constructs – manipulation as evidence, attentional traction, and retroactive contextualization – explaining how gestures externalize inference, steer attention, and license reinterpretation. Methodologically, we specify micro-analytic interaction scripts and propose interaction maps (states × actions) to model exploration topologies (i.e., the paths a reader can take through states and actions) without empirical datasets. Our core contribution is a typology mapping mechanism → information operation → cognitive function, clarifying why multi-step, consequence-bearing interactions support spatial reasoning, causal rehearsal, classification, and language development, whereas surface busyness risks extraneous load. We theorize print-AR 'hybrid bridges': AR acts as a registered lens that models hidden processes and scales while preserving reference to the page. Accessibility and ethics are constitutive – redundant multimodal cues, inclusive imagery, privacy-by-design, and sustainable, page-complete pathways – summarized by three maxims: constrain to enable, cue before you ask, and augment, never gate. The framework offers design-to-cognition principles and a transferable vocabulary that subsequent studies can operationalize across developmental stages and cultural editions, aligning aesthetics, cognition, and ethics in post-digital picture book literacy.

KEYWORDS

Epistemic objects; interactive picture books; multimodal literacy.

CITATION

Mašát, Milan, "The Book Body as Interface in Interactive Picture Books (2010-2025). A theoretical framework for material and hybrid meaning-making in post-digital children's literacy", *JIB*, 5 (April 2026): 153-168. DOI: 10.57579/2026.8.

1. Opening the Interface: Framing Interactive Picture Books (2010-2025)*

The exploration of interactive picture books as epistemic objects in children's reading from 2010 to 2025 reveals a marked evolution in design and engagement strategies. Contemporary picture books, rich in meta fictive devices, invite children to co-construct narratives and move fluidly between the story world and real-world knowledge (Serafini & Tompkins 2015). Interactivity, as a theoretical construct, can be understood as a principle that deepens dialogic reading and narrative participation through purposeful, story-relevant affordances (Vanderschantz, Timpany & Wright 2020).

By 'hands that read' we mean that interpretation in interactive picture books is not only ocular and verbal but also haptic and procedural: the hand performs epistemic work. In movable and hybrid picture books, meaning is staged through sequences of touch – cue, action, reveal, and re-framing – so reading becomes a guided choreography of inquiry rather than passive reception. Across a continuum – from materially engineered print artifacts to augmented and fully digital formats – varied levels of interactivity promote distinct modes of language development and conceptual understanding (Batarello Kokić 2015). Theoretical accounts emphasize that highly interactive titles are intentionally designed for multiple readings and are sensitive to developmental and cognitive dimensions of reading, foreground interpretive co-presence rather than physical co-reading (Jiao 2023). The rise of hybrid, multimodal 'book-objects', including tactile and kinetic experiments in the Brazilian context, exemplifies how physical interaction may extend narrative coherence and activate reflective engagement (dos Santos Feres & Mattos 2023).

Framed as the book body as interface, this article conceptualizes the physical and hybrid constitution of picture books as a locus where meaning-making is enacted rather than merely represented. It examines:

- 1) how materiality and kinetics contribute to narrative and conceptual structures;
- 2) which design affordances invite exploration, inference, and cognitive recall;
- 3) how interactive frameworks conceptually mediate the reader's understanding of time, scale, and causality.

For readers beyond literary theory (makers, educators, curators, and conservators), the article provides (1) a transferable vocabulary for describing how mechanisms 'do meaning', (2) a typology linking mechanism to information operation and cognitive function, and (3) practice-oriented maxims for design and evaluation. The aim is to make the interactive picture book legible as a knowledge interface rather than a novelty object.

Methodologically, the analysis proceeds through close theoretical and semiotic examination of interaction scripts – flaps, windows, volvelles, pull-tabs, layering, and embedded digital interfaces (QR/AR) – as textual and material signs that generate interpretive depth. This yields a typology of interactive features by cognitive function and introduces a conceptual vocabulary – 'manipulation as evidence', 'attentional traction', 'retroactive contextualization' – to describe their role in constructing meaning. The discussion also considers conceptual limits such as cognitive load, ecological sustainability of augmented materials, and copyright, ultimately proposing a model of interactivity that aligns aesthetic design with a theory of sustainable, post-digital literacy.

* A generative AI language model (used as an assistive tool, e.g., ChatGPT) was employed only for stylistic polishing, language editing, terminology harmonization, outlining section headings and a typology table "skeleton", and drafting alternative phrasing for the abstract and conclusion for Author review. AI was not used to generate empirical data, invent or select citations, make interpretive claims, or create figures. All figures and tables are author-generated. No non-public or sensitive data were entered into AI tools. All websites cited in this article were last accessed on 12 March 2026.

2. Theoretical Framework

2.1. Visual Culture and Multimodality

This paper adopts a visual-culture and multimodality lens in which meaning is co-produced across verbal, visual, spatial, and tactile modes. Multimodality, as articulated by Prinsloo, frames meaning making as the orchestration of semiotic resources within specific social and material contexts (Prinsloo 2024). We therefore treat layout, scale, salience, sequencing, and typography – together with page-turn timing – as constitutive of interpretation rather than ancillary to it. Because interactive picture books unfold in time as well as space, their rhetoric is inherently procedural: flaps, die-cuts, transparencies, and rotations regulate the order of appearance of evidence. Within visual-culture studies, such devices are theorized not as ornaments but as epistemic structuring: they choreograph attention, comparison, and hypothesis formation.

2.2. ‘Epistemic Objects’ and Situated Knowing

We conceptualize interactive picture books as epistemic objects – artefacts that generate questions and remain, in principle, open because new aspects emerge through engagement (Praet 2023). This aligns with situated and distributed views of cognition in which knowing is enacted in tasks and settings rather than housed exclusively ‘in the head’. *Material Engagement Theory (MET)* further suggests that material culture (tools, texts, interfaces) participates in cognition, shaping and extending cognitive routines (Overmann 2023). Under this theoretical stance, pop-up mechanisms, volvelles, pull-tabs, layering, and AR overlays are formal operations that externalize steps of inference: they stabilize comparisons, make counterfactuals manipulable, and anchor temporal reasoning in visible transformation. Literary precedents – e.g., discussions of embodied reading in highly visual poetic forms such as Anne Carson’s – support treating material and visual structures as scaffolds of knowledge acquisition (Van Praet 2023). More broadly, cognition is embedded in multimodal, material ecologies (Borkent 2023; Murphy 2023).

2.3. Materiality, Kinetics, Embodied Reading

Taking the book body as interface foregrounds how material features and kinetics mediate attention and memory. In theoretical terms, friction, resistance, weight, and translucency function as operators that modulate pacing and cognitive load. We name three analytical constructions to be used throughout the paper:

- 1) manipulation as evidence – bodily actions (lifting, rotating, sliding) are treated as epistemic moves that disclose or test propositions;
- 2) attentional traction – design affordances that grip and steer attention long enough for comparison and inference to occur;
- 3) retroactive contextualization – later openings (e.g., overlays) reframe earlier representations, expanding the hypothesis space and prompting re-interpretation.

AR layers are theorized as hybrid bridges that maintain reference to the printed substrate while affording zoom, animation, or alignment; their epistemic value depends on registration with the page and on congruence with print constraints.

2.4. Synthesis

P1: In interactive picture books, meaning making is procedurally organized by material and visual operations (Prinsloo 2024).

P2: As epistemic objects, these books’ structures situated inquiry by inviting open-ended manipulation (Van Praet 2023; Overmann 2023).

P3: Embodied interaction (materiality + kinetics) provides stable handles for attention, memory, and inference, which we model via manipulation as evidence, attentional traction, and retroactive contextualization (Borkent 2023).

These propositions ground the article's subsequent typology and terminology without relying on empirical datasets; the argument remains rigorously theoretical.

3. Corpus & Methods

This section does not present empirical sampling or results; it formulates a design blueprint for how a corpus and methods could be constituted to interrogate the book body as interface. The aim is to make the analytic pathway explicit – so that future studies can implement, adapt, or contest it – while remaining on a strictly theoretical plane.

3.1. Corpus (2010-2025): Rationale and Selection Principles

The corpus is conceptually specified as author books and international editions (2010-2025) that foreground material interactivity: pull-tabs, flaps and windows, volvelles/rotating wheels, layered transparencies, leporello/accordion folds, and hybrid titles with QR/AR augmentations. The rationale is twofold. First, 'toy-book' mechanics such as pull-tabs exemplify tactile interactivity that restructures attention and expectation in young readers (Silva & Martins 2018). Second, the semiotic resources of movable picture books – compositional layout, engineered kinetics, paratextual cues – are theorized to recruit engagement and scaffold narrative inference (Serafini & Moses 2023). Inclusion is guided by formal rather than topical criteria: an item belongs if its meaning-making relies on a non-trivial material or hybrid operation that sequences or conditions access to information. Exclusions are likewise formal: purely typographic variants without interactive affordances and purely screen-native apps without a print anchor are outside scope. Within the 2010–2025 horizon, the corpus is stratified by mechanism (e.g., pull-tab vs. volvelle) and by hybridity (print-only vs. print + AR) to enable comparative reasoning about how different affordances organize cognition. This conceptual corpus also anticipates cross-cultural comparability by incorporating international editions, acknowledging that paratext and manufacturing conventions modulate affordance legibility.

3.2. Methods: Micro-Analysis of “Interaction Scripts”

Methodologically, the article advances a micro-analytic program for describing how interactive script reading features are. An interaction script is defined as a procedural chain – cue → action → reveal → re-framing – that binds perception, gesture, and interpretation. For analytic tractability, the unit of analysis is the interaction turn (one completed mechanism operation and its immediate semiotic consequences). Each turn is annotated along four axes:

- 1) affordance cueing (salience, iconography, haptic contrast);
- 2) gesture specification (lift, slide, rotate, fold, scan);
- 3) information operation (addition, subtraction, substitution, alignment, animation);
- 4) cognitive function is mapped to theoretical constructs (manipulation as evidence, attentional traction, retroactive contextualization).

To make such scripts query able in principle, we adapt techniques from corpus linguistics as methodological metaphors: a structured, machine-readable representation (e.g., TEI-like markup) could be indexed and queried with concordance-style tools to surface recurring patterns across mechanisms and editions (Chua 2022). While the primary objects are material, the logic of tokenization, annotation, and pattern search remains applicable: instead of lexical tokens, we index mechanism tokens (e.g., VOLVELLE[SUBSTITUTION+ALIGNMENT]) and retrieve their co-occurrence with paratextual prompts or AR overlays. In parallel, AR-enabled items are

modelled as hybrid bridges whose digital behaviors (zoom, rotation, animation) are coded only insofar as they remain registered to the printed substrate; this preserves the analytic focus on the page-anchored interface.

Finally, we acknowledge a diachronic dimension: the rise of multimodal and hybrid picture books is itself a transformation in narrative form, where design and materiality become central carriers of meaning (Serafini 2024). A theoretically specified corpus – stratified by mechanism and hybridity – thus supports comparative claims about how different interactive operations structure attention, memory, and inference, without presupposing any dataset or measurement protocol. In sum, the proposed corpus and micro-analytic method remain agnostic about outcomes yet make operationally clear how future empirical research could demonstrate the epistemic roles of interactive features theorized in this article.

4. Analysis

4.1. Materiality & Kinetics as Meaning Carriers

We begin with the claim that movable picture books mean through movement. Pop-ups, leporello/accordion folds, flaps, windows, pull-tabs, volvelles, layered transparencies, and hybrid print-AR overlays are not ‘add-ons’; they are formal operators that organize perception and inference in time. Following the multimodal account sketched above, we treat each operator as implementing a mechanic of discovery: a patterned sequencing of cue → embodied action → reveal → re-framing. The pop-up – by compressing a three-dimensional scene into a latent state and releasing it at the page turn – converts a routine action into an epistemic event (emergence). The leporello reframes narrative linearity as spatial unfolding: when extended, it affords panoramic comparison and long-range alignment (e.g., before/after, micro/macro). Pull-tabs and wheels (volvelles) enact substitution, alignment, or controlled animation; by coupling a precise gesture to a visible transformation, they stabilize causal attributions (“if I slide/rotate, *then* × changes”). In all such cases the book’s body carries meaning because of material resistance, friction, hinge strength, and opacity pace disclosure and modulate salience.

Two theoretical consequences follow. First, kinetic rhetoric: the ordering reveals functions like syntax for attention. Sequential disclosure regulates hypothesis space – what can be inferred now versus later – and supports retroactive contextualization, the re-reading of earlier elements considering new openings. Second, embodied indexicality: gestures index referents; the place where the hand travels become a deictic anchor (“this window corresponds to that interior”). These properties clarify why movable features are consistently associated, in literature, with heightened engagement and exploratory behavior (Serafini & Moses 2023; Silva & Martins 2018). Yet our point is theoretical: engagement is not merely motivational but structural. Because the artefact requires action to disclose information, it externalizes steps of reasoning that purely verbal texts leave internal.

To formalize the analysis, we model each mechanism as a triad Mechanism × Information Operation × Cognitive Function. For instance:

- 1) Pop-up × Addition/Depth Articulation × Spatial inference (emergent 3D structure supports occlusion reasoning);
- 2) Leporello × Accumulation/Alignment × Temporal and scalar mapping (simultaneous access to distant panels enables comparison across time/scale);
- 3) Pull-tab × Substitution/Causation × Counterfactual rehearsal (repeated sliding rehearses “what changes/what stays” under controlled variation);
- 4) Volvele × Rotation/Permutation × Category formation (combinatorics make part–whole and feature–instance relations manipulable);

5) Transparency overlay × Revision × Retroactive contextualization (overlay rewrites earlier scene while keeping the prior state visible for comparison);

6) Print-AR bridge × Registration/Animation × Mechanistic reasoning (animation tied to print coordinates maintains reference while modelling hidden processes).

These pairings are not empirical claims; they are analytic commitments that allow us to speak precisely about how materiality and kinetics carry meaning beyond illustration.

4.2. Design Affordances for Exploration & Inference

If mechanisms are operators, affordances are how operators become legible and usable. We define three families of affordances that recruit exploratory cognition:

1) cueing affordances: visual and haptic signposts (die-cut edges, iconography, texture contrast, micro-typographic pointers) that announce ‘where’ and ‘how’ to act without verbal instruction. Good cueing yields attentional traction – a grip on attention long enough for comparison and inference to happen;

2) constraint affordances: engineered limits (stop-points, discrete positions, one-way folds) that quantize action and protect cognitive load. Constraints make meaningful states finitely enumerable; a wheel that clicks into three positions teaches through contrastive alignment instead of open-ended twiddling;

3) bridge affordances: features that maintain continuity across layers – most saliently in print-AR pairings. Proper registration ensures that digital behaviors (zoom, rotation, animation) index back to the page rather than displacing it; the page remains the primary interface, the device a lens.

Within this landscape we adopt the term *explanate* for designed invitations to exploratory manipulation that yield novel patterns of interpretation (Somanath, Sharlin & Sousa 2014; Kiliç 2021). Explanates are not simply ‘play’; they are structured opportunities for abductive reasoning – trying out a gesture to see which hypothesis about the scene becomes most coherent. On our account, a book furnishes stronger explanates when (a) cueing is legible, (b) constraints discretize states, and (c) bridges preserve reference across layers. Such design supports the micro-analytic constructs introduced earlier: manipulation as evidence (actions produce inspectable traces), attentional traction (attention is held by tactility + visual payoff), and retroactive contextualization (new states illuminate prior ones).

A theoretical corollary concern maps of interaction. We propose representing a title’s design as a graph where nodes are states (panels, overlays, AR scenes) and edges are permitted actions (lift, slide, rotate, scan). Edge labels carry information operations (addition, substitution, alignment). Graph structure then predicts cognitive affordances: high out-degree zones invite exploration; strongly connected components foster looped rehearsal; bridges between components trigger global re-framing. This formalism gives shape to what the literature calls engagement (Serafini & Moses 2023; Silva & Martins 2018) and clarifies design choices that produce productive exploratory behavior (Somanath, Sharlin & Sousa 2014).

Crucially, explanates have developmental grain. For early readers, coarse, binary mechanisms (closed/open; before/after) stabilize basic contrasts. For older children, multi-state devices (three-position tabs; multi-ring wheels) make combinatorics and conditionality manipulable. Across ages, the guiding principle is the same: affordances should constrain to enable. Over-abundant branching without cueing degrades attentional traction; over-tight constraint removes space for hypothesis testing.

4.3. Pedagogical Scenarios (time, scale, causality)

The figure below (**Fig. 1**) schematically illustrates the basic mechanisms referenced in the scenarios below (lepurello, pull-tab, volvelle, flap/window). Although the article is theoretical, pedagogy enters here as scenario design: generic task patterns that align with the affordances above. The goal is not to prescribe classroom practice but to indicate how material interactions can mediate conceptual learning.

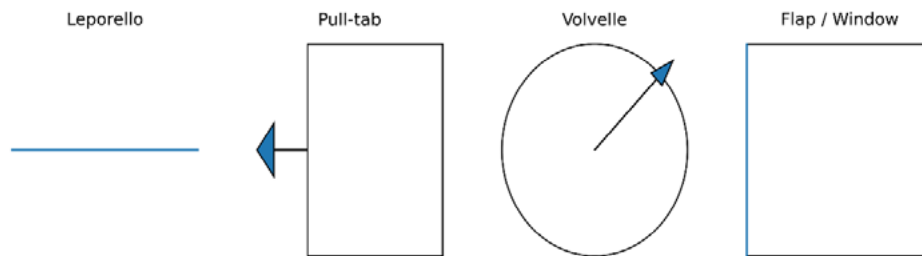


Fig. 1. || Schematic illustration of the basic interactive mechanisms referenced in the pedagogical scenarios: lepurello, pull-tab, volvelle, and flap/window.

A. Time (ordering, simultaneity, revision):

- *Scenario A1: Sequentialising events.* A flap-series reveals phases of a process (e.g., seed → sprout → plant). The interaction map enforces a left-to-right order; each reveal licenses retroactive contextualization (the sprout reframes the seed drawing). Theoretical outcome: children practice temporal indexing by binding gestures to event boundaries. Example mechanism: a three-flap sequence with a repeated icon cue (e.g., seed → sprout → plant), where each opening creates a stable, inspectable state that can be named and compared.
- *Scenario A2: Simultaneous timelines.* A lepurello displays parallel strands; folding/unfolding sets the window of simultaneity. Theoretical outcome: comparison across panels supports temporal alignment without cognitive overload because physical extent constrains the field of view. Example mechanism: a lepurello that can be folded to align non-adjacent panels, allowing readers to compare 'earlier' and 'later' states within a single visual field.

B. Scale (micro/macro, part/whole):

- *Scenario B1: Scalar zoom.* Transparency overlays add magnified detail at fixed location; AR, when registered to the page, affords dynamic zoom. Theoretical outcome: by preserving page reference, bridges foster scalar mapping – children co-index macro forms with microstructures rather than replacing one with the other. Example mechanism: a transparency overlay registered to a printed diagram, where micro-detail aligns precisely with a macro scene, preserving a shared reference frame.
- *Scenario B2: Volvelle classification.* Rotating windows permute features; discrete click-stops quantize categories. Theoretical outcome: combinatorial manipulation externalizes feature selection, supporting the formation and testing of categories under visible constraint. Example mechanism: a volvelle with discrete click-stops that cycles through category combinations, making classification reversible and explicit.

C. Causality (mechanism, dependency, counterfactual):

- *Scenario C1: Pull-tab mechanism model.* Sliding causes a visible change (e.g., cloud moves → rain appears). Theoretical outcome: a tight gesture–outcome coupling stabilizes causal attributions and invites counterfactual rehearsal (“what if I slide back?”). Example mechanism: a pull-tab with two hard stops that transforms a before/after state, enabling explicit cause/

effect comparison.

- *Scenario C2: Multi-step reveal.* A pop-up sequence requires two coordinated actions (open + lift) to expose a hidden linkage. Theoretical outcome: serial dependency makes causal chaining tangible: one action prepares conditions for another. Example mechanism: a two-step reveal (open + lift) that withholds a relation until coordinated action is completed.

Within these scenarios, language development is theorized not as a parallel track but as articulation of action. Because gestures leave inspectable states, children can anchor utterances in shared referents (“now the wheel shows...”, “after we open this...”). This is consistent with work framing movable books as harnessing semiotic resources to capture attention and scaffold narrative participation (Serafini & Moses 2023; Kiliç 2021) and with observations that tactile interactivity enhances the reading experience (Silva & Martins 2018). From a theoretical perspective, such tasks cultivate three families of skill:

- 1) inferential skills: abductive updates under sequential disclosure; managing uncertainty via retroactive contextualization;
- 2) spatial skills: alignment, rotation, and occlusion reasoning made manipulable by overlays, leporello, and wheels;
- 3) language skills: deictic precision, temporal/aspectual marking, and conditional constructions emerging from talk organized around visible state changes.

The construct of attentional traction clarifies why structured tasks work: cueing + payoff keep attention ‘parked’ long enough for comparison to complete. Conversely, construction also predicts failure modes. Over-rich AR layers that drift from print registration erode traction; ambiguous cueing invites aimless manipulation. The design remedy is to restore strong bridges and crisp constraints. This echoes the call for design elements that promote active discovery and learning in educational contexts (Kiliç 2021; Park & Tak 2023) while remaining compatible with a print-anchored interface. Finally, we theorize economies of attention: good interaction maps ration novelty across the book. Early spreads invest in teaching the code (clear cueing, generous payoff); mid-book spreads vary operation within constraint (same mechanism, new content); late spreads exploit retroactive payoff, prompting readers to revisit earlier states with newly learned operators. This economy aligns with the literature’s emphasis on “exploration that teaches” (Serafini & Moses 2023; Somanath, Sharlin & Sousa 2014) and grounds our terminology (manipulation as evidence, attentional traction, retroactive contextualization) in a system-level view of design.

Materiality and kinetics carry meaning by scripting discovery; design affordances convert mechanics into legible explanates; and theoretically framed pedagogical scenarios show how time, scale, and causality can be mediated through the book-body interface. The result is not a classroom recipe but an analytic vocabulary and set of form–function mappings that future empirical work can test, refine, or contest (Serafini & Moses 2023; Silva & Martins 2018; Somanath, Sharlin & Sousa 2014; Kiliç 2021; Park & Tak 2023).

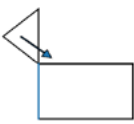

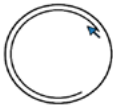



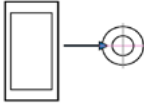
5. Typology of Interactive Features by Cognitive Function

We propose a typology that maps mechanism → cognitive function → learning potential, treating movable/AR features as cognitive artefacts whose form organizes inquiry. In line with research on interactivity, mechanisms that require multi-step, consequential manipulation tend to align with higher-order operations (analysis, synthesis, evaluation), whereas single-click or purely decorative actions risk staying at perceptual or recall levels (Ismail 2024). The typology is therefore *graded*: it values affordances that externalize reasoning (manipulation as evidence) yet cautions that complexity can overshoot and induce extraneous cognitive load if cueing and constraint are

weak (García Mora et al. 2022; Song et al. 2014). Following work on interactive exhibits, we also foreground interaction style – how users are guided to act – as a determinant of cognitive yield (Trybulec & Iłowiecka-Tańska 2023). Here ‘interactive exhibits’ refers to hands-on installations in children’s museums and science centers, ranging from table-top stations for individual or small-group manipulation to room-scale environments requiring bodily navigation.

5.1. Core Table (Mechanism → Function → Learning Potential)

Reading note. Entries indicate *dominant* functions; in practice, mechanisms stack (e.g., a pull-tab under a transparency yields rehearsal plus revision).

Mechanism (print/AR)	Cognitive function (dominant)	Learning potential (if well-cued & constrained)
Flap / Window 	Selective disclosure; contrastive comparison	Grounds attentional traction; supports observation → inference updates; cultivates evidential talk (“after opening, we now see...”)
Pull-tab (slide) 	Causal rehearsal; state transition	Stabilizes cause/effect claims by coupling gesture → outcome; enables counterfactual testing (“if I slide back...”)
Volvelle (wheel/permute) 	Classification; feature permutation; conditional reasoning	Externalizes category formation; rehearses combinatorics under discrete stops; strengthens metalinguistic vocabulary (“when feature × aligns...”)
Pop-up (depth emergence) 	3D spatial inference; occlusion reasoning	Makes hidden relations visible; supports part–whole mapping and perspective shift; anchors narrative salience to depth cues
Leporello / Accordion 	Alignment across distance; timeline/scale mapping	Enables simultaneous comparison (time, scale, parallel plots); fosters global coherence building
Transparency / Overlay 	Revision; retroactive contextualization	Visually ‘rewrites’ prior state while preserving reference; trains iterative interpretation and error correction
AR registered to page (animation/zoom/rotation) 	Mechanistic modeling; scalar zoom	Creates hybrid bridges that maintain page reference while revealing hidden process or micro-detail; supports analogy building without disembodiment from print
Haptic contrasts (texture/relief/weight)	Attentional gating; pacing	Tactile salience modulates rhythm and load; marks priority features and ‘affordance hotspots’
Collaborative pointer spaces (margins/tabs designed for co-reading)	Distributed reasoning; explanation-seeking talk	Channels dialogic moves (ask, predict, justify); approximates benefits seen in “collaborative tools” for higher-order synthesis (Ismail 2024)

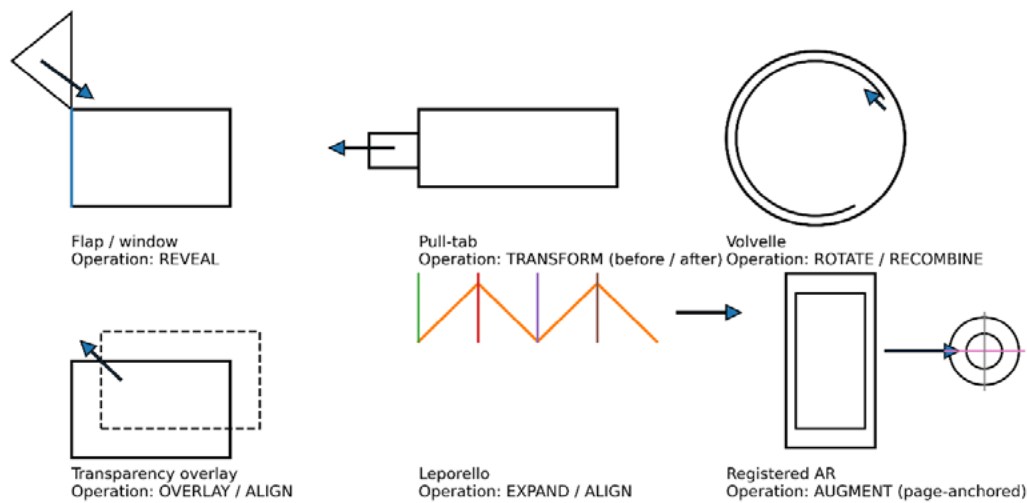


Fig. 2. || Icon legend for Table 1: mechanism and information operations

5.2. Key terms (concise definitions)

- **Manipulation as evidence** – bodily actions (lift/slide/rotate) produce inspectable states that can be cited in claims (“because sliding revealed Y..”).
- **Attentional traction** – sustained, directed attention achieved by clear cueing and a salient payoff, long enough for comparison/inference to complete.
- **Retroactive contextualization** – later states reframe earlier ones while preserving reference (overlays, staged reveals).
- **Constraint affordance** – engineered limits (click-stops, hinge range) that quantize states and reduce search space, lowering extraneous load.
- **Explanate** – a designed invitation to exploratory manipulation that yields novel interpretive patterns (cf. Somanath, Sharlin & Sousa 2014; Kiliç 2021).
- **Hybrid bridge** – print ↔ AR coupling where digital behavior remains registered to the page, preserving the page as the primary interface.
- **Registration fidelity** – precision of AR alignment with print; predictor of cognitive yield (mis-registration inflates extraneous load).
- **Interaction map** – a graph of states (nodes) and permitted actions (edges) that models a book’s exploratory topology.

5.3. Typological Principles (Design-to-Cognition Alignment)

- 1) Depth of operation over surface busyness. Multi-state, consequence-bearing mechanisms approximate the epistemic payoffs of simulations and collaborative environments (Ismail 2024).
- 2) Constrain to enable. Discrete states, stop-points, and one-way folds reduce extraneous load and make contrasts legible (García Mora et al. 2022; Song et al. 2014).
- 3) Cue before you ask. Iconic handles, die-cut edges, and haptic signals should precede text-based instruction; cueing is the precondition of attentional traction.
- 4) Preserve reference in hybridity. AR should index back to the page (hybrid bridges); free-floating overlays risk disembodiment and cognitive drift.

5) Design for interaction styles. Provide collaborative pointer spaces and teacher-legible ‘turns’ so that co-reading talk can appropriate mechanisms as cognitive artefacts (Trybulec & Hłowiecka-Tańska 2023).

A typology that weights mechanism, cognitive function, and learning potential clarifies why some interactions teach while others merely entertain. It also operationalizes a design stance: engineer affordances that externalize reasoning, ration novelty, and keep attention anchored to the page, thereby converting interactivity into durable understanding.

6. Accessibility & Ethics

In a print-anchored, hybrid ecology, accessibility and ethics are constitutive design constraints, not afterthoughts. We treat them as properties of the book body as interface across tactile print and page-registered AR.

Accessibility begins with redundant signaling: every operative element should be legible through at least two channels (visual + haptic or auditory). Raised edges, textured grips, and varied resistance create tactile contrasts that orient readers who rely on touch; clear arrowheads, notch cues, and consistent iconography stabilize navigation for those who benefit from strong wayfinding (Marwah, Thakur & Tanwar 2024; Sarangam 2022). In AR layers, the analogue is persistent on-page anchors and breadcrumb overlays that maintain reference to the print substrate. Reduce-motion options, flicker-safe animation, and user-controlled pacing extend this logic to neurodiverse readers.

Representation is an ethical affordance. Inclusive visual vocabularies – skin-tone ranges, gender expression diversity, disability visibility, multilingual labels – raise the likelihood that readers will recognize themselves and others in the narrative space. Crucially, inclusion must coincide with role equity (characters with agency, not token presence) and with alt-text/description layers in AR for non-visual access (Sarangam 2022).

AR can widen or narrow access. An equity-first hybridity principle follows:

- 1) the print object must remain complete in its core learning path; AR augments, it does not gate content;
- 2) all AR functions should degrade gracefully to low-tech tasks that achieve the same cognitive ends;
- 3) assets are device-light and offline-capable to mitigate bandwidth and hardware disparities (Regenbrecht, Zwanenburg & Langlotz 2022).

Where AR gathers interaction data, the minimal ethical baseline is data minimisation, transparent on-page notices, on-device processing by default, and no third-party tracking. Classroom modes should avoid personal identifiers and store nothing beyond the current session unless informed consent is obtained (Regenbrecht, Zwanenburg & Langlotz 2022). Because the primary interface is a children’s book, non-collection is the norm; exceptions demand explicit justification.

AR’s material and energy costs must be weighed against its pedagogical benefits. Two design propositions follow: ‘longevity’ (open standards, stable identifiers, printed fallback if servers retire) and ‘lightness’ (compressed assets, low-compute animations, bounded session times). Positively, AR can support environmental literacy by staging resource cycles or real-time eco-data overlays that link local action to global systems – if executed with modest footprints (Biswas & Ghosal 2025).

Accessibility and ethics cohere in three principles:

- 1) redundancy with reference – multimodal cues that always index back to the page;
- 2) augment, never gate – AR adds pathways without creating dependency or inequality;

3) minimize harm, maximize legibility – privacy-first pipelines, sustainable assets, and inclusive imagery as default settings.

Under these conditions, AR can extend the epistemic affordances of the book body – through haptic contrasts, clear navigation, and equitable representation – while respecting the ethical horizon of children’s reading (Marwah, Thakur & Tanwar 2024; Sarangam 2022; Regenbrecht, Zwanenburg & Langlotz 2022; Biswas & Ghosal 2025).

7. Discussion

In practical terms, the framework can be used as a checklist for designing, describing, and exhibiting interactive picture books so that actions reliably produce interpretable states. Treating the book body as interface clarifies how knowledge is mediated through form. On our account, movable mechanisms (flaps, tabs, volvelles, leporella, pop-ups) operate as formal operators that sequence cue → action → reveal → re-framing, externalizing steps of inference (manipulation as evidence) and stabilizing attention (attentional traction). AR, when registered to the page, functions as a hybrid bridge: it extends the same operators – zoom, rotation, animation – without disembodiment of the reader from the tactile substrate. Thus, print and AR are not rivals but co-actors in a single epistemic ecology: the page remains the primary reference frame; the device behaves as a controlled lens.

This framing also explains why the book endures as a vital educational technology: it evolves by incorporating multimodal affordances while retaining its material indexicality. In our terms, good designs ration novelty, quantize states (constraints), and preserve reference across layers (bridges), thereby enabling retroactive contextualisation without cognitive spillage. The synergy matters pedagogically: print gives haptic pacing and shared deixis (co-reading talk can point to this flap, that window), while AR can model hidden processes or scales that print alone cannot, provided registration fidelity is high.

Finally, the framework is compatible with emerging AI–AR collaborations: AI may act as an interaction-aware explainer that reads the same maps of states and edges we ascribe to books, proposing next inquiries or verbalizing contrasts, while AR visualizes them. None of this displaces critical reading; rather, it scaffolds how readers evaluate claims by tying language to visible, manipulable states – an alignment of engagement with analytic judgment.

7.1. Implications for Practice

Select and exhibit titles by form–function fit:

- 1) prefer mechanisms with discrete states and clear cueing (strong traction, low extraneous load);
- 2) program displays to teach the code early (demonstration spread), then vary within constraint;
- 3) in hybrid shows, on-page anchors and offline-capable assets, AR augments, never gates core content.

Design tasks that bind talk to action:

- 1) use the time-scale-causality scenarios: flap sequences for temporal indexing; leporello for alignment across distance; pull-tabs/volvelles for causal rehearsal and classification;
- 2) make interaction maps explicit (“What states?”, “What actions?”). Prompt because-talk anchored in states (“After we open X, what changed?”);
- 3) in AR, keep the page visible; ask students to co-index digital behaviors to printed loci.

Engineer exploratory interaction without overload:

- 1) cue before you ask (iconic handles, haptic grips); constrain to enable (click-stops, one-way folds);
- 2) build hybrid bridges with high registration fidelity; provide reduce-motion and caption/alt-text layers;
- 3) align with sustainable education: lightweight assets, durable identifiers, print-complete pathways.

Across roles, the maxim holds: externalize reasoning, preserve reference, ration novelty – so interactivity teaches, not just entertains.

Practical description template (one sentence + one action):

- (1) mechanism + operation: “A [mechanism] is used to [reveal/transform/align/rotate]...”;
- (2) reader action: “The reader is cued to [lift/slide/rotate]...”;
- (3) epistemic payoff: “This action makes [difference/state/relationship] inspectable for comparison or explanation”.

Example 1 (pull-tab, causality). A pull-tab is used to transform a scene from state A to state B (before/after), making a cause/effect relation visible. The reader is cued to slide the tab via a handle and arrow icon. This action makes the change inspectable for comparison and causal explanation, enabling the reader to cite the result as evidence.

Example 2 (volvelle, classification). A volvelle is used to rotate and recombine categories, producing multiple discrete combinations in a fixed window. The reader is cued to rotate the disc via a finger notch and circular arrow. This action makes category relations inspectable for classification and revision, enabling the reader to justify choices by pointing to a specific state.

Example 3 (transparency overlay, scale/alignment). A transparency overlay is used to align micro-detail with a macro scene by overlaying a registered layer on the base page. The reader is cued to layer the sheet by corner marks and an icon cue. This action makes scale relations inspectable for comparison, enabling the reader to explain how details map onto the larger system. For practical uptake, Figures 1–2 provide a minimal visual vocabulary (mechanism schematics and icon legend) that can be reused in design briefs, cataloguing, and exhibition labels.

7.2. Limitations & Future Research

This is a theoretical account: we specify constructs (traction, manipulation as evidence, retroactive contextualization), a typology (mechanism → function → potential), and design principles; we do not report empirical outcomes. Future work should (i) test the interaction-map model against usage traces and discourse in co-reading; (ii) examine developmental grain – how discrete vs. continuous mechanisms align with age-related capacities; (iii) quantify registration fidelity and its effect on load and inference in print-AR pairings; (iv) compare cross-cultural editions where manufacturing and paratext alter affordance legibility; and (v) evaluate AI-AR orchestration for transparency, privacy, and equity. Such studies would not replace the theory but probe its boundaries, refining how the book body interface mediates knowledge across time, scale, and causality.

8. Conclusion

Conceiving the book body as interface reframes interactive picture books as ‘epistemic objects’ that do not merely depict knowledge but *stage* it through material and hybrid operations. Across 2010-2025, movable and AR-augmented forms show how mechanics of discovery – sequenced cue → action → reveal → re-framing – externalize reasoning and pace attention. Our theoretical

mapping of mechanism \times information operation \times cognitive function, together with the construct's manipulation as evidence, attentional traction, and retroactive contextualization, explains why some interactions teach while others distract. The accompanying design maxims – constrain to enable, cue before you ask, and preserve reference in hybridity – articulate conditions under which interactivity supports inference, language, and spatial reasoning.

Print and AR, properly coupled, form a single epistemic ecology: the page remains the reference frame; AR acts as a registered lens that models hidden processes and scales without disembodiment of the reader. Accessibility and ethics are constitutive here: redundant cues, inclusive imagery, privacy-by-design, and sustainable, page-complete pathways ensure that augmentation adds without gating. While our account is purely theoretical, it offers a micro-analytic vocabulary and interaction maps that future empirical work can operationalize across developmental stages and cultural editions. Ultimately, post-digital picture books demonstrate how aesthetics, cognition, and ethics can be deliberately aligned: by asking hands to move and eyes to compare, they render inference visible and bind imagination to evidence.

In this sense, “hands that read” names a literacy of guided touch: actions are epistemic rather than decorative. The proposed vocabulary can therefore be applied not only in criticism but also in cataloguing, conservation notes, exhibition labels, and design briefs.

8.1. Glossary of key terms (for makers, educators, curators)

Attentional traction: sustained, directed attention achieved through clear cueing plus a meaningful payoff – long enough for comparison, inference, or language work.

Constraint: engineered limits (hard stops, discrete positions, one-way folds) that prevent ambiguous actions and make states comparable.

Cue: a perceptible invitation to act (handle, notch, arrow, icon, texture) that tells the reader what to do next.

Interaction map: a simple model of a book's exploration space as states \times actions (what can be done, and what changes when it is done).

Manipulation as evidence: a designed action (lift/slide/rotate) produces an inspectable state that readers can point to and use in explanation (“When we open this, we can see...”).

Registered AR: augmented content aligned to a printed locus so that digital information remains co-indexed to the page (the page stays the reference frame).

Retroactive contextualization: later interaction changes how earlier pages are interpreted; meaning updates backward once new information is revealed.

References

- BATARELO KOKIĆ, Ivana. 2015. "Nove razine interaktivnosti dječjih slikovnica". *Školski vjesnik. Časopis za pedagoški teoriju i praksu*, 64, 3: 377-398. <https://hrcak.srce.hr/file/222636>.
- BISWAS, Anirban, and Indrajit Ghosal. 2025. "Augmented Reality as a Catalyst for Environmental Awareness and Sustainability Education". In *Advances in Computational Intelligence and Robotics Book Series*, 113-144. DOI: [10.4018/979-8-3373-1892-9.ch005](https://doi.org/10.4018/979-8-3373-1892-9.ch005).
- BORKENT, Marjolein. 2023. *Multimodality and Cognition*. Oxford: Oxford University Press. DOI: [10.1093/oso/9780197509784.003.0003](https://doi.org/10.1093/oso/9780197509784.003.0003).
- CHUA, Shi Min. 2022. "Compiling and analysing a large corpus of online discussions to explore users' interactions". *Applied Corpus Linguistics*, 2, 2: 100017. DOI: [10.1016/j.acorp.2022.100017](https://doi.org/10.1016/j.acorp.2022.100017).
- GARCÍA MORA, Juan José, María E. Patiño-Jaramillo, and Sebastián J. Moreno Jiménez. 2022. "Contextualización cognitiva de la interactividad". *Human Review*, 12, 4: 1-19. DOI: [10.37819/revhuman.v12i4.1054](https://doi.org/10.37819/revhuman.v12i4.1054).
- ISMAIL, Ahmed. 2024. "Exploring the Levels of eLearning Interactivity: A Review of Research Literature". *Journal of Ecohumanism*, 3, 7: 2997-3024. DOI: [10.62754/joe.v3i7.4693](https://doi.org/10.62754/joe.v3i7.4693).
- JIAO, Yifan. 2023. "Picturing interactivity: design exploration of a highly interactive picturebook". In *IASDR 2023. Life-Changing Design*, edited by De Sainz Molestina et al., Politecnico di Milano, October 9-13. DOI: [10.21606/iasdr.2023.695](https://doi.org/10.21606/iasdr.2023.695).
- KILIÇ, Ayşegül. 2021. "Three dimensional movable books for children". *Journal of Arts*, 6, 2: 179-184. DOI: [10.26809/JOA.6.2.08](https://doi.org/10.26809/JOA.6.2.08).
- MARWAH, Riju, Jyotin Singh Thakur, and Pranav Tanwar. 2024. "Augmented Reality Assistive Technologies for Disabled Individuals". *arXiv preprint arXiv:2409.02053*. DOI: [10.48550/arxiv.2409.02053](https://doi.org/10.48550/arxiv.2409.02053).
- MURPHY, Keith M. 2023. "Multimodality". In *A New Companion to Linguistic Anthropology*. Edited by Alessandro Duranti, Rachel George and Robin Conley Riner, 443-460. DOI: [10.1002/9781119780830.ch24](https://doi.org/10.1002/9781119780830.ch24).
- OVERMANN, Karenleigh A. 2023. "Materiality in Cognition". In *The Materiality of Numbers: Emergence and Elaboration from Prehistory to Present*, 176-194. Cambridge: Cambridge University Press.
- PARK, Su Yeon, and Jeonghwa Tak. 2023. "빌 코터(Bill Cotter)의 「절대 누르면 안 돼!」와 「절대 만지면 안 돼!」의 그림책 물질성 읽기" [*Reading materiality in Bill Cotter's picture books "Don't Push the Button!" and "Don't Touch This Book!"*]. *Han'gug Yu'a Gyoyug Yeon'gu*, 25, 3: 50-72. DOI: [10.15409/riece.2023.25.3.3](https://doi.org/10.15409/riece.2023.25.3.3).
- PRINSLOO, Mastin. 2024. "Multimodality and Trans-Semiotics". In *Handbook of Plurilingual and Intercultural Language Learning*, 139-151. DOI: [10.1002/9781394165957.ch10](https://doi.org/10.1002/9781394165957.ch10).
- REGENBRECHT, Holger, Sander Zwanenburg, and Tobias Langlotz. 2022. "Pervasive Augmented Reality. Technology and Ethics". *IEEE Pervasive Computing*, 21: 84-91. DOI: [10.1109/MPRV.2022.3152993](https://doi.org/10.1109/MPRV.2022.3152993).
- PAULA, Anabel Medeiros Azerêdo de, Beatriz dos Santos Feres, and Margareth Silva de Mattos. 2023. "The Contemporary Children's Object-Book: Hybridization and Multimodality." *Pensares em Revista*, 28: 157-183. DOI: [10.12957/pr.2023.77881](https://doi.org/10.12957/pr.2023.77881).

SARANGAM, Anand. 2022. "Analyzing the Challenges and Opportunities for Making VR/AR Experiences Accessible to Individuals with Visual, Auditory and Mobility Impairments". *Journal of Engineering and Applied Sciences Technology*: 1-4. DOI: [10.47363/jeast/2022\(4\)e150](https://doi.org/10.47363/jeast/2022(4)e150).

SERAFINI, Frank. 2024. "The Instantiation of Narrative in Picturebook Forms". *Libri et Liberi*, 13, 1: 9-26. DOI: [10.21066/carcl.libri.13.1.1](https://doi.org/10.21066/carcl.libri.13.1.1).

SERAFINI, Frank, and Laura Moses. 2023. "An Analysis of the Semiotic Resources of Contemporary Movable Picturebooks". *Bookbird: A Journal of International Children's Literature*, 61, 2: 26-36. DOI: [10.1353/bkb.2023.0021](https://doi.org/10.1353/bkb.2023.0021).

SERAFINI, Frank, and Felicia Tompkins. 2015. "Books in / As / through Books". *The Reading Teacher*, 68, 5: 344-346. DOI: [10.1002/TRTR.1324](https://doi.org/10.1002/TRTR.1324).

SILVA, Sara Reis da, and Diana Martins. 2018. "Tirar, descubrir e interpretar: una caracterización del libro Pull-the-tab". *AILIJ Anuario de Investigación en Literatura Infantil y Juvenil*, 16: 141-152. DOI: [10.35869/AILIJ.V01I16.1345](https://doi.org/10.35869/AILIJ.V01I16.1345).

SOMANATH, Sowmya, Ehud Sharlin, and Mario Costa Sousa. 2014. "'Explorances' or why (some) physical entities help us be more creative". *ACM Symposium on Applied Perception*: 132. DOI: [10.1145/2628257.2628351](https://doi.org/10.1145/2628257.2628351).

SONG, Hee Sun, Maria V. Pusic, Matthew W. Nick, Umut Sarpel, Jan L. Plass, and Adina Kalet. 2014. "The cognitive impact of interactive design features for learning complex materials in medical education". *Computer Education*, 71: 198-205. DOI: [10.1016/J.COMPEDU.2013.09.017](https://doi.org/10.1016/J.COMPEDU.2013.09.017).

TRYBULEC, Marcin, and Ilona Hłowiecka-Tańska. 2023. "How do interactive exhibits at a science center become cognitive artifacts?". November. DOI: [10.21203/rs.3.rs-3664996/v1](https://doi.org/10.21203/rs.3.rs-3664996/v1).

VANDERSCHANTZ, Nicholas, Clare Timpany, and Kristy Wright. 2020. "Enhancing Interactivity: How has design exploration of physically and intellectually interactive picturebooks enhanced shared reading?". In *Proceedings of DRS*, 513-534. DOI: [10.21606/DRS.2020.192](https://doi.org/10.21606/DRS.2020.192).

VAN PRAET, Hilde. 2023. "Touching the screen: on visual textuality and knowledge construction in Anne Carson's poetry". *Textual Practice*. DOI: [10.1080/0950236x.2023.2259354](https://doi.org/10.1080/0950236x.2023.2259354).