Journal of Electronic Research and Application, 2026, Volume 10, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

Railway of Emotions: Visualizing Per-Character Sentiment in Narrative Text

Xiaoyun Zhu

Yunnan Open University, Kunming 650500, China

Abstract: The readers of long-form narratives find it challenging to track character emotions, differentiate emotional states, and stay engaged. Current sentiment visualization systems either aggregate emotions across characters or use disconnected diagrams requiring visual switching. We introduce SentiViz, a railway-themed system that employs a unified metaphor to represent real-time, per-character emotions during narrative reading, where characters are trains, emotions are trackside scenes, and plot structure is a train map with intersecting tracks. It uses a layered update strategy ambient lighting changes instantly for emotional awareness, while scene content transitions gradually to minimize distraction. A pilot study with ten participants reading two stories (with and without SentiViz) showed significant improvements in identifying main emotions ($p = 0.016$) and understanding emotional shifts ($p = 0.031$), with effect sizes of +1.40 and +1.50 on a 5-point scale. The system also scored significantly higher in aesthetics ($p = 0.008$) and recommendation likelihood ($p = 0.047$). Exploratory analysis indicated greater benefits for low-confidence readers. This work demonstrates that unified metaphors with real-time, per-character sentiment visualization can effectively support emotional interpretation during narrative reading, particularly for ESL learners and similar groups.

Keywords: SentiViz; Per-character sentiment; Railway metaphor; Emotional interpretation; Real-time visualization; AIGC Imagery

Online publication: March 31, 2026

1. Introduction

Long-form narrative reading requires readers to maintain an evolving mental model of who is present, what is happening, and how characters feel as events unfold [1–3]. In practice, readers often struggle to track multiple characters and their relationships, notice subtle emotional cues, and recognize when emotions shift across scenes, partly because working memory can hold only a limited number of meaningful elements at once [4–6]. Engagement also varies widely across individuals, and lower engagement can directly reduce comprehension and memory during reading [7–8].

Existing narrative and sentiment visualizations often add cognitive burden rather than reduce it: they split information across disconnected views (e.g., entities, relations, sentiment timelines), forcing readers to mentally

integrate "who feels what, when, and toward whom." When a representation does not match how people reason about a task, performance can suffer, especially when multiple encodings must be combined [9–12]. In the emotion-literature space, many analyses emphasize a single global arc instead of per-character trajectories, and systems that provide affective cues typically do not integrate character-specific emotion with interaction context in a way that supports in-the-moment reading [13,14].

We introduce SentiViz, a visualization system that takes a metaphor-first approach to reduce this integration cost. SentiViz uses a unified railway metaphor, where characters are trains, emotions are rendered as trackside scenes, character relationships appear in an in-cabin view, and plot structure is summarized as a train-map of intersecting routes. Using one coherent visual language across views helps readers form a single mental model instead of switching between unrelated diagrams, and prior work suggests that well-chosen metaphors can improve comprehension and search [15,16]. Unlike post-hoc analysis tools, SentiViz is synchronized with the reading cursor so emotions and interactions update as the reader progresses through the text.

1.1. Research question

Can real-time visualization of per-character sentiment during narrative reading improve as follows:

(1) Understanding of story plots;

(2) Interpretation of character emotions (including shifts over time);

(3) Reading engagement compared to text alone?

1.2. Contributions

The contributions are as listed:

(1) SentiViz: A unified railway-themed interface for narrative reading that visualizes per-character emotions and plot structure with coordinated outside, in-cabin, and train-map views;

(2) AIGC imagery pipeline: A text-to-image workflow that generates scene and character visuals from narrative prompts to enrich context during reading;

(3) Pilot evaluation: A within-subject study comparing SentiViz with a text-only baseline on plot understanding, emotional interpretation, and engagement.

2. Literature review

Prior work related to SentiViz falls into three strands as follows:

(1) Sentiment visualization dashboards for social/personal text;

(2) Psychology-grounded emotion analysis for literature;

(3) Creative multi-view storytelling systems that use generative imagery.

2.1. Sentiment dashboards for social and personal text

A common approach visualizes sentiment with module-based dashboards composed of familiar charts such as word clouds, timelines, and stream/river views. For example, SentiSand integrates coordinated views (e.g., word cloud, ThemeRiver, and summary charts) to support exploration of polarity trends over time [17] (**Figure 1**). While effective for retrospective analysis of personal posts, these systems emphasize aggregated sentiment and require users to integrate heterogeneous chart types; they also typically do not model narrative entities (characters) or interactions central to story understanding.

Figure 1. SentiSand system user interface. (a) Page navigation: it can switch between overview and details (b) Selection of time period (c) Data presented in pie chart proportions (d) Sentiment story of the selected time period, consisting of four modules (e) Report download functionality [17]

2.2. Psychology-grounded emotion models for literary analysis

Moving beyond polarity, systems such as EmotionLens analyze literary texts using psychology-inspired representations (e.g., valence-arousal) with coordinated views like affective word clouds and emotion planes [14]. This work provides principled emotion spaces and supports exploratory analysis across works and authors, but it is largely post-hoc and does not support real-time, per-character trajectories or directed emotions during in-situ reading.

2.3. Creative multi-view storytelling with AIGC imagery

Interactive visual storytelling systems suggest that coordinated views and generated imagery can improve engagement and accessibility. Jinling Fenghua, for instance, combines relationship networks, filters, and AIGC-generated portraits to explore cultural-heritage personages, works, and geo-events [18] (**Figure 2**). However, these views are not unified under a single metaphor and do not target character-centric emotion tracking synchronized with narrative reading.

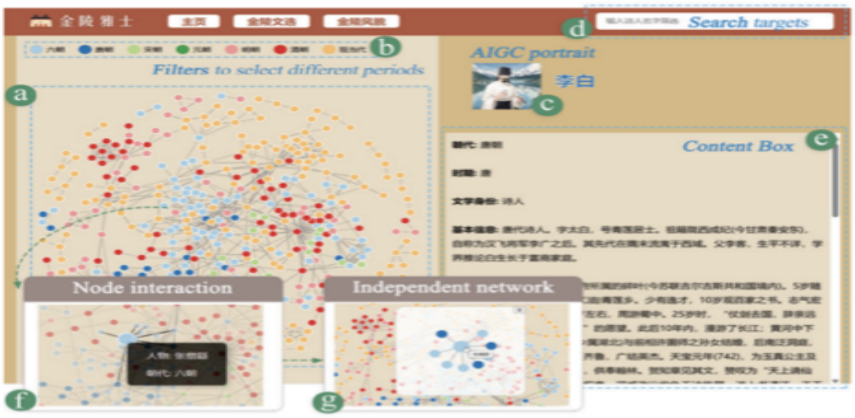

Figure 2. Jinling Fenghua. (a) Relationship network (b) Filters (c) AIGC portrait (d) Search box (e) Content box (f) Hover over a node (g) A highlighted personage network.

2.4. Positioning of SentiViz

SentiViz builds on coordinated multi-view interaction, psychology-inspired emotion representations, and

generative imagery, but targets a different task: supporting narrative reading in real time. Its key distinction is a metaphor-first railway design that unifies views and visualizes per-character emotions (including directed affect) synchronized with the reading cursor, reducing the need to mentally integrate disconnected representations.

3. Design

3.1. Design goals

Guided by our research question and prior work, SentiViz is designed to support narrative reading by improving emotion interpretation, plot understanding, and engagement:

3.1.1. R1: Support interpretation of character emotions

Readers should quickly identify each active character's emotion, compare emotions across characters, and notice emotional transitions as the story unfolds. Unlike tools that emphasize aggregate sentiment, our focus is per-character and supports emotions directed toward other characters [14].

3.1.2. R2: Support understanding of story plots

Readers should understand who participates in each story unit, how characters interact, and how key events progress over time. The challenge is providing both local context and global structure without forcing readers to integrate disconnected diagrams [12].

3.1.3. R3: Sustain reading engagement

The visualization should remain readable and enjoyable during continuous reading. Since engagement affects comprehension and memory, we aim for a metaphor that is intuitive and visually rich without distracting from the text [7,8].

3.2. Overview

SentiViz uses a unified railway metaphor to visualize narrative elements on a coordinated canvas: characters as trains, emotions as trackside scenes, relationships inside the cabin, and plot structure as a train-map. All views are synchronized with the reader's scroll position, enabling real-time emotion tracking while reading rather than post-hoc analysis.

3.2.1 Design alternatives

We explored metaphor-based designs (e.g., a growing tree or a garden) and abstract charts (timelines/heatmaps). Tree-like metaphors did not scale well to multi-character co-occurrence and interaction, while charts were precise but less engaging and required more decoding during reading. The railway metaphor naturally supports temporal progression, character "meetings" at shared stations, and familiar navigation patterns.

3.2.2. Visualization strategy

To augment reading rather than distract, we use layered updates: fast channels (ambient light/weather) change immediately to signal emotional shifts, while richer scene content transitions more gradually. The outside view supports continuous awareness; the in-cabin and train-map views update at story-unit boundaries and are accessed as needed.

3.2.3. Interface at a glance

The text reader drives the visualization state in real time. Users switch among three coordinated views as outlined (**Figure 3**):

(1) An outside view for continuous emotion cues;

(2) An in-cabin view for co-presence and directed emotions;

(3) A train-map view for global structure and navigation.

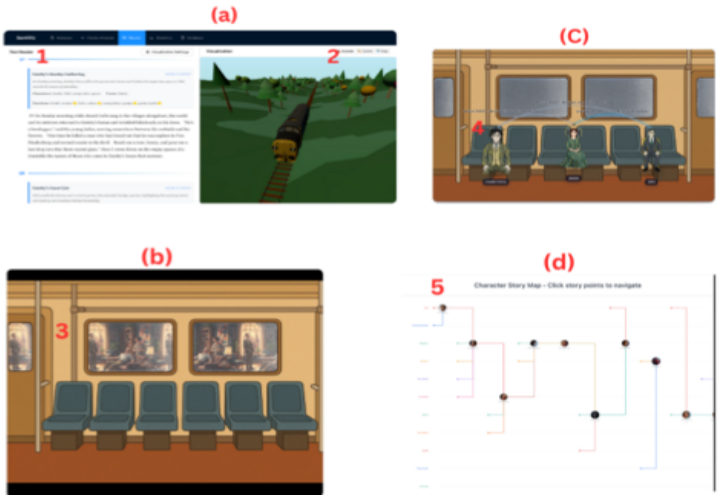

Figure 3. SentiViz overview with coordinated, railway-themed views driven by the reading cursor. (a) Reading-driven control: the text reader (1) updates the current scene in real time; the tab bar (2) switches among Outside / In-Cabin / Train-Map views. The Outside view encodes the current emotion using train speed, surrounding scenery, and weather. (b) In-Cabin (narrator only): no main character is present; window panels show images generated from the current passage, scrolling with the train so motion speed reflects emotion intensity. (c) In-Cabin (with characters): three characters are present; curves above seats depict emotions directed toward one another, and seat distance encodes relationship valence (closer = more positive). (d) Train-Map of storylines: each row is a character; stations (circles) mark story points and include a representative image. A station appears on a character's row when that character is the focus at that point, and routes intersect when characters participate in the same story point.

3.2.4. Innovation

SentiViz unifies per-character emotion, directed affect between characters, and plot structure under one metaphor and keeps them synchronized with the reading cursor. We also encode emotion categories with explicit, learnable scene templates rather than only abstract charts.

3.2.5. Data flow

Given uploaded text, we did as follows:

(1) Segment it into story units;

(2) Extract characters and per-unit emotions (including directed emotions);

(3) Generate per-unit imagery (thumbnails/window panels);

(4) Render the results across coordinated railway-themed views driven by the active unit.

3.3. Details and contributions

We adopted the OCC appraisal framework as the conceptual basis for emotion categories [19]. We used a fixed set of 23 labels as the interface between analysis and visualization. Early trials with purely generated scenes were visually diverse but did not reliably communicate the intended emotion; we therefore use handcrafted scene templates aligned to the OCC labels to keep emotion cues explicit and learnable.

3.4. Outside view

The outside view shows a train moving through trackside scenes that reflect the current emotion category in real time (**Figure 3a**). Emotion is encoded through as follows:

(1) The surrounding environment (scene templates);

(2) Atmospherics (light, weather, visibility);

(3) Apparent motion (speed).

To avoid abrupt disruption, we generate scenery in scroll-aligned chunks: ambient light/weather updates immediately on emotional changes, while newly generated chunks transition smoothly into view.

3.4.1. Outside view templates

Each OCC label maps to a handcrafted scene template; **Figure 4** shows eight examples.

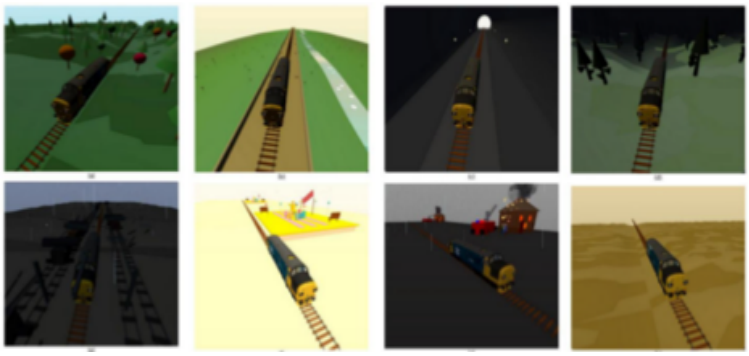

Figure 4. Outside-view templates aligned with OCC labels. Each panel shows the train passing a scene whose visual state corresponds to the current emotion. (a) Neutral: Forest with balanced light and medium openness. (b) Joy: Bright day with a sparkling river and open horizon. (c) Hope: Tunnel with light visible at the end. (d) Fear: Dark forest with limited visibility. (e) Distress: Railyard in adverse conditions. (f) Happy-for: Person being praised on a stage under warm light. (g) Pity: Someone's house on fire, conveying shared sorrow. (h) Resentment: Barren desert viewed from a detached vantage. Scenes are concise, repeatable templates; we handcraft one per label to keep the mapping learnable and explicit.

3.4.2. Design rationale

This view primarily supports R1 by making shifts immediately visible during reading via fast atmospheric cues, while still providing categorical understanding through scene templates. The familiar travel metaphor and controlled motion support R3 by keeping the visualization engaging without requiring readers to interpret dense charts.

3.5. In-cabin view: Characters and directed emotions

The in-cabin view emphasizes who is present in the current story unit and how they feel toward each other (**Figure 3b–c**). When no main character is present, it shows the narrator's perspective; when characters are present, it renders them seated with interpersonal cues.

3.5.1. Windows

Scroll-synchronized panels display images generated from the current passage to maintain narrative grounding. Prompts include the current story summary and the outside view's color/atmosphere so imagery stays stylistically coherent across views.

3.5.2. Seats and links

Characters appear with nameplates; seat distance encodes relationship valence, and curves above seats encode directed emotions between pairs.

3.5.3. Design rationale

This view supports R1 and R2 by making co-presence and directed affect explicit at the moment they occur, reducing the need to infer relationships purely from text. AIGC imagery provides contextual anchoring and supports R3 by adding visual variety without changing the core metaphor.

3.6. Train-map view: Storylines and navigation

The train-map provides an overview of plot structure and supports navigation between story units (**Figure 5**). Each row represents a character; each story unit is a station; and stations appear on multiple rows when multiple characters participate, making intersections and co-occurrence explicit. Hover reveals a title/summary, and click jumps the reader to the corresponding text position while synchronizing the other views. The map primarily supports R2 by externalizing narrative structure and character intersections, and secondarily supports R1 by providing image-based landmarks that help readers recall emotional context across the story.

Figure 5. In-cabin view. (a) Narrator-only: No main character is present; window panels show images generated from the current passage to keep the scene anchored to the text. (b) Characters present: Three characters are seated with nameplates; curves above the seats encode emotions directed toward one another, and seat distance encodes relationship valence (closer = more positive). The view updates in real time as the reader scrolls, supporting interpretation of character emotions and understanding of plot dynamics.

4. Implementation

4.1. Stack

The React web app client uses Three.js for in-cabin, outside, and train-map views. A Python FastAPI server handles uploads and offline pre-analysis. The analysis uses spaCy (character extraction), facebook/bart-large-mnli (zero-shot OCC label assignment), OpenAI (LLM utilities, embeddings), and Google Gemini (text-to-image). A graph builder organizes story units, character participation, and emotions. Results are stored in MongoDB; the

client downloads a compact scene index with precomputed results and media for local rendering.

4.2. Rationale

A web app simplifies deployment and integrates well with hosted ML services. BART provides stable zero-shot classification, while LLMs offer concise summaries. On 400 sentences re-labeled from GoEmotions to OCC, BART achieved a Macro-F1 of 0.3020 vs. 0.2497 for GPT-4.1, leading us to adopt BART for label assignment. Gemini's edit mode ensures style-consistent images. Pre-analysis at upload avoids server round-trips, keeping interactions responsive.

4.3. Pre-analysis pipeline

Uploaded text is as outlined:

(1) Segmented into story units with characters and relevant text extracted by GPT-4.1;
(2) Assigned OCC labels per narrator/character and directed emotion labels with BART;
(3) Generated titles/summaries (LLM) and thumbnail images (Gemini) per unit;
(4) Compiled into a scene index with unit boundaries, character sets, labels, titles/summaries, and media URLs.

4.4. Runtime synchronization (client-side)

The reader's scroll position, tracked by an intersection observer, maps to the active unit, applying corresponding labels and metadata to all views.

4.5. Challenges and solutions

The challenges and solutions are as detailed:

(1) Label quality: Zero-shot classification struggled with rare emotions. We improved stability with curated prompts, normalization rules, and post-hoc consistency checks;
(2) Pre-analysis latency: Long texts were slow to process. We parallelized labeling and image generation, streaming results progressively (text first, thumbnails soon after);
(3) Visual consistency: To align imagery with the outside view, prompts included the color theme and OCC label; deterministic seeds maintained character continuity;
(4) Reproducibility notes: We release as follows:
    (i) The 23-label list and label→scene template mapping;
    (ii) Prompt templates for imagery;
    (iii) Scene index data schema with examples;
    (iv) Environment details (model names/versions, API parameters).

These enable re-implementation with alternative providers while preserving behavior in **Figure 3**.

5. Evaluation

5.1. Methodology

We conducted a pilot study with ten users to evaluate whether visualizing character sentiment enhances story plot understanding, character emotion interpretation, and reading engagement. Each participant read two stories, experiencing both conditions, one with the sentiment visualization tool and one without, to enable direct within-subject comparison.

Ten participants (n = 10) were recruited with demographics and reading characteristics shown in **Table 1**. Emotional understanding was assessed using six custom five-point Likert scale items (1 = strongly disagree, 5 = strongly agree), measuring text clarity, main emotion identification, subtle emotion recognition, emotional shift perception, character emotion distinction, and emotional relationship understanding. Cognitive workload was evaluated using two NASA-TLX subscales (mental demand and perceived performance, rated 1–10). User experience was measured with five Likert items (1–5) assessing engagement, frustration (reverse-scored), aesthetics, recommendation likelihood, and overall reward.

Participants read two 2,000-word narratives written by our team: (1) Alan's medieval fantasy journey to Black Spire, and (2) Joren's science fiction journey to Krayven's asteroid fortress. Both stories featured a single protagonist on a challenging quest with emotionally charged scenes. Story and tool condition order were counterbalanced using a Latin Square design to control for order effects.

After each reading (15–20 minutes), participants completed two objective comprehension assessments: factual comprehension (10 multiple-choice questions per story on plot events, character actions, settings, and details) and emotional comprehension (10 questions per story on emotional states, motivations, changes, and interpersonal dynamics). They also completed the subjective questionnaires. Participants then re-examined an emotionally complex scene (10 minutes) and answered additional questions. After both stories, they provided open-ended feedback on both conditions.

The two test conditions were as follows:

(1) Control condition: Participants read plain text without visualizations;

(2) Visualization condition: The sentiment visualization interface was displayed alongside the text. Both conditions used the same web-based platform.

Sessions were conducted individually in a quiet lab on desktop computers with 24-inch monitors. Given the small sample size (n = 10) and ordinal rating scale data, non-parametric Wilcoxon signed-rank tests were used to compare conditions, with statistical significance set at $\alpha = 0.05$.

Table 1. Participant demographics (n = 10)

Aspects	Distribution
Age	
18–21	9 (90%)
25–30	1 (10%)
Gender	
Male	7 (70%)
Female	3 (30%)
Reading confidence (1–5)	
High (5 out of 5)	5 (50%)
Other	5 (50%)
Weekly reading (outside study)	
> 5 hours/ week	3 (30%)
< 1 hour/ week	6 (60%)
Other	1 (10%)
Learning/ reading disabilities	None

5.2. Results

The pilot study suggests the sentiment visualization tool is effective. Participants showed significant improvements in two key emotional understanding measures when using the tool, along with higher ratings for visual appeal and recommendation likelihood. Objective comprehension scores remained high across conditions, indicating potential ceiling effects.

Statistical comparisons used two-tailed Wilcoxon signed-rank tests ($p < 0.05$). For each measure, mean (M), standard deviation (SD), Wilcoxon statistic (W), and p-value are reported.

The overall results are as summarized:

(1) Emotional understanding: The tool significantly improved identifying main emotions ($p = 0.016$, +1.40 points) and understanding emotional shifts ($p = 0.031$, +1.50 points) on the 5-point scale. Distinguishing character emotions showed a strong trend ($p = 0.063$). Other measures favored the tool but lacked significance (**Table 2**);

(2) Cognitive workload: The tool showed trends toward reduced mental demand and improved perceived performance, though differences were not statistically significant (**Table 3**);

(3) User experience: The tool received significantly higher ratings for aesthetics ($p = 0.008$) and recommendation likelihood ($p = 0.047$). Engagement, frustration, and reward ratings showed no significant differences (**Table 4**);

(4) Objective comprehension: No significant differences emerged between conditions for factual or emotional comprehension. Both conditions performed consistently high (> 8/10), suggesting ceiling effects may have obscured differences (**Table 5**).

An exploratory analysis examined effectiveness by reading confidence (**Table 6**). The five participants (50%) rating confidence below 5/5 showed larger effect sizes (+1.60 to +1.80 points for emotional understanding items) than the full sample, though none reached significance (all $p > 0.10$) due to small subsample size. These results suggest the tool may benefit lower-confidence readers, but larger-scale testing is needed.

Open-ended feedback revealed four main themes. Distracting visual elements (40% of participants) included concerns about interface clutter and unclear purposes of secondary features (e.g., cabin and map). Technical performance issues (30%) involved lag and scene loading delays disrupting focus. Interaction design preferences included a suggestion for mouse-click navigation instead of scrolling. One participant noted story similarity reduced the second story's challenge. Positively, one participant highlighted improved readability due to text chunking, and another suggested adding ambient sounds for immersion.

Table 2. Emotion understanding likert scale results (1–5 scale)

Item	With	Without	Diff	p	Sig
Text clarity	3.70 (1.34)	3.10 (0.99)	+0.60	0.109	ns
Main emotion ID	4.50 (0.71)	3.10 (1.29)	+1.40	0.016	*
Subtle emotions	3.70 (1.06)	3.00 (0.82)	+0.70	0.223	ns
Emotional shifts	4.50 (0.53)	3.00 (1.41)	+1.50	0.031	*
Distinguish chars	4.25 (0.71)	3.12 (0.83)	+1.12	0.063	ns
Char. relationships	4.00 (0.94)	3.60 (0.84)	+0.40	0.344	ns

Note: "With" and "Without" columns show mean rating scores with standard deviation in parenthesis. Diff = difference between conditions; *$p < 0.05$; ns = not significant.

Table 3. NASA task load index results (1–10 scale)

Measure	With	Without	Diff	p	Sig
Mental demand	6.50 (2.32)	7.60 (1.58)	-1.10	0.156	ns
Performance	6.40 (2.67)	5.70 (2.36)	+0.70	0.344	ns

Note: ”With” and ”Without” columns show mean scores with standard deviation in parentheses. Lower Mental Demand = less cognitive load (better); Higher Performance = better perceived success; ns = not significant

Table 4. User experience scale results (1–5 scale)

Dimension	With	Without	Diff	p	Sig
Engagement	2.20 (0.79)	2.70 (1.34)	-0.50	0.375	ns
Frustration	2.60 (1.17)	2.40 (1.17)	+0.20	0.828	ns
Aesthetics	3.90 (0.57)	2.10 (0.99)	+1.80	0.008	**
Recommendation	3.50 (0.97)	2.40 (1.26)	+1.10	0.047	*
Reward	3.30 (1.16)	2.90 (1.10)	+0.40	0.344	ns

Note: ”With” and ”Without” columns show mean ratings with standard deviation in parentheses. Diff = difference between conditions; * $p < .05$; ** $p < .01$; ns = not significant.

Table 5. Comprehension test performance (out of 9–10 points)

Test type	With	Without	Diff	p	Sig
Factual: Story 1	8.20 (1.10)	9.00 (1.00)	-0.80	0.250	ns
Factual: Story 2	9.40 (0.55)	9.60 (0.55)	-0.20	1.00	ns
Emotional: Story 1	9.20 (0.84)	8.60 (1.52)	+0.60	0.750	ns
Emotional: Story 2	7.80 (1.30)	7.40 (2.30)	+0.40	1.00	ns

Note: ”With” and ”Without” columns show mean test scores with standard deviation in parentheses. No significant differences were found between conditions (all $p > .05$); ns = not significant.

Table 6. Key results for low-confidence readers (n = 5, confidence scores 2–3)

Measure	With	Without	Diff	p	Sig
Main emotion ID	4.20 (0.84)	2.60 (1.34)	+1.60	0.125	ns
Emotional shifts	4.20 (0.45)	2.40 (1.52)	+1.80	0.250	ns
Subtle emotions	3.80 (1.30)	2.80 (0.84)	+1.00	0.312	ns
Distinguish chars.	4.00 (0.82)	2.75 (0.96)	+1.25	0.500	ns
Aesthetics (UES)	4.00 (0.71)	2.20 (1.30)	+1.80	0.125	ns
Recommendation (UES)	3.80 (0.84)	2.80 (1.64)	+1.00	0.250	ns

6. Discussion

Our research investigated whether visualizing evolving character sentiment improves understanding, interpretation, and engagement. Results show partial success, where significant gains in emotional interpretation and aesthetic

appeal, but mixed outcomes for plot comprehension and engagement.

6.1. Interpretation of character emotions (Success)

The visualization significantly improved identifying main emotions ($p = 0.016$, +1.40 points) and understanding emotional shifts ($p = 0.031$, +1.50 points), with a strong trend for distinguishing character emotions ($p = 0.063$). This supports our design principle: layered ambient updates (immediate lighting changes with gradual scene transitions) made emotions salient without disrupting reading flow. The railway metaphor with OCC-mapped scenes successfully facilitated real-time emotional interpretation.

6.2. Understanding of story plots (Mixed)

No significant differences emerged in comprehension tests, likely due to ceiling effects (both conditions scored > 8/10). The tests may not capture the tool's potential benefit during reading rather than post-hoc recall. In addition, more complex narratives with harder questions are needed to properly assess this aspect.

6.3. Engagement (Paradox)

The tool received significantly higher ratings for aesthetics ($p = 0.008$) and recommendation ($p = 0.047$), yet engagement scores were paradoxically lower (though non-significant). Qualitative feedback identified technical lag (30% of participants) as the cause, disrupting reading flow and contradicting our augmentation principle. Optimizing system performance is critical for future iterations.

6.4. Low-confidence readers

Exploratory analysis revealed larger effect sizes for lower-confidence readers (improvements of +1.60 to +1.80 points), though statistical significance was not reached (n = 5). This suggests the tool may particularly benefit readers struggling with emotional interpretation, such as ESL learners. A larger study targeting this population could confirm this hypothesis.

6.5. Comparison to prior work

Our unified metaphor approach outperformed prior systems. Unlike EmotionLens's aggregate visualization, our per-character tracking significantly improved emotional interpretation ($p < 0.05$), validating the importance of character-level granularity [14]. Compared to SentiSand's disconnected views, our railway metaphor received higher aesthetic ratings, supporting our hypothesis that coherent metaphors improve user experience [17]. Unlike Jinling Fenghua's static portraits, our dynamic, scene-specific AIGC imagery enhanced aesthetic appeal while maintaining narrative coherence [18]. These comparisons suggest unified metaphors and real-time, character-specific visualization offer meaningful advantages for narrative understanding.

6.6. Design implications

Effective elements included the unified railway metaphor, layered ambient updates, AIGC imagery, and the train-map for plot overview. Refinements are needed to eliminate technical lag, simplify the interface by hiding secondary features (cabin details, map) by default, and provide clearer guidance on view usage. Moreover, balancing visual richness and reading flow requires selective display and user control.

6.7. Limitations

The small sample (n = 10) limited statistical power, especially for subgroup analyses. Ceiling effects in comprehension tests prevented detection of potential differences. Technical lag disrupted the experience for 30% of participants. Both stories shared similar adventure structures despite different genres. The short-term study could not assess long-term engagement or learning effects. Implementation issues included occasional emotion misclassifications due to untrained OCC models, growing pre-analysis time with text length, and increased storage requirements for generated assets. The system remains a research prototype requiring optimization for production use.

6.8. Future work

Priorities include as follows:

(1) System optimization to eliminate lag;

(2) Interface simplification based on user feedback;

(3) A larger study (n > 30) with diverse participants and harder comprehension tests;

(4) Tailoring the tool for ESL learners and other populations struggling with emotional interpretation;

(5) Evaluation with different narrative genres and longer texts to assess scalability and sustained engagement.

7. Conclusion

This paper introduces SentiViz, a railway-themed visualization system designed to enhance understanding of story plots, interpretation of character emotions, and reading engagement by dynamically visualizing evolving sentiment for individual characters in narrative texts. The system employs a unified metaphor, representing characters as trains, emotions as trackside scenes, and plot structure as a train-map, with all views synchronized in real-time as readers scroll. The layered update strategy, immediate ambient lighting changes coupled with gradual scene transitions, highlights emotional shifts without disrupting reading flow. Results from our pilot study (n = 10) demonstrate the visualization's effectiveness in improving emotional interpretation. Participants showed significant gains in identifying main emotions (+1.40 points, $p = 0.016$) and understanding emotional shifts (+1.50 points, $p = 0.031$) when using the tool versus text alone. The system also received significantly higher ratings for visual aesthetics ($p = 0.008$) and recommendation likelihood ($p = 0.047$), successfully fulfilling its core purpose of supporting real-time emotional interpretation during reading. While comprehension tests showed no significant differences, this was likely due to ceiling effects (both conditions scored above 8/10), and the tests may not capture benefits occurring during reading rather than post-hoc recall. Technical lag reported by 30% of participants undermined engagement scores despite high aesthetic ratings, highlighting the importance of system performance for the intended augmentation experience. Our findings suggest that sentiment visualization with unified metaphors is a valuable addition to narrative reading, particularly for improving emotional understanding. This is especially true for readers with lower confidence in their reading ability, who showed even larger effect sizes (improvements of +1.60 to +1.80 points), though statistical significance was not reached due to the small subsample size. The system holds particular promise for ESL learners and similar populations who may struggle with interpreting character emotions in text.

Disclosure statement

The author declares no conflict of interest.

References

[1] Zwaan R, Langston M, Graesser A, 1995, The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. Psychological Science, 6(5): 292–297.

[2] Graesser A, Singer M, Trabasso T, 1994, Constructing Inferences During Narrative Text Comprehension. Psychological Review, 101(3): 371–395.

[3] Zacks J, Speer N, Reynolds J, 2009, Segmentation in Reading and Film Comprehension. Journal of Experimental Psychology: General, 138(2): 307–327.

[4] Cowan N, 2001, The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity. Behavioral and Brain Sciences, 24(1): 87–114.

[5] Gernsbacher M, 1990, Language Comprehension as Structure Building. Hillsdale, NJ: Lawrence Erlbaum Associates.

[6] Trabasso T, van den Broek P, 1985, Causal Thinking and the Representation of Narrative Events. Journal of Memory and Language, 24(5): 612–630.

[7] Hidi S, Renninger K, 2006, The Four-Phase Model of Interest Development. Educational Psychologist, 41(2): 111–127.

[8] Schraw G, Lehman S, 2001, Situational Interest: A Review of the Literature and Directions for Future Research. Educational Psychology Review, 13(1): 23–52.

[9] Vessey I, 1991, Cognitive Fit: A Theory-based Analysis of the Graphs vs. Tables Literature, Proceedings of the 12th International Conference on Information Systems (ICIS), 562–573.

[10] Shah P, Hoeffner J, 2002, Review of Graph Comprehension Research: Implications for Instruction. Educational Psychology Review, 14(1): 47–69.

[11] Cleveland W, McGill R, 1984, Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. Journal of the American Statistical Association, 79(387): 531–554.

[12] Larkin J, Simon H, 1987, Why a Diagram is (Sometimes) Worth Ten Thousand Words. Cognitive Science, 11(1): 65–100.

[13] Vishnubhotla K, Hammond A, Hirst G, et al., 2024, The Emotion Dynamics of Literary Novels, arXiv, https://arxiv.org/abs/2403.02474

[14] Wang B, Shi Q, Wang X, et al., 2025, Emotion-Lens: Interactive Visual Exploration of the Circumplex Emotion Space in Literary Works via Affective Word Clouds. Visual Informatics, 2025(9): 84–98.

[15] Baldonado M, Woodruff A, Kuchinsky A, 2000, Guidelines for Using Multiple Views in Information Visualization, Proceedings of the Working Conference on Advanced Visual Interfaces (AVI), 110–119.

[16] Ziemkiewicz C, Kosara R, 2008, The Shaping of Information by Visual Metaphors. IEEE Transactions on Visualization and Computer Graphics, 14(6): 1269–1276.

[17] Xie Y, Jing J, Saha T, et al., 2024, Sentisand: Visualizing Individual Sentiments on Social Media through Storytelling, SSRN Preprint, https://doi.org/10.2139/ssrn.4772802

[18] Xie A, Xie Y, Liu Y, et al., 2025, Jinling Fenghua: Unfolding Cultural History of the Jinling Context via Visual Storytelling, Proceedings of the 28th International Conference on Computer Supported Cooperative Work in Design (CSCWD).

[19] Steunebrink B, Dastani M, Meyer J, 2009, A Formal Model of Emotions: Integrating Qualitative and Quantitative Aspects. KI -Kunstliche Intelligenz, 23(3): 18–25.