Generative AI-Driven Optimization of Event Narrative Experience Design: A VR Experimental Study Based on Dynamic and Static Narrative

Renjingyi Zhang

School of Innovation and Design, City University of Macau, Macau

Abstract. This study investigates how generative artificial intelligence (AI) enhances interactive narrative experiences in virtual reality (VR) by comparing dynamic AI-driven storytelling with pre-scripted static narratives. A controlled experiment (N = 40) was conducted to evaluate user experience across six dimensions, including immersion, personalization, emotional resonance, and overall VR experience. Analytic hierarchy process (AHP) and one-way ANOVA were employed to assess weighted importance and group differences. Results indicate that dynamic AI-driven narratives significantly outperform static narratives in immersion, personalization, emotional resonance, and overall VR experience. Gender-based differences were also observed in affective responses. Correlation analyzes further reveal strong associations between immersion, emotional resonance, and perceived VR experience quality. These findings demonstrate that generative AI can meaningfully enhance narrative engagement and user-centered experience design in immersive environments, providing empirical support for adaptive storytelling as a design strategy in human–computer interaction.

Keywords: Dual narrative, Generative AI, Analytic Hierarchy Process, Personalized Experience, Emotional Resonance.

1. Introduction

Against the backdrop of the rapid development of digital technology and virtual reality (VR) technology, the way of acquiring information and experiencing the world is undergoing profound changes. From early 360 -degree panoramic videos to today's highly interactive immersive environments, VR -based interactive storytelling has undergone significant evolution [1]. However, in the process of this technological innovation, a fundamental contradiction has become increasingly prominent: there is always an irreconcilable tension between the fixed narrative framework carefully designed by designers and the ever-growing personalized experience needs of users.

Traditional interactive narrative (static narrative) follows a linear structure, and users can only experience content along a preset path. Although this method ensures the integrity and logical coherence of the narrative [2], it greatly restricts the user's right to choose. After repeated experiences, users are prone to aesthetic fatigue because the content is highly predictable and lacks variation [3]. This phenomenon has spurred a strong demand for personalized narrative experiences worldwide - users are no longer satisfied with passively accepting stories, but hope to become co-builders of the narrative. For example, Richards et al. found through comparative experiments that primary school students using personalized VR teaching performed significantly better than those using traditional teaching methods [4].

VR technology has expanded the spatial dimension of interactive narrative. Scholars have focused on narrative elements, processes, and structures to construct digital narrative frameworks and formulate design principles. For example, Mary Raul Ryan [5] proposed a three-level interactive narrative model (external interaction layer, exploratory interaction layer, ontological interaction layer) and extracted a design process to enhance immersion. Xu Congyao [6] and others analyzed the interactive narrative system in the digital display of cultural heritage and proposed a four-stage design process—cultural decoding, context modeling, technology integration, and experience evaluation. They also proposed three design principles—cultural authenticity, narrative coherence, and interactive transparency, providing methodological support for the digital transformation of this field.

In recent years, the rise of generative AI has provided a new approach to solving the above contradictions. Its dynamic content generation capabilities, such as ChatGPT, Claude3, Gemini, and DeepSeek, can create diverse narrative paths in real time based on user input, and theoretically provide a unique experience for each user [7]. However, this technology also faces a key challenge: how to maintain the internal logic and emotional depth of the narrative while giving users creative freedom?

there is a lack of systematic discussion on the collaborative mechanism between generative AI and dual narratives, and existing results mostly analyze the technology or narrative itself in isolation; secondly, compared with the mature drama theory evaluation system of static narrative [8], dynamic narrative lacks complete quality standards, making it difficult to balance personalization and artistry; finally, previous VR narrative research has focused on game scenes, using preset 3D/360 -degree videos while ignoring interactivity [9]; or Smith's emotion computing model did not fully consider the dynamic balance problem [5].

Therefore, this study intends to adopt a mixed research design, collect data through questionnaires and controlled experiments, and use the Analytic Hierarchy Process (AHP), analysis of variance, and correlation analysis for quantitative verification. The aim is to systematically explore the dynamic balance mechanism of generative AI in interactive narratives. The core research questions include: (1) Can dynamic narratives break through the limitations of traditional narrative content solidification? (2) How to balance the freedom and narrative coherence of dynamic narratives? (3) Which factors have the most significant impact on user experience? (4) How to maintain stable emotional resonance when users participate in dynamic narratives?

2. Literature Review

2.1 Interactive Narrative

With the rapid development and continuous progress of digital technology, the theory of interactive narrative has also emerged and has now become an important branch for studying narrative phenomena in the new information age. The concept of interactive narrative first appeared in the game field and was proposed by the famous game designer Chris Crawford in 2004 [8]. In his view, interactivity is the core content of interactive narrative. Interactivity refers to "a cyclical process between two or more active individuals, in which each party takes turns listening, thinking and speaking, forming a certain form of dialogue" [10].

After interactive narratives were integrated into research on virtual reality and user experience, especially in fields such as information and interaction design, the influence factors and design principles of interactive narratives in the dimension of interactive perception experience have gradually become the focus of research. For example, scholar Young [11] proposed a binary model of generative narrative in virtual reality space. This binary model separates the story and narrative, so as to generate a more coherent narrative. In terms of user experience, scholar Thie [12] took the interactive 3D virtual reality device "Beyond Manzanar" as an example and focused on exploring the dramatic possibilities of the emotional state caused by the interaction of users in the virtual reality environment. These studies provide a direct theoretical basis for the dual narrative architecture of this study, but have not yet solved the problem of balancing personalization and coherence of dynamic narratives based on generative AI.

According to the literature, there are two main lines in the study of interactive narrative. The first is the narratological path, represented by Ryan, which focuses on how interactivity changes the narrative structure and proposes the "spectrum of interactivity" [5]. The second main line is the computational path, represented by Mateas & Stern, which focuses on how to achieve narrative generation through computational models. Therefore, this study attempts to bridge these two traditions and regards generative AI as a new type of computational narrative engine. Its probability-based language model essence forms an interesting correspondence with the "possible world" theory of classical narratology.

2.2 Static Narrative

Static narrative refers to a design narrative without plot development and change, and the content of the "narration" is often expressed as a concept of "reason" [13]. Under the design narrative type of static narrative, the design object itself already has the function of narration and comes with its own narrative content. The development of the design object can, in a sense, promote the development of the narrative content. Static narrative is an interactive experience with a pre-determined narrative structure. Its core feature is the closedness of narrative possibilities. Users can only move within the boundaries set by the creator, just like walking in a carefully designed narrative maze - although the path can be freely chosen, the position of the wall and the number of exits have been fixed [5]. At the same time, this is also an implicit narrative method. The creator uses aesthetic means to create literary content, bringing about the development and improvement of design in the spatial dimension. Static narrative can be divided into two types according to the fixed nature of its content: "multiple things and one event" and "one event and one thing". The narrative content is unique and remains stable [13]. However, the rigid framework of static narrative also brings obvious limitations - users can only be passive recipients and cannot participate in the creation or change of narrative content, which leads to the problem of "formulaic" when experiencing the same content repeatedly [4]. This study believes that these limitations of static narrative are the key driving force for the development of dynamic narrative, but existing research has failed to fully explain how to maintain the artistic value of narrative while breaking the fixedness.

2.3 Dynamic Narrative

Dynamic narrative, compared with traditional static narrative, breaks through the fixed content. Its core lies in breaking the original rigid framework and constructing an adaptable and growing story system. Dynamic narrative refers to design narrative with plot development and changes. The behavior of "narrating" is often a "event" story [13]. This is a more explicit plot narrative. It redefines the relationship between creators and users. Dynamic narrative transforms users from "readers" to "participants". Through the algorithm system, it responds to users' decisions and emotional states in real time, and promotes the development and improvement of design in the time dimension. The content of the narrative will achieve a greater extension than the design object itself. Dynamic narrative is also a narrative with multiple plot changes. It pursues innovative development of "event" story content, which can be divided into two types: "multiple things and multiple events" and "one thing and multiple events" [13]. This study pays special attention to the paradigm shift brought about by generative AI. On the one hand, generative AI expands the possible boundaries of dynamic narrative. On the other hand, it redefines the relationship between creators and users, which is also an area that current research has not yet touched.

In this study, "generative AI-driven" specifically refers to using a Large Language Model (LLM) based on the Transformer architecture as a narrative engine. Unlike traditional rule-based or branching interactive narrative systems, this system leverages the few-shot learning capability and contextual understanding mechanism of LLM to transform user interactions into coherent narrative text output in real time through carefully designed prompts, thereby achieving non-predefined dynamic generation of narrative paths.

2.4 Comparison of Technological Approaches for Generative AI -Driven Interactive Narrative

Compared to branching narratives, the classic implementation of branching narratives is based on finite state machines and tree structures. The narrative possibilities are limited by the exhaustive enumeration of pre-written branches, resulting in a combinatorial explosion in design complexity as the number of options increases [20]. This study uses the probabilistic generation of large language models to expand the narrative space from discrete nodes to a semantic continuum. Unlike procedural content generation methods, which rely heavily on manually compiled domain-specific rule bases [21], generative AI obtains general narrative knowledge through pre-training on massive corpora, significantly reducing the threshold for engineering narrative logic. Compared to planner-based narrative systems, such as those using STRIPS-type planners or intent-driven architectures to generate and coordinate character actions, these systems require researchers to formalize complete action domains, preconditions, and narrative target states [11]. This study only needs to set high-level narrative constraints and style guidance through prompts to maintain basic coherence while flexibly responding to users' open-domain input. Crucially, traditional methods mostly regard "interactivity" as a mechanism for accessing and selecting pre-made content [5], while this study places the AI generator at the core of the real-time narrative loop, so that each user interaction directly and uniquely triggers a new content generation, thereby realizing a paradigm shift from "story navigation" to "story symbiosis". This architectural difference is not only about implementation efficiency, but also fundamentally reshapes the possibility of narrative generation, the role of the creator, and the creative relationship of human-machine collaboration.

3. Research Design

In order to answer the research questions, this study adopted a mixed research approach, in which researchers conducted quantitative and qualitative research at the same stage of the research process [9]. The interactivity of the two narratives was compared by measuring participants' choices and emotional responses as well as their level of immersion.

3.1 Participants

This study recruited 40 participants (21 males and 19 females) from southeastern China on March 17, 2025. Participants were selected through campus recruitment. Participants were divided into two study groups. The average age of the participants was 25.5 years. 20 participants were randomly assigned to the experimental group (using a dynamic narrative system based on generative AI), which employed a large language model (ChatGPT) and immersive virtual reality support; 20 participants were randomly assigned to the control group (using a traditional fixed narrative system), which used a pre-set storyline and conducted the experiment using an Apple iPad, as shown in Figure 1.

The experiment used OpenAI's GPT-4 Turbo model (GPT -4-0125-preview version) as the narrative generation engine. Using this version ensured the uniqueness of the model weights, providing a reproducible technical benchmark for this study. This model was chosen based on its: ① 128K context window, which maintains consistency in long narratives; ② enhanced reasoning ability, capable of handling complex causal relationships; ③ multimodal understanding foundation, facilitating future expansion. API call parameter settings: temperature=0.7, max_tokens=500.

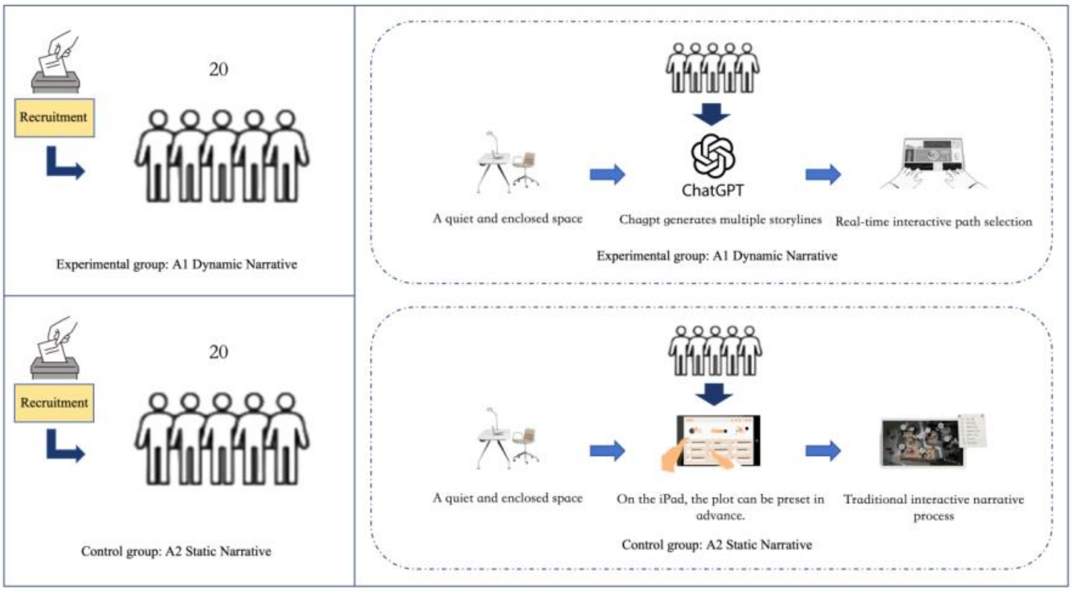

Figure 1. Design of the dual narrative control experiment

3.2 Experimental Procedure

As shown in Figure 2, the experimental procedure consisted of three phases, lasting approximately 60 minutes. Before the experiment, participants received training on the operation of the experimental equipment, and researchers explained the interactive logic of the dynamic narrative system. During the preparation phase (approximately 15 minutes), participants were asked to complete Worksheet 1, which included a pre-test of their narrative comprehension abilities and a questionnaire on their virtual reality usage experience. In the next 30 minutes, participants engaged in an immersive narrative experience. The experimental group experienced a dynamic narrative system supported by generative AI, with multiple storylines generated by GPT-4 and allowing real-time interaction and path selection. The control group experienced a linear, static narrative system with pre-defined plots, where participants could not choose the plot development. In the final 15 minutes, all participants were required to complete Worksheet 2, conducting quantitative and qualitative assessment tasks, including a narrative experience questionnaire and an emotional resonance survey.

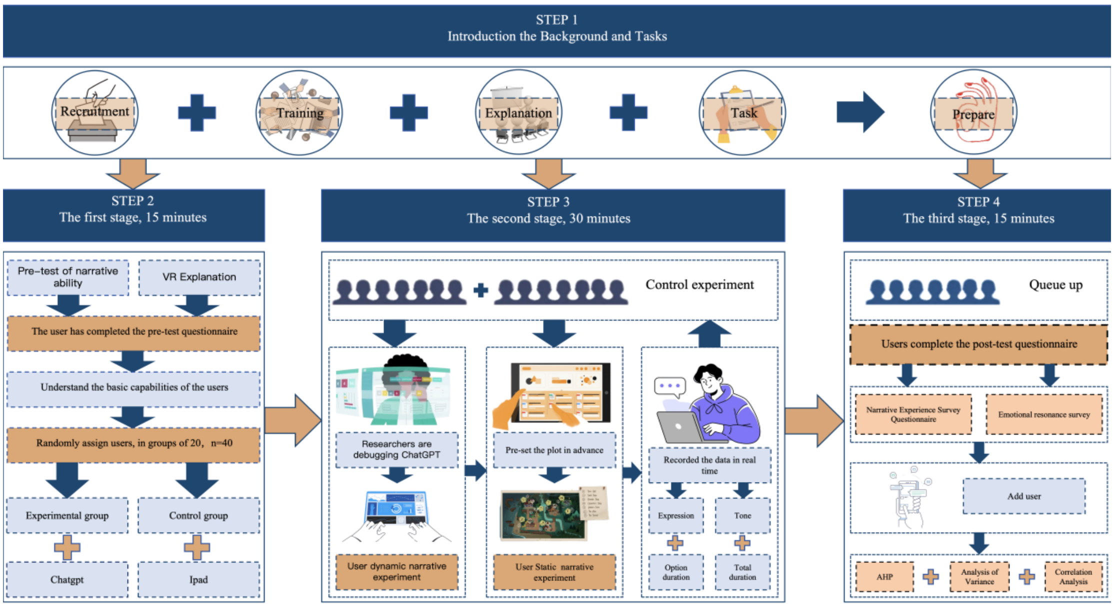

Figure 2. Experimental Procedure

3.3 Measuring Tools

The measurement tools used in this study are two worksheets. Worksheet 1 was used before the operation and theoretical training, and included a pre-test of the participants' narrative comprehension ability (as shown in Table 1) and a questionnaire on the participants' virtual reality usage experience (as shown in Table 2). The questionnaire items in Table 1 are based on Young's [11] narrative cognition framework, which includes six dimensions such as "event sequencing" and "causal reasoning", and are summarized into six options; the questionnaire items in Table 2 refer to Cavus N and Ibrahim D's [14] mobile learning device experience scale, and are adjusted to form six items. Worksheet 2 was used after the experiment and included a narrative experience questionnaire and an emotional resonance survey. In addition to the basic scale, the experiment recorded the decision-making pattern data of each participant, including behavioral indicators such as average decision time, path selection diversity, and number of backtracking, which were used for subsequent user characteristic analysis.

Table 1. Pre-test of Narrative Comprehension Ability

serial number	Questions	Rating (1-7)
Q1	I can clearly determine the chronological order of events in a story.	1 2 3 4 5 6 7
Q2	When I read or watch a story, I can quickly grasp the cause-and-effect relationships within it.	1 2 3 4 5 6 7
Q3	I can distinguish the structural differences between dynamic and static narratives.	1 2 3 4 5 6 7
Q4	I can accurately identify the beginning, climax, and ending of a story.	1 2 3 4 5 6 7
Q5	I can imagine multiple possible directions for the story.	1 2 3 4 5 6 7
Q6	I won't get confused or lost by complex story structures.	1 2 3 4 5 6 7

Table 2. Questionnaire on Virtual Reality User Experience

serial number	Questions	Rating (1-7)
Q1	I am familiar with VR (virtual reality) technology.	1 2 3 4 5 6 7
Q2	I have used VR devices for entertainment, learning, or social activities on numerous occasions.	1 2 3 4 5 6 7
Q3	I understand the operation methods of VR devices (gesture control, spatial positioning, etc.).	1 2 3 4 5 6 7
Q4	Every time I use VR devices, I can quickly adapt to the environment.	1 2 3 4 5 6 7
Q5	I feel immersive in the VR environment.	1 2 3 4 5 6 7
Q6	I am willing to continue using VR technology for immersive learning or story experiences in the future.	1 2 3 4 5 6 7

To ensure the validity and scientific nature of the experimental data, two structured pre-test questionnaires were designed and implemented, and the Likert seven-point scale was used for scoring [7], where 1 represents “very unfamiliar” and 7 represents “very familiar”, in order to assess the participants’ basic cognitive level and proficiency in using the technology.

In the post-experiment evaluation phase, a narrative experience questionnaire (as shown in Table 3) and an emotional resonance survey were designed and implemented. The questionnaire was based on the AHP analytic hierarchy process and six evaluation criteria were set in advance: immersion, personalized experience, emotional resonance, system usability, content diversity, and technical implementation difficulty. The relative importance of these six criteria was compared and scored pairwise [14]. The scoring standard adopted the "1-9 scale" proposed by Satty [15], thereby constructing a criterion-level judgment matrix. Next, the participants also need to score the relative performance of the "dynamic narrative" and "static narrative" systems under each criterion dimension (as shown in Table 4), thereby constructing a scheme-level judgment matrix.

Table 3. Questionnaire on the “Criterion Level” of Narrative Experience

Serial Number	Comparison items	Rating (1-9)
Q1	Immersive experience vs. personalized experience	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q2	Immersion vs. Emotional Resonance	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q3	Immersion vs. System Usability	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q4	Immersion vs. Content Diversity	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q5	Immersive experience vs. technical difficulty	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q6	Personalized experience vs. emotional resonance	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q7	Personalized experience vs. system usability	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q8	Personalized experience vs. content diversity	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q9	Personalized experience vs. technical difficulty of implementation	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q10	Emotional resonance vs. system usability	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q11	Emotional resonance vs. content diversity	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q12	Emotional resonance vs. technical difficulty	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q13	System availability vs. content diversity	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q14	System availability vs. technical implementation difficulty	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9
Q15	Content diversity vs. technical implementation difficulty	1 ,3, 5, 7, 9, or 1/3, 1/5, 1/7, 1/9

Table 4. Narrative Experience "Solution Layer" Scoring Table

Guidelines	Score
immersion	Dynamic narrative 1, 3, 5, 7, 9; Static narrative 1/3, 1/5, 1/7, 1/9
Personalized experience	Dynamic narrative 1, 3, 5, 7, 9; Static narrative 1/3, 1/5, 1/7, 1/9
Emotional resonance	Dynamic narrative 1, 3, 5, 7, 9; Static narrative 1/3, 1/5, 1/7, 1/9
System Availability	Dynamic narrative 1, 3, 5, 7, 9; Static narrative 1/3, 1/5, 1/7, 1/9
Content diversity	Dynamic narrative 1, 3, 5, 7, 9; Static narrative 1/3, 1/5, 1/7, 1/9
Technical difficulty	Dynamic narrative 1, 3, 5, 7, 9; Static narrative 1/3, 1/5, 1/7, 1/9

In addition, for the score explanation of the "1-9 scale", "1" means equally important, "3" means slightly important, "5" means obviously important, "7" means very important, "9" means extremely important, and "1/x" means that the compared party is more important than the compared party, to the same degree [15].

4. Data Analysis and Results

4.1 Data Analysis of Dual Narrative

In order to explore how to balance the “openness” and “logical coherence” of dynamic narrative, this study adopts the analytic hierarchy process, which is a multi-criteria decision-making method that combines qualitative and quantitative approaches. This method can quantitatively transform and analyze the user’s emotional issues in the decision-making process, thereby improving the objectivity and accuracy of decision-making. It is suitable for decision-making problems with many evaluation criteria, complex structures, and difficulty in quantification [16].

Based on the data obtained from the two tables, and combined with AHP, a three-layer structure model of dual narrative was constructed (as shown in Figure 3).

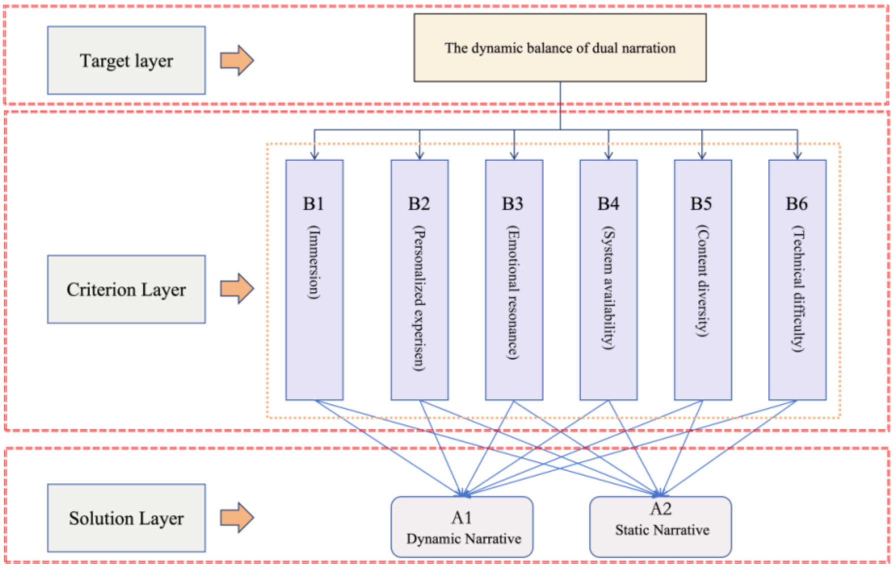

Figure 3. Dual Narrative Three-Layer Structure Model

For this model, the target layer is "dynamic balance of dual narratives". The criteria layer contains six dimensions: "immersion", "personalized experience", "emotional resonance", "system usability", "content diversity" and "technical implementation difficulty", named B1, B2, B3, B4, B5 and B6. The scheme layer is divided into "dynamic narrative" and "static narrative", named A1 and A2 [17].

Taking the criteria layer in the dual narrative three-layer structure model as an example: First, participants compare the six elements of this layer in pairs and score them according to the scale. Then, the average value of the judgment matrix results is taken to establish the judgment matrix (as shown in Table 5). Finally, the data is normalized and the weight ω of each element is calculated.

Table 5. Weights of Criterion-Level Elements in Dual Narrative Contrast

index	B1	B2	B3	B4	B5	B6	weight ω
B1	1	3	5	7	4	6	0.403
B2	1/3	1	3	5	2	4	0.246
B3	1/5	1/3	1	3	1/2	2	0.142
B4	1/7	1/5	1/3	1	1/4	1/2	0.077
B5	1/4	1/2	2	4	1	3	0.072
B6	1/6	1/4	1/2	2	1/3	1	0.060

After determining the weights ω of each element, to ensure the rationality of the hierarchical model structure and the usability of the judgment matrix, a consistency check is needed on the results of the judgment matrix. If the consistency ratio CR < 0.1, it indicates that the consistency check has been passed. The check steps are as follows:

(1) Find the largest eigenvalue λ_{max}.

$$\lambda_{\max} \approx \frac{1}{n} \sum_{i=1}^{n} \left(\frac{(Aw)_i}{w_i} \right)$$

In the formula, n is the order; A is the judgment matrix; and ω is the weight vector.

(2) Calculate the consistency ratio (CR).

$$CR = \frac{CI}{RI}$$

In the formula, CI is the consistency index $CI = \frac{\lambda_{\max} - n}{n-1}$; RI is the random consistency index corresponding to the order n (as shown in Table 6).

Table 6. Random Consistency Indicators

Order n	1	2	3	4	5	6	7	8	9	10
RI value	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

Taking the criteria layer elements in the dual narrative comparison requirement as an example, the calculated CR=0.013, which is less than 0.1, indicating that it passes the consistency test. Otherwise, the judgment matrix needs to be corrected.

After passing the consistency test, the weight values of each element in the criteria layer are multiplied by the weight values of each element in the scheme layer to obtain the comprehensive weight value of each requirement in the scheme layer. Then, the requirements are sorted according to the comprehensive weight value to obtain the requirement priority [15].

The higher the priority, the more important the need. Therefore, in the comparative study of the dynamic balance of dual narratives, the needs of high priority should be met as much as possible (as shown in Table 7).

Table 7. Demand Ranking

Demand Factors	Overall weight	Demand Priority
Dynamic Narrative	0.632	1
Static Narrative	0.368	2

This study takes a dual narrative system based on virtual reality technology as the research object, designs a three-stage experimental process, and conducts a systematic evaluation of users' narrative understanding, immersive experience and system optimization strategies through quantitative questionnaires, behavioral data recording and AHP hierarchical analysis [18]. According to the calculated results, compared with "static narrative", "dynamic narrative" is more effective in improving participants' "immersion", "personalized experience" and "emotional resonance"; the AHP results also further confirm that "immersion" and "personalized experience" are key criteria affecting narrative design optimization. And the comprehensive weight analysis shows that the "dynamic narrative" scheme has a higher weight and higher priority in most indicators.

4.2 Analysis of Variance

To accurately assess the advantages of dynamic narrative over static narrative and explore potential differences in user groups, this study conducted an analysis of variance (ANOVA). First, a one-way ANOVA was used to compare the performance of dynamic narrative (A1) and static narrative (A2) across six user experience metrics. The results were calculated using SPSS 22.0 (F is the mean, P is significance, and η^2 is the effect size) (as shown in Table 8). Before data analysis, the Shapiro-Wilk test ($p > 0.05$) and Levene test ($p > 0.05$) were used to verify that the data conformed to the assumptions of normal distribution and homogeneity of variance.

Table 8. Comparison of user experience scores between dynamic and static narratives (n=40)

Evaluation indicators	A1	A2	F	P	η^2
B1	5.82±0.71	4.13±1.02	28.64	< 0.001	0.428
B2	6.35±0.63	3.97±1.15	45.27	< 0.001	0.542
B3	5.91±0.68	4.05±0.93	32.18	< 0.001	0.471
B4	4.76±0.85	4.52±0.77	1.12	0.297	0.032
B5	6.12±0.59	3.45±1.21	38.91	< 0.001	0.506
B6	3.24±0.92	3.05±0.88	0.67	0.418	0.024

The results showed that dynamic narratives significantly outperformed static narratives in terms of immersion ($F=28.64$, $p<0.001$, $\eta^2=0.428$), personalized experience ($F=45.27$, $\eta^2=0.542$), emotional resonance ($F=32.18$, $\eta^2=0.471$), and content diversity ($F=38.91$, $\eta^2=0.506$) (all $p<0.001$, $\eta^2>0.14$ being a large effect). There was no significant difference between the two groups in system usability and technical implementation difficulty ($p>0.05$), indicating that the technology platform itself did not interfere with differences in experience.

Next, a two-way ANOVA was used to organize the above data to examine the interaction effects of different groups and genders on four indicators: "immersion" (Figure 4), "personalized experience" (Figure 5), "emotional resonance" (Figure 6), and "content diversity" (Figure 7). The horizontal axis represents the groups: dynamic narrative (A1) and static narrative (A2); the vertical axis represents the Likert 7-point scale score; purple represents males and blue represents females.

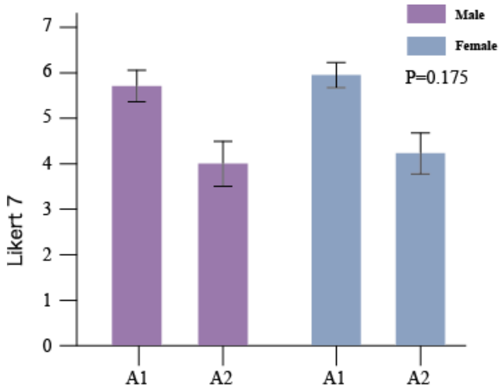

Figure 4. Interaction effect of group and gender on immersion

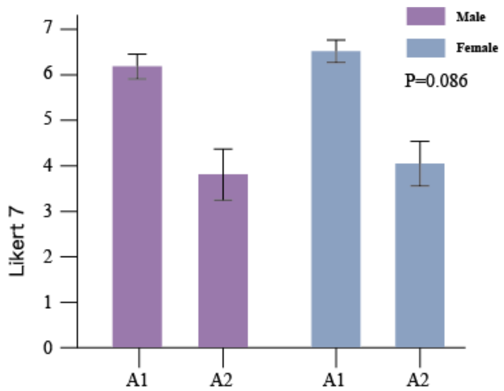

Figure 5. Interaction effect of group and gender on personalized experience

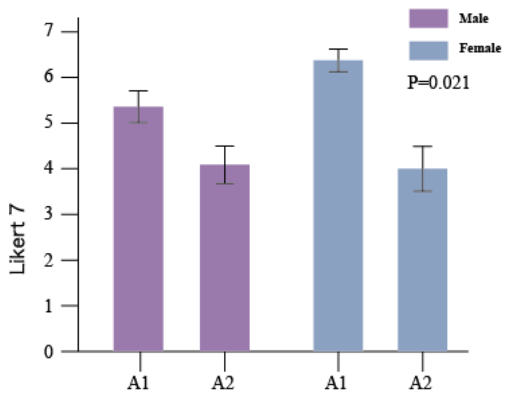

Figure 6. Interaction effect of group and gender on emotional resonance

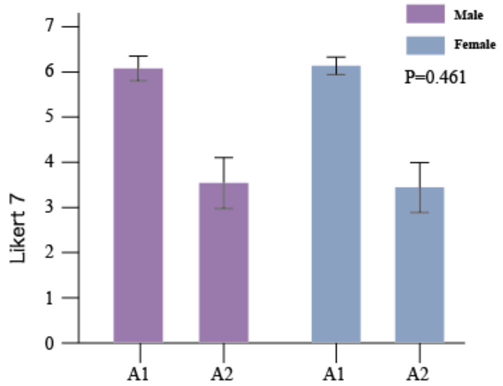

Figure 7. Interaction effect of group and gender on content diversity

The results showed (as shown in Table 9) that dynamic narrative had a significant main effect on all four indicators (all $p < 0.001$), but gender only moderated emotional resonance ($p = 0.021$). Women scored 1.02 points higher than men in emotional resonance in dynamic narrative (6.42: 5.40), which may be related to their higher emotional sensitivity to narrative details [13]. The gender differences in other indicators were not significant, indicating that the core advantages of dynamic narrative are universal.

Table 9. Group and Gender Scores for the Four Indicators

index	Group	male	female
B1	A1	5.40±0.73	6.42±0.51
	A2	4.10±0.85	4.00±1.01
B2	A1	5.72±0.75	5.92±0.67
	A2	4.05±1.10	4.21±0.94
B3	A1	6.22±0.65	6.48±0.59
	A2	3.88±1.20	4.06±1.10
B5	A1	6.08±0.62	6.16±0.56
	A2	3.52±1.18	3.38±1.24

4.3 Correlation Analysis

To investigate and discuss the intrinsic relationships among various dimensions of user experience, this study used Pearson correlation coefficients to analyze the correlation between core indicators (immersion, personalized experience, emotional resonance, system usability, and content diversity) and user characteristics (VR usage experience and narrative ability). Data was collected using a Likert 7-point scale and, after Shapiro-Wilk testing, was confirmed to conform to a normal distribution ($p > 0.05$), satisfying the preconditions for parameter testing. As shown in Figure 8, a heatmap of user experience indicator correlations was created based on the data; red indicates a strong positive correlation ($r > 0.7$), and blue indicates a negative correlation.

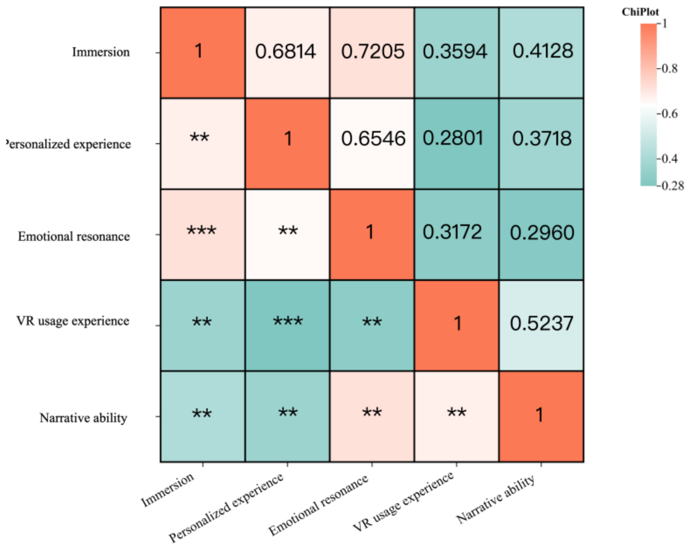

Figure 8. Heatmap of correlation between user experience metrics (*: p < 0.05; **: p < 0.01; ***: p < 0.001)

The results (as shown in Table 10) show that immersion and emotional resonance (r=0.7205) are highly synergistic, indicating that improving immersive design can directly enhance emotional connection; VR experience is moderately correlated with immersion (r=0.35), suggesting that tutorials should be added for beginners; personalized experience is not related to VR experience (r=0.28, p=0.09), indicating that the personalized advantages of dynamic narrative have universality.

Table 10. Correlation Matrix among User Experience Metrics (n=40)

Index	Immersion	Personalized experience	Emotional resonance	VR usage experience	Narrative ability
Immersion	1	0.6814	0.7205	0.3594	0.4128
Personalized experience	0.6814	1	0.6546	0.2801	0.3718
Emotional resonance	0.7205	0.6546	1	0.3172	0.2960
VR usage experience	0.3594	0.2801	0.3172	1	0.5237
Narrative ability	0.4128	0.3594	0.7205	0.6814	1

5. Conclusion and Outlook

5.1 Conclusion and Discussion

This study utilizes a dual narrative mechanism mediated by generative AI, based on a VR platform, and employs technological means to construct a comparative experimental framework of dynamic and static narratives. Through a hybrid research approach, it systematically evaluates key experiential dimensions such as "immersion," "personalized experience," and "emotional resonance." The calculations yield the following important findings.

First, the model constructed through AHP hierarchical analysis points out that "immersion" and "personalized experience" are the two high-weight dimensions that should be prioritized in the optimization process of interactive narrative systems, and these two points should be the primary focus. In contrast, static narratives, due to their inherent limitations, despite having a story structure that combines completeness and logic, are unable to meet the current development trends of "decentralization" and "personalization" in interactive narratives.

Secondly, in the four core dimensions of immersion (η^2= 0.428, p<0.001), personalized experience (η^2=0.542, p<0.001), emotional resonance (η^2=0.471, p<0.001), and content diversity (η^2=0.506, p<0.001), the dynamic narrative system significantly outperformed the static narrative system. Furthermore, effect size analysis showed that the improvement in personalized experience was the most significant, $\eta^2 > 0.5$, confirming the unique value of generative AI in providing personalized content. Furthermore, through two-way ANOVA, female users scored significantly higher than male users in emotional resonance in dynamic narratives. This difference is likely due to women's high emotional sensitivity to narrative details, consistent with previous research findings on gender and empathy.

Finally, correlation analysis revealed a strong coupling relationship between immersion and emotional resonance ($r=0.72$, $p<0.001$), while VR usage experience showed a moderate correlation with immersion ($r=0.35$, $p=0.028$).

Therefore, generative AI provides a novel approach to interactive storytelling. Its supported dynamic narrative mechanism not only overcomes the limitations of fixed content in traditional narratives at the technical level, but also achieves deeper interaction and emotional connection between the user and the narrative content at the user experience level. The findings of this paper not only provide a new evaluation model for the optimized design of dual-narrative systems, but also offer direction for future exploration of digital storytelling integrating AI and VR.

5.2 Research Limitations and Boundary Conditions

The conclusions of this study need to be understood under specific boundary conditions.

Firstly, in this study, 40 participants from universities in eastern China were randomly assigned to groups. However, the demographic homogeneity of this group limited the cross-cultural and cross-age applicability of the findings, and the generalization of the conclusions to a wider population requires careful verification. Furthermore, while the experimental group used VR headsets and the control group used tablet devices, although statistical tests in our analysis confirmed no differences between groups in "system availability" and "technical implementation difficulty," the inherent baseline differences in immersion between the two media could still serve as confounding variables, affecting the pure comparison of core indicators such as "immersion."

Secondly, the core intervention in this study is dynamic narrative generation based on the GPT- 4 Turbo model (gpt-4-0125-preview version). This technology itself has an inherent risk of content coherence—large language models may experience contextual forgetting or logical jumps in long narrative sequences. Although we have constrained it through cueing engineering, the long-term artistic consistency of its generated content still needs to be examined.

At the same time, while a short 30-minute experience can effectively capture the initial experience and novelty effect, it cannot reveal the interaction fatigue or changes in the perception of narrative depth that users may experience after long-term, repeated use of a dynamic narrative system.

Therefore, the conclusions of this study are mainly applicable to plot-driven, moderately complex VR narrative experiences. For narrative types that heavily rely on character depth or thematic consistency, the applicability of generative AI solutions may require different design frameworks and evaluation criteria.

5.3 Future outlook

In recent years, the frequent appearance of concepts such as "artificial intelligence", "narrative intelligence", "virtual narrative", "5G technology" and "metaverse" indicates that artificial intelligence is profoundly reshaping the research paradigm of interactive narrative. Since 2021 was hailed as the "first year of metaverse", related technologies have evolved rapidly: in March 2023, OpenAI released the multimodal pre-trained model GPT-4; in February 2024, the video generation model Sora was launched, further promoting the development of AIGC. At present, AIGC has been widely used in film and television production, game design and art creation. With advanced algorithms such as deep learning, image generation and loop generation models, it has not only improved the personalization and interactivity of narrative, but also provided richer content support and expression forms for cross-media and multimodal narratives, and expanded the user's freedom and sense of participation in virtual situations. In the future, artificial intelligence will continue to advance the boundaries of interactive narrative research in terms of intelligent content generation, construction of virtual narrative space, dynamic interaction mechanism and narrative dissemination form, while providing new technical paths and theoretical support for virtual space design and user experience optimization in the field of public cultural services.

References

[1] Zhao, X., & Peng, D. (2022). Research on digital narrative mode of archives in digital empowerment environment. Archival Science Research, (5), 67–73.

[2] Baldoni, S., Hadj Sassi, M. S., Carli, M., et al. (2024). Definition of guidelines for virtual reality application design based on visual attention. Multimedia Tools and Applications, 83(16), 49615–49640. https://doi.org/10.1007/s11042-024-18672-7

[3] Koenitz, H., Ferri, G., Haahr, M., et al. (2015). Interactive digital narrative: History, theory and practice.

[4] Richards, D., Lupack, S., Bilgin, A. A. B., et al. (2023). Learning with the heart or with the mind: using virtual reality to bring historical experiences to life and arouse empathy. Behaviour & Information Technology, 42(1), 1–24. https://doi.org/10.1080/0144929X.2022.2148621

[5] Ryan, M. L. (2012). Narrative as virtual reality 2: Revisiting immersion and interactivity in literature and electronic media.

[6] Xu, C., & Xu, X. (2022). Digital display design of cultural heritage from the perspective of interactive narrative. Packaging Engineering, 43(12), 241–249.

[7] Chai, Y., & Yin, Y. (2024). Immersive experience design of scenic area public facilities from the perspective of interactive narrative. Packaging Engineering, 45(16), 403–412.

[8] Green, D., Hargood, C., & Charles, F. (2021). Use of tools: UX principles for interactive narrative authoring tools. Journal on Computing and Cultural Heritage, 14(3), 1–25. https://doi.org/10.1145/3450341

[9] Huang, X., & Macgilchrist, F. (2024). From physical feelings to empathy: An immersive virtual reality approach to facilitate physical empathy. Computers and Education Open, 7, 100215. https://doi.org/10.1016/j.caeo.2024.100215

[10] Kitromili, S., Hargood, C., Millard, D., et al. (2022). Locative authoring: evaluating the StoryPlaces authoring tool. In International Conference on Interactive Digital Storytelling (pp. 483–498). Springer International Publishing. https://doi.org/10.1007/978-3-031-22224-2_36

[11] Young, R. M. (2007). Story and discourse: A bipartite model of narrative generation in virtual worlds. Interaction Studies, 8(2), 177–208. https://doi.org/10.1075/is.8.2.04you

[12] Thiel, T. (2004). Beyond Manzanar: Creating dramatic structure in Ergodic narratives. In International Conference on Technologies for Interactive Digital Storytelling and Entertainment (pp. 246–251). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-30183-1_24

[13] Yang, X. (2019). A Preliminary Discussion on Dual Narratives and Their Typological Characteristics in Design. Chinese Literature and Art Review, (6), 47–53.

[14] Cavus, N., & Ibrahim, D. (2017). Learning English using children's stories in mobile devices. British Journal of Educational Technology, 48(2), 625–641. https://doi.org/10.1111/bjet.12421

[15] Yang, Y. (1999). Research on the adjustment method of inconsistency judgment matrix in the analytic hierarchy process. Operations Research and Management, (3), 12–16.

[16] Wang, Y., Sun, L., & Yang, X. (2025). Research on AI-assisted barrier-free intelligent wheelchair design based on FBS and AHP. Packaging Engineering, 46(4), 172–182.

[17] Pan, C., Pan, Z., & Zhao, D. (2024). Research on the design of elderly care service robots driven by user needs. Packaging Engineering, 45(22), 41–55.

[18] Xiang, B., Huang, G., & Feng, J. (2011). Fuzzy comprehensive evaluation of the health of the water-based system in the Dongjiang River Basin based on AHP method. Hydropower Energy Science, 29(10), 1–4.

[19] Zhao, X., & Peng, D. (2023). A Study on Digital Narrative Experience in Archives Based on Embodied Interactive Narrative. Archival Science Research, (5), 88–97.

[20] Lai, J., & Ranjan, R. (2026). Applying Jobs-To-Be-Done Framework to Design AI Applications for Women. In HCI International 2025 – Late Breaking Papers, Part II (pp. 133). Springer Nature. https://doi.org/10.1007/978-3-031-93405-8_11

[21] Mateas, M., & Stern, A. (2003). Façade: An experiment in building a fully-realized interactive drama. In Game Developers Conference (Vol. 2, pp. 4–8).