https://doi.org/10.37420/j.cer.2025.016

Examining the Cognitive Communication Effects of AI Text-to-Video in Animated News

Zihao Wang*
School of intelligent science and technology, Xi 'an University of Science and Technology, China
***Corresponding author E-mail: 2464658798@qq.com**

Abstract:
With the rapid advancement of generative artificial intelligence technology, the application of AI Text-to-Video in the field of news dissemination is becoming increasingly widespread, posing challenges to traditional animated news formats. This study employs experimental methods and in-depth interviews to explore the differences in user cognitive effects, emotional attitudes, and media preferences between AI Text-to-Video in animated news and traditional animated news. The findings reveal that AI Text-to-Video in animated news holds an advantage in initial cognitive effects and emotional responses, whereas traditional animated news performs better in long-term memory retention. Additionally, user preferences for the two news formats tend to be neutral, indicating their functional complementarity in news production. In the future, news production should emphasize " human-machine collaboration," with journalists transitioning from "content producers" to " planners and reviewers." Simultaneously, optimizing AI emotional models is essential to enhance the authenticity and diversity of news.

Keywords:
AI Text-to-Video, animated news, cognitive communication, emotional response, media preference

1 Introduction

In the process of animated news production at the scene, journalists often face the challenge of arriving after the news event has concluded, making it difficult to obtain timely information. Simultaneously, with the rise of Manuel Castells' network society, the internet has transformed into a fluid space, significantly promoting the widespread adoption of distributed network technologies. As a result, the latest updates from news scenes are often released first by "grassroots journalists." To address this, media outlets such as Xinhua News Agency and The Beijing News have explored the "staged reenactment" (a method of recreating news events through animation or live reenactments) model of animated news, using animation or live reenactments to help audiences clearly understand the actual situation of news scenes.

However, with the explosive growth of generative artificial intelligence technology, the application of AI Text-to-Video in the media industry has achieved continuous breakthroughs, challenging the traditional "staged reenactment" model of animated news. This raises the question: What are the differences in cognitive communication effects between traditional animated news and AI Text-to-Video in animated news? How can news media leverage AI Text-to-Video to enhance news production efficiency?

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

2 Literature Review

2.1 Cognitive communication effects

Cognitive communication effects represent a research field that integrates audience cognitive theory and communication effects theory. Domestic research primarily focuses on theoretical discussions, with limited empirical studies. Keqin Lin and Jingping Zeng argue that cognitive communication emphasizes that all human cognitive activities are based on bodily experiences of the external world. Jingjing Han and Haoqing Yu further point out that cognitive communication falls within the realm of media psychology, focusing on how the brain and mind receive and process information in communication contexts. With the advancement of mobile internet technology, many scholars have conducted research on cognitive communication in the context of new media technologies. For example, Zhaoyang Wang and Jiejie Wei, through controlled experiments, found that stronger emotional attitudes among mobile short video news users lead to better cognitive effects, and differences in content arrangement also influence users' cognitive effects. In contrast, studies by foreign journals are predominantly empirical. Kexin Song, through a survey on public trust in experts, revealed that factors such as experts' professional knowledge, reputation, communication skills, and the transparency of the online environment influence public cognitive patterns. While scholars both domestically and internationally have conducted extensive research on cognitive communication effects, most studies focus on theoretical development and integration with psychology, leaving the questions raised in this study unanswered.

2.2 Animated news phenomenon

Based on a systematic review of existing literature, academic research on the animated news phenomenon began in 2010. Domestic scholar Yanbin Sun first defined animated news in his article *"animated news is Just Misplaced"*, describing it as a news reporting model that presents newspaper-reported news in animated form. Shuhua Wang supplemented the definition, proposing that animated news uses dynamic and visual animations to reconstruct and interpret the process of news events, enhancing the vividness of news reporting. Additionally, domestic scholar Mengli Xie pointed out that early animated news primarily relied on hand-drawn animations, while modern animated news is created using software such as AE and 3DMax. Wenqian Xu also highlighted new characteristics of animated news, including fragmentation, entertainment value, storytelling, and readability. Despite the extensive domestic research on animated news, most studies focus on its content characteristics, challenges, current status, and narrative expression, with limited exploration of its communication effects.

2.3 AI Text-to-Video

As an emerging field, research on AI Text-to-Video by domestic and international scholars is still in its infancy. A search on CNKI using "AI Text-to-Video" as the keyword, limited to the field of news and media, yields only 38 relevant papers. Anbin Shi and En Zheng argue that AI Text-to-Video, as an extension of the GPT-4 model in the field of visual communication, represents an advanced machine learning model capable of generating video content through AI. Sha Xun suggests that AI Text-to-Video is a technology that uses artificial intelligence to transform user-input text or images into video content, extending users' sensory and emotional thinking. Fengming Wen and Yuqun Wen explored the role of AI Text-to-Video in various stages of

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

short video content planning, shooting, post-production, and presentation. Although the development of this communication technology has entered a fast lane, scholars' focus remains primarily on technological development and application, making it difficult to effectively address the questions raised in this study.

2.4 Research questions

A review of domestic and international literature reveals that existing research primarily focuses on single areas such as cognitive communication effects, animated news, and AI Text-to-Video, lacking comprehensive studies that integrate all three.

Based on this, the study aims to address the following questions:

Q1:Compared to traditional animated news, does AI Text-to-Video in animated news elicit stronger emotional reactions?

Q2:Compared to traditional animated news, does AI Text-to-Video in animated news promote deeper cognitive effects among users?

Q3:Compared to traditional animated news, is AI Text-to-Video in animated news more preferred by users?

Q4:Compared to traditional animated news, what are the advantages and disadvantages of AI Text-to-Video in animated news in the news production process?

3 Research Hypotheses and Methodology

3.1 Elaboration of research ideas

This study adopts a positivist research paradigm, combining experimental methods (Experiment 1 and Experiment 2) and in-depth interviews (Experiment 3) to compare the effects of AI Text-to-Video in animated news and traditional animated news on users' cognitive effects, emotional attitudes, and media preferences. It explores the advantages and disadvantages of AI Text-to-Video in animated news in the news production process, verifies research hypotheses through data analysis, and ultimately draws conclusions.

3.2 Research hypotheses

Based on the questions raised above, this study proposes the following research hypotheses:

H1: There is a significant difference in users' cognitive effects between AI Text-to-Video in animated news and traditional animated news, with AI Text-to-Video in animated news yielding better cognitive effects.

H2: There is a significant difference in users' emotional attitudes between AI Text-to-Video in animated news and traditional animated news, with AI Text-to-Video in animated news eliciting stronger emotional responses.

H3: There is a significant difference in users' media preferences between AI Text-to-Video in animated news and traditional animated news, with AI Text-to-Video in animated news being more favored by users.

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

3.3 Experimental method

3.3.1 Experimental environment

In this experiment, participants can use their own smartphones or computers to watch videos. One day before the experiment, the experimenter will schedule the experiment time with participants to ensure stable internet connectivity, a relaxed and focused state of mind, and a quiet and comfortable environment to avoid external interference and noise. During the experiment, participants are divided into three WeChat groups, with eight participants in each group. The experiment is conducted over eight days, with the first part on Day 1 and the second part on Day 8.

3.3.2 Participants

To test the hypotheses, this study employs a controlled experimental method, recruiting participants through Xiaohongshu, QQ groups, and personal referrals. To avoid gender bias, the recruited participants are evenly split between males and females. Before the experiment, participants are randomly assigned to Groups A, B, and C. The experiment includes 12 males and 12 females, totaling 24 participants, with each group consisting of 4 males and 4 females.

3.3.3 Experimental design

To accurately assess participants' cognitive levels and abilities and minimize bias caused by individual differences, a baseline cognitive ability test is administered to the 24 participants before the formal experiment. The test covers five areas: attention, memory, logical reasoning, language ability, and mathematical ability, with a total of 5 questions worth 1 point each. Participants scoring below 4 points are excluded. In this experiment, all 24 recruited participants passed the test and are eligible for the formal experiment.

The experiment consists of two parts. The first part is the initial test (Experiment 1), where Group A learns about the event through other channels without watching a video, Group B watches AI Text-to-Video in animated news, and Group C watches traditional animated news. Subsequently, a questionnaire is used to evaluate each group's cognitive effects and emotional attitudes toward the news details. The second part is the crossover test (Experiment 2), conducted one week after the initial test (Experiment 1). In this phase, Group C watches AI Text-to-Video in animated news, and Group B watches traditional animated news. The questionnaire is used again to assess changes in cognitive effects and media preferences, further verifying the differences in cognitive communication effects between the two news formats.

Considering event exposure and memory costs, the stimulus materials for this experiment are selected to have low event exposure and low memory requirements to avoid cognitive interference and memory demands affecting the results. Therefore, this experiment uses a social news story titled "*Man Trapped at the Bottom of a Well in Shaanxi Rescued After 4 Hours"* from April 2024 as the stimulus material. This news story contains complete elements and reliable sources, meeting the experimental requirements. The traditional animated news material is sourced from Jiupai News, affiliated with Changjiang Daily, while the AI Text-to-Video in animated news is produced using the audiovisual content AI model "Runway."

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

3.3.4 Questionnaire design

The questionnaire is divided into four sections.

The first section is the baseline cognitive ability test, designed to assess participants' foundational cognitive levels and abilities, ensuring they meet the experimental requirements. The test content has been mentioned earlier and is applicable to the initial test (Experiment 1).

The second section is the memory test, aimed at analyzing the differences in users' cognitive effects between AI Text-to-Video in animated news and traditional animated news, with a focus on users' memory retention of event-related information. It consists of 10 questions, each worth 1 point, with a total score of 10. Incorrect answers receive no points.

The third section is the user emotional attitude assessment, designed to measure users' emotional responses after watching the news and evaluate the differences in emotional attitudes between AI Text-to-Video in animated news and traditional animated news. It includes 10 questions, with evaluation indicators such as sympathy,admiration, tension, moved, concern, warmth, trust, recognition, relief, and inspiration. These two sections are applicable to both the initial test (Experiment 1) and the crossover test (Experiment 2).

The fourth section is the user media preference assessment, designed to measure users' preferences for different news formats and analyze the differences in media preferences between AI Text-to-Video in animated news and traditional animated news. It consists of 10 questions, with evaluation indicators including attractiveness, comprehensibility, innovation, interest stimulation, information dissemination capability, visual impact, viewing experience, future usage intention, and social media applicability. This section is applicable to the crossover test (Experiment 2). Both the third and fourth sections use a 5-point Likert scale, where 1 indicates "strongly disagree" and 5 indicates "strongly agree." Higher scores indicate more positive emotional attitudes or stronger media preferences.

3.4 In-depth interview method

3.4.1 Sampling method

The sampling method for this interview is purposive sampling, selecting representative news professionals as interviewees. The sampling criteria include: (1) covering media workers with different years of experience to ensure diversity in interview results; (2) selecting news professionals from different fields and positions to reflect diverse perspectives on news production and dissemination practices. Ultimately, five interviewees meeting the sampling criteria were identified through professional networks and other means.

3.4.2 Interviewees

Based on the findings from the initial test (Experiment 1) and the crossover test (Experiment 2), the in-depth interview (Experiment 3) selected five media industry professionals as interviewees. These respondents have extensive experience in the media industry and can provide in-depth analysis and evaluation from a professional perspective on the cognitive communication effects, emotional responses, media preferences, and media management of AI Text-to-Video in animated news compared to traditional animated news. The demo-

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

graphic characteristics of the interviewees are shown in Table 1.

Table 1 Demographic Characteristics of Interviewees

Number	Age	Gender	Occupation	Duration of employment
A01	25	Male	Advertising Editor at Provincial Media	Less than 5 years
A02	35	Female	Publicity Officer at University	5–10 years
A03	35	Male	Former News Professional	More than 10 years
A04	23	Male	Short Video Journalist at Provincial Media	Less than 5 years
A05	28	Female	Editor and Reviewer at National Media	Less than 5 years

3.4.3 Interview process

The interviews were conducted in a semi-structured manner, covering topics such as the interviewees' basic information, their perspectives on the cognitive communication effects of AI Text-to-Video in animated news versus traditional animated news, audience emotional responses, media preferences, and media management. The interviews were conducted via telephone, with each interview lasting approximately 45 minutes. The interview process took place from February 8, 2025, to February 10, 2025. Open-ended questions were used to guide the interviewees into in-depth discussions, ensuring the richness and depth of the interview content. All interviews were recorded and transcribed into a 6,085-word verbatim transcript for subsequent analysis.

3.4.4 Ethical considerations

The entire interview process strictly adhered to academic ethical standards. First, before the interviews, each interviewee was provided with a detailed explanation of the research purpose, interview content, and how the interview text would be used, and informed consent was obtained. Second, the privacy of the interviewees was respected and protected, with all interview data anonymized.

4 Research Findings

4.1 Primacy effect and sensory priority: AI Text-to-Video in animated news holds initial cognitive advantages

AI Text-to-Video in animated news demonstrates significant cognitive advantages during initial information reception, enhancing users' memory effects through the primacy effect, multi-sensory stimulation, and reduced cognitive load. According to the memory test results from the initial test (Experiment 1) (Table 2), Group B, which was first exposed to AI Text-to-Video in animated news, achieved the highest memory score (8.5 points), significantly higher than Group C's score of 7.625 points when first exposed to traditional animated news. This result indicates that AI Text-to-Video in animated news has a notable cognitive advantage during initial information reception. This phenomenon can be explained by the primacy effect theory, which

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

posits that individuals have stronger memory encoding capabilities for information encountered for the first time, especially during the early stages of information presentation, where novel and unique information leaves a deeper impression. During the interviews, interviewees unanimously agreed that the novel format and multi-sensory stimulation of AI Text-to-Video in animated news quickly capture attention and enhance memory effects during initial exposure.

"When I first watched AI Text-to-Video in animated news, it engaged multiple senses, which activated more brain regions and enhanced memory." (Interviewee A05) The visual impact and fragmented narrative structure of AI Text-to-Video reinforce users' initial cognitive experiences through multi-sensory stimulation, aligning with the "sensory priority" logic of communication. When information is presented simultaneously through multiple sensory channels such as vision and hearing, it activates multiple brain regions, enhancing information encoding and storage. As Interviewee A05 stated: "In my work, I have conducted similar data analyses, and the results showed that users spent 35% more time on AI Text-to-Video compared to traditional news, with emotional keywords like 'shocking' appearing 69% more frequently in the comments section."

Individuals' cognitive resources are limited, and when information presentation reduces cognitive load, users can process information more effectively. In fact, the initial cognitive advantage of AI Text-to-Video in animated news does not solely stem from the technology itself but is achieved by optimizing user cognitive effects through short duration and strong visuals that reduce cognitive load. Interviewee A04 noted, "AI visuals are more impactful, with customizable camera movements and styles that align with short video users' habits, enhancing memory points and emotional resonance." This also creates "cognitive novelty points" for users. "Audiences are already familiar with AI through various media reports, and their first exposure to AI Text-to-Video effectively activates this cognitive memory. Novelty, curiosity, and exploration can all serve as descriptors for the optimal performance of memory retention and emotional identification." (Interviewee A03)

Table 2 Memory Test Results

Group	Experiment 1 (Average Score)	Experiment 2 (Average Score)	News Type (Experiment 1 / Experiment 2)
A	6.625	Not Participated	—
B	8.5	9.25	AI Traditional
C	7.625	8.125	Traditional AI

4.2 Cognitive schema and humanized narrative: the sustained memory advantage of traditional animated news

Traditional animated news holds an advantage in deep memory consolidation due to its coherent narrative, high information density, and humanized emotional expression. These features can activate users' cognitive schemas. Although AI Text-to-Video in animated news demonstrates clear advantages in initial cognition, the memory test results from the crossover test (Experiment 2) (Table 2) show that Group B, after switching to traditional animated news, further improved their memory score to 9.25 points, while Group C, after switch-

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

ing to AI Text-to-Video in animated news, only slightly increased to 8.125 points. This indicates that traditional animated news has a stronger advantage in sustained memory retention.

According to the Elaboration Likelihood Model (ELM), users' deep processing of information relies on logical coherence and the activation of cognitive schemas. Cognitive schemas are frameworks formed through long-term experience that help users quickly understand and memorize new information. traditional animated news, with its coherent narrative structure and rich information density, triggers users' central route processing, engaging their "slow thinking" mechanisms and connecting with existing cognitive schemas, thereby consolidating long-term memory. Interviewee A05 noted, "traditional animated news is more coherent, making it easier for audiences to activate cognitive schemas and associate information." Interviewee A04 added, "traditional animated news is often based on in-depth investigations, multi-source verification, and contextual supplementation, resulting in higher information density." Using flood reporting as an example, they pointed out that traditional animated news requires a combination of "on-site visits and software production," with far greater detail accuracy than AI Text-to-Video, making it easier to form long-term memories in users' minds.

Additionally, the "humanized" characteristics of traditional animated news in emotional expression and narrative style are also key factors in its sustained memory advantage. During the interviews, Interviewee A03 emphasized, "AI Text-to-Video often lacks depth in emotional expression and narrative support, resulting in superficial memory retention." Interviewee A01 also mentioned, "AI news tends to stay on the surface of events, while the human touch in traditional news makes it more empathetic and trustworthy, naturally leading to sustained advantages." Essentially, this sustained advantage results from the combined effects of "content authority" and "narrative humanization." However, its drawbacks, such as long production cycles and high costs, cannot be ignored, especially in breaking news scenarios where immediacy is crucial. Therefore, Interviewee A05 suggested that formulaic reporting in fields like finance and sports could be fully automated using AI, freeing up human resources for in-depth investigations.

4.3 Cognitive dissonance and emotional models: divergent characteristics in user emotional identification

AI Text-to-Video in animated news elicits stronger emotional resonance during initial exposure, but traditional animated news has a more lasting emotional impact. This difference primarily stems from cognitive dissonance, generational differences among users, and the mechanical nature of AI's emotional expression. The emotional test results from the initial test (Experiment 1) and the crossover test (Experiment 2) (Table 3) show that Group B scored the highest in both tests, with a particularly high score (44.875) when first exposed to AI Text-to-Video in animated news, which slightly decreased (43.5) after switching to traditional animated news. In contrast, Group C had lower emotional identification scores, with an initial score of 36.125 when using traditional animated news and no significant improvement (36.5) after switching to AI Text-to-Video in animated news. This suggests that AI Text-to-Video in animated news may be more effective at eliciting initial emotional resonance, while traditional animated news has a stronger sustained impact.

This difference can be explained by cognitive dissonance theory, which posits that when new information conflicts with existing beliefs, users adjust their attitudes or behaviors to reduce psychological discomfort. The stronger emotional response of Group B to AI Text-to-Video in animated news compared to traditional

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

animated news may stem from the novelty of the technology and its low cognitive load, while Group C's low identification may reflect their entrenched media usage habits. Additionally, during the interviews, Interviewee A01 observed that older users criticized AI news as "fake," while younger users found it "interesting." This generational difference aligns with the characteristics of "early adopters" and "laggards" in the Diffusion of Innovations theory. Younger users, as "early adopters," are more receptive to the novelty and technological advantages of AI, while older users, as "laggards," tend to rely on traditional news formats.

Interviewee A02 pointed out, "Previously, using AI to generate content about architecture or policy introductions didn't resonate well with our audience. They found traditional animated news more relatable, and I believe AI news lacks a stylistic model." Interviewee A04 attributed the mechanical nature of AI's emotional expression to its reliance on training data, stating, "It might be limited by algorithms and material libraries, making it seem mechanical." This mechanical emotional expression struggles to evoke deep emotional resonance, especially in news events requiring complex emotional narratives. The divergence in user emotional identification ultimately reflects differences in the acceptance threshold for AI's emotional models. Therefore, optimizing the emotional expression capabilities of AI Text-to-Video in animated news and enhancing its acceptance across different user groups are crucial. "In the future, we need to optimize emotional models by incorporating elements like dialect narration and regional cultural features to bridge generational and regional gaps." (Interviewee A01)

Table 3 Emotional Test Results

Group	Experiment 1 (Average Score)	Experiment 2 (Average Score)	News Type (Experiment 1 / Experiment 2)
A	40.375	Not Participated	—
B	44.875	43.5	AI → Traditional
C	36.125	36.5	Traditional → AI

4.4 Use and gratification and functional complementarity: the underlying logic of neutral media preferences

Users show no significant preference for either AI Text-to-Video in animated news or traditional animated news, as their choices are based on specific needs and usage contexts. This reflects the instrumental rationality of news consumption and the "functional complementarity" of the media ecosystem in the intelligent communication environment. The media preference evaluation results (Table 4) show that the preferences of Group B and Group C for AI Text-to-Video in animated news and traditional animated news are similar, indicating that users' media preferences are neutral and not inherently biased toward one form of news. As emphasized by the Uses and Gratifications Theory, users' choice of media form depends on their specific needs rather than the inherent characteristics of the media form itself. Interviewee A01 noted that government

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

clients prefer the authority of traditional formats, while cultural and tourism clients pursue the visual appeal of AI. This differentiated demand suggests that users choose news formats based on specific usage scenarios and functional needs rather than an inherent preference for a particular media form.

Furthermore, the neutral tendency in users' media consumption reflects a shift toward instrumental rationality in news consumption behavior. In a mediatized society, users switch media based on communication contexts and individual needs, demonstrating enhanced audience agency. "WeChat users lean toward traditional formats, while Douyin users prefer AI. Post-00s users seek a hybrid experience of 'information + entertainment.' " (Interviewee A05) This diversity in media choices not only reflects the diversity of user needs but also reveals a media ecosystem characterized by "functional complementarity" rather than "substitution" in the intelligent communication environment. In other words, AI and traditional media are not mutually exclusive but instead leverage their respective strengths in different scenarios based on user needs.

From the perspective of media ecology, the diversity and complementarity of media forms are key to the healthy functioning of the media ecosystem. As Interviewee A03 observed, "More media outlets are joining the trend of adopting AI news, somewhat reflecting a 'if you can't beat them, join them' mentality." This essentially reflects the adaptation and transformation of media in response to changes in users' media consumption habits. Such changes are both a passive choice in the face of media ecosystem reconstruction and an active attempt to embrace user-centric practices.

Table 4 Media Preference Evaluation Results

Group	Average Score
B	28.5
C	28.625

4.5 Human-Machine collaboration and ethical challenges: reconstructing the news production paradigm

While AI Text-to-Video technology can reduce costs and improve efficiency, it also faces limitations such as authenticity risks, homogenization, and insufficient personalized expression. In the future, a "human-machine symbiosis" news production paradigm needs to be established. During the interviews, interviewees generally acknowledged the cost-saving and efficiency-enhancing value of AI Text-to-Video technology. For example, Interviewee A01 stated, "In our work, it saves at least 50% of manpower." The introduction of new technology can indeed significantly improve production efficiency and optimize resource allocation. However, interviewees also expressed concerns about the authenticity risks and homogenization pitfalls of AI in news production. Interviewee A04 pointed out, "In my work, I've found that AI often generates fabricated information, such as misrepresenting domestic events as foreign ones, which can lead to errors or misinformation." This not only damages the credibility of news but may also mislead public perception.

Additionally, Interviewee A05 mentioned, "The formulaic and templated nature of AI Text-to-Video in an-

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

imated news can lead to a ‘one-size-fits-all’ approach. For example, during Spring Festival reports, multiple media outlets’ AI videos used elements like red colors, lanterns, and floating data, resulting in excessive homogenization.” This phenomenon undermines the uniqueness and diversity of news and may lead to audience esthetic fatigue. Furthermore, AI Text-to-Video still has significant limitations in personalized expression and stylistic storytelling. For instance, Interviewee A02 noted, “AI lacks distinct personal styles.” Interviewee A03 also believes that AI technology often struggles with videos requiring unique styles.

Within the framework of Actor-Network Theory, the differences in user cognitive effects between AI Text-to-Video in animated news and traditional animated news are essentially manifestations of technological rationality versus humanistic rationality. Their competition is not a zero-sum game. In the actor-network, there is an “obligatory passage point,” meaning that the primary actor imposes rules on the secondary actor, and only by passing through this point can the secondary actor's needs be met and problems solved. Therefore, a “human-machine symbiosis” news production paradigm must be established, creating a dynamic mutual construction process between technology and social actors. In this regard, Interviewee A04 suggested, “The role of journalists should be restructured, shifting from ‘content producers’ to ‘planners and reviewers,’ leveraging the efficiency advantages of AI while ensuring the authenticity and objectivity of news through human review.” Additionally, Interviewee A05 emphasized the need to consider users’ psychological and emotional needs, aligning values to reduce negative experiences such as the “uncanny valley effect.” Interviewee A01 recommended, “Establish AI material training positions and form specialized AI news teams to enhance efficiency.”

5. Conclusion and Outlook

Currently, AI Text-to-Video, as a multimodal communication form, is triggering a paradigm shift in the field of news dissemination. This study, through a combination of experiments and in-depth interviews, thoroughly analyzed and explored the differences in cognitive communication effects, emotional responses, and media preferences between AI Text-to-Video in animated news and traditional animated news. The research found that AI Text-to-Video in animated news demonstrates significant cognitive advantages during initial information reception, primarily due to its novel format and multi-sensory stimulation. However, traditional animated news holds a stronger advantage in deep memory consolidation, closely related to its systematic narrative and content authority. Users’ emotional responses to the two news forms differ, and their media preferences are neutral, indicating that user choices depend more on specific contexts and individual needs. Although AI Text-to-Video technology offers significant cost-saving and efficiency-enhancing advantages in news production, it also faces challenges such as authenticity risks and homogenization.

Future research could further optimize the emotional models of AI Text-to-Video to enhance its emotional expression capabilities and user acceptance. Meanwhile, news media organizations should actively explore a “human-machine symbiosis” news production paradigm, redefining the role of journalists from “content producers” to “planners and reviewers” to better adapt to the media ecosystem in the intelligent communication environment.

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

References

[1].Shi Anbin & Zheng En. (2024). Entering “Integrated Reality”: The Reshaping of the News Media Industry by Text-to-Video Technology. Media Observation, (04), 27-36. (in Chinese)

[2].Xun Sha. (2024). Research on the Communication Practice of AI Text-to-Video in Human-Computer Relationships from the Perspective of Media Equation Theory. Communication and Copyright, (21), 45-48. (in Chinese)

[3].Wen Fengming & Wen Yuqun. (2024). Research on the Path of Text-to-Video AI Empowering Short Video Content Production. Audio-Visual World, (03), 31-34. (in Chinese)

[4].Arkhipova Daria & Viidalepp Auli. (2023). Lotman’s semiotics of culture in the age of AI: analyzing the cultural dynamics of AI Text-to-Video art in the semiosphere. Semiotica, (255), 149-160.

[5].Wang Zhaoyang & Wei Jiejie. (2021). A Comparative Experimental Study on the Cognitive Effects of Mobile Short Video News Users. Journalism and Communication Review, (01), 13-25. (in Chinese)

[6].Song Kexin. (2024). Media and Discourse: A Study on the Relationship between Expert Discourse Authority and Popular Cognition from the Perspective of Network Communication. Communication & Education Review, (8).

[7].Lin Keqin & Zeng Jingping. (2023). Cognitive Strategic Communication: Related Topics, Core Demands, and Practical Paradigms. Journal of Xi'an Jiaotong University (Social Sciences Edition), (05), 101-112. (in Chinese)

[8].Han Jingjing & Yu Haoqing. (2024). The Development of Cognitive Communication Theory in the Chinese Context: Challenges, Applications, and Paths. Global Media Journal, (03), 74-91. (in Chinese)

[9].Zeng Zihan. (2022). Cross-cultural Empathic Communication of National Image: Analysis of International Communication Strategies and Effects of the Beijing Winter Olympic Games. Asian Social Science, (11), 1-1.

[10].Han Zhuoxuan. (2024). New Forms and Changes in Social Media Fatigue Behavior: A Case Study of “momo” on Xiaohongshu. New Media Research, (18), 23-27. (in Chinese)

[11].Zhang Yuhong & Chen Kefang. (2024). Digital Multimodal Empathy Narratives: A New Paradigm for International Communication of Local Culture in the Digital Intelligence Era. Scientific Journal of Humanities and Social Sciences, (5).

[12].Nie Zhiteng. (2022). Research on the Influence of News Communication Mode Change on Alleviating Audience’s Mental Anxiety under the Background of Media Integration. Psychiatria Danubina, (S2), 295-295.

[13].Wang Shuhua. (2010). Ethical Reflections on “Moving News” and Feasibility Exploration for Future Development. China Radio & TV Academic Journal, (03), 32-33. (in Chinese)

[14].Tony Liao, Elizabeth Rodwell & Debriunna Porter. (2024). Media frames, AI romantic relationships, and the perspectives of people in relationships; mapping and comparing news media themes with user perspectives. Information, Communication & Society, (12), 2314-2332.

[15].Barbie Zelizer. (2004). When facts, truth, and reality are God-terms: on journalism's uneasy place in cultural studies. Communication and Critical/Cultural Studies, (1), 100-119.

[16].Wang Zhaoyang & Yu Huilin. (2019). Emotional Preference Effects in News Short Video Communication: An Empirical Study Based on Pear Video's Social Section. Journalism and Communication Review, (03),

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)

42-55. (in Chinese)
[17].Cui Zeyu. (2022). Research on the Current Situation and Optimization Strategies of AI Anchors: A Case Study of Ziniu News AI Broadcasting Column. Television Technology, (11), 137-140. (in Chinese)
[18].Chen Jie. (2024). Impacts of Internet literacy and Internet contact on the communication effect of university students' ideological and political education in China. Acta psychologica, 104321 -104321.
[19].Gao May Hongmei. (2023). From the Traditionalists to GenZ: conceptualizing intergenerational communication and media preferences in the USA. Online Media and Global Communication, (3), 422-445.
[20].Allysson Viana Martins & André Fabrício Holanda da Cunha. (2013). Memory and Actor-Network Theory: Mediation in Websites of Estadão and Folha de S. Paulo. Verso e Reverso, (65), 108-116.
[21].Sun Yanbin. (2010). "Moving News" is Just Moving in the Wrong Place. Young Journalist, (01), 68-69. (in Chinese)
[22].Xie Mengli. (2022). Analysis of the Development of Newspaper Moving News in the New Media Era: A Case Study of "Moving News" by The Beijing News. Media Forum, (11), 89-91. (in Chinese)
[23].Xu Wenqian. (2019). My View on the Development Status of Moving News in China. Northern Media Research, (02), 42-43. (in Chinese)

© 2025 by the author(s); licensee Mason Publish Group, this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (http://creativecommons.org/licenses/by/4.0/)