

"Nakanohito Does Not Exist!": Hyperpersonal Interaction and Image Construction in Virtual Streamer Culture

Xinyu Zhu

*Southeast University-Monash University Joint Graduate School (Suzhou), Southeast University,
Suzhou, China*

xzhu0115@student.monash.edu

Abstract. Virtual streamers, as an emerging form of computer-mediated communication, exhibit unique communicative characteristics in their image construction mechanisms. This study applies hyperpersonal interaction theory and employs literature review and case analysis methods to construct a three-level analytical framework for virtual streamer image construction. The research findings reveal three distinct levels: At the first level of real body construction, "nakanohito" (the person inside) achieves selective transmission of bodily information through motion capture and other technologies, forming a "one-to-one" relational model distinct from voice actor logic. At the second level of virtual image construction, principles of body phenomenology are applied to transform real bodies into symbolic digital bodies, transmitting idealized information to audiences through carefully designed visual symbols. At the third level of media persona construction, two appreciation modes emerge: "media persona as real body" and "media persona as virtual image." The research conclusions indicate that virtual streamer image construction is a dynamic process involving multiple participants, where audience feedback plays a crucial role in hyperpersonal interaction, driving continuous optimization and evolution of the image. This study provides a theoretical framework for understanding media persona construction in the digital age and holds significant value for research on communication mechanisms of virtual media forms.

Keywords: virtual streamers, hyperpersonal interaction theory, body phenomenology, image construction, media persona

1. Introduction

The emergence of virtual streamers represents a fascinating intersection of digital technology and human communication, fundamentally challenging traditional notions of identity, embodiment, and media interaction. Since Kizuna AI's debut in December 2016, which marked the birth of the virtual streamer phenomenon, this new form of computer-mediated communication has rapidly evolved into a global cultural phenomenon that transcends geographical and linguistic boundaries.

Virtual streamers operate through a unique mechanism where real human performers, known as "nakanohito" (the person inside), animate digital avatars through motion capture technology, creating an intriguing paradox of presence and absence. Unlike conventional media personalities

who appear directly before audiences, virtual streamers exist in a liminal space between reality and virtuality, where authentic human expression merges with carefully crafted digital personas.

This phenomenon raises compelling questions about the nature of mediated interaction in digital environments. How do audiences form emotional connections with entities that explicitly acknowledge their constructed nature? What role does the hidden "real body" play in shaping audience perceptions and relationships? How do the dynamics of revelation and concealment operate within these mediated interactions?

The significance of studying virtual streamers extends beyond entertainment culture. These digital entities serve as harbingers of evolving human-computer interaction patterns, offering insights into how individuals navigate authenticity in increasingly digitized social spaces. As virtual and augmented reality technologies continue advancing, understanding the mechanisms underlying virtual streamer culture becomes crucial for anticipating future developments in digital communication and social interaction.

This research addresses these questions by applying hyperpersonal interaction theory to analyze virtual streamer image construction. Through systematic examination of representative cases and theoretical analysis, we seek to illuminate the complex interplay between technology, performance, and audience reception that defines this emerging media form.

2. Literature review

2.1. Conceptual definition of virtual streamers

The term "virtual streamer" originated in Japan and encompasses various related concepts including Virtual YouTuber, Virtual UP (VUP), and "persona operator." The concept was first introduced by Kizuna AI, widely recognized as the world's first virtual streamer, in her inaugural YouTube video posted in December 2016. In this groundbreaking video, she introduced herself saying "My name is Kizuna AI" and explicitly proposed the virtual streamer concept by stating "I am a virtual YouTuber, Kizuna AI" [1].

Currently, Chinese academia lacks a unified and clear definition of virtual streamers. From a technical perspective, scholars define virtual streamers as content creators who use virtual avatars for video production or live streaming. These streamers typically avoid appearing in their real human form, instead interacting with audiences through virtual character images (usually anime-style) using motion capture devices, facial expression capture software, Live2D, or 3D rendering technologies [2]. From the perspective of performance format and cultural consumption, other scholars conceptualize virtual streamers as abstract cultural consumption symbols constructed through artificial intelligence and virtual reality technologies, fulfilling audiences' spiritual, cultural, and entertainment needs through symbolic performances [3].

Virtual streamers can be categorized along multiple dimensions according to different classification standards (Table 1). From a technical standpoint, virtual streamers fall into two major categories: 2D and 3D. In terms of driving mechanisms, they can be divided into AI-driven and human-driven types. Regarding interaction patterns, virtual streamers can be further classified as real-time interactive or non-real-time interactive.

Table 1. Classification standards and categories of virtual streamers

Classification Standard	Category 1	Category 2
Technical Type	2D Virtual Streamers	3D Virtual Streamers
Driving Type	AI-driven	Human-driven
Interaction Type	Real-time Interactive	Non-real-time Interactive

Building upon hyperpersonal interaction theory, this study defines virtual streamers as digitized characters created through advanced computer graphics, motion capture, and deep learning technologies. These entities are driven by real human performers and possess the capability to engage in real-time interactive communication with human users on network platforms through 2D or 3D virtual avatars, while maintaining inherent consumption attributes.

2.2. Conceptual definition of nakanohito

The concept of "nakanohito" (中の人, literally "the person inside") originated in Japan, initially referring to performers inside mascot costumes. The term gradually expanded to encompass real actors inside character suits in tokusatsu productions and voice actors in various media. In 2016, J. Matsuda, driven by the vision of creating "an existence that lives on the same world line and timeline as us, without lies and with a sense of reality," collaborated with his team to create Kizuna AI [4].

Matsuda recognized that simply setting up a character as a "high school girl" would fail to answer fundamental questions about "why it should be two-dimensional," reducing the creation to something resembling mere animation. To address this challenge, the team incorporated a "nakanohito" into their project, ensuring that the virtual avatar transcended static character design to become a dynamic presence animated through real human motion capture and expression capture technologies. Since then, "nakanohito" has been used to refer to the actors, voice performers, or operators behind virtual streamer avatars, which is the definition adopted in this study.

2.3. Hyperpersonal interaction theory

Hyperpersonal interaction represents a theoretical concept initially proposed by Joseph B. Walther within the context of computer-mediated communication (CMC). In 1992, Walther presented a crucial perspective in his early research, arguing that interpersonal relationship effects in computer-mediated communication might not exhibit the negative tendencies that early experimental studies suggested. He specifically noted that despite the absence of nonverbal cues in computer-mediated communication, people could still form personalized impressions of others through textual information and develop relationships while expressing multidimensional relational information based on these impressions [5].

In 1995, Walther conducted experimental research exploring interpersonal relationship effects in computer-mediated communication. His findings revealed that as time progressed, participants in computer-mediated communication could compensate for the absence of nonverbal cues to some extent, developing interpersonal relationship dimensions such as intimacy, similarity, and relaxation through linguistic and textual cues [6]. Furthermore, Walther discovered that suspicion and deception in computer-mediated communication processes were perceptible, with suspicion manifested through nonverbal behaviors, though participants with varying degrees of suspicion exhibited different behavioral patterns [7].

These discoveries enriched the foundation of hyperpersonal interaction theory, demonstrating that interpersonal relationships in computer-mediated communication are dynamically evolving and influenced by individual behaviors and perceptions. In 1996, Walther formally introduced the term "hyperpersonal interaction" in his scholarly work, explaining that in computer-mediated communication environments, interacting parties establish and maintain interpersonal relationships through limited communication cues despite the absence of visual, auditory, and social identity cues. These relationships potentially become more intimate and cohesive than face-to-face communication, ultimately evolving into a form of intimacy that transcends conventional interpersonal boundaries [8].

3. Methodology

This research employs a qualitative approach combining literature review and case analysis methods to examine virtual streamer image construction mechanisms through the lens of hyperpersonal interaction theory.

The literature review component systematically surveys existing scholarly works on virtual streamers, computer-mediated communication, and hyperpersonal interaction theory. This approach allows for comprehensive theoretical foundation building while identifying gaps in current research on virtual streamer phenomena. Key theoretical frameworks examined include Walther's hyperpersonal interaction theory, Merleau-Ponty's body phenomenology, and media persona studies from film and media research traditions.

The case analysis method focuses on representative virtual streamers and significant events within virtual streamer culture. Primary cases examined include Kizuna AI as the pioneering figure in virtual streaming, the "Four Heavenly Kings" of early virtual streamers (Siro, Mirai Akari, Nekomasu, and Kaguya Luna), and contemporary examples such as A-SOUL members. Critical incidents analyzed encompass the 2019 Kizuna AI "four versions" controversy, various streamer "graduation" and identity transition cases, and audience reaction patterns to these events.

Data sources include publicly available video content from platforms such as YouTube and Bilibili, social media interactions, fan community discussions, and industry reports. The analysis examines both the technical aspects of virtual streamer presentation (motion capture, 3D modeling, Live2D animation) and the social dynamics of audience engagement (comments, donations, fan-created content).

4. Results

4.1. First level: construction of real body

Eight years have elapsed since Kizuna AI's debut, and today's virtual streamers have evolved into diverse forms while gradually distancing themselves from traditional voice actor paradigms. However, the real body remains fundamental to virtual image creation, consistently playing a pivotal role in hyperpersonal interaction. Through motion capture and facial recognition technologies, the "nakanohito's" real body transforms into dynamic virtual expressions—a conversion that operates both technically and strategically. Virtual streamers achieve precise control over fan perception by selectively concealing or revealing real body information.

What connections existed between early virtual streamers and voice actors? How does the concept of "nakanohito" relate to this history? Understanding these relationships requires examining the historical development to clarify the trajectory and comprehend the relationship between virtual

streamers and their "nakanohito," thereby furthering discussion of the first level of virtual streamer image construction—real body construction.

As previously mentioned, Kizuna AI is widely recognized as the first virtual streamer, not only because she debuted in December 2016 but also because she was the first to use the term "virtual streamer." However, she was not the only early virtual streamer. The 2017 cohort known as the "Four Heavenly Kings" included Siro [9], Mirai Akari [10], Nekomasu [11], and Kaguya Luna [12]. Among these five pioneers including Kizuna AI, four were believed to have experience as voice actors or in voice-over work. Only Nekomasu lacked such background, yet he played a crucial role in this discourse. Approximately one year before Kizuna AI's debut, he had uploaded videos featuring a pink-haired cat-eared character, stating: "Originally, this should be done by voice actors" [13]. This reveals his belief that virtual streamer performers should be voice actors, even before the term "virtual streamer" existed.

According to hyperpersonal interaction theory, individuals in computer-mediated communication environments construct and maintain interpersonal relationships through limited cues. Early virtual streamers heavily relied on voice actors for image creation, with voice actors using distinctive vocal performances to endow virtual characters with distinct personalities, leaving lasting impressions on audiences. This tight connection between virtual streamers and voice actors emerged because early motion capture and 3D modeling technologies remained imperfect, causing frequent interruptions and model clipping issues in virtual character movements. Kizuna AI exemplified this limitation—her first self-introduction video exhibited less fluid character movement [1]. Her first dance cover video was not released until one year after debut, with the title self-deprecatingly noting "now finally" while still displaying stuttering and model clipping phenomena [14].

Early virtual streamers faced technical limitations that restricted their physical expression capabilities, potentially explaining why many early "nakanohito" had voice acting backgrounds. In essence, voice actors played crucial roles in early virtual streamer development, significantly influencing image construction as part of the real body, while technology limited characters' physical expression.

Furthermore, early virtual streamer agencies operated using voice actor management logic, representing a strategic choice for interaction within specific media environments. Voice actor researcher Nozawa discovered through interviews that although voice actors acknowledge their existence as "nakanohito" behind anime characters, they maintain an otaku-like mentality of "nakanohito does not exist." Like fans, they consider characters as independent entities possessing their own voices, with voice actors merely helping characters vocalize. Nozawa used "effacement" to describe the phenomenon of voice actors hiding themselves during performance [15]. From Silvio's "animation" concept perspective [16], voice actors breathe life into anime characters by providing human voices and speech capabilities, making characters appear more vivid. Giving life to characters constitutes part of voice actor work, and they can voice multiple characters, reflecting the common "many-to-one" or "one-to-many" correspondence between characters and voice actors in the industry.

Many-to-one refers to situations where multiple characters are voiced by the same actor, such as Kobayashi Yu voicing numerous anime characters including Sakurazaki Setsuna from "Negima!," Sarutobi Ayame from "Gintama," and Urushibara Ruka from "Steins;Gate." One-to-many describes situations where the same character is portrayed by different actors, typically occurring when voice actors cannot continue their roles or due to extended animation timelines necessitating different voice actors. For example, Doraemon has experienced several voice actor changes: originally voiced by Tomita Kosei, then Nozawa Masako, followed by Oyama Nobuyo from 1979 to 2005, and

currently voiced by Mizuta Wasabi. Similarly, Mouri Kogoro from "Detective Conan" was originally voiced by Kamiya Akira, but due to certain circumstances, Kamiya announced his departure, with Koyama Rikiya as his successor. The underlying "voice actor logic" behind these "many-to-one" or "one-to-many" relationships operates through "effacement" and "animation."

This discussion of "voice actor logic" serves to contrast with virtual streamer "nakanohito" real bodies, enabling better understanding of virtual streamer image construction at the first level. Early virtual streamer agencies operated using voice actor management logic, but a representative 2019 incident highlighted fundamental differences between "nakanohito" and voice actors, revealing complex connections between virtual streamer images and their real bodies.

Kizuna AI, recognized as the first virtual streamer, began releasing videos titled "Would you believe if there were four Kizuna AIs?" on her YouTube channel starting May 25, 2019 [17], launching the "Daily Kizuna AI" video series and the "four Kizuna AIs" project. Based on appearance order and voice characteristics, fans numbered these four versions: "AI #1" represented the familiar and beloved Kizuna AI image voiced by Kasuga Nozomi, "AI #2" possessed a more mature onee-san vocal tone, "AI #3" was called "JK AI" with voice similar to AI #1 and gradually became the channel's main image, while "AI #4" specifically targeted the Chinese market, wearing Chinese-style clothing and speaking Chinese.

Initially, people assumed this was a humorous, playful project. However, as "AI #1" faded from public view citing busyness, and with her voice actor posting seemingly company-critical comments on Twitter, fans began suspecting "AI #1" had been replaced by management. Angry fans began unsubscribing and leaving numerous protest comments under videos. Eventually, under public pressure, the company renamed "AI #2" and "AI #3" as "Love-chan" [18] and "Aipii" [19] respectively, while "AI #4" became "AIChannel China Kizuna AI" [20] and moved to Bilibili for activities. Although this somewhat appeased fan outcry, Kizuna AI's popularity subsequently declined, making this incident a career turning point that ultimately led to her graduation concert in February 2022 and cessation of activities.

This incident fully demonstrates fans' crucial role in the interactive relationship between virtual streamers and "nakanohito." Fans construct emotional connections with virtual streamers based on limited information presented by the streamers. When companies attempted to apply "one-to-many" voice actor logic to handle virtual streamer-nakanohito relationships, they disrupted fans' existing cognitive frameworks and emotional expectations. Fans believe each virtual streamer character possesses uniqueness and continuity, preferring "one-to-one" models to maintain character image stability and emotional connections with themselves. Facing strong fan protests, the company adjusted its strategy, returning to the "one-to-one" model, indicating that fan feedback wields undeniable influence in virtual streamer image construction interaction processes.

From another perspective, "one-to-many" relationships in virtual streamer culture present a relatively ambiguous situation. Generally, when voice actors voice multiple characters, their identities remain public, allowing fans to clearly know which characters a particular voice actor has voiced. However, for virtual streamers, most "nakanohito" identities are kept confidential and not actively disclosed to the public. This anonymity makes it difficult for fans to determine exactly who the nakanohito are or whether they are simultaneously portraying other virtual streamers or maintaining other online identities.

Occasionally, some virtual streamers choose to return to streaming activities with new virtual images after graduation. For instance, the first and second generation "nakanohito" of Wenjing [21] continued streaming as Mayumi [22] and Mingqian Nailu [23] respectively after graduation. During this process, fan community continuity became quite apparent, as the new images' fan bases

essentially comprised followers of the previous characters. In other situations, fans typically rely on comparison and investigation to speculate about nakanohito identities. This information sometimes becomes casual conversation or discussion topics among fans, but occasionally leads to misinformation or negative impacts on virtual streamer images, causing some fans to strongly oppose discussing virtual streamer nakanohito information.

Through this phenomenon, we observe that although virtual streamer fan communities adopted the "nakanohito" concept from voice actor culture, voice actor industry logic gradually lost fan acceptance. Fans occupied dominant positions regarding virtual streamer-nakanohito relationship perceptions, potentially even influencing company decision-making. This suggests that relationships between "nakanohito" and virtual streamers were not fixed from the beginning but gradually formed through various contingent events during virtual streamer development, including company management strategies, fan feedback, and protest situations. Without fan protests, certain virtual streamers might still maintain states of being portrayed by multiple nakanohito.

Current perspectives on "nakanohito" and virtual streamer relationships show divergent states. Some fans maintain open attitudes toward discussing nakanohito topics and actively seek nakanohito identity information, while others refuse to discuss nakanohito existence. Some virtual streamers even claim they have no "nakanohito." Although "one-to-many" or "many-to-one" relationships from voice actor logic have decreased, the viewpoint that "nakanohito does not exist" has been accepted by some fans and virtual streamers. This reflects the complexity and diversity regarding character-performer relationships within virtual streamer culture.

In virtual streamer image construction, the first level of real body construction proves both foundational and crucial. The relationship between virtual streamers and "nakanohito" transcends mere technical presentation, representing a core element of image construction. Virtual streamer image construction often functions as a one-to-one relationship, highlighting the significance of "nakanohito" in shaping character personality and image. As hyperpersonal interaction theory demonstrates, individuals in computer-mediated communication can construct and maintain social relationships through carefully designed information exchange. For virtual streamers, "nakanohito" as actual carriers of real bodies serve as crucial cues for transmitting information to fans through their performances.

4.2. Second level: construction of virtual image

Computer-mediated communication allows individuals to construct and maintain social relationships through carefully designed information exchange methods, potentially surpassing traditional face-to-face communication [8]. In virtual streamer cultural contexts, these hyperpersonal interaction characteristics manifest particularly clearly in digital image construction—virtual images presented through digital technology become interactive media tools with audiences. Based on previous discussions, we understand that virtual streamer virtual images take various forms in contemporary media environments, including 3D models, Live2D, augmented reality (AR), virtual reality (VR), and VRChat platforms. These forms serve as communication bridges between virtual streamers and fans, all achieved through digital technology.

In other words, the virtual streamer images we observe are not the "nakanohito's" actual physical appearance but digital images processed and transformed by computers. This raises a fundamental question: "What kind of body is a digital image?" To explore this question, this study employs Maurice Merleau-Ponty's body phenomenology for analysis. We chose Merleau-Ponty's theory because of his unique insights into embodied existence—through thoroughly understanding

embodied bodies, we can grasp the distinctive characteristics of virtual streamers' digital bodies and virtual images.

Merleau-Ponty proposed the concept of "embodiment" aimed at transcending Western traditional mind-body dualism, elevating the flesh to ontological heights [24]. He emphasized that our bodily perception plays a central role in shaping individual subjectivity and constructing relationships with external things, rather than viewing the body merely as material existence. To understand this point, consider Merleau-Ponty's example of "left and right hands touching each other." In this example, when we touch our right hand with our left hand, initially the left hand actively perceives while the right hand becomes the perceived object. However, the right hand can also transform into the active party, perceiving the left hand. In this interactive process, both hands serve as perceiving subjects and perceived objects, alternately assuming active and passive roles [24].

Within philosophical domains, our understanding of the body can be distinguished into at least two conceptual levels. The first level treats the body as the subject of experience, enabling us to perceive the external world while serving as the foundation for our existence and life experience—termed the "phenomenal body." The second level treats the body as a material object, relatively passive and a perceptible entity—termed the "physical body" [25]. This distinction breaks the strict Western philosophical tradition of human-as-subject and thing-as-object dichotomy, challenging Cartesian mind-body dualism where the thinking "I" is considered the subject while "my body" is viewed merely as a pure physical entity.

Shi Ke explains that conceptually, the phenomenal body represents the "embodied" aspect of "body," while the physical body represents the "non-embodied" aspect of "form." Although we can understand these concepts analytically, they are not completely independent. Shi Ke illustrated this theoretical framework in Figure 1 [25].

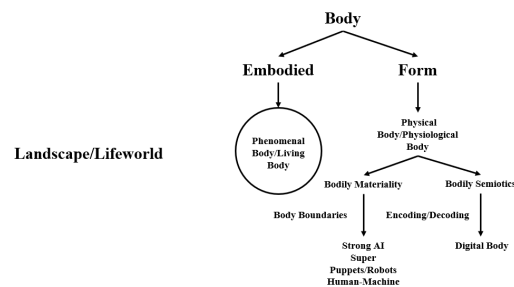


Figure 1. Shi Ke's theoretical framework

When discussing the physical body, Shi Ke proposed two dimensions: bodily materiality and bodily semiotics. Regarding bodily materiality, he drew from McLuhan's media theory, considering clothing and masks as body extensions in daily life. He further suggested that if "skin" no longer constitutes the absolute boundary of the body, then material entities possessing artificial intelligence, puppets, and robots might represent extreme manifestations of material non-embodied existence. However, for this study, bodily semiotics proves more important—how the body "represents" as symbols or images. These symbols and images are not physical objects but may be files, memory machines, or encodings within information flows. Therefore, in today's network era, encoding and decoding processes become part of non-embodied bodily practices, including 3D holographic images and "remote theater," all representing semiotic bodily manifestations [25].

Merleau-Ponty's "embodiment" theory reveals the body's central position as the core of perception and expression, while the "cue management" mechanism in hyperpersonal interaction theory further demonstrates that virtual streamers achieve selective encoding of limited social cues

by transforming real bodies into symbolic virtual images (such as motion capture data). This digital body construction also belongs to phenomenological "embodied extension" and represents strategic self-presentation by senders in hyperpersonal interaction.

The core significance of this process lies in creating a virtual character capable of movement, whether through 3D models or Live2D technology. Advanced technologies capture the "nakanohito's" facial expressions and body movements, then transform these movements into virtual character dynamic presentations through computer processing. Even though "nakanohito" typically do not publicly reveal their real bodies, their movements and expressions are captured and digitized, creating a symbolic virtual body that represents the "nakanohito's" movements and expressions. This virtual body transcends physical form constraints, becoming a digitally mediated existence that not only preserves the "nakanohito's" dynamic characteristics but also interacts with audiences through virtual images, creating new pathways for bodily expression and communication.

In this process, the "nakanohito's" embodied real body does not disappear but transforms into a symbolic body. As shown in Figure 2 [26], five virtual streamers interact through motion capture technology, with their virtual avatars embracing together. However, in virtual space, there are no physical boundaries of flesh, so achieving this movement actually relies on real bodies' vision and flesh to adjust movements and perceive their own bodies relative to others' bodies. As hyperpersonal communication theory explains, virtual streamers utilize computer-mediated communication to create virtual identities in cyberspace through symbolization and digitization, expressing themselves through digital images and movements.



Figure 2. A-SOUL first anniversary special live stream

In this process, information senders (virtual streamers) transform "nakanohito" characteristics into transmittable symbolic information through carefully designed virtual images and movements, while receivers (audiences) construct perceptions and impressions of virtual streamers based on this symbolic information.

Virtual image construction depends not only on real body existence but also achieves transcendent expression based on this foundation. In contemporary media environments, virtual streamer image forms are highly diverse, with digital technology advancements making virtual image construction possible, endowing virtual streamers with abundant expressiveness and interactivity. The virtual image construction process involves multiple technological and artistic integrations, such as motion capture and facial recognition technologies that transform "nakanohito" real movements and expressions into virtual image dynamic presentations. This transcends simple replication, creating new forms of bodily expression that endow virtual characters with unique personality and charm.

From the feedback mechanism perspective of hyperpersonal interaction theory, audience feedback plays significant roles in virtual image construction. Audiences provide feedback through watching virtual streamer live streams and expressing opinions on social media platforms. Virtual

streamer teams adjust virtual images based on this feedback. If audiences generally favor specific movements or expressions from virtual streamers, teams might increase such designs in subsequent content. If audiences express dissatisfaction with particular aspects of virtual images, teams will make corresponding adjustments. This feedback-based adjustment embodies dynamic communication in hyperpersonal interaction, enabling continuous virtual image optimization that better meets audience expectations, thereby strengthening emotional connections between virtual streamers and audiences.

Virtual streamer images encompass more than visual presentation—they serve as carriers transmitting emotions and information through symbols and images. At this level, virtual image construction transcends physical body limitations, becoming a symbolic form of existence. Virtual streamers' clothing styles, hairstyle designs, and specific movements all carry symbolic meanings. Through these symbolic representations, audiences can perceive character personalities, emotions, and story backgrounds, building deeper emotional connections with virtual streamers. This symbolic construction surpasses physical body restrictions.

In hyperpersonal interaction, audience understanding and interpretation of these symbols are influenced by various factors including cultural backgrounds and personal experiences. When designing virtual images, virtual streamer teams also consider target audience characteristics to ensure symbols accurately convey intended information, promoting effective interaction with audiences.

Virtual image construction represents a complex, multi-layered process based on real bodies that creates virtual identities transcending physical forms through combinations of digital technology and semiotics. These virtual identities significantly enhance interaction between virtual streamers and audiences within hyperpersonal interaction, providing rich expressive methods for virtual streamer image construction while continuously evolving with technological development and changing audience demands.

4.3. Third level: construction of media persona

The abstract part starts with “Abstract:”. Each paper must have an abstract. The abstract of each paper should be between 150 and 250 words. What exactly do we appreciate when watching virtual streamers' videos and live streams? The answer is straightforward—aside from subtitles, background music, and other elements, audiences appreciate the movements, the virtual streamers who engage in dialogue with them. This interactive mode demonstrates the "feedback loop" mechanism in hyperpersonal interaction theory: audiences participate in virtual streamer media persona construction through comments, donations, and other forms of engagement, while virtual streamers adjust their image strategies based on audience feedback, forming dynamic interactive cycles.

This leads to a second question: which virtual streamer is being appreciated? Is it the virtual streamer's real body or their virtual image? For instance, when we say a virtual streamer is laughing, who exactly is laughing? When we describe a virtual streamer as cute, who is actually cute? At this point, we realize this question proves more complex than the first. Virtual streamer movements in videos, except for special circumstances, are essentially generated by tracking the "nakanohito's" movements. Regarding the question of which virtual streamer is being appreciated, we might answer "the nakanohito," yet the nakanohito does not directly communicate with us. Audiences only communicate through specific forms of media and character performance. This character presented through media is called a "media persona" [27].

Through bodily or linguistic address (such as camera-facing poses, direct camera gazes, intimate conversations with audiences, and unique terms for addressing audiences), audiences develop

feelings of bidirectional social relationships with these personas. This special social relationship is termed "parasocial interaction (PSI)" [27]. However, the interaction mode between virtual streamers and audiences transcends traditional parasocial interaction boundaries. Through live streaming videos and social media platforms, audiences and virtual streamers establish novel bidirectional communication relationships. These relationships encompass not only audience unidirectional consumption of virtual streamer content but also real-time interaction through comments, likes, shares, and other methods. This interactive mode embodies core concepts in hyperpersonal interaction theory—in computer-mediated communication environments, people can construct and maintain social relationships through carefully designed information exchange methods. In virtual streamer live streams, audience participation transforms from passive states to active, bidirectional communication, strengthening connections based on hyperpersonal interaction.

Based on these insights and theoretical foundations, we now analyze virtual streamer media images in detail. Media image construction relates to film and media studies, with star image construction discussions emerging in star studies from the latter half of the 20th century [28]. Sociologist Dyer identified four factors constructing star images in comprehensive star research: first, construction by producers; second, construction through media activities including film appearances, talk shows, and magazine interviews; third, construction through critics' narratives surrounding stars; finally, star image construction through audience communication [29].

Dyer's star research revealed the multi-agent nature of image construction (including producers, audiences, and media), a perspective that resonates with feedback loop mechanisms in hyperpersonal interaction theory. In virtual streamer media persona shaping, fans participate in collaborative image construction through comments and fan-created content, extending the "receiver idealization" process in hyperpersonal interaction into dynamic negotiation mechanisms. Analyzing virtual streamers from this perspective clarifies differences between their real bodies, virtual images, and media personas. Virtual streamer media personas often transcend the personas' own intentions, continuously forming through complex interactions among producers and audiences in the aforementioned star image construction processes, fully embodying the dynamism and creativity of information exchange in hyperpersonal interaction.

The above discussion primarily characterized virtual streamers as constructed images—media personas—from media communication research and star studies perspectives. However, how virtual streamer media personas are appreciated remains unclear. When audiences watch virtual streamers, they see not actual persona images but the aforementioned virtual images. Here, we notice something important: virtual streamer images are always associated with character virtual images. For instance, in star cases, the persona faces and images audiences appreciate undoubtedly belong to the stars themselves. However, in virtual streamer cultural contexts, streamer images are virtual images, and audiences can only perceive persona real body movements converted to virtual image representations. Personas lack their own faces, with audiences judging persona appeal only through virtual image design aesthetics. This unique correspondence between persona real bodies and virtual images operates continuously, with character real bodies and virtual images frequently overlapping.

Beyond this correspondence, virtual streamer media persona appreciation diversity reveals image characteristics. Some virtual streamers share experiences and thoughts related to their real bodies, partially breaking virtual-reality boundaries, while others consistently maintain character performance in videos and social media presentations, appearing purely through virtual images. According to these virtual streamer behavioral categories, audience appreciation methods also change.

Differences in media persona appreciation can be divided into two categories. When real bodies appear significantly, audiences can appreciate virtual streamers as "media persona as real body." Such virtual streamers can discuss experiences related to their real bodies without issues, with their statements viewed as virtual streamer expressions based on the persona's actual experiences and thoughts, even with certain modifications. Conversely, in the latter case, audiences appreciate virtual streamers as "media persona as virtual image." Here, audiences value virtual streamers' role-playing as virtual images, appreciating them as virtual images that rarely discuss real body-related experiences. Additionally, during appreciation, if audiences fail to control mentions of real bodies, they cannot properly appreciate such virtual streamers.

Either appreciation mode varies by virtual streamer, but regardless, as media personas, they can be analyzed from hyperpersonal interaction perspectives. Using Kizuna AI as an example, she shared content resembling daily life anecdotes during early live streams. Although these contents' authenticity was questionable, fans viewed them as real experience sharing from the "nakanohito" behind Kizuna AI's virtual image, appreciating Kizuna AI as "media persona as real body." Fans actively participated in discussing these experiences as if communicating with real people, bringing Kizuna AI closer to fans and strengthening emotional connections.

Using A-SOUL member Ava as an example, she consistently maintains her energetic girl virtual image setting in live streams and social media platforms. Her words, actions, and personality characteristics closely center around this virtual image, rarely involving real body-related content. When audiences appreciate Ava, they mostly view her as "media persona as virtual image," immersing themselves in her constructed virtual world. Audiences develop affection for her represented virtual image through her cute behaviors and energetic dances during live streams. If Ava suddenly broke this virtual image during streams by mentioning excessive real body information, some audiences might feel disconnected, affecting their appreciation experience. This suggests that in hyperpersonal interaction, tacit information exchange modes exist between virtual streamers and audiences, with both parties continuously adjusting perceptions and expectations during interactions.

Taking "media persona as virtual image" appreciation with obvious differences between real bodies and virtual images as an example, audiences discuss whether virtual image personalities are cute or emotions moving—in other words, they appreciate virtual image personalities and emotions as virtual streamer media personas. Using Severus Snape from the Harry Potter film series as comparison, this character was portrayed by actor Alan Rickman. Snape represents a complex character in the Harry Potter story world, with personality shaped by various events including unresolved feelings for Lily Potter, complex attitudes toward Harry, and secret activities at Hogwarts. Snape is viewed as having deep, contradictory, and psychologically conflicted personality traits—characteristics that are fictional creations by J.K. Rowling.

Alan Rickman brought these psychological contents vividly to screen through masterful acting skills. If the actor's performance failed to accurately capture Snape's personality traits or deviated from original character development, the character presentation might be considered unsuccessful. Audience understanding and feelings about Snape largely depend on how Rickman conveyed these fictional psychological contents through body language, facial expressions, and dialogue.

However, virtual streamers lack specific characters to portray because generally, their virtual images do not exist in specific story worlds nor possess attributable emotions. Virtual images lack story worlds and characteristics formed within story worlds. When virtual streamers are appreciated as virtual image media personas, they are not playing roles in specific story worlds. What is appreciated is neither performed virtual images nor "nakanohito" real bodies themselves, but

partially inherited media personas that can be described as "performed" media personas. Virtual streamers do not play story world characters but perform media personas by continuously creating stories.

Therefore, virtual streamer image components can be divided into three levels: real body, virtual image, and media persona. Real bodies, though hidden, provide the foundation for image construction; virtual images serve as direct presentation carriers to audiences, closely connected and overlapping with real bodies; media personas are shaped through information exchange and feedback between producers and audiences in hyperpersonal interaction based on the former two, being appreciated as real body or virtual image media persona forms. "Nakanohito" perform virtual images while engaging in hyperpersonal interaction with audiences. In this process, both parties continuously adjust and adapt, ultimately presenting unique, continuously evolving media persona images.

5. Discussion

This study's three-level analytical framework demonstrates that virtual streamer image construction operates as a complex process challenging traditional boundaries between authenticity and performance, presence and absence. The "one-to-one" relationship model between virtual streamers and their "nakanohito" represents a fundamental departure from voice actor practices, indicating that hyperpersonal interaction theory's emphasis on selective self-presentation becomes particularly pronounced in virtual streaming contexts.

Applying Merleau-Ponty's body phenomenology to digital embodiment reveals how virtual streamers navigate the paradox of physical absence and experiential presence. The transformation of real bodies into symbolic digital forms creates new modalities of corporeal expression transcending traditional spatial and temporal limitations, enabling what we term "distributed presence"—being-with-others through technological mediation while maintaining authentic intersubjective connections.

The feedback loop mechanisms highlight the collaborative nature of image construction, where audiences function as active co-creators in ongoing persona development. This extends hyperpersonal interaction theory beyond dyadic communication to encompass community-based meaning-making processes. The distinction between "media persona as real body" and "media persona as virtual image" appreciation modes reveals sophisticated cognitive processes through which audiences navigate virtual-real boundaries, suggesting that digital literacy encompasses both technical skills and phenomenological competencies.

Virtual streamers problematize traditional notions of genuine self-expression by demonstrating that constructed, mediated identities can generate authentic emotional responses and meaningful social connections. The analysis also reveals how technological affordances shape but do not determine communicative practices, with specific technology uses emerging through dynamic negotiations between performers, audiences, and platform constraints.

6. Conclusion

This study examined virtual streamer image construction through hyperpersonal interaction theory, developing a three-level analytical framework illuminating the interplay between real bodies, virtual images, and media personas. The research demonstrates that virtual streamers constitute emerging forms of human-computer interaction offering insights into digitally mediated social relationships.

The findings contribute to communication theory by extending hyperpersonal interaction concepts to account for community-based meaning-making processes while revealing how body phenomenology principles operate within digital environments. Virtual streamer culture has evolved beyond voice actor paradigms to establish unique "one-to-one" relationship models driven primarily by fan feedback and emotional investment.

These findings have important implications for understanding digital identity, online community formation, and human-computer interaction. As virtual and augmented reality technologies develop, virtual streamer culture mechanisms may preview broader shifts in how individuals construct and maintain social relationships through digital mediation.

However, this study has limitations. The analysis focused primarily on Japanese and Chinese cases, limiting generalizability across cultural contexts. Future research might examine diverse cultural settings, employ complementary quantitative analyses of audience engagement patterns, and conduct longitudinal studies tracking relationship development over time. As AI capabilities advance and motion capture becomes more accessible, boundaries between human-driven and AI-driven virtual personas may blur further, creating new analytical challenges.

Future research should investigate psychological and social implications of sustained engagement with virtual personas, particularly regarding parasocial relationships, emotional well-being, and social skill development. Cross-platform comparative studies could illuminate how technological affordances shape virtual streamer practices, while broader implications for media industries, digital labor, and intellectual property frameworks deserve scholarly attention.

In conclusion, virtual streamers exemplify the dynamic, collaborative, and technologically mediated nature of contemporary digital communication. Their image construction processes reveal sophisticated mechanisms through which individuals and communities navigate increasingly complex relationships between virtual and real, individual and collective, authentic and performed aspects of digital social life. As we move toward more immersive digital environments, insights from virtual streamer culture will prove increasingly valuable for understanding human communication in networked societies.

References

- [1] A.I.Channel. (2016, December 1). [Self-introduction] Nice to meet you! I'm Kizuna AI ㄥ(ㄴ'ㄹ) [Video]. YouTube. <https://www.youtube.com/watch?v=NasyGUeNMTs>
- [2] Chen, X., & Liu, T. (2022). Fan differentiation, migration and resistance in the context of platform "breaking the circle": An investigation based on virtual streamer fan groups. *Journalism & Communication Research*, 6, 22-38.
- [3] Fu, Q. (2021). Encoding of idol symbols: Research on artificial intelligence virtual idol consumption culture. *Learning and Practice*, 2, 132-140.
- [4] J. Matsuda. (2022, March 13). Kizuna AI really exists. note. <https://note.com/jmatsuda2501/n/nc323452afa3f>
- [5] Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, 1, 52-90.
- [6] Walther, J. B. (1995). Relational aspects of computer-mediated communication: Experimental observations over time. *Organization Science*, 2, 186-203.
- [7] Burgoon, J. K., Buller, D. B., Dillman, L., et al. (1995). Interpersonal deception: Effects of suspicion on perceived communication and nonverbal behavior dynamics. *Human Communication Research*, 2, 163-196.
- [8] Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 1, 3-43.
- [9] Siro Channel. (2017, August 12). [Self-introduction] Nice to meet you! I'm Siro! [001] [Video]. YouTube. https://www.youtube.com/watch?v=fLC5TE_KYcw
- [10] Mirai Akari Project. (2017, October 25). [New project] Launch of the Mirai Akari Project [Video]. YouTube. <https://www.youtube.com/watch?v=I8ZBYReOEsg>

- [11] Kemomimi Kingdom “Fox.” (2015, December 18). Tutorial: Using Live2D from now on [Video]. YouTube. <https://www.youtube.com/watch?v=goiTm4t5DVM>
- [12] Kaguya Luna Official. (2017, December 10). [Self-introduction] Kaguya Luna’s special skills are amazing!!!! [Video]. YouTube. <https://www.youtube.com/watch?v=TeKTVFgw1hM>
- [13] Kemomimi Kingdom “Fox.” (2015, December 18). Tutorial: Using Live2D from now on [Video]. YouTube. <https://www.youtube.com/watch?v=goiTm4t5DVM>
- [14] A.I.Channel. (2017, June 2). [Now dancing] Hare Hare Yukai [#119] [Video]. YouTube. <https://www.youtube.com/watch?v=HnQSdlbw3es>
- [15] Nozawa, S. (2016). Ensoulment and effacement in Japanese voice acting. *Media Convergence in Japan*, 3, 169-199.
- [16] Silvio, T. (2010). Animation: The new performance? *Journal of Linguistic Anthropology*, 2, 422-438.
- [17] A.I.Channel. (2019, May 25). Would you believe it if I said there are four Kizuna AIs? #1 [Video]. YouTube. <https://www.youtube.com/watch?v=cWwohK0mgqA>
- [18] Love-chan. (2025, August 16). Virtual YouTuber Wiki: Love-chan. FANDOM. <https://virtualyoutuber.fandom.com/wiki/Love-chan>
- [19] Aipii. (2025, August 16). Virtual YouTuber Wiki: Aipii. FANDOM. <https://virtualyoutuber.fandom.com/wiki/Aipii>
- [20] AIChannel China Kizuna AI. (2025, August 16). AIChannel China Kizuna AI personal homepage. Bilibili. <https://space.bilibili.com/484322035>
- [21] Wenjing_QianNiao Official. (2025, August 16). Wenjing_QianNiao Official personal homepage. Bilibili. <https://space.bilibili.com/667526012>
- [22] Mayumi_mayumi. (2025, August 16). Mayumi_mayumi personal homepage. Bilibili. <https://space.bilibili.com/1265605287>
- [23] Mingqian Naillv. (2025, August 16). Mingqian Naillv personal homepage. Bilibili. <https://space.bilibili.com/2132180406>
- [24] Merleau-Ponty, M. (1962). *Phenomenology of perception*. Routledge & Kegan Paul.
- [25] Shi, K. (2017). *From Grotowski to holographic images: Embodiment and disembodiment in live art*. Chongqing University Press.
- [26] Jabu Jabu. (2021, December 11). [A-SOUL 1st Anniversary] 2021.12.11 Special Anniversary Livestream! [Recording] [Video]. Bilibili. <https://www.bilibili.com/video/BV1Ka411k7rM>
- [27] Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry*, 3, 215-229.
- [28] Hayward, S. (2013). *Cinema Studies: The key concepts*. Routledge.
- [29] Dyer, R., & McDonald, P. (2019). *Stars*. Bloomsbury Publishing.