

# Understanding Vtuber Live Streaming: Exploration of Psychological Attributes of Viewers

Ying Xu, Nan Niu \*

School of Business and Management, Jilin University, Changchun, China

\* Corresponding author: nanelle@163.com

**Abstract.** As an emerging communication practice, there are many research gaps in the field of Vtubing (Vtubing refers to the Vtuber live streaming, and Vtuber refers to the virtual streamer), including the lack of psychological attribute analysis. Through in-depth qualitative interviews, this study comprehensively explores the key psychological attributes of viewers when watching Vtubing, including perceived persona attractiveness, perceived appearance attractiveness, perceived voice attractiveness, perceived reliability, perceived anthropomorphism, immersion, psychological distance, and imagination. This study provides suggestions for Vtuber owners to design and manage Vtubers with important theoretical and practical significance.

**Keywords:** Vtuber, Vtuber live streaming, psychological attributes.

## 1. Introduction

As a new Internet trend, live streaming has gained tremendous popularity around the world with its advantages of high synchronization and high interactivity [1-2]. Anyone worldwide can watch live streaming through Youtube, Twitch, Bilibili, and Tiktok. These platforms allow anyone with a webcam and an internet connection to become a streamer, exporting live content and interacting with online viewers in real time [3]. Recent advances in motion capture and artificial intelligence technologies have enabled streamers to represent their appearances with 2D or 3D animated avatars. Such streamers embodying computer-animated avatars are often called virtual streamers (Vtubers) [3]. They have rapidly gained international popularity since their debut in 2016 [4], are highly sought after by young people, and have produced many excellent achievements. For example, the Vtuber Gawr Gura surpassed 1 million subscribers in only 41 days on Youtube [3]. A Chinese entertainment company, Yuehua Entertainment, created a virtual idol group called “A-SOUL,” which had a record of 12 million in popularity value in single live streaming [5]. Vox, a world-renowned Vtuber, had its first live streaming in China with revenues of 1.11 million RMB and nearly 40,000 payers [6]. In short, Vtubers are leading a new live-streaming craze around the world.

Nevertheless, Vtubing is still a relatively new phenomenon - thus, there is currently a lack of theoretical research on it. To the best of our knowledge, it is unclear exactly what psychological attributes viewers develop when watching Vtubing. Supposing Vtuber owners cannot understand this issue in depth, it will be challenging to design appropriate Vtubers and, in turn, to achieve the desired Vtubing effects. Therefore, the research goal of this study is to comprehensively explore the key psychological attributes of viewers when watching Vtubing.

## 2. Method

From the technology-driven perspective, current Vtubers include mainly “AI technology-driven Vtubers” and “real person-driven Vtubers.” AI technology-driven Vtubers are powered by artificial intelligence, like most AliMe Avatars in Taobao live streaming rooms. Real person-driven Vtubers are 2D or 3D animated virtual avatars powered by motion capture devices that can provide performances in live streaming and are controlled by Nakanohitos, where the Nakanohitos are the actors performing behind the avatars who provide voices and control Vtubers’ movements [4]. Compared to AI technology-driven Vtubers, real person-driven Vtubers are more common because of their low requirements in terms of technologies and costs. Thus, this study conducted semi-

structured interviews with real person-driven Vtubers as research objects.

**2.1. In-depth interview design and process**

We selected China as the empirical context for the interviews because it is one of the most important markets for Vtubing [3-4]. We reached the potential “target objects” through “the V forum” on Baidu Tieba, “the Vtuber forum” on Baidu Tieba, and “the Vtuber super-topic” on Weibo, which are the largest Vtuber discussion boards in China, in terms of both the number of active users and threads.

The purpose of the interviews was to identify the key psychological attributes of viewers when they watch Vtubing. The interviews primarily focused on viewers’ psychological perceptions when watching Vtubing. Interviewees were also guided to think about the differences, strengths, and weaknesses between Vtubers and real-person streamers, which can tap into salient psychological attributes. The interviews were conducted via audio calls on September 5-7, 2022, and each interview lasted approximately 30 minutes, with an honorarium of RMB 10 for each interviewee.

**2.2. Background information of the interviewees**

There were 20 interviewees, including six males and fourteen females, aged 17-27, and viewed Vtubing mainly through Bilibili, Youtube, and Twitch. The results of the descriptive statistical analysis of interviewees are shown in Table 1.

**Table 1.** Descriptive statistical analysis of the interviewees

ID	Ages	Genders	Years of watching Vtubing	Times of watching Vtubing each week	Viewing platforms
No1	18	Male	2	3-4	Bilibili
No2	18	Female	0-1	10	Bilibili
No3	17	Male	3	4-5	Bilibili
No4	18	Female	3-4	2-3	Bilibili, Youtube
No5	17	Female	0-1	7	Bilibili
No6	18	Female	6-7	1-2	Bilibili, Tiktok
No7	18	Female	0-1	3-4	Bilibili, Youtube
No8	17	Female	1-2	7	Bilibili
No9	20	Female	0-1	1	Bilibili
No10	20	Male	2	1-2	Bilibili
No11	20	Female	3-4	7	Bilibili, Youtube, Twitch
No12	20	Female	0-1	2-3	Bilibili, Youtube
No13	18	Female	0-1	7	Youtube, Twitch
No14	21	Male	0-1	2-4	Bilibili
No15	21	Male	1	2-3	Bilibili
No16	18	Female	0-1	2-3	Bilibili
No17	20	Female	0-1	7	Bilibili, Youtube
No18	20	Female	1	7	Bilibili, Youtube, Twitch, Twitter
No19	20	Female	0-1	2-3	Bilibili, YouTube
No20	27	Male	4-5	2-7	Bilibili

**3. Result**

By compiling all the interview recordings, a transcript with more than 100,000 words was formed, and several key psychological attributes of viewers when watching Vtubing were summarized as follows.

**3.1. Perceived persona attractiveness**

Each Vtuber has a persona, which is its fundamental attribute. The persona of Vtubers refers to the accepted or pre-designed manner in which the Nakano Hitos play roles [7], including many aspects

such as gender, age, identity, personality, expression, and action. Since Vtubers exist in the virtual world with infinite possibilities, the personas of Vtubers are varied and even beyond the imagination of people, such as a divine messenger who likes to drink milk and an eagle-bodied demon who likes to play games. Combining the concepts of persona and attractiveness [8], perceived persona attractiveness refers to the extent to which viewers perceive that personas can bring them a pleasant experience and convey a good impression, i.e., the degree to which Vtubers' personas are attractive to viewers. In the interviews, No4, No6, and No17 supported the attractiveness of Vtubers' personas. For example, No17 indicated, "the Vtuber's persona appeals to my preferences, so I am satisfied with it."

### 3.2. Perceived appearance attractiveness

Undoubtedly, the appearances of streamers can play crucial roles in the effectiveness of live streaming [9]. Appearance attractiveness describes how attractive or likable a person's appearance features are [10]. Simply put, perceived appearance attractiveness refers to the extent to which viewers perceive Vtubers' appearances to be attractive and can be categorized into different aspects, such as cuteness and beauty attractiveness [11]. First, unlike real-person streamers, Vtubers are presented as cartoon images, and cuteness is one of the most typical characteristics of cartoon images. In the interviews, when asked, "What attracts you to view Vtubing? What are the advantages of Vtubers over real-person streamers?" many interviewees stated that the cartoon images of Vtubers are pretty cute, such as No4, No5, No6, No10, No14, No20. Second, beauty is also a typical characteristic of Vtubers. 14 out of 20 interviewees praised the beauty of Vtubers. As an example, No17 thought that a beautiful image was the most significant advantage of Vtubers compared to real-person streamers. We believe that this finding cannot mean that real-person streamers are not beautiful, but the well-designed images of Vtubers are more in line with people's pursuits of beauty.

### 3.3. Perceived voice attractiveness

Live streaming can bring viewers both visual and auditory levels of sensory enjoyment. However, because the fictional appearances of Vtubers cause viewers to pay more attention to voices when watching Vtubing than live streaming by real-person streamers, the voice is undoubtedly one of the most important forms of media expression in Vtubing and a key influencing factor for the effectiveness of Vtubing. Perceived voice attractiveness refers to the degree to which viewers perceive Vtubers' voices to bring them a pleasant experience and convey a good impression [8]. In simple terms, it refers to the extent to which Vtubers' voices attract viewers. When asked, "What attracts you to view Vtubing? What are the advantages of Vtubers over real-person streamers?" many interviewees mentioned Vtubers' voice attractiveness, such as No5, No9, No12, No16, No17, and No19. Even No19 commented, "I think the voice is ranked first among Vtubers' all attributes." In addition, in the field of Vtubing, Nakanohitos are committed to fully exploiting the communication values of voices to create a unique experience for viewers, and even many professional voice actors serve as Nakanohitos. For example, "the Nakanohito may be a professional voice actor (No15)"; "most Vtubers' voices are more professional than those of real-person streamers (No20)". In short, Vtubers generally have excellent and professional voices.

### 3.4. Perceived reliability

The interview results show that viewers can perceive higher reliability when watching Vtubing than real-person live streaming. Reliability refers to the quality of trustworthiness or consistently good performance and is highly correlated with safety, controllability, and usefulness [12]. There are multiple reasons to explain the high reliability of Vtubing. First, Vtubers tend to have pre-determined and near-perfect personas compared to real-person streamers. In order to maintain their attractiveness to viewers, some Nakanohitos strictly follow these original personas instead of taking liberties [4]. As No13 said, "a Vtuber with a well-determined persona will not surprise me. But a real-person streamer is unpredictable". Second, Vtubers are separated from Nakanohitos, which reduces the

effects on Vtubers to a certain extent (No8, No10). Third, some interviewees clearly indicated that Vtubers could bring them a sense of security, such as “the Vtuber’s distance from me makes me feel at ease (No20)”. For these reasons, scandalous incidents for Vtubers occur less than for real-person streamers, making Vtubers more reliable.

### **3.5. Perceived anthropomorphism**

Vtubers are virtual characters supported by technologies, so anthropomorphism is one of the essential features of Vtubers. Anthropomorphism refers to the tendency to assign human characteristics, intentions, and behaviors to non-human objects [13-14]. Perceived anthropomorphism refers to the extent to which viewers perceive non-human objects as anthropomorphic. At the appearance level, the degree of anthropomorphism varies across Vtubers, including “relatively high-level anthropomorphic images (No8, No9),” “relatively medium-level anthropomorphic images, such as humans with wings (No4), humans with tails (No5),” and “relatively low-level anthropomorphic images, such as spoons and water pipes.” In general, the degree of appearance anthropomorphism is influenced by the personas of Vtubers (No4). At the behavior level, Vtubers' behaviors are anthropomorphic because Nakanohitos shape their expressions and movements through motion capture technologies. However, due to the difference in the technological level of different motion capture devices, Vtubers exhibit different degrees of behavior anthropomorphism. In general, Vtubers with 3D avatars have a higher degree of behavior anthropomorphism than Vtubers with 2D avatars [15].

### **3.6. Immersion**

Immersion is defined as the feeling of being entirely absorbed by the mediated world and ignoring the real or ordinary world [16], such as participants feeling that they are “really there,” “forgetting the outside world,” and “losing the sense of time” [17-19]. The sense of immersion when watching Vtubing was confirmed by the interviewees, such as “I felt fully immersed, enjoyed the atmosphere, and got a sense of comfort (No1)”; “Sometimes I immersed in a virtual scenario constructed by the Vtuber (No9)”; “when I watch Vtubing, I felt a sense of immersion, and I might put myself in the shoes of the Vtuber and advance the plot together with it, and then I would empathize with it (No18).” Compared to real-person streamers, viewers usually see Vtubers as characters in a near-perfect virtual world and, therefore, less bound by real-world social norms [4]. This virtual world helps viewers release the stress of real life and gain pleasure, so viewers place their emotions in this virtual world and deeply immerse themselves in this wonderful world.

### **3.7. Psychological distance**

Psychological distance is a subjective experience that refers to the subjectively perceived distance of Vtubers from viewers [20]. Consistent with the study by Lu et al. [4], the interviewees reported feeling more distant from Vtubers than real-person streamers. This sense of distance is understandable given that Vtubers are presented virtually and often have personas far away from the real world. Unlike previous literature suggesting that shorter distances are more beneficial [21-23], several interviewees supported the positive impacts of psychological distance. For example, No4 mentioned, “it would make me feel better if we had the right distance.” No20 thought that this distance would make people feel especially at ease.

### **3.8. Imagination**

One of the key features of virtual reality is imagination [24-25], and Vtubing is essentially a manifestation of the fusion of the virtual and real worlds [4]. Therefore, imagination is also one of the key features of Vtubing. Imagination comes in part from the virtuality of Vtubers. Virtuality stimulates viewers’ curiosities. For instance, No3 wondered what kind of person the Vtuber is in reality. This curiosity further stimulates viewers to perceive and imagine things that do not exist creatively. As No11 said, “because Vtubers are detached from reality, I felt that I am watching a

virtual world with no limitations of reality so that I can imagine infinitely.” No19 also stated, “compared to real-person streamers, Vtubers have animated images and carry out more role-playing activities. Thus, Vtubers can stimulate my imagination.”

## 4. Conclusion

### 4.1. Theoretical implications

Firstly, to the best of our knowledge, this study is one of the first studies on Vtubing. From the novel psychological perspective, this study comprehensively explores the key psychological attributes of viewers when watching Vtubing, which deeply expands the limits of understanding Vtubing and opens up new research ideas for the field of live streaming. Secondly, the current study simultaneously discovers the persona, appearance, and voice attractiveness of Vtubing, which theoretically expands and enriches the dimensions of attractiveness.

### 4.2. Practical implications

Understanding various key psychological attributes can support Vtuber owners in designing and managing Vtubers.

Firstly, Vtuber owners should be aware that with the increasing number of Vtubers, novel and exciting personas are more likely to make Vtubers stand out. As No4 commented during the interview, “some personas are impossible to exist in real life, so I feel them interesting and attractive.” Vtuber owners should also determine their target viewers in advance and design their personas accordingly to appeal to different preferences.

Secondly, in addition to making the appearances of Vtubers beautiful, Vtuber owners can also design baby-like facial features such as “big eyes, high and prominent forehead, round cheeks, and small mouth” for Vtubers to foster their level of cuteness.

Thirdly, voice attractiveness is not only limited to the attractiveness of the sound, but also the tone of voice, the speed of speech, the clarity of pronunciation, and the emotion the voice contains. At the same time, the voices of Vtubers should be consistent with their images and personas as much as possible because the overall coordination will be more conducive to promoting the immersion and imagination of viewers when watching Vtubing.

Fourthly, although Vtubers are more reliable than real-person streamers, this does not mean they are free from the risks of scandalous incidents. Vtuber owners need to regulate the behaviors of Nakanohitos strictly. For example, strict regulations can be introduced to prevent scandals from happening.

Fifthly, Vtuber owners should comprehensively determine the degree of anthropomorphism of Vtubers according to their needs, costs, and personas of Vtubers.

Sixthly, “ideal design elements” need to be added to live streaming rooms to make Vtubing a “perfect world,” enhancing viewers’ immersion.

Seventhly, Vtuber owners need to maintain the mystery of Vtubers to keep a moderate distance between viewers and Vtubers, which will help foster viewers’ imagination.

### 4.3. Limitations and future studies

This study has certain limitations. First, this study only considered the perceptions of 20 Chinese interviewees, limiting the findings’ diversity. Second, future research can conduct empirical studies based on the key psychological attributes found in this study, which will be a key research direction in the future.

## Acknowledgements

This paper was supported by the Jilin Province Philosophy and Social Science Planning Fund Office in China [grant numbers 2022W5].

## References

- [1] Zhao, Q., Chen, C. D., Cheng, H. W., & Wang, J. L. (2018). Determinants of live streamers' continuance broadcasting intentions on Twitch: A self-determination theory perspective. *Telematics and Informatics*, 35 (2), 406 - 420.
- [2] Zhou, F., Chen, L., & Su, Q. (2019). Understanding the impact of social distance on users' broadcasting intention on live streaming platforms: A lens of the challenge-hindrance stress perspective. *Telematics and Informatics*, 41, 46 - 54.
- [3] Tang, M. T., Zhu, V. L., & Popescu, V. (2021, October). AlterEcho: loose Avatar-Streamer coupling for expressive VTubing. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 128-137). IEEE.
- [4] Lu, Z., Shen, C., Li, J., Shen, H., & Wigdor, D. (2021, May). More kawaii than a real-person live streamer: understanding how the otaku community engages with and perceives virtual YouTubers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1 - 14).
- [5] CRI Online. (2022). A woman's group called A-SOUL is highly popular, attracting strong cooperation from famous brands. Retrieved November 27, 2022, from <https://baijiahao.baidu.com/s?id=1736853399201637706&wfr=spider&for=pc>.
- [6] People's Convergence Media. (2022). The live streaming debut of Vtuber "vox" in China, revenue over a million RMB in nearly 2 hours. Virtual digital characters raised the popularity of the new "blue ocean". Retrieved November 28, 2022, from <https://baijiahao.baidu.com/s?id=1732585240223115586&wfr=spider&for=pc>.
- [7] Dion, D., & Arnould, E. (2016). Persona-fied brands: managing branded persons through persona. *Journal of Marketing Management*, 32 (1-2), 121 - 148.
- [8] DeGroot, T., & Motowidlo, S. J. (1999). Why visual and vocal interview cues can affect interviewers' judgments and predict job performance. *Journal of Applied Psychology*, 84 (6), 986.
- [9] Zhu, L., Li, H., Nie, K., & Gu, C. (2021). How Do Anchors' Characteristics Influence Consumers' Behavioural Intention in Livestream Shopping? A Moderated Chain-Mediation Explanatory Model. *Frontiers in Psychology*, 12.
- [10] Sokolova, K., & Kefi, H. (2020). Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions. *Journal of retailing and consumer services*, 53, 101742.
- [11] Kuraguchi, K., & Kanari, K. (2020). Face inversion effect on perceived cuteness and pupillary response. *Frontiers in Psychology*, 11, 558478.
- [12] Barua, Z., Aimin, W., & Hongyi, X. (2018). A perceived reliability-based customer satisfaction model in self-service technology. *The Service Industries Journal*, 38 (7-8), 446 - 466.
- [13] Kim, S., & McGill, A. L. (2011). Gaming with Mr. Slot or gaming the slot machine? Power, anthropomorphism, and risk perception. *Journal of Consumer Research*, 38 (1), 94 - 107.
- [14] Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, 114 (4), 864.
- [15] Miao, F., Kozlenkova, I. V., Wang, H., Xie, T., & Palmatier, R. W. (2022). An emerging theory of avatar marketing. *Journal of Marketing*, 86 (1), 67 - 90.
- [16] Fang, J., Chen, L., Wen, C., & Prybutok, V. R. (2018). Co-viewing experience in video websites: The effect of social presence on e-loyalty. *International Journal of Electronic Commerce*, 22 (3), 446 - 476.
- [17] Carù, A., & Cova, B. (2006). How to facilitate immersion in a consumption experience: Appropriation operations and service elements. *Journal of Consumer Behaviour: An International Research Review*, 5 (1), 4 - 14.
- [18] Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, 66 (9), 641 - 661.
- [19] Hudson, S., Matson-Barkat, S., Pallamin, N., & Jegou, G. (2019). With or without you? Interaction and immersion in a virtual reality experience. *Journal of Business Research*, 100, 459 - 468.

- [20] Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological review*, 117 (2), 440.
- [21] Moore, S. G., & McFerran, B. (2017). She said, she said: Differential interpersonal similarities predict unique linguistic mimicry in online word of mouth. *Journal of the Association for Consumer Research*, 2 (2), 229 - 245.
- [22] Itani, O. S., Kalra, A., Chaker, N. N., & Singh, R. (2021). “Because you are a part of me”: Assessing the effects of salesperson social media use on job outcomes and the moderating roles of moral identity and gender. *Industrial Marketing Management*, 98, 283 - 298.
- [23] Sikora, R. T., & Chauhan, K. (2012). Estimating sequential bias in online reviews: a Kalman filtering approach. *Knowledge-Based Systems*, 27, 314 - 321.
- [24] Chittaro, L., & Ranon, R. (2007). Web3D technologies in learning, education and training: Motivations, issues, opportunities. *Computers & Education*, 49 (1), 3 - 18.
- [25] Huang, H. M., Liaw, S. S., & Lai, C. M. (2016). Exploring learner acceptance of the use of virtual reality in medical education: a case study of desktop and projection-based display systems. *Interactive Learning Environments*, 24 (1), 3 - 19.