



## OPEN ACCESS

## EDITED BY

Sergi Bermúdez i Badia,  
University of Madeira, Portugal

## REVIEWED BY

Itaru Kuramoto,  
The University of Fukuchiyama, Japan  
Will Grant,  
Australian National University, Australia

## \*CORRESPONDENCE

Hiroshi Sakuma  
✉ [hiroshi.sakuma@acm.org](mailto:hiroshi.sakuma@acm.org)

## SPECIALTY SECTION

This article was submitted to  
Human-Media Interaction,  
a section of the journal  
Frontiers in Computer Science

RECEIVED 13 September 2022

ACCEPTED 20 February 2023

PUBLISHED 23 March 2023

## CITATION

Sakuma H, Hori A, Murashita M, Kondo C and  
Hijikata Y (2023) YouTubers vs. VTubers:  
Persuasiveness of human and virtual presenters  
in promotional videos.  
*Front. Comput. Sci.* 5:1043342.  
doi: 10.3389/fcomp.2023.1043342

## COPYRIGHT

© 2023 Sakuma, Hori, Murashita, Kondo and  
Hijikata. This is an open-access article  
distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# YouTubers vs. VTubers: Persuasiveness of human and virtual presenters in promotional videos

Hiroshi Sakuma<sup>1,2\*</sup>, Ao Hori<sup>3</sup>, Minami Murashita<sup>3</sup>, Chisa Kondo<sup>3</sup>  
and Yoshinori Hijikata<sup>3</sup>

<sup>1</sup>Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan, <sup>2</sup>Graduate School of Engineering Science, Osaka University, Osaka, Japan, <sup>3</sup>School of Business Administration, Kwansai Gakuin University, Nishinomiyu, Japan

With the recent advances in motion tracking technologies and three-dimensional computer graphics software, communication through avatars has become increasingly popular. Can avatars be sufficiently persuasive, when compared to traditional forms of interpersonal communication? What factors contribute to the persuasiveness of virtual influencers? Existing literature has studied the differences in persuasiveness between human and virtual speakers extensively, particularly in education. However, few studies have been conducted on product promotion. Therefore, in this study, we investigated the characteristics of persuasiveness for humans and virtual influencers, as well as the differences between them in this regard in a more modern and practical situation: product introduction videos used in influencer marketing. Specifically, we recruited participants to watch product introduction videos on YouTube, presented by either humans or avatars. The videos were similar, except for the appearance of the presenter. Before and after watching the videos, the participants were asked to complete a questionnaire about their willingness to purchase the products and the characteristics of presenters' persuasiveness. The results show that although promotion *via* avatars can increase the participants' willingness to purchase, human influencers were more persuasive. However, the virtual YouTubers (VTubers) were more persuasive in certain product domains. VTubers who can change their appearance to match the product domain have potential for future applications. We also attempted to construct a model of persuasiveness in this pragmatic context based on Dyson's persuasiveness rating scale and the overall impression of the video. The degree of persuasiveness was found to be related to the presenters' likability, whether the presenter was a human or an avatar, the degree of familiarity between the presenter and the audience, the presenters' trustworthiness, and the quality as well as the entertainment level of the video. This model may be helpful for running successful promotions on YouTube. Our findings verify that avatars can be fairly persuasive in some situations, including promotional videos. These findings contribute to the future development of communication through avatars.

## KEYWORDS

persuasiveness, avatar, influencer, promotion, VTuber, YouTube, virtual beings

## Introduction

Since the establishment of the online video-sharing platform YouTube in 2005, “YouTubers”—people who post and stream videos on the platform—have become increasingly popular. More recently, with advances in three-dimensional computer graphics (3DCG) software, virtual YouTubers (VTubers), who post and stream videos using 3DCG avatars that are similar to characters in anime, have also gained widespread popularity (Liudmila, 2020). The 3DCG software captures the facial expressions and movements of VTubers and maps them into a 3D model, thus animating the VTubers’ avatars and enabling them to record videos with natural-looking 3D animations.

YouTube’s Culture and Trends Report noted that, as of October 2020, VTubers were garnering over 1.5 billion views per month (Allocca, 2020). According to the ranking of earnings on YouTube, known as Super Chat, in 2021, 8 of the top 10 YouTube video contributors worldwide were VTubers (Playboard, 2022). Currently, “Gawr Gura” has more than 4 million subscribers, making them the VTuber with the largest following. In particular, VTubers have the advantage of controlling avatars that are character-like and do not have to show their own faces or reveal their true physical appearance; however, it is difficult for them to show delicate facial expressions and movements. Thus, it is not clear in what situations videos made by VTubers using avatars can be as convincing to viewers as those performed by real people (YouTubers), or in what ways they differ.

In conventional corporate advertising, companies promote their products directly using mass media advertising, such as television, newspapers, and magazines. In contrast, YouTubers and VTubers introduce the product from the users’ perspective, sharing their experiences with products and services with the users to build a more intimate and personal relationship (Freeman and Chapman, 2007). This type of influencer marketing alters consumer behavior by disseminating information on social media (Brown and Hayes, 2008; Jin et al., 2019; Hudders et al., 2020; Vrontis et al., 2021). Therefore, researchers have been interested in studying factors affecting influencer marketing, such as perceived credibility (Xiao et al., 2018).

In particular, product promotions have been on the rise on video-sharing platforms, such as YouTube and TikTok, as videos provide more information than text on Twitter or photos on Instagram. Unlike text and photos, videos can convey changes in facial expressions and movements, as well as provide detailed instructions on how to use a product. Although the persuasive effect of such videos has been attracting considerable attention, no specific research has been conducted on this topic. However, from a marketing perspective, investigating the persuasive effect of such videos is imperative, as it can serve as a guideline for how YouTubers (humans) and VTubers (avatars) can be utilized, and what presenter types should be used to effect changes in purchasing decisions when introducing products. Further, such investigation is also meaningful in terms of studying persuasion and in practical situations.

In the field of marketing, numerous studies have investigated the effects of corporate advertising on consumers’ willingness to purchase diverse products (Krugman, 1965; Park and Young, 1986).

Recent marketing studies have also shown that TV, print, and other advertising, as well as celebrity endorsements, influence purchase intentions (Arshad and Aslam, 2015). Moreover, advertising entertainment, advertising familiarity, social imaging, and advertising spending influence purchase behavior (Haider and Shakib, 2018). Research on smartphone advertising has also shown the importance of contextual advertising and other types of advertising that are location- and time-specific (Lee et al., 2017).

Prior research has studied persuasion in the context of purchase decisions on websites and other sources (Hopkins et al., 2004). Moreover, several studies have suggested that the use of avatars representing companies and products on websites can improve attitudes toward products and improve user satisfaction, as well as willingness to purchase (Choi et al., 2001; Holzwarth et al., 2006). However, prior studies have not examined the differences in persuasiveness between human and avatar presenters in product promotional videos using experimental designs that are similar to real environments.

Therefore, this study investigates the persuasive effects of using YouTubers or VTubers in product promotional videos (i.e., videos introducing specific products) using an experimental design that is similar to the actual video promotion and viewing environment. Specifically, we attempt to determine whether YouTubers or VTubers exhibit persuasive effects, such as influencing purchase decisions, identify which factors contribute to their persuasiveness, and highlight the differences between them. This study clarifies the differences in persuasive effects between humans and avatars in the modern and practical setting of product promotion through YouTube videos. As the existing evidence suggests that avatar attractiveness affects the favorability perceived by users, and favorability affects persuasiveness, we also explore the role of favorability in this study (Keeling et al., 2010; Khan and Sutcliffe, 2014).

We must note that prior research has investigated the impact of interaction with salespeople using avatars on product promotion. Various studies have attempted to determine whether avatars (in the form of interactive agents on websites) can affect purchase decisions (McGoldrick et al., 2008; Keeling et al., 2010). For example, multimodal interactions between users and avatars providing product information have been shown to enhance the enjoyment of the online shopping experience (Jin and Bolebruch, 2009). Moon et al. (2013) have demonstrated that interactions between users and salespeople and peers in virtual stores can increase users’ social presence, shopping enjoyment, positive attitudes toward brands, and willingness to purchase. Nevertheless, this study examines the impact on purchase decisions in the practical setting of video promotion.

Persuasion is also one of the key themes in psychology, with various studies investigating persuasion in many domains, not limited to purchase decisions. For example, research has been conducted on the differences in persuasiveness between humans and avatars in the field of education. Some studies have shown that non-verbal expressions are also important for the persuasive ability of robots (Chidambaram et al., 2012). The influence of eye gaze has also been studied in terms of persuasive strategies using robots (Ham et al., 2015). In terms of learning effectiveness in expressive education, several studies have investigated whether

study participants change their minds when exposed to a lecture (Zanbaka et al., 2006). In these studies, the experimental setting was such that participants were directly persuaded regarding a single solution to a problem on which they were divided (Jamy, 2015; Baxter et al., 2017; Hashemian et al., 2019).

The current study, however, uses content that includes entertainment elements in addition to the persuasive content, that is, product promotion. In other words, not all of the video content is related to persuasion. Additionally, in this study, we use professional-quality videos that are similar to actual YouTube promotional videos.

Against this backdrop, we intend to see if we can be persuasive through our avatars by studying whether VTubers (avatars) can be used to promote products on YouTube. Because this study uses a practical experimental environment, it is significant as an empirical study in marketing and as a study of persuasion in psychology.

To determine whether YouTubers (human appearance) or VTubers (avatar appearance) are more persuasive when promoting products, which factors contribute to this difference, and in what way, a human YouTuber wearing a motion-capture suit under his clothes filmed a product promotion video. By utilizing the captured motion, we were able to produce a product promotional video for the VTuber. We asked a group of viewers to watch these videos with the same audio, composition, and other conditions, except for the presenters' appearance, and compared the differences in persuasiveness through a questionnaire. In the questionnaire, based on previous research on persuasion, users were asked about the impressions they had of the videos and presenters after viewing the promotional videos (Mullennix et al., 2003).

Our study aims to answer the following two research questions:

1. How do YouTubers and VTubers influence their persuasiveness and viewers' purchase decisions when promoting products using videos, and what are the differences between them?
2. What are the mechanisms through which the impressions about the promotional videos and video contributors influence their persuasiveness (i.e., persuasiveness structural model)?

## Materials and methods

### Participants

Using a social media application (Twitter), we recruited 318 participants—mostly students from Kwansei Gakuin University and Osaka University—without gender segregation. The cases of participants excluded from the study are discussed later.

### Research design

We employed a between-subjects experimental design. Immediately after submitting the application, the respondents were asked to complete a pre-questionnaire to gauge their state of mind. The participants were then randomly divided into groups to watch a product promotional video presented by either a YouTuber or a VTuber. Afterwards, the participants in each group watched promotional videos for two different product categories (tapioca

drinks and game apps). After viewing the videos, they were asked to complete a post-questionnaire about their impressions of the presenter and the video content, as well as their willingness to purchase the product. The participants were paid a gratuity of 1,500 Yen.

This study was reviewed and approved by the Research Ethics Review Committee of Kwansei Gakuin University's "Behavioral Research on Human Subjects." Informed consent was obtained from the participants by means of written informed consent forms.

## Materials

For the experiment, product promotional videos were created that differed only in the appearance of the presenter (i.e., a human or an avatar). We created videos for two product categories (i.e., tapioca drinks and game apps) because they are familiar to young people. We struck a balance by selecting the two products from different product categories: food and beverages (tapioca drinks) and entertainment (game apps). As VTubers are avatars, they cannot really consume tapioca drinks. The intention was also to check if this would make a difference.

Specifically, the YouTuber introduced the product in a filming studio, which was then edited to create a video of a human (YouTuber) introducing the product. However, this presenter was wearing a tracking suit, and his body position and movements were recorded in the same chronological order by the tracking system in the filming studio. In addition, a 3DCG model of an avatar of a character based on this presenter was created in advance. By moving and recording this avatar model in the same manner as actual human movements, a product promotional video involving the avatar (VTuber) was also created. To make it viewer-friendly, the avatar was designed by a well-known professional Japanese character designer. This was because existing literature demonstrates that avatars with realistic human appearances may seem "creepy" (Tinwell et al., 2011). The videos were also edited with the help of a major VTuber studio.

To ensure that the videos would not look out of place when posted on YouTube as actual YouTuber and VTuber videos, the product promotional videos were produced by a professional team and studio that actually produces and delivers YouTuber and VTuber videos. Perception Neuron Pro was used as the tracking suit, Unity was used as the software to manipulate the 3DCG models, and Adobe Premiere was used for video editing.

The two product promotional videos differed only in the appearance of the presenter. However, the content of speech, audio, and video composition were identical, as shown in Figure 1. The length of the videos was approximately 9 and 6 min for the tapioca drink and game app, respectively.

## Procedure

As noted above, participants were recruited through social media. They were then asked to complete a pre-survey generated on Survey Monkey. Following this, after a period of 1–2 weeks, they were asked to watch the source video. Immediately



following the viewing, participants were asked to complete a post-questionnaire. The following subsection describes the content of the questionnaire.

## Questionnaire summary

The pre-questionnaire included items measuring participant demographics and their willingness to purchase the tapioca drinks and game apps. The post-questionnaire did not ask for any information about the user but asked the same questions about their willingness to purchase the products, using exactly the same format as in the pre-questionnaire. Both the VTuber and YouTuber groups responded to the same questionnaire.

The post-questionnaire was more voluminous than the pre-questionnaire. Dyson's persuasiveness rating scale was employed as the primary rating instrument (Mullennix et al., 2003). This persuasiveness rating scale measures effectiveness of the product promotion, perception toward the message, and perception toward the presenter. To compare with the synthesized persuasiveness index that was calculated later, the perceived persuasiveness toward the presenter was directly evaluated using one question item.

The participants were also asked about their overall impression, including the favorability felt toward the presenter, perceived trustworthiness of the presenter, eye contact felt with the presenter, closeness between the presenter and the participants, and qualities of the product promotional video.

The post-questionnaire response time was measured to determine if the entire video was viewed appropriately. This

included the time spent watching the video and the minimum response time to the questionnaire. Further, we included a brief set of questions to ascertain whether the video was watched. These questions were designed to exclude respondents who either did not watch the video or did not take the video seriously.

## Questionnaire details

The pre-questionnaire asked for information about the user (sex, personality traits, anime viewing preferences, and familiarity with VTubers and YouTubers).

The purchase decision was examined by ranking the products the participants would like to purchase. For each tapioca drink and game app, seven different products were prepared. Participants were asked to rank the products in the order in which they would like to purchase them. They were asked to rank the seven products in the pre-questionnaire and to repeat the process in the post-questionnaire to measure how the rankings varied. This was based on a questionnaire used in an existing agent persuasion study (Ogawa et al., 2009).

To assess persuasiveness, we used Dyson's persuasiveness rating scales, which are used as a measure of an agent's persuasiveness (Mullennix et al., 2003). Effectiveness of the product promotion was rated on a 9-point Likert scale for multiple adjective pairs provided to the participants for each subscale. Perception toward the message and the presenter was rated on a 7-point Likert scale. Each adjective pair is shown below (adjective pairs marked with an asterisk "\*" are reversal items).



Effectiveness of the product promotion: Bad—Good, Foolish—Wise, Negative—Positive, Beneficial—Harmful, Convincing—Unconvincing, Effective—Ineffective.

Perception toward the message: Flamboyant—Conservative, \*Stimulating—Boring, Vague—Specific, Unsupported—Supported, Complex—Simple, \*Convincing—Unconvincing, Boring—Interesting.

Perception toward the presenter: Unintelligent—Intelligent, \*Straightforward—Evasive, \*Active—Inactive, \*Qualified—Unqualified, \*Sincere—Insincere, Meek—Forceful, Incompetent—Competent, \*Honest—Dishonest, Unassertive—Assertive, Uninformed—Informed, Untrustworthy—Trustworthy, Timid—Bold, Loud Voice—Soft-Spoken Voice, Deep Voiced—Squeaky Voiced, Fast Speaking—Slow Speaking, Heavy Accent—Faint Accent, Talked Too Long—Did not Talk Long Enough, Heavy Nasality—Faint Nasality, Monotone—Lively.

The overall impression included questions on favorability felt toward the presenter, the perceived trustworthiness of the presenter, eye contact felt with the presenter, and closeness between the presenter and the participants. Each item was evaluated directly using one question, as provided below, following which the responses were obtained on a 7-point Likert scale.

- How favorable was your impression of the presenter?
- How trustworthy did you think the presenter was?
- To what extent did you feel that the presenters looked at you when they talked to you?

Closeness refers to the degree of similarity between the participant and the presenter, as perceived by the participants. To evaluate closeness, we employed the Inclusion of Other in the Self Scale (Aron et al., 1992). This scale indicates the degree of overlap between representations of self and others, as indicated by the overlap of the two circles. In this study, the assessment was obtained using a 7-point Likert scale.

To estimate impressions of the videos, we included the following questions about the likability, completeness, and interestingness of each viewed video. The responses were obtained using a 7-point Likert scale.

- How much did you like the product promotional video that you watched?
- How good was the quality of the video for product promotion?
- How interesting was the content of the product introduction video?

## Data analysis

For the actual analysis, participants (those who watched the videos till the end and responded the questions) were filtered using the following procedure. First, we selected the respondents who spent more than 20 min, or at least longer than the length of the video, answering the questionnaire. We selected 20 minutes as the threshold based on the results of time measurements with a pilot sample of approximately 10 people. Twenty minutes is slightly longer than the minimum time required to have watched all the videos. Then, the respondents who correctly answered questions

that could be easily answered if they had watched the video (e.g., the episode played in the game app video, the flavor of the drink featured in the tapioca drink video) were picked.

In this study, Dyson's measure of persuasiveness consisted of three categories: effectiveness of the product promotion, perception toward the message, and perception toward the presenter. Cronbach's alpha coefficients were used to confirm consistency within these measures. For the analysis of Research Question 1, we conducted a two-factor analysis of variance (ANOVA) between participants in the YouTuber and VTuber groups and the pre- and post-questionnaire. For subsequent analyses, these measures were combined using principal component analysis to create a synthesized persuasiveness index. The validity of this index was confirmed by checking the contributions of the principal components, as well as by correlating them with the overall impression of the presenter's persuasiveness, which had been answered beforehand.

In addition, the impressions respondents had of the videos and presenters for the tapioca drink and the game app were obtained separately. Therefore, we could verify whether the impressions significantly differed by product category. Specifically, we tested the possibility that the participants might have thought that the avatar was not consuming the tapioca drink, thus affecting the results. Using the cosine similarity measure, consistency (similarity) was calculated to evaluate the consistency of the respondents' impressions of the presenter in each video. Taking the responses to each impression item as a vector value, the inner product of the vector of impressions from the tapioca drink video and the vector of impressions from the game app video was divided by their norm. If the measure was close to 1, then the respondents had the same impression, regardless of the video content. Conversely, if it was close to 0, the respondents' impressions varied greatly, depending on the video content. Cosine similarities were determined for each participant and their means were calculated.

To use Dyson's measure of persuasiveness (i.e., the synthesized persuasiveness index) as the objective variable in the multiple regression analysis of Research Question 2, its principal components had to be valid. The explanatory variables included the overall impressions (the presenter's favorability, presenter's trustworthiness, presenter's eye contact, closeness with the presenter, likability of the video, completeness of the video, and interestingness of the video) and whether a VTuber or a YouTuber was featured in the video.

## Results

As mentioned above, 318 participants were initially recruited. Then, to filter the data, we only included in the analysis those who had responded to both the pre- and post-questionnaires, which resulted in 248 participants for analysis. Following this, unserious respondents were excluded from the analysis, and the number of participants was reduced. Specifically, we excluded those who responded to the post-questionnaire in less than 20 min (13 respondents) and those who gave incorrect answers to simple questions measuring whether they had watched the videos properly (39 respondents). In the end, 196 respondents were included in the analysis.

TABLE 1 Changes in rankings of the two product domains.

	Tapioca drinks (M ± SD)		Game apps (M ± SD)	
	Pre-ranking	Post-ranking	Pre-ranking	Post-ranking
VTuber	4.247 ± 2.220	2.660 ± 1.559	3.371 ± 1.841	3.670 ± 2.228
YouTuber	3.717 ± 2.137	3.162 ± 1.846	3.869 ± 2.048	3.000 ± 2.000

M, median; SD, standard deviation; VTuber, Virtual YouTuber.

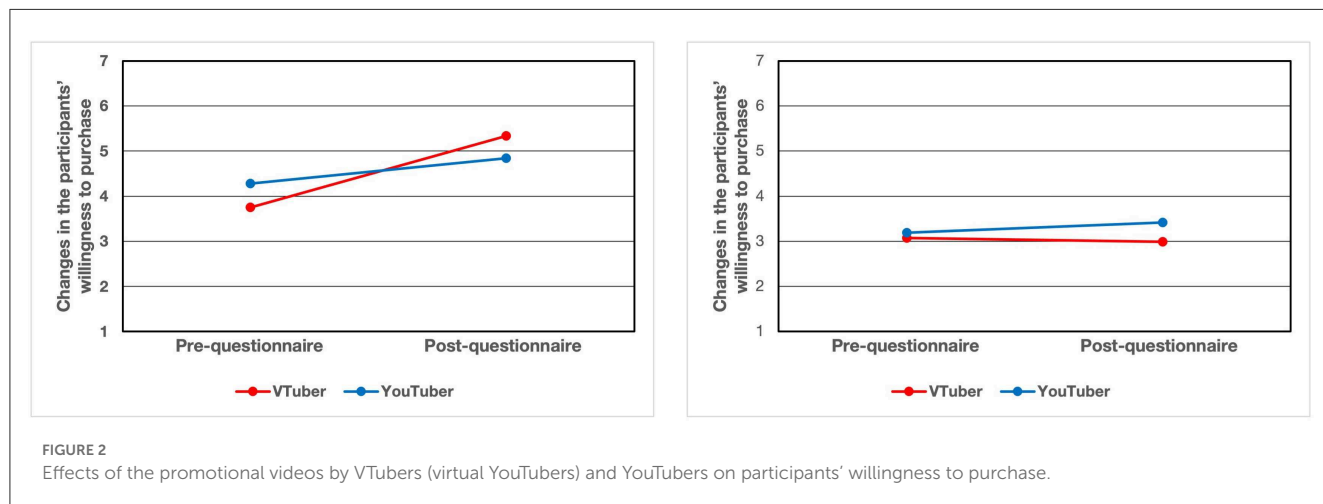


FIGURE 2

Effects of the promotional videos by VTubers (virtual YouTubers) and YouTubers on participants' willingness to purchase.

In addition, we checked the consistency of the main evaluation measure of this study: Dyson's measure of persuasiveness ratings. Specifically, we checked the Cronbach's alpha coefficients for each measure across participants in the YouTuber and VTuber groups and for each product category (tapioca drinks and game apps). The values were greater than 0.7 under both conditions, confirming the consistency of the responses.

Next, we discuss the results for each research question.

1. How do YouTubers and VTubers influence their persuasiveness and viewers' purchase decisions when promoting products using videos, and what are the differences between them?

First, we measured the effect of the product promotional videos on the respondents' willingness to purchase. The participants were asked to rank several product groups, including those promoted in the videos, according to their willingness to purchase, both before and after watching the videos. For each product, the change in ranking was calculated by subtracting the pre- from the post-ranking. We averaged the change in rankings for each participant and used ANOVA to compare the results of the VTubers and YouTubers. Although the mean was higher for YouTubers (median ± standard deviation: 0.712 ± 1.787) than for VTubers (0.644 ± 1.629), we found no significant differences in the variation in the rankings [ $F_{(0, 195)} = 0.076, p = 0.7829$ ].

Then, we analyzed the differences by product category. As illustrated in Table 1 and Figure 2, the participants' rankings of the tapioca drinks and game apps were analyzed using ANOVA to measure the differences between the VTuber and YouTuber groups and before and after viewing. For the tapioca drinks, the results showed a main effect for pre- and post-ranking, with a significant increase in purchase intent ranking [ $F_{(1, 194)} = 44.4, p$

< 0.001]. There was also an interaction effect [ $F_{(1, 194)} = 10.3, p < 0.005$ ]. Then, a back-test showed that the changes in rankings for participants in both VTuber [ $F_{(1, 194)} = 48.7, p < 0.001$ ] and YouTuber groups were significant [ $F_{(1, 194)} = 5.96, p = 0.016$ ]. By contrast, there was no main effect for game apps. However, there was an interaction effect [ $F_{(1, 194)} = 8.85, p < 0.005$ ]. Further, a back-test demonstrated that participants in the YouTuber group experienced more changes in rankings compared to those in the VTuber group [ $F_{(1, 194)} = 9.8, p < 0.005$ ].

In terms of persuasion details, the respondents were asked about their impressions of the promotion in the videos they watched, the content of the messages, and the presenters, with 6, 7, and 19 items, respectively. The detailed data of the persuasiveness rating scale are shown in Table 2. The ANOVA revealed that VTubers sounded more conservative in their messages than YouTubers. Additionally, we found that the YouTubers' messages were supported more than that of the VTubers, and that the YouTube presenters' speech did not seem like it was longer than that of the VTubers'. We then synthesized indicators of persuasiveness to ascertain and identify the differences in persuasiveness between the VTuber and YouTuber groups, and to serve as one objective variable in the multiple regression analysis. We combined the respondents' impressions of multiple items (32 items) in three categories (i.e., effectiveness of the product promotion, perception toward the message, perception toward the presenter) into a single index, as shown in Figure 3. Specifically, a principal component analysis was conducted to synthesize the impressions held about both videos and summarize the impressions held about these categories. The contribution of the first principal component (the synthesized persuasiveness index) was 0.833, which was sufficiently representative. Meanwhile,

TABLE 2 Data from the persuasiveness rating scales.

	Presenter (M ± SD)		Contrast	
	VTuber	YouTuber	F	p
<b>Effectiveness of the product promotion</b>				
Good	4.740 ± 1.333	4.566 ± 1.296	0.78	ns
Wise	4.186 ± 1.334	4.333 ± 1.287	0.62	ns
Positive	4.773 ± 1.351	5.040 ± 1.279	2.00	ns
Harmful	4.866 ± 1.503	4.495 ± 1.449	2.99	<sup>+</sup> p < 0.10
Unconvincing	4.660 ± 1.338	4.596 ± 1.449	0.10	ns
Ineffective	4.402 ± 1.352	4.495 ± 1.431	0.22	ns
<b>Perception toward the message</b>				
Conservative	3.660 ± 1.243	3.192 ± 1.032	8.16	**p < 0.01
Boring	3.804 ± 1.462	3.838 ± 1.454	0.03	ns
Specific	3.876 ± 1.501	4.040 ± 1.524	0.57	ns
Supported	3.691 ± 1.417	4.182 ± 1.720	4.70	*p < 0.05
Simple	5.402 ± 0.991	5.222 ± 1.069	1.47	ns
Unconvincing	4.381 ± 1.280	4.182 ± 1.438	1.04	ns
Interesting	4.103 ± 1.696	3.778 ± 1.703	1.78	ns
<b>Perception toward the presenter</b>				
Intelligent	3.928 ± 1.667	4.121 ± 1.423	0.75	ns
Evasive	4.907 ± 1.437	5.242 ± 1.102	3.33	<sup>+</sup> p < 0.10
Inactive	4.010 ± 1.396	3.889 ± 1.449	0.35	Ns
Unqualified	4.299 ± 1.507	4.141 ± 1.470	0.54	ns
Insincere	5.051 ± 1.271	4.879 ± 1.008	1.10	ns
Forceful	3.670 ± 1.146	3.889 ± 1.413	1.40	ns
Competent	4.155 ± 1.230	4.364 ± 1.185	1.45	ns
Dishonest	5.268 ± 1.312	5.475 ± 1.001	1.52	ns
Assertive	3.577 ± 1.299	3.768 ± 1.347	1.00	ns
Informed	4.062 ± 1.314	4.323 ± 1.582	1.56	ns
Trustworthy	4.271 ± 1.373	4.404 ± 1.490	0.64	ns
Bold	3.784 ± 1.245	3.818 ± 1.201	0.04	ns
Soft	3.784 ± 1.160	3.980 ± 1.442	1.09	ns
Squeaky	3.804 ± 1.012	3.990 ± 0.732	2.15	ns
Slow	5.196 ± 1.154	5.404 ± 1.053	1.72	ns
Unaccented	5.897 ± 1.272	5.586 ± 1.231	2.00	<sup>+</sup> p < 0.10
Not Long	4.134 ± 1.660	4.626 ± 1.508	4.68	*p < 0.05
Less nasal	4.773 ± 1.702	4.657 ± 1.478	0.26	ns
Lively	2.959 ± 1.399	3.051 ± 1.553	0.19	ns

M, median; SD, standard deviation; VTuber, virtual YouTuber. <sup>+</sup>p < 0.10, \*p < 0.05, \*\*p < 0.01.

the contribution of the second principal component was only 0.053, which mainly accounted for the respondents' impressions of the presenters. The loadings of the promotional videos, message, and presenters on the persuasiveness index were 0.726, 0.553, and 0.409, respectively. For the synthesized persuasiveness

index, we used the average scores of the tapioca drinks and game apps.

Meanwhile, by considering each item for each video as a vector (32-dimensional vector with 32 items as elements in three categories), we could calculate how close (i.e., consistent) the

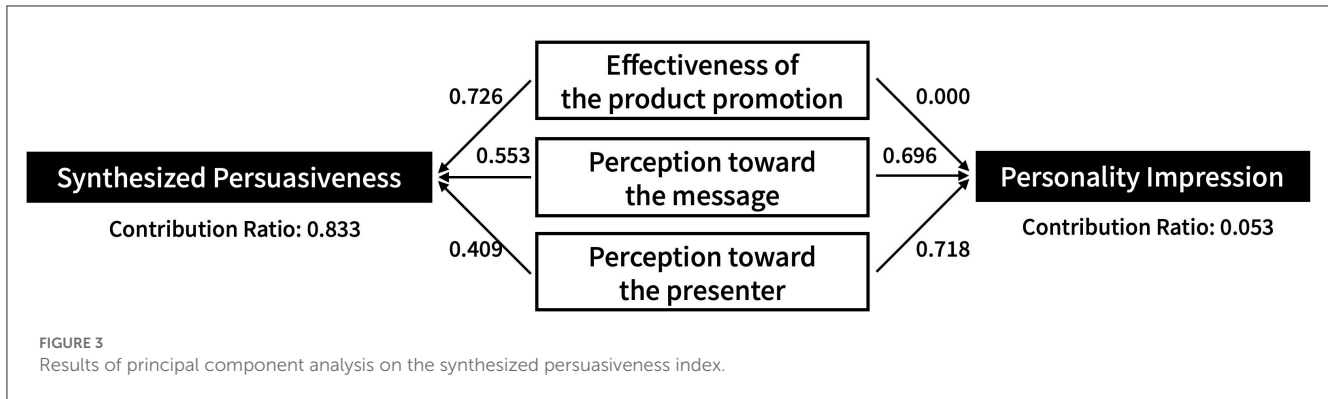


FIGURE 3 Results of principal component analysis on the synthesized persuasiveness index.

TABLE 3 Cosine similarity between the two domains.

	Effectiveness of the product promotion (M ± SD)	Perception toward the message (M ± SD)	Perception toward the presenter (M ± SD)
VTuber	0.936 ± 0.062	0.939 ± 0.043	0.951 ± 0.040
YouTuber	0.953 ± 0.041	0.947 ± 0.0034	0.956 ± 0.033

M, median; SD, standard deviation; VTuber, virtual YouTuber.

impressions formed based on the tapioca drinks video were to the impressions created based on the game apps video, in terms of cosine similarity. The cosine similarity was calculated as the inner product of the vector of impressions formed based on the tapioca drink video and the vector of impressions created based on the game app video, divided by their respective norms. For each participant, it is possible to determine whether each vector group of impressions perceived in the tapioca drinks video matches each vector group of impressions perceived in the game apps video. The cosine similarities were mostly close to 1, as shown in Table 3. As the impressions formed on both videos are very similar, their average can be used to create a measure of persuasiveness. However, there was a difference between the YouTuber and VTuber groups in terms of consistency of their impressions about the two promotional videos, with the YouTuber group being more consistent in their perceptions than the VTuber group [ $F_{(1, 194)} = 4.68, p = 0.032$ ]. The perceptions about the message and presenters showed no differences in consistency.

Figure 4 shows the differences between the VTuber and YouTuber groups on the synthesized persuasiveness index, with the YouTuber group showing significantly more perceived persuasiveness [ $F_{(1, 194)} = 7.31, p = 0.0075$ ]. The overall evaluation also included an item directly measuring the presenters' perceived persuasiveness. The correlation coefficient between this item and the synthesized persuasiveness index was 0.70, implying a high correlation. Additionally, the correlation coefficient with the aforementioned ranking—that is, change in the willingness to purchase—was 0.45, indicating a correlation trend.

2. What are the mechanisms through which the impressions about the promotional videos and video contributors influence their persuasiveness (i.e., persuasiveness structural model)?

With the synthesized persuasiveness index as the objective variable, we conducted multiple regression analysis using the following

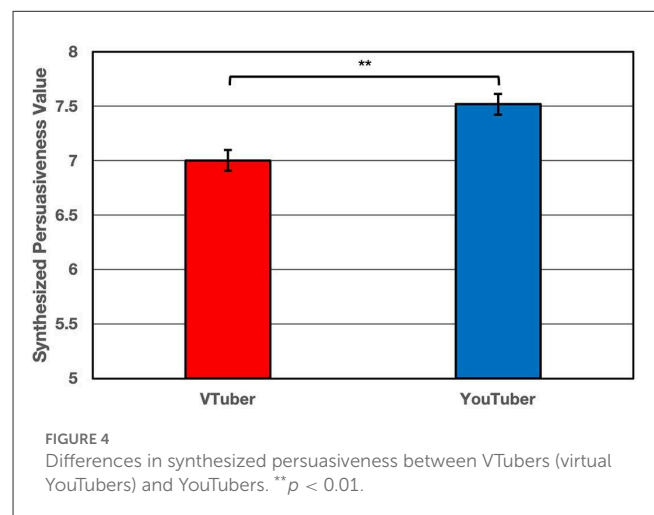
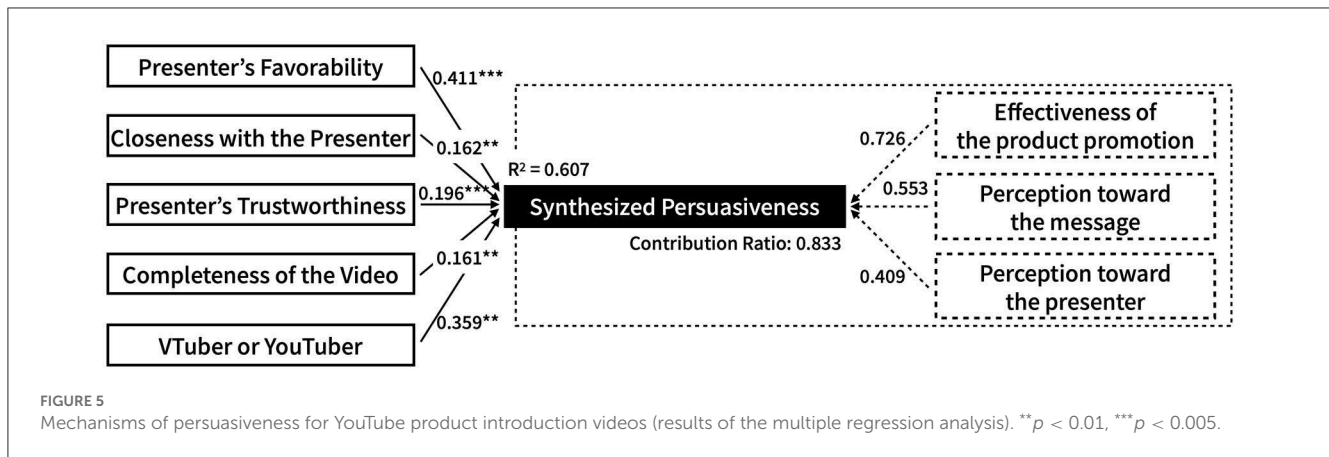


FIGURE 4 Differences in synthesized persuasiveness between VTubers (virtual YouTubers) and YouTubers. \*\* $p < 0.01$ .

explanatory variables: the overall impressions (the presenters' favorability, presenters' trustworthiness, presenters' eye contact, closeness with the presenter, likability of the video, completeness of the video, interestingness of the video) and whether the presenter was a VTuber or a YouTuber. The results of the multiple regression analysis revealed that persuasiveness was explained by the participants' favorability toward the presenter, closeness with the presenter, presenters' trustworthiness, completeness of the video, and presenter type (a VTuber or YouTuber), as illustrated in Figure 5.

Particularly influential was the favorability of the video contributor (presenter; coefficient: 0.41), followed by presenter type (a VTuber or YouTuber; coefficient: 0.36); the higher the favorability of the presenter, the more persuasive. The model is well-represented with an adjusted coefficient of determination of 0.61.





For the purposes of subsequent discussion, we also analyzed the differences for each item of the overall evaluation. The results are listed in Table 4. For the aforementioned indicators, there were no significant differences between the YouTuber and VTuber groups in the presenters' favorability [ $F_{(1, 194)} = 0.001$ ,  $p = 0.971$ ], presenters' trustworthiness [ $F_{(1, 194)} = 1.857$ ,  $p = 0.175$ ], completeness of the video [ $F_{(1, 194)} = 0.969$ ,  $p = 0.326$ ], and interestingness of the video [ $F_{(1, 194)} = 0.203$ ,  $p = 0.653$ ]. Meanwhile, the presenter's eye contact and closeness with the presenter were significantly higher among respondents in the YouTuber group than those in the VTuber group, with [ $F_{(1, 194)} = 17.7$ ,  $p < 0.001$ ] and [ $F_{(1, 194)} = 17.7$ ,  $p < 0.001$ ], respectively.

## Discussion

### Research question 1

First, referring to Ogawa et al.'s (2009) study on product promotion by robots, we conducted an experiment to examine the changes in purchase decisions. We found a main effect for the changes in the willingness to purchase tapioca drinks, with a significant improvement in ranking. We also identified an interaction effect, with a significant change in ranking for viewers of both VTubers and YouTubers when back-testing was conducted. In contrast, there was no main effect for the game app; however, there was an interaction effect, with the YouTuber group reporting significantly greater fluctuations in ranking than the VTuber group. For the game app video, the results of the analysis of individual items also showed that the respondents formed more positive impressions about the promotional videos and the content of the messages presented by human YouTubers compared to VTubers (avatars).

When the averages of the tapioca drinks and game videos are compared, the average for YouTubers is higher. However, the results of the change in the ranking for tapioca drinks is greater for VTubers than for YouTubers. This suggests that some products are better or worse in certain domains than others. However, one possible problem with the experimental design is that the questionnaire for the ranking changes was administered after viewing the product introduction video for THE ALLEY

(the target brand), which may have led the participants to believe that the experimenter expected an improvement in THE ALLEY's ranking. It is also possible that it would have been difficult for the participants to sort through the pictures of each brand of tapioca drink and the text of its characteristics and ask them about their attitudes toward the ambiguous sensation of taste. We plan to analyze the changes in attitude and behavior induced by persuasion by conducting further experiments in the future.

However, this does not mean that human influencers are always effective in persuasion, while virtual ones (avatars) are ineffective. Indeed, our results showed that promotional videos presented by both humans and avatars can cause a change in purchase intent depending on the product category (or video content).

Further, Dyson's measures of persuasiveness, which assessed the impressions about the effectiveness of the product promotion, perception toward the message, and perception toward the presenter, were synthesized using principal component analysis. The contribution ratio of the synthesized persuasiveness index was 0.833, indicating good representation of persuasiveness. This was used as an evaluation index for persuasiveness in the multiple regression analysis described below. The second principal component loaded heavily on the impression about presenters; however, its contribution ratio was 0.053, indicating that it could not represent persuasiveness to a great degree. For comparison, the overall evaluation also directly explored impressions about persuasiveness, and the correlation coefficient with the persuasiveness index was highly correlated at 0.70. When this persuasiveness index was used to compare the VTuber and YouTuber groups, respondents in the YouTuber group were more significantly persuaded about the product. In other words, humans have greater persuasive power than avatars.

Previous studies have compared the persuasive power of humans and avatars and found that virtual characters can be similarly persuasive (Zanbaka et al., 2006, 2007). In particular, they pointed out that androids can be as persuasive as humans (Ogawa et al., 2009). Using the YouTube environment, our results do not differ significantly from theirs. However, we show that differences are affected by the content of the video and the experimental environment setting.

The loadings of the effectiveness of the product promotion, perception toward the message, and perception toward the

TABLE 4 Ratings for overall impression.

	The presenter's favorability (M ± SD)	The presenter's eye contact (M ± SD)	The closeness with the presenter (M ± SD)	The presenter's trustworthiness (M ± SD)	The completeness of the video (M ± SD)	The interestingness of the video (M ± SD)
VTuber	4.10 ± 1.49	3.45 ± 1.32	1.81 ± 0.79	4.46 ± 1.32	4.07 ± 1.32	3.78 ± 1.57
YouTuber	4.11 ± 1.51	4.26 ± 1.36	2.27 ± 1.28	4.46 ± 1.17	4.26 ± 1.34	3.69 ± 1.39

M, median; SD, standard deviation; VTuber, virtual YouTuber.

presenter on the synthesized persuasiveness index were 0.726, 0.553, and 0.409, respectively; this indicates that the quality of the video, the message articulated, and the viewer's impression about the presenter, in that order, affect persuasiveness.

In addition, we used cosine similarity to estimate the consistency of the respondents' impressions about videos involving the two product categories: tapioca drinks and game apps. Most of the cosine similarities were close to 1, as shown in Table 3, indicating that the impressions formed based on the two videos are very similar. Thus, the average of the tapioca drinks and game apps could be used to create a measure of persuasiveness. However, there were differences in the consistency of impressions about the promotion videos, with the VTuber group experiencing less consistency in the impressions about the two video promotions. It is possible that the impression of a conservative explanation was helpful in introducing the tapioca drink, while the impression of a well-explained and reasoned explanation was helpful in introducing the game. One possible reason is the fact that the respondents experienced a less informative facial impression from the avatar compared to that from the human presenter.

## Research question 2

In this study, the persuasiveness structural model and the underlying mechanisms were examined through multiple regression analysis using seven items as explanatory variables: the overall impressions (the presenter's favorability, presenter's trustworthiness, presenter's eye contact, closeness with the presenter, likability of the video, completeness of the video, and interestingness of the video), and the presenters' appearance (human or avatar). Indeed, the objective variable was an index of persuasiveness that had been examined using a principal component analysis and other methods, and the validity of this index as being representative of persuasiveness was discussed in the previous section.

The results of the multiple regression analysis showed that persuasiveness was explained by the presenter's favorability (coefficient: 0.41), presenter type (VTuber or YouTuber; coefficient: 0.36), presenter's trustworthiness (coefficient: 0.20), closeness with the presenter (coefficient: 0.16), and completeness of the video (coefficient: 0.16). As the coefficient of determination was 0.61, the model was considered to be reasonably well-represented.

Specifically, the presenter's favorability has the greatest impact on persuasiveness, which is consistent with previous studies (Keeling et al., 2010; Khan and Sutcliffe, 2014). Additionally, whether the presenter is a VTuber (avatars) or a YouTuber

(humans) also has a significant impact, with humans having more persuasive power. Persuasiveness is also likely to vary depending on trust in the presenter, degree of closeness to the presenter, and the quality of the video. Thus, we suggest that designing avatars with a high level of trustworthiness and closeness to the audience may increase persuasiveness. While it is difficult to create or change human appearance so that it is highly trustworthy and highly relatable, it is easy to change the appearance of avatars. Further, the viewers' degree of closeness to the presenter is significantly higher for YouTubers than for VTubers, suggesting that there is room for improvement in the future. What constitutes a reliable avatar, and what type of avatar one perceives as relatable are issues that should be investigated by future studies.

Some studies have found that people feel more favorability and trust toward virtual agents that mimic participants' head movements than those that do not (Verberne et al., 2013). Hence, in the future, presentations by avatars should be partially automated, with the possibility of generating on-the-fly videos that mimic the user and gradually change their behavior. Such innovations may aid in developing more persuasive promotional videos by avatars.

Owing to some technical aspects, the YouTuber made the audience feel that he was looking at them significantly more than the VTubers in terms of the presenter's eye contact. However, the impact on persuasiveness was limited.

## Limitations

For both humans and avatars, the study has a limitation in that only one male presenter was considered. As research has shown that women are more easily persuaded by male avatars and men are more easily persuaded by female avatars (Zanbaka et al., 2006), we intend to conduct further experiments with female YouTubers and VTubers.

Moreover, avatar designs were created by professional designers, with general digital avatars (anime-style avatars) familiar to Japanese participants. In the future, we intend to expand on this research using multiple presenters, as outside Japan, YouTubers and VTubers are in demand in different ways, and the results may vary.

## Conclusions

This study examined the characteristics of persuasiveness for human and avatar presenters and the differences between them in this regard, in the setting of product promotional videos on YouTube. Although the findings show that humans are more

persuasive than avatars, the persuasive effect can vary, depending on the product category. Further, it is possible that different avatar design techniques can increase persuasiveness.

Using a between-subjects experimental design, with the assistance of professional character designers and video creators, we created videos with exactly the same audio, angle of view, and composition for a YouTuber with a so-called human appearance and a VTuber using an avatar with a character-like 3DCG model. After viewing the videos, the participants were asked to complete a questionnaire about their impressions of the presenters and the videos related to persuasiveness, as well as overall impression measures, such as favorability and trustworthiness. Changes in willingness to purchase the products presented in the videos were also measured before and after the experiment.

Although there were differences depending on the product category, humans were more likely than avatars to alter participants' willingness to purchase. However, product promotions by avatars also influenced the willingness to purchase in the case of tapioca drinks. Regarding persuasiveness, the presenter's favorability and presenters' appearance (human or avatar) had a significant impact. The results also suggested that persuasiveness could be enhanced by designing avatars that are more trustworthy and closer to the audience. In this regard, future research should explore how to design a more persuasive appearance through variation in avatar appearance or using techniques that generate spontaneous movements by the avatars in response to the user.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by Kwansai Gakuin University. The patients/participants provided their written informed consent to participate in this study.

## References

- Allocca, K. (2020). *Analyzing Pop Culture with YouTube's Culture and Trends Report*. Available online at: <https://blog.youtube/culture-and-trends/analyzing-pop-culture-youtubes-culture-trends-report/>
- Aron, A., Aron, E. N., and Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *J. Pers. Soc. Psychol.* 63, 596–612.
- Arshad, M. S., and Aslam, T. (2015). *The Impact of Advertisement on Consumer's Purchase Intentions*. Available online at: <https://ssrn.com/abstract=2636927> (accessed July 12, 2022).
- Baxter, P., Ashurst, E., Read, R., Kennedy, J., and Belpaeme, T. (2017). Robot education peers in a situated primary school study: personalisation promotes child learning. *PLoS ONE* 12, e0178126. doi: 10.1371/journal.pone.0178126
- Brown, D., and Hayes, N. (2008). *Influencer Marketing: Who Really Influences your Customer?* London: Routledge.
- Chidambaram, V., Chiang, Y. H., and Mutlu, B. (2012). "Designing persuasive robots: how robots might persuade people using vocal and nonverbal cues," in *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, 293–300.
- Choi, Y. K., Miracle, G. E., and Biocca, F. (2001). The effects of anthropomorphic agents on advertising effectiveness and the mediating role of presence. *J. Interact. Mark.* 2, 19–32. doi: 10.1080/15252019.2001.10722055
- Freeman, B., and Chapman, S. (2007). Is "YouTube" telling or selling you something? Tobacco content on the YouTube video-sharing website. *Tob. Control.* 16, 207–210. doi: 10.1136/tc.2007.020024
- Haider, T., and Shakib, S. (2018). A study on the influences of advertisement on consumer buying behavior. *Bus. Stud. J.* 9.

## Author contributions

HS designed the experiments, compiled the data, wrote the first draft of the paper, performed the analyses, and contributed to fundraising. AH and MM performed the experiments, performed the analyses, and contributed to preparing the manuscript. CK supported the experiments and data analyses. YH designed the experiments and contributed to the experimental design and preparing the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was mostly supported by the Masason Foundation and partially supported by JST, CREST Grant Number JPMJCR20D4, Japan.

## Acknowledgments

We thank Mogumo and MUGENUP for designing the avatars used in the experiments and Punch Entertainment (Vietnam) Co., Ltd. for modeling the avatars and creating the Unity package. We thank Activ8 for their help in creating the YouTuber and VTuber videos.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Ham, J., Cuijpers, R. H., and Cabibihan, J. J. (2015). Combining robotic persuasive strategies: the persuasive power of a storytelling robot that uses gazing and gestures. *Int. J. Soc. Robot.* 7, 479–487. doi: 10.1007/s12369-015-0280-4
- Hashemian, M., Mascarenhas, A. S., Santos, P. A., and Prada, R. (2019). “The power to persuade: a study of social power in human-robot interaction,” in *Proceedings of the 28th IEEE International Conference on Robot and Human Interactive Communication*, 1–8.
- Holzwarth, M., Janiszewski, C., and Neumann, M. M. (2006). The influence of avatars on online consumer shopping behavior. *J. Mark.* 70, 19–36. doi: 10.1509/jmkg.70.4.019
- Hopkins, C. D., Raymond, M. A., and Mitra, A. (2004). Consumer responses to perceived telepresence in the online advertising environment: the moderating role of involvement. *Mark. Theory* 4, 137–162. doi: 10.1177/1470593104044090
- Hudders, L., De Jans, S., and De Veirman, M. (2020). The commercialization of social media stars: a literature review and conceptual framework on the strategic use of social media influencers. *Int. J. Advert.* 40, 327–375. doi: 10.1080/02650487.2020.1836925
- Jamy, L. (2015). The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *Int. J. Hum. Comput.* 77, 23–37. doi: 10.1016/j.ijhcs.2015.01.001
- Jin, S. A., and Bolebruch, J. (2009). Avatar-based advertising in second life. *J. Interact. Advert.* 10, 51–60. doi: 10.1080/15252019.2009.10722162
- Jin, S. V., Muqaddam, A., and Ryu, E. (2019). Instafamous and social media influencer marketing. *Mark. Intell. Plan.* 37, 567–579. doi: 10.1108/MIP-09-2018-0375
- Keeling, K., McGoldrick, P., and Beatty, S. (2010). Avatars as salespeople: communication style, trust, and intentions. *J. Bus. Res.* 63, 793–800. doi: 10.1016/j.jbusres.2008.12.015
- Khan, R. F., and Sutcliffe, A. (2014). Attractive agents are more persuasive. *Int. J. Hum. Comput.* 30, 142–150. doi: 10.1080/10447318.2013.839904
- Krugman, H. E. (1965). The impact of television advertising: learning without involvement. *Public Opin. Q.* 29, 349–356.
- Lee, E. B., Lee, S. G., and Yang, C. G. (2017). The influences of advertisement attitude and brand attitude on purchase intention of smartphone advertising. *Ind. Manag. Data Syst.* 117, 1011–1036. doi: 10.1108/IMDS-06-2016-0229
- Liudmila, B. (2020). “Designing identity in VTuber era,” in *ConVRgence (VRIC) Virtual Reality International Conference Proceedings*.
- McGoldrick, P. J., Keeling, K. A., and Beatty, S. F. (2008). A typology of roles for avatars in online retailing. *J. Mark. Manag.* 24, 433–461. doi: 10.1362/026725708X306176
- Moon, J. H., Kim, E., Choi, S. M., and Sung, Y. (2013). Keep the social in social media: the role of social interaction in avatar-based virtual shopping. *J. Interact. Advert.* 13, 14–26. doi: 10.1080/15252019.2013.768051
- Mullennix, J. W., Stern, S. E., Wilson, S. J., and Dyson, C. L. (2003). Social perception of male and female computer synthesized speech. *Comput. Hum. Behav.* 19, 407–424. doi: 10.1016/S0747-5632(02)00081-X
- Ogawa, K., Bartneck, C., Sakamoto, D., Kanda, T., Ono, T., and Ishiguro, H. (2009). “Can an android persuade you?” in *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication (Toyama)*, 553–557.
- Park, C. W., and Young, S. M. (1986). Consumer response to television commercials: the impact of involvement and background music on brand attitude formation. *J. Mark. Res.* 23, 11–24.
- Playboard (2022). *Most Super Chatted Channels in Worldwide*. Available online at: <https://playboard.co/en/youtube-ranking/most-superchatted-all-channels-in-worldwide-total> (accessed July 12, 2022).
- Tinwell, A., Grimshaw, M., Nabi, D. A., and Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Comput. Hum. Behav.* 27, 741–749. doi: 10.1016/j.chb.2010.10.018
- Verberne, F. F., Ham, J., Ponnada, A., and Midden, C. H. (2013). “Trusting digital chameleons: the effect of mimicry by a virtual social agent on user trust,” in *Proceedings of the International Conference on Persuasive Technology*, 234–245.
- Vrontis, D., Makrides, A., Christofi, M., and Thrassou, A. (2021). Social media influencer marketing: a systematic review, integrative framework and future research agenda. *Int. J. Consum. Stud.* 45, 617–644. doi: 10.1111/ijcs.12647
- Xiao, M., Wang, R., and Chan-Olmsted, S. (2018). Factors affecting YouTube influencer marketing credibility: a heuristic-systematic model. *J. Media Bus. Stud.* 15, 188–213. doi: 10.1080/16522354.2018.1501146
- Zanbaka, C., Goolkasian, P., and Hodges, L. F. (2006). “Can a virtual cat persuade you? The role of gender and realism in speaker persuasiveness,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1153–1162.
- Zanbaka, C. A., Ulinski, A. C., Goolkasian, P., and Hodges, L. F. (2007). “Social responses to virtual humans: implications for future interface design,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1561–1570.