

# Your Avatar Seems Hesitant to Share About Yourself: How People Perceive Others' Avatars in the Transparent System

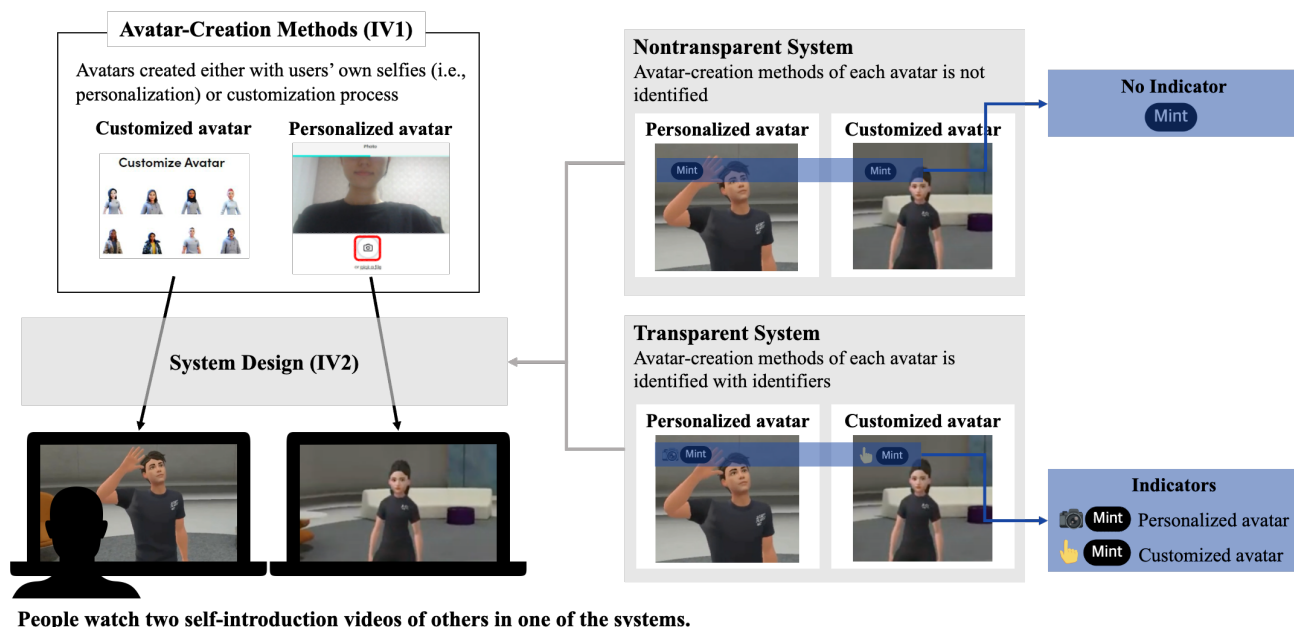
Yeonju Jang\*  
yj376@cornell.edu  
Department of Information Science,  
Cornell University  
Ithaca, New York, USA

Taenyun Kim\*  
kimtaeny@msu.edu  
Department of Media and  
Information, Michigan State  
University  
East Lansing, Michigan, USA

Huisung Kwon†  
huisungk@kaist.ac.kr  
Department of Industrial Design,  
KAIST  
Daejeon, Republic of Korea

Hyemin Park†  
hyemin525@g.skku.edu  
Interdisciplinary Course for Service  
Design Convergence, Sungkyunkwan  
University  
Suwon, Gyeonggi, Republic of Korea

Ki Joon Kim  
kj.kim@cityu.edu.hk  
Department of Media and  
Communication, City University of  
Hong Kong  
Hong Kong



**Figure 1: Study Overview.** IV1 is an avatar-creation method. Avatars were created with either users' own selfies (i.e., personalization) or a customization process. IV2 is the system design. In the nontransparent system, the avatar creation methods of each avatar are not identified. In the transparent system, the avatar-creation method of each avatar is identified with identifiers. The camera symbol indicator signifies personalized avatars, and the finger symbol indicator signifies customized avatars.

\*Both authors contributed equally to this research.

† Author names are in alphabetical order.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

CHI '24, May 11–16, 2024, Honolulu, HI, USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0330-0/24/05

<https://doi.org/10.1145/3613904.3642568>

## ABSTRACT

In avatar-mediated communications, users often cannot identify how others' avatars are created, which is one of the important information they need to evaluate others. Thus, we tested a social virtual world that is transparent about others' avatar-creation methods and investigated how knowing about others' avatar-creation methods shapes users' perceptions of others and their self-disclosure. We conducted a 2x2 mixed-design experiment with system design (nontransparent vs. transparent system) as a between-subjects and avatar-creation method (customized vs. personalized avatar) as a

within-subjects variable with 60 participants. The results revealed that personalized avatars in the transparent system were viewed less positively than customized avatars in the transparent system or avatars in the nontransparent system. These avatars appeared less comfortable and honest in their self-disclosure and less competent. Interestingly, avatars in the nontransparent system attracted more followers. Our results suggest being cautious when creating a social virtual world that discloses the avatar-creation process.

## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; *Laboratory experiments.*

## KEYWORDS

avatar-mediated communication, receiver, sender, transparent system, self-disclosure

### ACM Reference Format:

Yeonju Jang, Taenyun Kim, Huisung Kwon, Hyemin Park, and Ki Joon Kim. 2024. Your Avatar Seems Hesitant to Share About Yourself: How People Perceive Others' Avatars in the Transparent System. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3613904.3642568>

## 1 INTRODUCTION

In avatar-mediated communication (AMC), avatars are the first interface that users encounter when interacting with others [30]. Users create avatars to express their identities by making their avatars look like them or explore them by creating avatars completely different from them [21, 35, 36]. On the other hand, when users see someone else's avatar, they are one of the early cues that help the users form others' first impressions [17, 29].

Understanding this, many companies and researchers are exploring various ways to create avatars for users to express themselves better or to have more agency in creating procedures. Recently to enhance users' representation, their focus has been on advancing developing technologies for creating personalized and photorealistic avatars, driven by the belief that such avatars will facilitate interactions and engagements in AMC. For instance, Apple's recent unveiling of *Vision Pro*<sup>1</sup> offers the option to create avatars that resemble the users' own faces through scanning [2]. Other commercial products like *Spatial.io*<sup>2</sup> and *Zepeto*<sup>3</sup> also use graphic engines to predict avatar vectors and render images from pre-defined 3D assets [44, 53]. In addition, various companies also provide avatar-creation methods that users can customize their appearance based on what they want, as users may also wish to create avatars that are different from themselves [36].

To fully understand how people use avatars in a virtual social world, it is important to study how avatars affect the way others see users and how they impact users' own image. While it is comparably well-studied how users' avatars influence their attitudes and behaviors (e.g., Proteus Effect; See [42] for the review and meta-analysis), it is rather understudied how others see the one behind

their avatars (cf. [14, 28]), which is a significant research gap in our understanding.

In AMC, compared to face-to-face (FtF) interaction, the user's true identity is often hidden behind the avatars [14]. Thus, users often rely on even small cues to judge others because there is not much contextual information in the virtual environments [51]. Hence, whether one personalizes the avatars using their own face or customizes them by choosing various appearance options would also be an important source for users to evaluate others.

However, users do not have the capability to identify which method others have used to create their avatars if the system does not allow them. This may not be beneficial in terms of system transparency because it limits the information that users need to have to evaluate others. We also found a lack of research on how users' perceptions of others change when they can identify how others' avatars are created, particularly in the context of multiple avatar-creation methods available in the system.

Thus, our study investigated the impact of avatar-creation methods and the transparent system that can identify others' avatar-creation methods on users' perceptions of others and their self-disclosure. We found that users with personalized avatars in a transparent system received less favorable evaluations from others compared to those with customized avatars or in the non-transparent system. We also found that users are more likely to engage in future interactions when the avatar creation method is not known.

These findings were somewhat unexpected, especially for those who believe that a transparent system and an avatar that resembles the user would facilitate interaction in AMC. We suggest that users have focused on the minimal cue of avatar personalization and assumed that the senders would be reluctant to self-disclose because they themselves would also be hesitant to talk about themselves when the anonymity is low in an online environment in light of the hyperpersonal model [51] and the theory of mind [4, 7, 40]. Furthermore, our study provides practical implications for developers to find an optimal balance in transparency about avatar creation, as it can influence user perceptions and interactions within virtual platforms.

## 2 RELATED WORKS

### 2.1 Self-Disclosure in AMC

In AMC, just like FtF interactions, users talk about themselves when they meet for the first time [17, 29]. Indeed, self-disclosure is one of the important factors that lead to social attractiveness [49] and future interaction [9], which is no exception in AMC. However, self-disclosures in AMC are also different from those in FtF interactions, as users are less comfortable talking about topics that harm their anonymity, such as their actual name and more private details about themselves [48]. Extant research has demonstrated that avatar design influences their self-disclosure via anonymity [16, 17, 20], especially in terms of similarity between users and avatars. For instance, Ichino and colleagues [17] showed that users disclose themselves more when they use avatars that look different from themselves compared to the ones that look similar to themselves. Additionally, Kang and Yang [20] showed that if avatars look similar to an individual's physical self, they tend to disclose themselves

<sup>1</sup><https://www.apple.com/apple-vision-pro/>

<sup>2</sup><http://spatial.io/>

<sup>3</sup><http://web.zepeto.me>

less, and when they represent their ideal self, they are willing to disclose more about themselves.

The online disinhibition effect may explain the underlying mechanism of why avatar creation methods might impact self-disclosure via anonymity. According to the online disinhibition effect [47], online users can choose to modify or conceal their identities, which enables them to be anonymous more easily. This anonymity allows users to express themselves more freely than they might in FtF interactions. In AMC, avatars serve as identities that manage the level of users' anonymity. If users opt for avatars mirroring their physical features, their identities become less anonymous. Conversely, selecting avatars that differ significantly enhances anonymity. This concept of visual anonymity, defined as the extent to which an individual's physical appearance remains unobservable [45], acts as a mediating factor in how the avatar-creation method influences users' self-disclosure. Hooi and his colleagues [16] support this notion, revealing that the personalization of avatars influences users' self-disclosure, mediated by their anonymity. However, the benefits of users' anonymity, being open to others, come with a tradeoff where the users may be perceived as less trustworthy by the recipient of their self-disclosure [22]. This dynamic underscores the delicate balance between the freedom of expression that anonymity affords and the potential skepticism it may generate regarding the authenticity or reliability of the shared content.

There are several seminal papers that point out the mutual influence of receiver and sender in virtual environment [41, 51]. Especially, the hyperpersonal model posits that users in virtual environments often rely on slight cues to evaluate others, as there is not much contextual information compared to FtF interaction, and sometimes make premature judgments [51]. For instance, in the experiment conducted in the study of Machneva and colleagues [28], individuals formed first impressions of others' trustworthiness solely based on their avatars. Interestingly, there was agreement upon which avatar users seemed trustworthy and which did not, although they were inaccurate about who is truly trustworthy. Also, an interview conducted in the study by Freeman and Maloney [13] showed that users evaluate others based on their avatars by bringing stereotypes and prejudice they already have in the real world.

However, compared to the abundance of literature on how avatar designs affect users (e.g., Proteus Effect [42]) and their self-disclosure (e.g., [16, 17, 20, 48]), less attention has been given to avatar design influencing others' perceptions of users and their self-disclosure. To our knowledge, no studies investigate how receivers perceive senders and their self-disclosure in AMC. However, it is highly likely that the users' (i.e., senders) choice of avatars influences how others (i.e., receivers) think about them and their self-disclosure. To fill this research gap, this study will investigate how people perceive others and others' self-disclosure based on the information about their avatar from the receivers' perspective.

## 2.2 Avatar as System Design

Avatars can be created through different methods. One of the most common ways to create avatars is through customization. In this method, users get to choose their avatar's attribute in a set of options allowed to them. Research consistently indicates that offering users the freedom to tailor their avatar's appearance enhances

their sense of presence, embodiment, and enjoyment compared to generic avatars [10, 12]. Consequently, customization options are now broadly acknowledged as the better standard in avatar design [23].

Recently, there has been growing attention and expectations on personalization options for avatars in both industry and academia [15, 43, 50]. Recent advancements in computer graphics have enabled the creation of photorealistic and personalized avatars to be more convenient and accessible [43]. Users can create their avatars by scanning themselves or using ordinary camera photos across platforms, from the social virtual world to virtual reality and augmented reality. Such personalized avatars are found to enhance users' embodiment and self-identification [15].

It is important to note that personalization is different from photorealism. Despite the frequent overlap in terminology [15], personalization refers to embedding individual traits in avatars, while photorealism pertains to the quality and fidelity of avatar representation [15]. Avatars can be photorealistic without reflecting individual traits, and conversely, they can embody individual traits in a less realistic manner. This dichotomy can intensify the perceived gap between senders who are aware of how their avatars are created and receivers who form perceptions of the avatars.

From the receiver's perspective, evaluating the level of anonymity and self-representation embodied by an avatar with different avatar-creation methods can be challenging, as these characteristics are not always overtly observable [45]. While some personalized avatar-creation methods produce highly realistic avatars that are easily distinguishable from avatars created through customization using pre-existing design elements, some methods generate avatars that have a similar level of abstraction to those created through customization, even after it is personalized. In this case, users may struggle to discern whether they are interacting with a personalized or customized avatar. This can restrict the information users require to assess others.

System design of platforms can help reduce such uncertainty experienced by receivers regarding senders' avatar-creation method and moderate anonymity and self-disclosure [19, 48]. Although it is not widely applied to AMC, platforms can be designed to reveal or conceal the sender's identity, which in turn leads to a different level of self-disclosure. For example, online community sites that only allow real-name usage reduce anonymity levels, affecting self-disclosure behaviors. On the other hand, platforms that provide anonymity through user IDs enable individuals to feel less burdened and engage in greater self-disclosure, detached from their real-world identity [27].

Similarly, system design indicating the avatar-creation methods of each avatar would influence how receivers perceive senders and their self-disclosure within virtual environments. Extensive research has delved into avatar design from the standpoint of individual choices (e.g., why individuals choose a particular avatar [13]) and technical capabilities (e.g., methods to create realistic avatars [44, 53]); however, it remains relatively unexplored to consider the system design of avatars with its impact on users' perception. This study extends the scope of avatar design from individual choices across customization and personalization to encompass system-level design considerations. In light of these considerations, we tested a transparent system that allows users to

identify two distinct avatar-creation methods—personalized or customized avatars and explored receivers’ perception toward senders and their self-disclosure when the system design is transparent about the avatar-creation method of the senders with the following research questions.

- **RQ1:** How does the transparent system design of the avatar-creation method influence how receivers perceive senders and their self-disclosure?
- **RQ2:** How does the avatar-creation method influence how receivers perceive senders and their self-disclosure?
- **RQ3:** How does the avatar-creation method influence receivers’ perceptions of senders and their self-disclosure, contingent upon transparent system design?

### 3 METHOD

We conducted a 2 x 2 mixed-design experiment with an avatar-creation method (customized vs. personalized avatar) as a within-subjects and system design (nontransparent vs. transparent system) as a between-subjects variable to explore how avatar-creation method and its system designs influence receivers’ perceptions and interpretations of others’ self-disclosure (Figure 1). All self-introduction videos were generated by senders under the same experimental conditions and were not perceived differently across conditions in the preliminary study (See Section 3.4 for details). Both the preliminary study and main experiment were approved by the University IRB in South Korea, and participants were compensated for their participation with a gift card equivalent to 5,500 won (4.15 US dollars).

We opted for a social virtual world where users have the opportunity to engage with strangers in public spaces. The chosen setting allows users to unexpectedly encounter strangers in public areas. The experiment focused on participants watching self-introduction videos of strangers to observe how they would perceive these unfamiliar individuals. The selection of scenarios involving user-stranger interactions was deliberate for two main reasons: 1) the use of avatar plays a pivotal role when users would like to maintain anonymity in online interaction [48], and 2) the initial information exchange with strangers is crucial in forming a first impression, with the avatar serving as one of the primary pieces of information users receive about the strangers.

#### 3.1 Participants

A total of 76 participants completed the study through an online recruitment platform in South Korea. All participants received informed consent before beginning the study. Two attention-check questions were included, and sixteen participants who did not successfully answer these questions were excluded. As a result, 60 participants were included in the final analysis with 30 each in non-transparent and transparent system conditions. No a priori power analysis was conducted due to the exploratory nature of this study. However, post-hoc power analysis with observed small to medium effect size ( $\eta_G^2 = .052$ , Cohen’s  $F = 0.23$ ) indicates that a sample size of 58 would be sufficient with 80 % power. The demographic data of 13 participants was not recorded due to technical issues. The participants’ average age was 32.61 years old (*Standard Deviation*

(*SD*) = 15.95). Among them, 53.2 % ( $n = 25$ ) were female, 42.6% ( $n = 20$ ) were male, 4.3% ( $n = 2$ ) preferred not to say.

#### 3.2 Procedure

We conducted a mixed-design experiment (2x2) to investigate the influence of avatar-creation methods and system designs on receivers’ perceptions and interpretations of others’ self-disclosure (Figure 1). Each participant was randomly assigned to either the nontransparent or transparent system design and viewed two self-introduction videos of the senders whose avatars were either manually customized by them (i.e., customized avatar-creation method) or created automatically by their selfie (i.e., personalized avatar-creation method). The order of these two videos, representing the two avatar-creation methods, was randomized.

Prior to viewing the self-introduction videos, all participants watched two explanation videos tailored to their respective experimental conditions (Figure 2). The first video introduced the concept of the social virtual world, illustrating how avatars can interact within virtual spaces, engage in activities like chatting or following others, and watch self-introduction videos on hypothetical profiles designed for the experiment. Only the participants in the transparent system condition watched the avatars with a small badge positioned in front of their usernames, indicating an avatar-creation method used for each avatar in the first video. The badge was not shown to the participants in the nontransparent condition during the video. The second video explained how users could create their avatars by either customizing their avatars through various options (customized avatars) or uploading their selfies (personalized avatars).

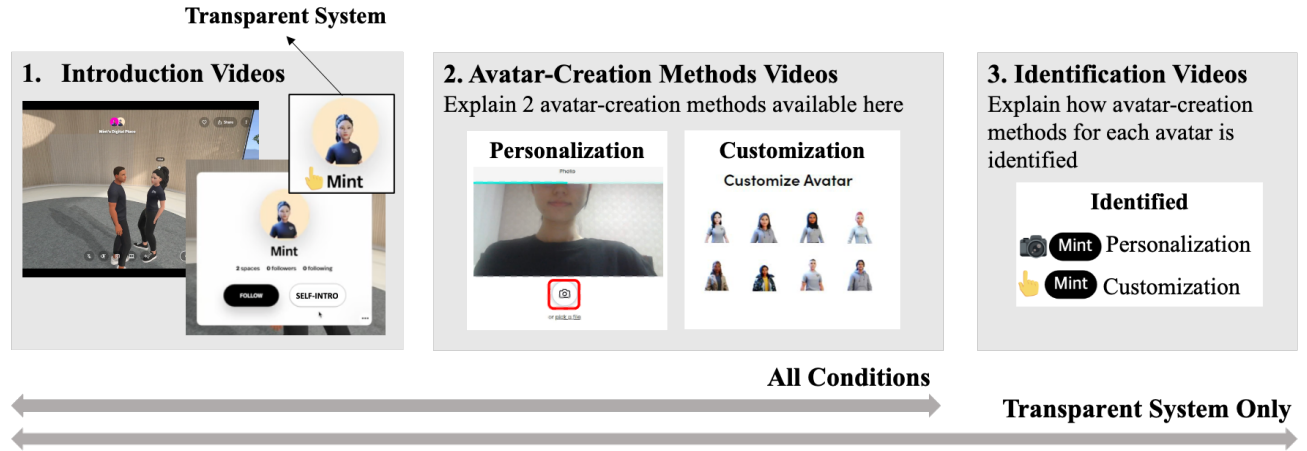
Participants assigned to the transparent system condition watched an additional video about the system design. They learned how to identify the senders’ avatar-creation method through a small badge left to the sender’s name. A pointing finger-shaped badge indicates that the avatar was created through customization, whereas a camera-shaped badge signifies that the avatar was created through the personalization process.

Following the viewing of the senders’ self-introduction videos, participants completed questionnaires about the self-introduction videos and the senders. Both explanation videos and self-introduction videos used in this study are uploaded in *OSF*<sup>4</sup>.

#### 3.3 Measures

**3.3.1 Manipulation Check for System Design.** We did not ask participants about their awareness of the system’s transparency regarding avatar-creation methods. This choice was deliberate since asking about transparency to participants in the nontransparent system condition can prompt them to speculate about the avatar-creation method. We wanted to grant them the freedom to either actively guess the avatar-creation method or remain indifferent to it. To make sure participants in the transparent system condition were aware of the system’s transparency, they watched an additional video explaining the badges that identify avatar-creation methods. In addition, all profiles were displayed with the badges throughout the explanation videos and introduction videos of others to reinforce their awareness about the transparency.

<sup>4</sup>[https://osf.io/ft2ga/?view\\_only=123cc665d845423daeb9c02470cfb381](https://osf.io/ft2ga/?view_only=123cc665d845423daeb9c02470cfb381)



**Figure 2: Explanation Videos for Conditions.** In the nontransparent system, participants watched 1. an introduction video and 2. an avatar-creation method video. In the transparent system, participants watched 1. an introduction video, 2. an avatar-creation method video, and 3. an identification video. In 1. an introduction video, participants were introduced to the social virtual world. In 2. an avatar-creation method video, participants were explained about two avatar-creation methods available here. In 3. an identification video, participants were explained about how avatar-creation methods for each avatar are identified.

**3.3.2 Attention Check.** After watching the videos, participants were asked about their perception of the sender and the sender’s self-introduction. During this questionnaire, we used two attention-check questions recommended by Abbey & Meloy [1] (“World War II happened before World War I” and “At some point in my life, I have had to consume water in some form.”) to make sure that participants were not distracted while they were answering the questionnaire.

**3.3.3 Anonymity.** Since previous works revealed the potential role of anonymity in the effect of the avatar-creation method on self-disclosure [16, 17, 20, 22], we measured anonymity (Cronbach’s alpha [ $\alpha$ ] = .94, Mean [ $M$ ] = 4.62,  $SD$  = 1.45) of the sender by using a scale developed by Wang [52]. We originally measured four items, but after dropping the items with internal consistency issues, only two were retained (i.e., “I didn’t know the counterpart in this video.” and “I can identify the counterpart in the video.”). Participants answered on a 7-point Likert scale ranging from 1 to 7 (1 = Strongly Disagree; 7 = Strongly Agree).

**3.3.4 Credibility.** We assessed credibility, considering that the anonymity of the sender may diminish the credibility of them perceived by the receiver [22]. Credibility has three dimensions: trustworthiness, competence, and goodwill [32]. We measured the credibility of the sender using a scale by McCroskey and Teven [32]. Participants were asked to answer how they felt about the sender on a semantic differential scale. *Trustworthiness* ( $\alpha$  = .90,  $M$  = 4.85,  $SD$  = 0.85) was measured on six items (e.g., dishonest - honest). *Competence* ( $\alpha$  = .90,  $M$  = 4.41,  $SD$  = 1.08) was measured on five items (e.g., unintelligent - intelligent). *Goodwill* ( $\alpha$  = .74,  $M$  = 4.24,  $SD$  = 0.77) was measured on five items (e.g., doesn’t care about me - care about me).

**3.3.5 Social Attraction.** Social attraction is identified as one of the primary outcomes resulting from self-disclosure [49]. Social

attraction ( $\alpha$  = .94,  $M$  = 4.50,  $SD$  = 1.33) of the sender was measured using a modified version of McCroskey and Richmond’s [33] scale. Participants were asked to rate their level of social attraction towards the sender on three specific items as follows: “I would like to have a friendly chat with them.”, “I think they could be friends of mine.”, “They are easy to get along with.”. Each of them was measured on a 7-point Likert scale (1 = Strongly Disagree; 7 = Strongly Agree).

**3.3.6 Self Disclosure.** Self-disclosure perceived by receivers was measured by the extent to which disclosed personal information was felt by the receivers (perceived breadth) and their perception of comfort and honesty in the senders’ self-disclosure.

*Perceived breadth of self-disclosure* ( $M$  = 2.95,  $SD$  = 1.81) of the sender was measured by asking participants to indicate whether the senders mentioned each of 10 topics of self-disclosure on a binary scale (mentioned - not mentioned). The topic was selected based on McCornack and Steven [31] three-layer framework. The topics include ‘Age or height information,’ ‘Hometown,’ ‘Major or occupation,’ ‘Personal habits,’ ‘Things I have done which I feel guilty about,’ ‘My deepest feelings,’ ‘What I like and dislike about myself,’ ‘What is important to me in life,’ ‘My worst fears’ and ‘Close relationships with other people.’

Upon the topics that participants reported that senders mentioned, they were asked to answer whether they thought that the senders were comfortable and honest in each topic of their self-disclosure.

*Perceived comfort of self-disclosure* ( $M$  = 4.82,  $SD$  = 1.36) was rated on a 7-point scale of whether the senders seemed comfortable talking about each topic (1 = Not at all comfortable; 7 = Completely comfortable). The score participants reported on each topic was averaged and made up a perceived comfortableness measure. If participants reported that senders did not mention any of the topics,

it was regarded that participants thought senders were "not at all comfortable" in their self-disclosure on every topic.

*Perceived honesty of self-disclosure* ( $M = 5.41$ ,  $SD = 1.26$ ) was rated on a 7-point Likert scale of whether the senders seemed honest about each topic (1 = Not at all honest; 7 = Completely honest). The score participants reported on each topic was averaged and made up a perceived honesty measure. If participants reported that senders did not mention any of the topics, it was regarded that participants thought senders were "not at all honest" in their self-disclosure on every topic.

**3.3.7 Willingness to Follow.** Finally, as an indicator of potential outcomes from self-disclosure [9] along with social attraction, willingness to follow the sender ( $M = 0.61$ ,  $SD = 0.49$ ) was measured by asking participants whether they would follow the sender after watching the self-introduction video on a binary scale (i.e., I would/would not follow the counterpart).

## 3.4 Stimuli

In order to enhance the external validity of stimuli that receivers would watch, we ran a preliminary study to obtain authentic self-introduction videos of users in a social virtual world platform (Section 3.4.1). A 2 x 2 between-design experiment with avatar-creation method (customized vs. personalized avatar) and system design (nontransparent vs. transparent system) with 55 participants as senders who introduced themselves with their avatars.

Then, to check whether video contents collected in four different conditions are perceived differently by receivers without any information about the system design, we additionally checked how receivers would perceive the self-introduction videos in four different conditions (Section 3.4.2). A total of 65 participants were recruited as receivers with a 2x2 mixed-design online experiment with an avatar-creation method (customized vs. personalized avatar) as a within-subjects variable and a system design (nontransparent vs. transparent system) as the between-subjects variable, same design as our main study.

### 3.4.1 Collecting Self-Introduction Videos.

**Participants.** A total of 55 participants took part in the study, recruited through an online platform and from one of the South Korean universities. Out of these, 33 participants submitted their self-introduction videos. Among these videos, six used personalized avatars in a transparent system, six with customized avatars in a transparent system, ten with personalized avatars in a nontransparent system, and 11 with customized avatars in a transparent system. We excluded videos that were shorter in length than Q1(29.25s), those that were not properly recorded such as showing the participants' faces or low-quality screens, or those that did not follow instructions on limitations of clothing options. Additionally, videos were excluded if participants did not pass the manipulation check or mentioned the study explicitly during their self-introduction (e.g., "I am a subject participating in the metaverse experiment"). We finally selected a total of eight videos, two for each condition. We included summary of detailed information for each video in Table 1.

**Procedure.** Participants were randomly assigned to one of four conditions and followed instructions step by step. In transparent system, participants were told that their avatar-created method would be identified by the system while in the nontransparent condition, they were told that their avatar-created method would not be identified by the system. Then, we asked participants to create their avatars in a designated method (customized vs. personalized avatar) and then introduce themselves with their avatars. Participants assigned to the personalized avatar-creation method created their avatars using their own selfies. In contrast, participants assigned to the customized avatar-creation method had the freedom to customize various aspects of their avatars as they desired. They could customize their avatar's facial features such as skin color, hair, eyes, and eyebrow shapes, regardless of their actual appearance. However, to maintain consistency between the two conditions, they were not allowed to directly customize clothes.

Then, they were encouraged to spend time in *spatial.io* with their avatars to become familiar with the virtual social world. We used *spatial.io* metaverse via web, which allows people to create avatars and virtual spaces and communicate with others (or with more than 700K monthly visitors). We selected this platform because it allows users to encounter one another in various virtual environments. Additionally, it utilizes *ReadyPlayerMe*<sup>5</sup> for avatar creation, which supports two distinct methods of avatar creation while maintaining a consistent level of avatar abstraction.

In the subsequent step, the participants recorded a self-introduction video through avatar and their voices in the metaverse to show to the new people they would meet there. They could move around the room and use a few animations (e.g., sending hearts, waving hands, jumping, dancing). At the end of the step, they uploaded the video and were asked about manipulation check questions and their self-disclosure. We asked whether they disclosed each of the ten personal information listed to measure breadth of self-disclosure on a binary scale (See details in Appendix A.4 below). Upon the topics that they disclosed, they were asked whether they were comfortable and honest on a 7-point Likert scale.

### 3.4.2 Evaluating Content Differences of Self-Introduction Videos.

**Participants and Procedure.** Sixty-five participants were recruited through the same platform with a similar process to the main study. Following the same screening process with the attention check questions, 59 participants were included in the baseline analysis, with an average age of 31.07 years ( $SD = 14.70$ ). 27 (45.8 %) were female, 29 (49.2%) were male, and 2 (3.3%) did not provide demographic information. Participants watched two self-introduction videos of one personalized avatar and one customized avatar that were recorded in either a transparent or nontransparent system. They were not informed about the avatar-creation methods and identification. All the videos were provided without identifiers, even if they were originally recorded in a transparent system.

**Analysis.** We conducted a 2-way mixed ANOVA with avatar-creation as a within-subjects variable and identification system as between-subjects variable for all analyses except for the following

<sup>5</sup><https://readyplayer.me/>

Conditions		Video		Demographic		Self-reported self-disclosure			
System design	Avatar-creation method	ID	Length (m:ss)	Gender	Age	Breadth	Comfort	Honesty	Descriptions***
Transparent	Personalized	A3	0:47	M	40	2	6	7	a, b, c
		A5	0:38	F	37	1	5	6	a, b
	Customized	B2	0:36	M	29	5	5.8	7	a, b, c, d, e, f
		B4	0:38	F	37	1	5	6	c, e
Nontransparent	Personalized	C4	0:48	M	25	3	5.3	7	d, g, h
		C8	0:39	M	24	NC**	NC	NC	NC
	Customized	D4	2:39	M	28	4	4.75	5	a, b, c, e, f
		D6	0:42	M*	25	2	6	6	a, b, e

**Table 1: Summary of Selected Self-introduction stimuli.** \*Participant D6 swapped the gender of the avatar, and \*\*responses about the self-disclosure of participant C8 were not collected. \*\*\*Details of self-disclosure are as follows: a. age or height information, b. hometown, c. major or occupation, d. my personal traits, e. what I like and dislike about myself, f. what is important to me in life, and g. my close relationships.

intention, which used a generalized linear mixed model with avatar-creation as a within-subjects variable and identification system as between-subjects variables given the binary nature of the variable.

Anonymity ( $\alpha = .57$ ,  $M = 4.80$ ,  $SD = 1.28$ ), trustworthiness ( $\alpha = .91$ ,  $M = 4.50$ ,  $SD = 1.33$ ), competence ( $\alpha = .87$ ,  $M = 4.50$ ,  $SD = 1.33$ ), goodwill ( $\alpha = .71$ ,  $M = 4.50$ ,  $SD = 1.33$ ), social attractiveness ( $\alpha = .95$ ,  $M = 4.22$ ,  $SD = 1.39$ ), perceived breadth ( $M = 2.86$ ,  $SD = 1.56$ ), comfortableness ( $M = 4.89$ ,  $SD = 1.16$ ) and honesty ( $M = 5.26$ ,  $SD = 1.02$ ) of self-disclosure, follow ( $P = 0.53$ ,  $SD = 0.50$ ) was all not significant (all  $p > .09$ ).

## 4 RESULTS

To see if receivers' perceptions, attitudes, and behavior vary depending on the conditions, we conducted a 2-way mixed ANOVA with avatar-creation methods as a within-subjects variable and transparent system design as a between-subjects variable for all analyses except for the following intention. For the following intention, we used a generalized linear mixed model with avatar-creation as a within-subjects variable and identification system as between-subjects variables given the binary nature of the variable—either receiver chooses or chooses not to follow the sender (See Table 2 and Figure 3 for Results). We also applied Bonferroni correction to adjust for type 1 error when performing multiple t-tests as a post-hoc analysis after a significant main effect of ANOVA.

### 4.1 Anonymity

The results of anonymity showed no significant main effect of the transparent system design ( $F[1, 58] = 0.16$ ,  $p = .69$ ; *Generalized Partial Eta* [ $\eta_G^2$ ] = .002) nor avatar-creation method ( $F[1, 58] = 0.68$ ,  $p = .16$ ,  $\eta_G^2 < .001$ ). There was also no interaction effect among the factors ( $F[1, 58] = 0.99$ ,  $p = .32$ ,  $\eta_G^2 = .002$ ). These results answer our RQs that neither transparent system design (RQ1) nor avatar creation method (RQ2) influenced senders' anonymity perceived by receivers as well as interaction (RQ3).

Additionally, to ascertain any potential differences between the videos within each condition, a pairwise t-test with Bonferroni correction was conducted for each condition on anonymity to see if there was any difference between the videos. The results showed

no significant differences between the videos in each condition ( $p = 1.0$ ).

### 4.2 Credibility

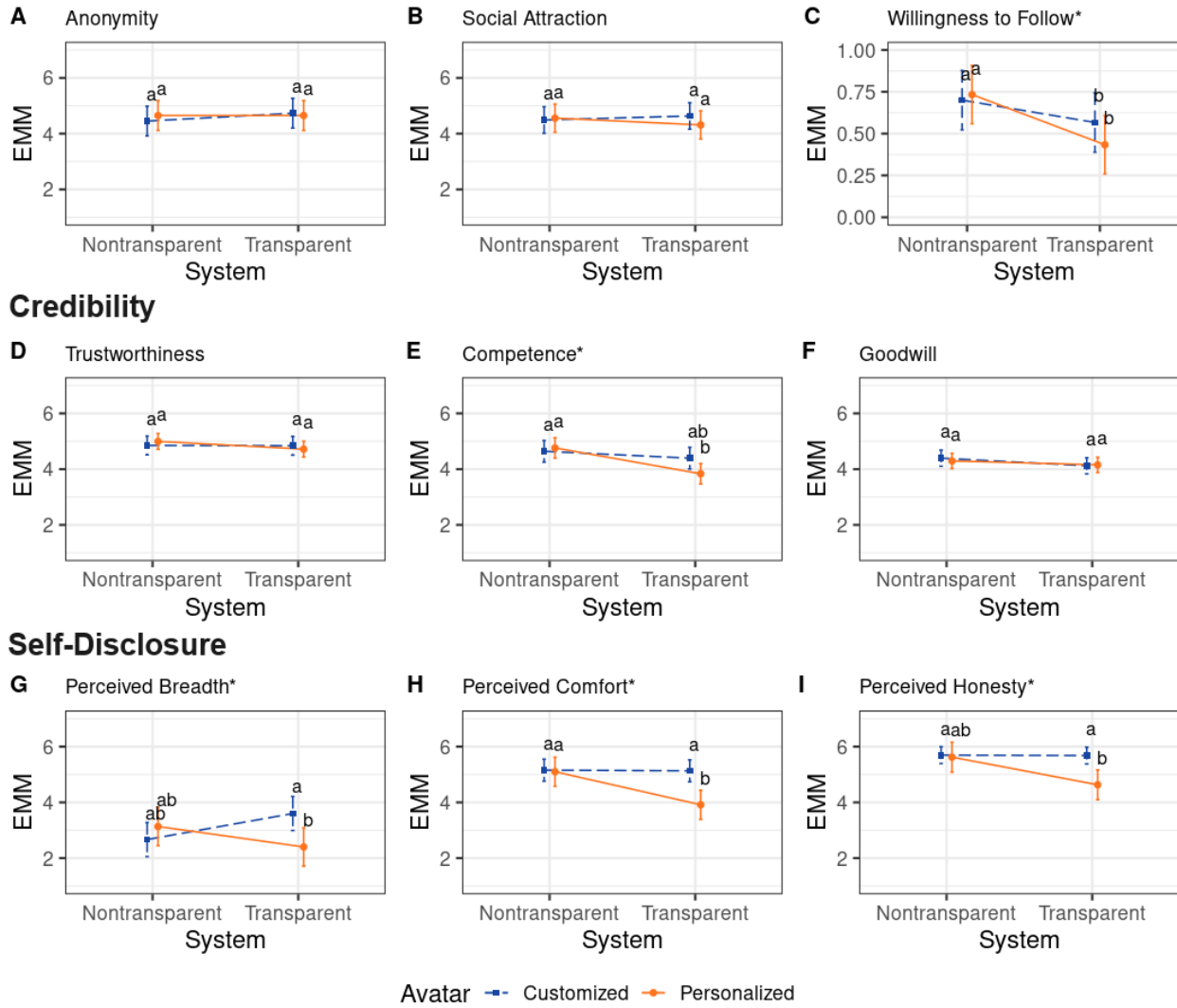
Regarding the receivers' credibility of the senders, none of the main or interaction effects were significant concerning trustworthiness and goodwill (all  $p > .24$ ). In terms of competence, there was a significant interaction effect between the transparent system design and avatar-creation method ( $F[1, 58] = 4.73$ ,  $p = .034$ ,  $\eta_G^2 < .027$ ).

Post-hoc tests with Bonferroni corrections revealed that receivers perceive the senders with the personalized avatars in the transparent system ( $M = 3.83$ , 95%  $CI[3.47, 4.20]$ ) as significantly less competent than those with the customized ( $M = 4.64$ , 95%  $CI[4.25, 5.03]$ ,  $p = .022$ ) or personalized avatars ( $M = 4.76$ , 95%  $CI[4.39, 5.13]$ ,  $p = .004$ ) in the nontransparent system. Other pairs were all non-significant (all  $p > .08$ ).

There was also a main effect of the transparent system design ( $F[1, 58] = 7.41$ ,  $p = .009$ ,  $\eta_G^2 < .077$ ), as receivers perceived the competence of the senders in nontransparent system ( $M = 4.70$ , 95%  $CI[4.39, 5.01]$ ) to be significantly higher than that of the senders in a transparent system ( $M = 4.11$ , 95%  $CI[3.81, 4.42]$ ). The main effect of the avatar-creation method was not significant ( $F[1, 58] = 1.98$ ,  $p = .17$ ,  $\eta_G^2 < .012$ ). These results answer our RQs that perceived competence, among dimensions of credibility, was the factor that was influenced by the transparent system design (RQ1), but not by avatar creation method (RQ2), while the avatar creation method influenced perceived competence when the system design was transparent (RQ3).

Additionally, to ascertain any potential differences between the videos within each condition, a pairwise t-test with Bonferroni correction was conducted for each condition on trustworthiness, competence, and goodwill to see if there was any difference between the videos. The results showed no significant differences between the videos in each condition on trustworthiness (all  $p > .12$ ) and competence (all  $p > .46$ ). However, in terms of goodwill, there was a significant difference between A5 ( $M = 4.56$ ,  $SD = 0.76$ ) and A3 ( $M = 3.69$ ,  $SD = 0.35$ ), the videos of the senders using personalized avatars in the transparent systems ( $t[22] = 4.16$ ,  $p = 0.000$ , *Cohen's d* = 1.46). Videos in other conditions were not significantly different





**Figure 3: Interaction plot results of Post-hoc test with Bonferroni corrections, indicating which groups have statistically significant differences. Shared letters denote no significant difference, while distinct letters indicate significant group differences. Asterisk(\*) on the constructs indicates that they were significant main or interaction effects. EMM = Estimated Marginal Means**

(all  $p > .08$ ). Thus, the result of goodwill should be interpreted with caution.

### 4.3 Social Attractiveness

Regarding social attraction, there was no significant main effect of the transparent system design ( $F[1, 58] = 0.03, p = .86, \eta_G^2 < .001$ ). Also, no main effect of the avatar-creation method was found ( $F[1, 58] = 0.39, p = .54, \eta_G^2 < .002$ ). The interaction effect among the factors was also not significant ( $F[1, 58] = 0.90, p = .35, \eta_G^2 = .005$ ). These results answer our RQs that neither transparent system design (RQ1) nor avatar creation method (RQ2) influenced receivers' social attractiveness toward senders as well as interaction (RQ3).

Additionally, to ascertain any potential differences between the videos within each condition, a pairwise t-test with Bonferroni correction was conducted for each condition on social attractiveness to see if there was any difference between the videos. The results showed no significant differences between the videos in each condition ( $p = .49$ ).

### 4.4 Self-disclosure

**4.4.1 Perceived Breadth of Self-disclosure.** In terms of the perceived breadth of the self-disclosure, an interaction between the transparent system design and avatar-creation method was found ( $F[1, 58] = 10.40, p = .002, \eta_G^2 = .054$ ). The main effect of the transparent system design ( $F[1, 58] = 0.07, p = .79, \eta_G^2 < .001$ ) was not significant.



Variables		Main Effect of Transparent System (RQ1)	Main Effect of Avatar-Creation Method (RQ2)	Interaction Effect (RQ3)
Perceived Anonymity		NS	NS	NS
Credibility	Trustworthiness	NS	NS	NS
	Competence	** (T < NT)	NS	* (T-P < T-C, NT-C)
	Goodwill <sup>†</sup>	NS	NS	NS
Social Attractiveness		NS	NS	NS
Self-Disclosure	Perceived Breadth <sup>†</sup>	NS	NS	** (T-P < T-C)
	Perceived Comfort	* (T < NT)	** (P < C)	* (T-P < T-C, NT-C, NT-P)
	Perceived Honesty	* (T < NT)	** (P < C)	* (T-P < T-C, NT-C)
Willingness to Follow		* (T < NT)	NS	NS

**Table 2: Summary of results.** NS denotes nonsignificant results, while Asterisk (\*) indicates significant results (\* at the  $\leq .05$  level and \*\* at the  $\leq .01$  level). T stands for Transparent System, NT for Nontransparent System, P for Personalized Avatars, C for Customized Avatars. Variables with <sup>†</sup> require cautious interpretation due to video differences.

.001) and avatar-creation method ( $F[1, 58] = 2.01, p = .16, \eta_G^2 = .011$ ) was not significant. The post-hoc analysis with Bonferroni corrections showed that the receivers perceived senders with personalized avatars in the transparent system ( $M = 2.40, 95\% CI[1.71, 3.09]$ ) disclosed significantly more about themselves than those with customized avatars in the transparent system ( $M = 3.60, 95\% CI[2.99, 4.21], p = .010$ ). Other pairs were not significantly different (all  $p > .15$ ). These results answer our RQs that the perceived breadth of self-disclosure by the avatar creation method only when the system design was transparent (RQ3), while there was no difference depending on the transparent system design (RQ1) or avatar creation method (RQ2).

Furthermore, to make sure to identify any potential differences between the videos within each condition, a pairwise t-test with Bonferroni correction was conducted for each condition on perceived breadth to see if there was any difference between the videos. The results showed a significant difference between B2 ( $M = 4.37, SD = 1.74$ ) and B4 ( $M = 2.27, SD = 1.49$ ), the videos of the senders using customized avatars in the transparent systems ( $t[24] = 3.49, p = 0.008, \text{Cohen's } d = 1.27$ ). Also, there was a significant difference between D4 ( $M = 3.31, SD = 1.14$ ) and D6 ( $M = 1.93, SD = 1.27$ ), the videos of senders using customized avatars in the nontransparent systems ( $t[26] = 3.13, p = 0.016, \text{Cohen's } d = 1.15$ ). Videos in other conditions were not significantly different (all  $p > .06$ ). Thus, the result of perceived breath should be interpreted with caution.

**4.4.2 Perceived Comfort of Self-disclosure.** In terms of the perceived comfort of the self-disclosure, there was a 2-way interaction of the transparent system design and avatar-creation method ( $F[1, 58] = 6.13, p = .016, \eta_G^2 = .051$ ) as well as the main effect of transparent system design ( $F[1, 58] = 7.02, p = .010, \eta_G^2 = .056$ ) and avatar-creation ( $F[1, 58] = 7.40, p = .009, \eta_G^2 = .061$ ). The result showed that senders using personalized avatars ( $M = 4.51, 95\% CI[4.14, 4.88]$ ) were perceived as less comfortable in their self-disclosure than those in customized avatars ( $M = 5.14, 95\% CI[4.86, 5.42]$ ). Also, senders in transparent systems were perceived ( $M = 4.52, 95\% CI[4.2, 4.85]$ ) less comfortable in their self-disclosure than those in nontransparent systems

( $M = 5.13, 95\% CI[4.8, 5.45]$ ). The post-hoc test revealed that receivers perceive that the senders with personalized avatars in the transparent system ( $M = 3.91, 95\% CI[3.39, 4.44]$ ) were significantly less comfortable in disclosing themselves than those with personalized avatars in the nontransparent system ( $M = 5.10, 95\% CI[4.58, 5.62], p = .013$ ), customized avatars in transparent ( $M = 5.13, 95\% CI[4.74, 5.52], p = .003$ ) or nontransparent system ( $M = 5.16, 95\% CI[4.76, 5.55], p = .002$ ). These results answer our RQs that the perceived comfort of self-disclosure was influenced by both the transparent system design (RQ1) and avatar creation method (RQ2). The effect of the avatar creation method was also contingent upon the system design (RQ3).

Additionally, to see if there are any potential differences between the videos within each condition, a pairwise t-test with Bonferroni correction was conducted for each condition on perceived comfort to see if there was any difference between the videos. The results showed no significant differences between the videos in each condition ( $p = .45$ ).

**4.4.3 Perceived Honesty of Self-disclosure.** In terms of the perceived honesty of the self-disclosure, a 2-way interaction of the transparent system design and avatar-creation method ( $F[1, 58] = 6.02, p = .017, \eta_G^2 = .042$ ) was found. The main effect of transparent system design ( $F[1, 58] = 4.64, p = .035, \eta_G^2 = .044$ ) and avatar-creation method ( $F[1, 58] = 7.98, p = .006, \eta_G^2 = .055$ ) was also significant. The result demonstrated that senders using personalized avatars ( $M = 5.13, 95\% CI[4.75, 5.51]$ ) were perceived as less honest in their self-disclosure than those in customized avatars ( $M = 5.69, 95\% CI[5.48, 5.90]$ ) and those in transparent systems ( $M = 5.16, 95\% CI[4.82, 5.49]$ ) were less honest in their self-disclosure than those in nontransparent systems ( $M = 5.66, 95\% CI[5.33, 5.99]$ ). The post-hoc test revealed that receivers perceive that the senders with personalized avatars in the transparent system ( $M = 4.63, 95\% CI[4.09, 5.17]$ ) conditions were significantly less honest in disclosing themselves than those with customized avatars in the transparent system ( $M = 5.68, 95\% CI[5.38, 5.98], p = .003$ ) or those with customized avatars in the nontransparent system ( $M = 5.70, 95\% CI[5.40, 6.00], p = .006$ ). These results answer our RQs that the perceived honesty of self-disclosure was influenced by both the

transparent system design (RQ1) and avatar creation method (RQ2). The effect of the avatar creation method was also contingent upon the system design (RQ3).

Furthermore, to identify any potential differences between the videos within each condition, a pairwise t-test with Bonferroni correction was conducted for each condition on perceived honesty to see if there was any difference between the videos. The results showed no significant differences between the videos in each condition ( $p = .20$ ).

## 4.5 Following Intention

The generalized linear mixed model was fitted with a binomial distribution using the logit link function. The model's goodness of fit is represented by an AIC of 163.6 and a BIC of 177.5. The result showed a significant main effect of the transparent system design ( $\beta = -0.46$ , 95% CI  $[-0.83, -0.09]$ ,  $p = .017$ ). The main effect of the avatar-creation method ( $p = .63$ ) or interaction effect ( $p = .37$ ) was not significant. The result showed that receivers were 2.54 times less likely to follow the senders in a transparent system ( $P = 0.50$ , 95% CI  $[0.36, 0.64]$ ) compared to those in the nontransparent system ( $P = 0.72$ , 95% CI  $[0.60, 0.84]$ ). These results answer our RQs that the transparent system design influences the following intention (RQ1), but not the avatar creation method (RQ2) or their interaction (RQ3).

Additionally, to make sure there were no potential differences between the videos within each condition, a pairwise t-test with Bonferroni correction was conducted for each condition with the following intention to see if there was any difference between the videos. The results showed no significant differences between the videos in each condition ( $p = .54$ ).

## 5 DISCUSSION

### 5.1 Knowing that Senders Used Personalized Avatar Colors the Receivers' Perception

Our study showed that the receivers' attitudes toward the sender and their self-disclosure were colored by their knowledge of how the sender's avatars were created. Only in the transparent system did receivers perceive that the senders using personalized avatars were less broad, less comfortable, and less honest about self-disclosure than those who identified using customized avatars.

While the online disinhibition effect [47] provides an explanation about why senders would self-disclose more when using customized rather than personalized avatars, it cannot explain why receivers perceived the senders are disclosing themselves less broadly, less comfortably, and less honestly. In order to understand these results, we need other psychological theories or communication models that focus on the receivers' perception and interpretation of online interactions.

According to the hyperpersonal model [51], individuals in virtual environments often focus on subtle cues, given the reduced context compared to FtF interactions. These minimal cues can greatly influence impression formation. In our study, avatars, particularly how they were created, acted as these crucial cues, and receivers might have focused on these avatar cues to make judgments about senders' self-disclosure. Indeed, Machneva and colleagues [28] suggest that receivers shape the first impression of the senders' solely based

on their avatars, although inaccurate. Moreover, while evaluating others by their avatars, they bring their pre-existing expectation or knowledge they already have in face-to-face communication [13].

In face-to-face communication, the most accessible pre-existing expectation or knowledge individuals can refer to is how they themselves think or behave. The theory of mind (ToM) posits that individuals can predict how others will behave by understanding and anticipating the thoughts, beliefs, and emotions of others, essentially placing themselves in another's position [4, 7, 40]. Drawing from ToM, receivers in our study might have applied their own feelings and experiences regarding personalized avatars. If they believed that a personalized avatar would make them less comfortable with self-disclosure as it resembles the sender too much, they might have projected that belief onto the sender.

Previous research supports this projection. Studies have indicated that individuals are more reluctant to disclose their personal information when they feel easily identifiable [6, 8, 16]. Notably, if the avatar closely mirrors one's real appearance, they are more hesitant to disclose themselves [16, 17, 19]. Thus, it stands to reason that receivers believed senders with personalized avatars would be less honest, as they themselves might feel the same discomfort in a similar situation.

However, it does not guarantee that the receivers' beliefs are correct, even though they use their own feelings to understand the senders' situation. When evaluating others based on their avatars, receivers often bring their own stereotypes and prejudices [13]. People also have egocentric biases in understanding others' situations and have difficulty acknowledging that someone else's perceptions can be different from, rather than similar to, their own. Although people initially adjust their thoughts about others to take into account their situation more, they often stop this adjustment once they reach a plausible estimation [11]. As a result, receivers often do not obtain enough accuracy in estimating others' characteristics based on their avatars. For instance, Machneva and colleagues [28] showed that while individuals' evaluation of trustworthiness based on others' avatars had consensus among users, their accuracy of predicting actual behavior was very poor.

### 5.2 Not Knowing Sender's Avatar Creation Method Leads to More Future Interaction

Our study also showed that receivers were less inclined to follow senders when the method behind the avatar creation was known compared to when it was undisclosed. This can potentially be attributed to the fact that we sometimes allure to the unknown in online space [34, 41]. While it is generally understood that additional information can boost attractiveness, this effect is modulated by the quality and desirability of the information [41]. Specifically, when certain details are perceived as unnecessary or even unwanted, their disclosure might negatively affect user behavior.

For instance, when an anonymous gaming YouTuber known as DREAM revealed his face, he encountered a wave of derogatory comments on social media, with users trending phrases such as "He's Ugly" and "#PutTheMaskBackOn." As a result, DREAM's YouTube views had significantly decreased, with his average sitting around five million views compared to his previous tens of millions. Eventually, DREAM had to wear his mask again to alleviate the

public outrage. It appears that the mask has become an essential element of his persona, and the public is unwilling to separate him from it [18]. The story of the YouTuber DREAM reveals how revealing certain previously concealed details, like his face, can negatively impact public perception and engagement, as his audience preferred his masked, mysterious persona. This is consistent with our findings. As a result of our study, it is possible that avatars with undisclosed creation methods may have gathered more followers, as they were allured to the unknown in digital environments.

### 5.3 Practical Implication

**5.3.1 Dilemma of Avatar Personalization.** Our study also provides some practical implications for applying personalized avatars. Our findings indicate that the fact that others use personalized avatars may make users perceive that others are less comfortable and honest about self-disclosure. Thus, if creating a personalized avatar is the sole method of avatar creation on a virtual platform, users' self-disclosure will not be perceived as comfortable and honest by others. It will potentially result in a negative impact on their interactions.

Thus, platform developers are encouraged to provide both personalized and customized avatar creation options to consider bidirectional social platforms. With a customized avatar option, individuals are more likely to have comfortable interactions with others. Although some studies report a positive impact of personalized avatars on users' experiences [43], our finding that receivers responded less favorably to personalized avatars should be taken into account, given the bidirectional nature of the interaction.

Furthermore, our study provides design guidelines for applying avatars in the Metaverse<sup>6</sup>. The Metaverse will transcend being merely an entertainment space to encompass a comprehensive realm that includes education [24], healthcare [3], and daily life, among others. Particularly in fields where building trust is crucial, our research suggests to users of the Metaverse that it may be more advantageous to opt for customized avatars over personalized ones.

**5.3.2 Dilemma of Transparency via System.** Our study also provides practical implications for how to design transparency in the platform, especially in the Metaverse. Our research indicates that transparency, while generally being valued to enhance trust and interaction quality [5, 39], can sometimes influence user perceptions negatively when applied to certain situations. In our study, receivers viewed the senders behind the avatar differently in the transparent system. They perceived senders with personalized avatars as less competent, less comfortable, and less honest about their self-disclosure. This means that senders can sometimes be victims of the prejudice accompanied by the avatar-creation method.

While being open and clear about the people who are in interactions can be beneficial to the receivers in a way that they are

fully told about the information they need to judge others online, it also highlights an important dilemma. If transparency operates in a way that penalizes the senders for no reason, its value comes into question. In a virtual platform, senders sometimes do not want to reveal themselves fully, and receivers, too, do not want to know the senders fully [34, 41]. This is because knowing too much can diminish the little mystery or surprise they have about the senders. Thus, for platform developers, it is not merely about providing complete transparency. Instead, finding the right balance with transparency is crucial.

A potential approach to designing transparency between the sender and receivers via a system could be giving users the autonomy to decide whether they want to disclose details about their avatar creation method. Instead of transparency of the individual's avatar-creation method being determined by a system, the system can provide options to users on whether and when they would like to reveal their avatar-creation methods. This way, users can decide for themselves what they are comfortable with. This aligns with prior findings that users' desire to be open varies across various types of relationships, ranging from superficial to deep connections [38, 39].

However, this approach should be implemented with caution. In this context, choosing not to share their avatar-creation method could also inadvertently work as cues, subjecting the senders to potential prejudice that they are hiding something because they do not reveal their avatar-creation methods.

Therefore, if the designers of Metaverse have decided to prioritize the transparency of the avatar-creation method, they should consider allowing users to choose whether or not to disclose how their avatar was created. Moreover, the designers should let the users know the potential influence of their choice of disclosure (or non-disclosure) on receivers' perception. In this way, designers will be able to find the optimal balance between transparency and user agency. This approach could grant users greater autonomy and positively influence individual expression and trust-building within the Metaverse. In summary, while transparency remains a commendable principle, its application needs to be nuanced, ensuring that it complements user preferences and enhances their online experience.

## 6 LIMITATIONS & FUTURE WORKS

Our study presents several limitations. Firstly, while we tried to select video stimuli that could capture the uniqueness of each condition, some differences between the videos within the conditions were observed, particularly in goodwill and the perceived breadth of self-disclosure in our nested design. This indicates that the video stimuli may not completely capture how senders might behave in each circumstance. Importantly, when the video stimuli were compared without any explanation about the manipulations, no differences in the receivers' perceptions of the sender were found. This suggests that the manipulations had a greater impact than the video content itself.

Secondly, we conducted our research on a hypothetical social virtual world platform, which may lead to concerns about ecological validity. The nature of interaction in our study was unidirectional; receivers only viewed pre-recorded self-introductions of the

<sup>6</sup>The term Metaverse, first coined by Neal Stephenson [46] and initially gained attention through platforms such as Second Life, surged again with immersive technologies (e.g., VR, AR, XR), notably illustrated by Facebook's rebranding to Meta. Despite the academic world still converging towards a unified definition of the Metaverse, as evidenced by the ongoing discussions referenced by Xu and colleagues [54], there's a general agreement that its essence lies in merging virtual technologies with the physical world [48]. Park and colleagues [37] have also contributed to this discourse by examining the Metaverse workspace functioning without immersive hardware. In this line of understanding of the Metaverse, our findings on avatar personalization and system transparency can provide significant design suggestions in the Metaverse while also acknowledging the need to consider variations in immersive experiences.

senders. Such a dynamic might not truly represent the interactive nature of real-world virtual platforms. Thus, future studies should explore our findings in more natural and bidirectional interactions to more accurately capture real-world scenarios. Also, our study only focused on strangers without prospecting future interaction. However, social virtual worlds vary in their interaction types, scale, and accessibility [26]. For example, acquaintances tend to interact in social virtual worlds such as Gather Town where support conference-type interactions. As results may vary in existing relationships where future interactions are expected, future studies need to explore how receivers' perceptions change when future interactions are expected.

Furthermore, our study was about a web-based social virtual world where users interact using avatars. It is worth noting that if users were to engage in a more immersive virtual reality (VR) experience utilizing advanced hardware like head-mounted displays (HMDs), they might experience enhanced physical, social, and self-presence—"a psychological state in which virtual (para-authentic or artificial) [social actors, physical objects, and selves] are experienced as if [they are real] in either sensory or nonsensory ways [25]." In such a setting, both the sender and receiver might perceive and behave differently during their interactions. Therefore, future studies should explore the applicability of our findings in immersive VR settings that employ HMDs and similar devices.

Finally, it is important to consider the potential confounding variables of our study. For instance, individuals more accustomed to AMC platforms and avatar customization (or personalization) might have different expectations and comfort levels, impacting how they perceive and engage with avatars in both transparent and nontransparent systems. Similarly, prior experience in creating avatars could affect users' evaluations of others' avatars as well as their own self-disclosure behaviors. Furthermore, individuals' demographic information, such as gender and age, can influence how they self-disclose in the digital environment. While these questions were not addressed in this study, the diverse experiences and backgrounds of social virtual world users are also important factors to consider in studying AMC. Therefore, our findings must be interpreted with caution and call for future studies to investigate the effects of familiarity and prior experience with AMC on the evaluations of self and others in AMC, as well as differences in gender and age.

## 7 CONCLUSION

In conclusion, this study offers valuable insights into AMC, indicating that even small factors like creating an avatar can have a big impact on how people perceive and interact with each other online. The study also shows that designers and developers should be careful when making a social online world that is transparent about the avatar-creation method. The findings provide a foundation for future studies on the complex dynamics of online communication, sharing personal information, and the psychological aspects of anonymity.

## ACKNOWLEDGMENTS

We thank Juyeob Lee, Seungjoon Son, Taebin Lim, and Yejin Lee from Sungkyunkwan University for helping organize and proctor

experiments. We also thank anonymous reviewers for providing valuable feedback to improve this paper.

## REFERENCES

- [1] James D Abbey and Margaret G Meloy. 2017. Attention by design: Using attention checks to detect inattentive respondents and improve data quality. *Journal of Operations Management* 53 (2017), 63–70.
- [2] Apple. 2023. Introducing Apple Vision Pro. <https://www.youtube.com/watch?v=TX9qSaGXFyg>
- [3] Gaurang Bansal, Karthik Rajgopal, Vinay Chamola, Zehui Xiong, and Dusit Niyato. 2022. Healthcare in metaverse: A survey on current metaverse applications in healthcare. *Ieee Access* 10 (2022), 119914–119946.
- [4] Simon Baron-Cohen, Alan M Leslie, and Uta Frith. 1985. Does the autistic child have a "theory of mind"? *Cognition* 21, 1 (1985), 37–46.
- [5] Md Momen Bhuiyan, Hayden Whitley, Michael Horning, Sang Won Lee, and Tanushree Mitra. 2021. Designing transparency cues in online news platforms to promote trust: Journalists' & consumers' perspectives. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–31.
- [6] Stephanie Booth-Kewley, Jack E Edwards, and Paul Rosenfeld. 1992. Impression management, social desirability, and computer administration of attitude questionnaires: Does the computer make a difference? *Journal of Applied Psychology* 77, 4 (1992), 562.
- [7] Lindsey J Byom and Bilge Mutlu. 2013. Theory of mind: Mechanisms, methods, and new directions. *Frontiers in human neuroscience* 7 (2013), 413.
- [8] Cathlin V Clark-Gordon, Nicholas D Bowman, Alan K Goodboy, and Alyssa Wright. 2019. Anonymity and online self-disclosure: A meta-analysis. *Communication Reports* 32, 2 (2019), 98–111.
- [9] Valerian J Derlega, Sandra Metts, Sandra Petronio, and Stephen T Margulis. 1993. *Self-disclosure*. Sage Publications, Inc.
- [10] Nicolas Ducheneaut, Ming-Hui Wen, Nicholas Yee, and Greg Wadley. 2009. Body and mind: a study of avatar personalization in three virtual worlds. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1151–1160.
- [11] Nicholas Epley, Boaz Keysar, Leaf Van Boven, and Thomas Gilovich. 2004. Perspective taking as egocentric anchoring and adjustment. *Journal of personality and social psychology* 87, 3 (2004), 327.
- [12] Isabel Fitton, Christopher Clarke, Jeremy Dalton, Michael J Proulx, and Christof Lutteroth. 2023. Dancing with the Avatars: Minimal Avatar Customisation Enhances Learning in a Psychomotor Task. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [13] Guo Freeman and Divine Maloney. 2021. Body, avatar, and me: The presentation and perception of self in social virtual reality. *Proceedings of the ACM on human-computer interaction* 4, CSCW3 (2021), 1–27.
- [14] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Alexandra Adkins. 2020. My body, my avatar: How people perceive their avatars in social virtual reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
- [15] Rebecca Fribourg, Ferran Argelaguet, Anatole Lécuyer, and Ludovic Hoyet. 2020. Avatar and sense of embodiment: Studying the relative preference between appearance, control and point of view. *IEEE transactions on visualization and computer graphics* 26, 5 (2020), 2062–2072.
- [16] Rosalie Hooi and Hichang Cho. 2014. Avatar-driven self-disclosure: The virtual me is the actual me. *Computers in Human Behavior* 39 (2014), 20–28.
- [17] Junko Ichino, Masahiro Ide, Hitomi Yokoyama, Hiroto Asano, Hideo Miyachi, and Daisuke Okabe. 2022. "I've talked without intending to": Self-disclosure and Reciprocity via Embodied Avatar. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–23.
- [18] Moises Mendez II. 2023. What to know about dream concealing his identity again. <https://time.com/6286572/youtube-dream-face-mask-back-on/>
- [19] Adam N Joinson. 2001. Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European journal of social psychology* 31, 2 (2001), 177–192.
- [20] Hye-seung Kang and Hee-Dong Yang. 2004. The effect of anonymity on the usage of avatar: Comparison of Internet relay chat and instant messenger. *AMCIS 2004 Proceedings* (2004), 327.
- [21] Hye-Seung Kang and Hee-Dong Yang. 2006. The visual characteristics of avatars in computer-mediated communication: Comparison of Internet Relay Chat and Instant Messenger as of 2003. *International Journal of Human-Computer Studies* 64, 12 (2006), 1173–1183.
- [22] Ruogu Kang, Stephanie Brown, and Sara Kiesler. 2013. Why do people seek anonymity on the internet? Informing policy and design. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2657–2666.
- [23] Anya Kolesnichenko, Joshua McVeigh-Schultz, and Katherine Isbister. 2019. Understanding emerging design practices for avatar systems in the commercial social vr ecology. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 241–252.

- [24] Bokyoung Kye, Nara Han, Eunji Kim, Yeonjeong Park, and Soyoung Jo. 2021. Educational applications of metaverse: possibilities and limitations. *Journal of educational evaluation for health professions* 18 (2021).
- [25] Kwan Min Lee. 2004. Presence, explicated. *Communication theory* 14, 1 (2004), 27–50.
- [26] Jie Li, Vinoba Vinayagamoorthy, Julie Williamson, David A Shamma, and Pablo Cesar. 2021. Social VR: A new medium for remote communication and collaboration. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–6.
- [27] Xiao Ma, Jeff Hancock, and Mor Naaman. 2016. Anonymity, intimacy and self-disclosure in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 3857–3869.
- [28] Maria Machneva, Anthony M Evans, and Olga Stavrova. 2022. Consensus and (lack of) accuracy in perceptions of avatar trustworthiness. *Computers in Human Behavior* 126 (2022), 107017.
- [29] Divine Maloney, Samaneh Zamanifard, and Guo Freeman. 2020. Anonymity vs. familiarity: Self-disclosure and privacy in social virtual reality. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*. 1–9.
- [30] Tony Manninen and Tomi Kujanpää. 2007. The value of virtual assets: the role of game characters in MMOGs. *International Journal of Business Science & Applied Management (IJBSAM)* 2, 1 (2007), 21–33.
- [31] Steven McCornack. 2014. *Interpersonal communication and you: An introduction*. Macmillan Higher Education.
- [32] James C McCroskey and Jason J Teven. 1999. Goodwill: A reexamination of the construct and its measurement. *Communications Monographs* 66, 1 (1999), 90–103.
- [33] Linda L McCroskey, James C McCroskey, and Virginia P Richmond. 2006. Analysis and improvement of the measurement of interpersonal attraction and homophily. *Communication Quarterly* 54, 1 (2006), 1–31.
- [34] David Myers. 1987. "Anonymity is part of the magic": Individual manipulation of computer-mediated communication contexts. *Qualitative Sociology* 10, 3 (1987), 251–266.
- [35] Kristine L Nowak. 2013. Choosing Buddy Icons that look like me or represent my personality: Using Buddy Icons for social presence. *Computers in Human Behavior* 29, 4 (2013), 1456–1464.
- [36] Kristine L Nowak and Jesse Fox. 2018. Avatars and computer-mediated communication: a review of the definitions, uses, and effects of digital representations. *Review of Communication Research* 6 (2018), 30–53.
- [37] Hyanghee Park, Daehwan Ahn, and Joonhwan Lee. 2023. Towards a Metaverse Workspace: Opportunities, Challenges, and Design Implications. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–20.
- [38] Sameer Patil and Jennifer Lai. 2005. Who gets to know what when: configuring privacy permissions in an awareness application. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 101–110.
- [39] Michael Pirson and Deepak Malhotra. 2011. Foundations of organizational trust: What matters to different stakeholders? *Organization science* 22, 4 (2011), 1087–1104.
- [40] David Premack and Guy Woodruff. 1978. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences* 1, 4 (1978), 515–526.
- [41] Stephen A Rains and Craig R Scott. 2007. To identify or not to identify: A theoretical model of receiver responses to anonymous communication. *Communication Theory* 17, 1 (2007), 61–91.
- [42] Rabindra Ratan, David Beyea, Benjamin J Li, and Luis Graciano. 2020. Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: a meta-analysis of the proteus effect. *Media Psychology* 23, 5 (2020), 651–675.
- [43] Anca Salagean, Eleanor Crellin, Martin Parsons, Darren Cosker, and Danaë Stanton Fraser. 2023. Meeting Your Virtual Twin: Effects of Photorealism and Personalization on Embodiment, Self-Identification and Perception of Self-Avatars in Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [44] Shen Sang, Tiancheng Zhi, Guoxian Song, Minghao Liu, Chunpong Lai, Jing Liu, Xiang Wen, James Davis, and Linjie Luo. 2022. Agileavatar: Stylized 3d avatar creation via cascaded domain bridging. In *SIGGRAPH Asia 2022 Conference Papers*. 1–8.
- [45] Craig R Scott and Stephen A Rains. 2020. (Dis)connections in anonymous communication theory: Exploring conceptualizations of anonymity in communication research. *Annals of the International Communication Association* 44, 4 (2020), 385–400.
- [46] Neal Stephenson. 1994. *Snow crash*. Penguin UK.
- [47] John Suler. 2004. The online disinhibition effect. *Cyberpsychology & behavior* 7, 3 (2004), 321–326.
- [48] Philipp Sykownik, Divine Maloney, Guo Freeman, and Maic Masuch. 2022. Something personal from the metaverse: goals, topics, and contextual factors of self-disclosure in commercial social VR. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [49] Jeffrey R Vittengl and Craig S Holt. 2000. Getting acquainted: The relationship of self-disclosure and social attraction to positive affect. *Journal of Social and Personal Relationships* 17, 1 (2000), 53–66.
- [50] Thomas Waltemate, Dominik Gall, Daniel Roth, Mario Botsch, and Marc Erich Latoschik. 2018. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE transactions on visualization and computer graphics* 24, 4 (2018), 1643–1652.
- [51] Joseph B Walther. 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication research* 23, 1 (1996), 3–43.
- [52] Sai Wang. 2020. The influence of anonymity and incivility on perceptions of user comments on news websites. *Mass Communication and Society* 23, 6 (2020), 912–936.
- [53] Shizun Wang, Weihong Zeng, Xu Wang, Hao Yang, Li Chen, Chuang Zhang, Ming Wu, Yi Yuan, Yunzhao Zeng, Min Zheng, et al. 2023. SwiftAvatar: Efficient Auto-Creation of Parameterized Stylized Character on Arbitrary Avatar Engines. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 6101–6109.
- [54] Jiangnan Xu, Konstantinos Papangelis, John Dunham, Jorge Goncalves, Nicolas James LaLone, Alan Chamberlain, Ioanna Lykourantzou, Federica L Vinella, and David I Schwartz. 2022. Metaverse: the vision for the future. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–3.

## A MEASURES

### A.1 Credibility

How do you feel about the counterpart?

#### Trustworthiness

	1	2	3	4	5	6	7	
Dishonest								Honest
Untrustworthy								Trustworthy
Dishonorable								Honorable
Immoral								Moral
Unethical								Ethical
Phoney								Genuine

#### Competence

	1	2	3	4	5	6	7	
Unintelligent								Intelligent
Untrained								Trained
Inexpert								Expert
Incompetent								Competent
Stupid								Bright

#### Goodwill

	1	2	3	4	5	6	7	
Doesn't care about me								Cares about me
Self-centered								Not Self-centered*
Unconcerned with me								Concerned with me
Insensitive								Sensitive
Not understanding								Understanding

### A.2 Social attraction

How do you feel about the counterpart?

(1 = Strongly disagree - 7 = Strongly agree).

- (1) I would like to have a friendly chat with her/him.
- (2) I think they could be a friend of mine.
- (3) They are easy to get along with.

### A.3 Anonymity

Please respond to the following questions. (1 = Strongly disagree - 7 = Strongly agree)

- (1) The counterpart in the video was anonymous to me.
- (2) I didn't know the counterpart in this video.
- (3) I can identify the counterpart in the video.\*
- (4) The actual identity of the counterpart in the video was hidden within their avatar.

### A.4 Self-disclosure

#### Breath of self-disclosure

Did the counterpart talk about the following topic? (Yes or No)

- (1) Age or height information
- (2) Hometown
- (3) Major or occupation
- (4) My personal habits
- (5) Things I have done which I feel guilty about
- (6) My deepest feelings
- (7) What I like and dislike about myself
- (8) What is important to me in life
- (9) My worst fears
- (10) My close relationships

#### Perceived comfort of self-disclosure

How comfortable did the counterpart seem when talking about the following topic?

(1 = Not at all comfortable - 7 = Completely comfortable)

#### Perceived honesty of self-disclosure

Did the counterpart seem truthful when discussing the following topic?

(1 = Not at all honest - 7 = Completely honest)

### A.5 Willingness to follow

Would you follow the counterpart?

- Yes, I would
- No, I wouldn't

\* These items were excluded in the final analysis.