JAIST Repository
https://dspace.jaist.ac.jp/

Title	Avatar Design for Bullying Prevention in the Metaverse: Avatar Appearances and the Presumption of Bullying
Author(s)	Kawakita, Teru; Kanai, Hideaki
Citation	AHFE International, 101: 52-58
Issue Date	2023-07
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/18813
Rights	Copyright (c) 2023 Authors. Teru Kawakita and Hideaki Kanai. AHFE International, Vol. 101, 2023, 52–58. This is an Open Access article distributed under the terms of Creative Commons Licence CC-BY [http://creativecommons.org/licenses/by/4.0/]. Original publication is available on AHFE Open Access via https://doi.org/10.54941/ahfe1002991.
Description	14th International Conference on Applied Human Factors and Ergonomics (AHFE) and the Affiliated Conferences 2023, July 20-24, 2023, San Francisco, California, USA

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Kansei Engineering, Vol. 101, 2023, 52–58
https://doi.org/10.54941/ahfe1002991

Avatar Design for Bullying Prevention in the Metaverse: Avatar Appearances and the Presumption of Bullying

Teru Kawakita and Hideaki Kanai

Japan Advanced Institute of Science and Technology, Nomi, Ishikawa, Japan

ABSTRACT

Several studies have demonstrated that individuals can be bullied because of their physical characteristics. We examined whether avatars can be victims or perpetrators of bullying based on their appearances. Participants were presented with four strong and four weak-impression avatars and asked to answer questions regarding impression, attractiveness, and presumption of bullying. The results reveal that the impressions of avatars are related to the presumption of bullying (p <.001). Individuals may hold the physical stereotypes of certain avatars. A simple solution for preventing potential harm in the metaverse is to not manipulate avatars with weak impressions. In addition, designers should avoid creating avatars with weak impressions, and individuals should recognize and control their stereotypes to avoid prejudice or discrimination.

Keywords: Kansei (affective) engineering, Appearance of avatar, Impression evaluation, Human-computer interaction, Presumption of bullying

INTRODUCTION

The metaverse is a computer-generated world with functions such as fully immersive, hyper spatiotemporal, self-sustaining, and economical (Wang et al., 2022). In the metaverse, individuals manipulate a digital representation of themselves called avatars, which serve as their virtual incarnation. This enables individuals to participate in a diverse range of social activities regardless of location and temporal constraints. Despite this technology's ability to enhance the human experience, online harassment (Wolak et al., 2007) and bullying persist. In particular, sexual harassment in the metaverse is a nascent area that the law can not yet fully address (Wiederhold, 2022). Similarly, bullying can occur in the metaverse. For example, a first-time user can be ostracized or ignored by others in a virtual game. Users of head-mounted displays (HMDs) may experience a heightened sense of immersion compared with those utilizing personal computer interfaces. If HMD users are bullied or harassed, the experience is particularly invasive to them. There are not enough laws in place to deal with bullying.

Some children repeatedly attack robots in the physical world (called robot bullying) (Kidokoro et al., 2015). In addition, non-human entities such as non-player characters (NPCs) and artificial intelligent-controlled characters

© 2023. Published by AHFE Open Access. All rights reserved.

within the metaverse can be negatively stereotyped or targeted by individuals with poor emotional regulation. We must consider manners for avoiding potential harm to individuals and non-human entities in the metaverse.

Bullying is caused by a variety of influences, including psychological and environmental factors. One cause of bullying is the impression of physical appearance. Ono and Hasegawa (2000) conducted a sociopsychological experiment in which participants identified bullying victims using facial photographs of 49 middle school students, and the results showed that people hold physical stereotypes of bullying victims. Individuals should not be bullied because of their physical appearance. However, children who are perceived as unattractive, overweight, or disabled are particularly vulnerable to bullying (Sweeting and West, 2001). Voss and Mulligan (2000) reported that short children tended to be bullied. The physical stereotypes of people with certain characteristics held by some individuals can lead to prejudice and discrimination. This type of discrimination based on appearance is commonly referred to as lookism (Tietje and Cresap, 2005). Individuals must not be subjected to discrimination based on unchangeable physical characteristics.

We considered how users and NPCs can avoid being bullied in the metaverse. A simple solution involves adjusting the appearance of avatars. In the metaverse, users can adopt all types of avatars, similar to changing clothes. By designing avatars that are less likely to be bullied, users can reduce their likelihood of being targeted based on their physical appearance. Therefore, avatar appearances and the possibility of bullying must be considered.

This study aims to explore the relationship between avatar appearances and bullying in the metaverse. It examines the avatar characteristics that are unlikely to be bullied. However, quantitatively examining whether bullying is caused by avatar appearances is difficult. Hence, we examined whether people hold physical stereotypes of avatar victims via impression evaluations.

METHOD

Participants

Forty-five graduate students participated in the evaluation ($M = 24.49$, $SD = 4.48$, forty males, four females, and one other). No participant had color vision abnormalities.

Design of Avatars

Male and female avatars were created using VRoid Studio. Kawakita et al., (2022) analyzed eye-tracking data to determine that people look closely at the faces of avatars. Therefore, the face elements of the avatars were carefully designed. The avatars were characterized by elements relating to strong or weak impressions based on three levels (Figure 1).

Glasses and red clothing are simply items on their own. However, a certain impression can be created by combining several items. As examples of the elements of weak impressions, Level 1 includes "fat + gloomy face," Level 2 includes "fat + gloomy face + glasses," and Level 3 includes "fat + gloomy face + glasses + downcast eyes." As examples of elements of strong

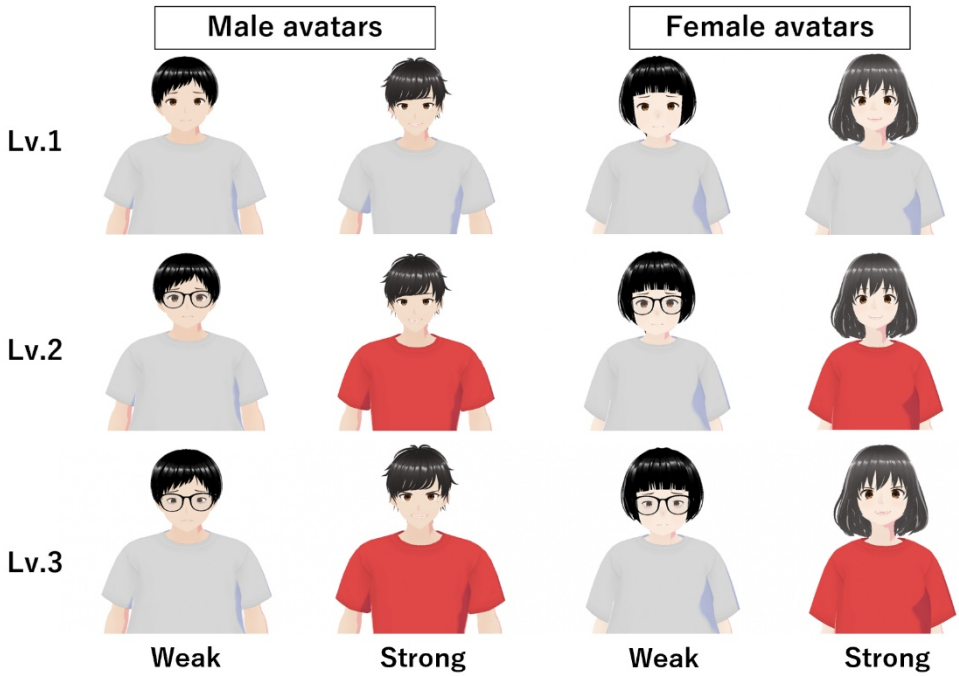

Figure 1: Three levels of strong and weak impressions.

impressions, Level 1 includes "good figure + cheerful face," Level 2 includes "good figure + cheerful face + red clothing (color code is #04848)," and Level 3 includes "good figure + cheerful face + red clothing + sharp face (eyebrows)."

Level 3 avatars rated well as either strong or weak impressions. Therefore, we added avatars with different hairstyles based on Level 3. Figure 2 shows the eight avatars used in this study. Each avatar was assigned a letter: A-H.

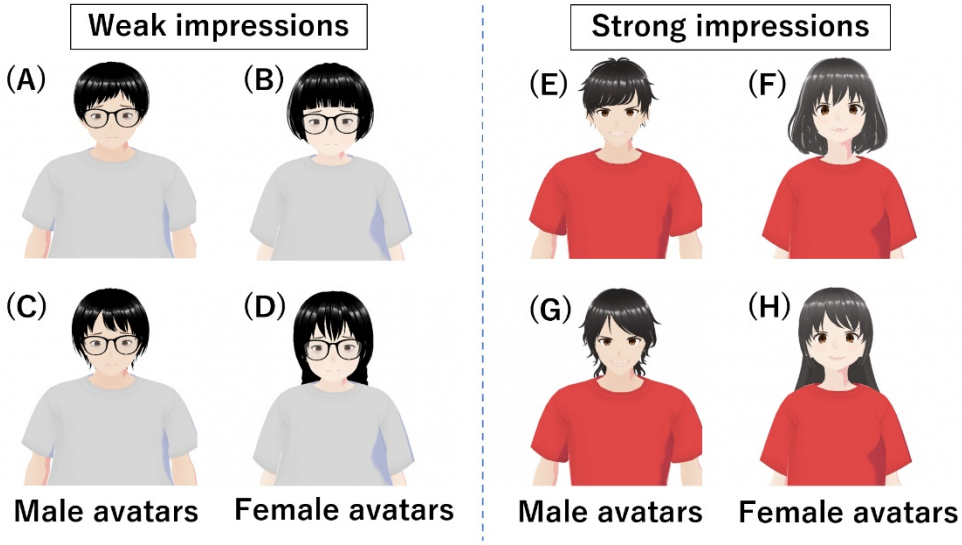

Figure 2: Weak and strong impressions.

Materials

A numerical rating scale (NRS) from 1 to 10 (10-point scale) was used for the evaluation. The survey items were as follows.

(1) Impression of avatars (weak to strong impression).

(2) Attractiveness of avatars (unattractive to attractive).

(3) The presumption of bullying (victim to perpetrator). In this study, the presumption of bullying meant that participants answered whether an avatar was a victim or perpetrator of bullying. Before responding to this question, participants were presented with the following situation:

"Some avatars were bullied by other avatars in the game. We will show you some avatars. Please guess if the avatar is a victim or perpetrator of bullying."

The rating of the presumption of bullying involved no criteria for determining which avatars were victims or perpetrators of bullying. Therefore, participants answered this question based solely on avatar appearance.

Procedure

This study used a within-subject design. Google Forms was used to record participants' responses. The avatars (Figure 2) were randomly presented utilizing Google Forms.

RESULTS

Table 1 presents the descriptive statistics of the avatars shown in Figure 2.

Table 2 presents the descriptive statistics of the male and female avatars for strong and weak impressions.

Table 1. Means for each avatar.

Item	A	B	C	D	E	F	G	H
Impression	2.71	2.76	2.89	2.67	7.13	6.82	7.44	6.42
	(1.46)	(1.65)	(1.61)	(1.67)	(2.03)	(1.96)	(2.08)	(1.92)
Attractiveness	2.98	3.42	2.96	3.11	5.8	6.56	5.33	6.02
	(1.47)	(1.57)	(1.36)	(1,67)	(1.91)	(1.94)	(1.89)	(1.8)
Bullying	3.07	2.84	3.02	2.78	7.42	7.16	7.8	7.09
	(1.86)	(1.85)	(1.79)	(1.94)	(1.84)	(1.85)	(1.97)	(1.9)

Parentheses indicate the standard deviation

Table 2. Means of male and female avatars: weak / strong.

	Weak avatars		Strong avatars	
	Male	**Female**	**Male**	**Female**
Impression	2.8 (1.53)	2.71 (1.65)	7.29 (2.05)	6.62 (1.94)
Attractiveness	2.97 (1.41)	3.27 (1.62)	5.57 (1.91)	6.29 (1.88)
Bullying	3.04 (1.82)	2.81 (1.88)	7.61 (1.91)	7.12 (1.87)

Parentheses indicate the standard deviation

Table 2 indicated that male avatars were rated slightly higher than female avatars in terms of the impression and presumption of bullying. Female avatars were rated slightly higher than male avatars in terms of attractiveness. Because the forty participants were male, they may have rated female avatars more favorably. However, the difference between male and female avatars was not statistically significant.

We hypothesize that the impression and attractiveness of the avatars relate to the presumption of bullying. A multilevel analysis was performed. The data-sets are 45 (participants) × 8 (avatars) = 360. There were no missing values. The model was as follows: [bullying ~ 1+impression+attractiveness+(1|avatars)]. $R^2 = .52$. The fixed-effect omnibus tests were as follows: impression of avatars ($p < .001$) and attractiveness of avatars ($p = .575$). Table 3 lists the fixed-effects parameter estimates. The results in Table 3 show that the impression of avatars significantly affects the presumption of bullying.

Table 3. Fixed effects parameter estimates.

	Estimate	**Standard error**	**95% Confidence interval**		*df*	*t*	*p*
			Lower	Upper			
Intercept	3.85	.69	2.5	5.21	9.74	5.57	<.001
Impression	.30	.06	.18	.41	357	5.09	<.001
Attractiveness	-.04	.06	-.16	.09	352.63	-.56	.575

The qualitative data from the participants were observed as follows.
Participants expressed that:

- The avatar's impression changes with the eyebrows, colors, and glasses.
- Different facial expressions changed the impression of the avatar.

DISCUSSION

This study aimed to explore the relationship between avatar appearances and bullying in the metaverse. It examined avatar characteristics that are unlikely to be bullied. The multilevel analysis revealed that the avatar impression can affect the presumption of bullying. Individuals tend to hold physical stereotypes of avatars and classify them as victims or perpetrators based solely on their appearance. These stereotypes do not immediately cause bullying. However, metaverse individuals should avoid using avatars with weak impressions to protect themselves from potential harm. In the metaverse, the probability of bullying arising from the physical stereotypes of the victim is not zero.

The multilevel analysis also indicated that the attractiveness of avatars does not significantly affect the presumption of bullying. Participants who are attracted to a particular avatar may be psychologically resistant to identifying the avatar as a victim or perpetrator.

Some users may prefer to manipulate avatars with weak impressions. Moreover, avatars are virtual incarnations, and some users may be uncomfortable using other avatars. In particular, users who associate their avatars more closely with themselves are more likely to be psychologically resistant to using other avatars. For such individuals, avoiding the use of avatars with weak impressions might not be a comprehensive solution. Ultimately, individuals who hold physical stereotypes of avatars must re-evaluate their presumptions. People must recognize their stereotypes and control them to avoid prejudice or discrimination. Several studies on perspective-taking have demonstrated that stereotypes can be changed (Yee and Bailenson, 2006; Banakou et al., 2016). Similar perspective-taking techniques may be applied to change the stereotypes of avatars in the metaverse.

As shown in Figure 1, a combination of several elements produces different impressions. Avatars with weak impressions are stereotyped as victims of bullying, and these carry potential risks. Therefore, this study shows that designers should be careful to not subconsciously build weak impressions into avatars to prevent potential risks.

LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

This study demonstrates that individuals tend to stereotype certain physical traits of avatars. However, stereotypes hold not been demonstrated to be associated with actual bullying or aggressive behavior. We plan to further research the possibility of the physical stereotypes of avatars causing bullying in the metaverse. A limitation of this study is the small number of female participants. Because all participants were presented with male and female avatars, we believe that gender differences affected the perception of the avatar's impressions. However, gender differences could not be verified because of an insufficient number of female participants.

RESEARCH ETHICS

This study was approved by Life Science Committee of Japan Advanced Institute of Science and Technology.

REFERENCES

Banakou, D., Hanumanthu, P. D., & Slater, M. (2016). Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias. *Frontiers in Human Neuroscience*, *10*, 601. https://doi.org/10.3389/fnhum.2016.00601

Kawakita, T., Onishi, G., Ishihara, S., Hashimoto, K., & Kanai, H. (2022). A Study on the Effects of Different Counselor Avatar Appearances and Seriousness of Client Problems. *Transactions of Japan Society of Kansei Engineering*, *21* (3), 267–274. (in Japanese).

Kidokoro, H., Suehiro, Y., Brscic, D., & Kanda, T. (2015). Robot Navigation Strategy for Escaping from Children's Bullying. *Transactions of Information Processing Society of Japan*, *56* (4), 1203–1216. (in Japanese).

Ono, T., & Hasegawa, Y. (2000). Physical Stereotypes of Ijime Victims. *The Japanese Journal of Experimental Social Psychology*, *40* (2), 87–94. (in Japanese).

Sweeting, H., & West, P. (2001). Being different: Correlates of the experience of teasing and bullying at age 11. *Research Papers in Education*, *16* (3), 225–246.

Tietje, L., & Cresap S. (2005). Is lookism unjust? The ethics of aesthetics and public policy implications. *Journal of Libertarian Studies*, *19* (2), 31–50.

Voss, L. D., & Mulligan, J. (2000). Bullying in school: Are short pupils at risk? Questionnaire study in a cohort. *British Medical Journal*, *320*, 612–613.

Wang, Y., Su, Z., Zhang, N., Xing, R., Liu, D., Luan, T. H., & Shen, X. (2022). A Survey on Metaverse: Fundamentals, Security, and Privacy. *IEEE Communications Surveys & Tutorials*. https://doi.org/10.1109/COMST.2022.3202047.

Wiederhold, B. K. (2022). Sexual Harassment in the Metaverse. *Cyberpsychology, Behavior, and Social Networking*, *25* (8), 479–480.

Wolak, J., Mitchell, KJ., & Finkelhor, D. (2007). Does online harassment constitute bullying? An exploration of online harassment by known peers and online-only contacts. *The Journal of Adolescent Health*, *41*(6), S51-S58.

Yee, N., & Bailenson, J. (2006). Walk a mile in digital shoes: The impact of embodied perspective-taking on the reduction of negative stereotyping in immersive virtual environments. *Proceedings of PRESENCE 2006: The 9th Annual International Workshop on Presence, Cleveland, Ohio, USA*.