

(2022). In *HCI International 2022 - Late Breaking Papers. Interaction in New Media, Learning and Games*. Lecture Notes in Computer Science, vol. 13517. Springer.

Gender and Culture Differences in Perception of Deceptive Video Filter Use

Susan C. Herring,^[0000-0001-7861-5206] Meredith Dedema,^[0000-0003-0203-8945] Enrique Rodriguez,^[0000-0002-8874-9336] and Leo Yang^[0000-0001-7332-5971]

¹ Indiana University, Bloomington IN 47405, USA
{herring, fndedema, enrodri, leoyang@indiana.edu}

Abstract. This study investigates how Augmented Reality (AR) beauty filters affect perception of the motives and trustworthiness of a stranger in online social interaction. One-on-one Zoom interviews were conducted with 44 video filter users from five different cultural backgrounds – China, India, South Korea, Spain, and the US – about their reactions to a hypothetical scenario intended to represent a deceptive use of beauty filters. The findings of quantitative and qualitative thematic analysis suggest that beauty filter use does not strongly affect assessments of trustworthiness among young video filter users, and that it is likely to be interpreted in different – and perhaps naïve – ways, such as a lack of self-confidence. Women express more compassion and understanding toward the “deceptive” stranger, while men express more negative judgment and distrust. Culture differences do not fall neatly along East-West lines, but rather vary for different themes. The findings contrast with evidence of actual risk of deception in online environments, particularly risk to women, who are more often targets of online dating deception.

Keywords: augmented reality, beauty filter, trust.

1 Introduction

It has recently become possible to modify one’s own and other peoples’ appearance in recorded videos and video chat using facial filters, augmented reality 3-D animations overlaid dynamically on the image of the face. Filters are enormously popular among young people on social media (Chua & Chang, 2016) and are rapidly gaining in popularity on video conferencing platforms such as Zoom. While filters can promote playful enjoyment, boost self-confidence, and facilitate identity exploration (e.g., Javornik et al., 2022), they also raise ethical concerns. The effects produced by face-transforming filters, such as those that change the user’s gender or age, overlap with those of deepfake videos and can be used to deceive, as for example in online dating site scams.¹ Beauty enhancement filters raise concerns about unrealistic beauty standards and increasing body dissatisfaction (e.g., Lyu et al., 2021) and can also be deceptive, for

¹ <https://slate.com/technology/2021/09/deepfake-video-scams.html>, last accessed 2022/05/26

example, if used to lure or entrap others.² More generally, difficult-to-detect filter use can exacerbate uncertainty and doubt in online interactions and promote a culture of skepticism or ‘disbelief as default’ (Gregory, 2019).

This study investigates how beauty filters affect the perception of the motives and trustworthiness of strangers in online social interaction through interviewing video filter users about their reactions to a hypothetical scenario intended to represent a deceptive use of beauty filters. Specifically, we ask: How does the use of a beauty filter that significantly changes one’s appearance affect trust in social interaction? Do participants’ answers to this questions vary according to their gender and cultural background?

2 Literature Review

2.1 Augmented Reality Filters

Augmented reality (AR) filters have become extremely popular with the general public in recent years, transforming the landscape of online communication (Fribourg et al., 2021). Some AR technologies are dedicated to improving the consumer experience; for example, by allowing consumers to virtually try on products such as virtual sunglasses (Yim & Park, 2019) and digital fashion (Xue et al., 2021). Other filters are intended for entertainment or for use on social media platforms. Filter effects can generally be situated along a spectrum of change from subtle beautification to cartoonification to grotesque distortion; some filters change the user’s gender or age; others exchange the user’s face with the face of another person. TikTok, for example, offers a wide range of facial modifications, from beauty enhancement to unrealistic distortions (Li, 2021). Some appearance enhancement apps that are popular in Asia even add filters directly to the camera of the user’s phone, such that any video recorded from the phone is automatically filtered, thereby extending filter use beyond social media.³ The increasingly normalized use of filters raises issues of authenticity and the potential for misrepresentation. Javornik et al. (2022) identified “ideal self-presentation,” “transformed self-presentation,” and “social interactions” as motivations for AR face filter use on Instagram. However, there is little research as yet on the social outcomes of video filter use. Most studies of online self-presentation focus either on photographs or selfies (e.g., Chae, 2017) or textual communication (e.g., Walther, 1996).

2.2 Filters and Self-Concept

AR filters serve as an important tool for online activities related to body satisfaction and self-esteem (Yu & Lee, 2020). Studies of teenage girls have found that using appearance-enhancing filters can boost the users’ mood and self-confidence, and that girls

² <https://timesofindia.indiatimes.com/blogs/voices/india-becoming-sextortion-capital-of-the-world/>, last accessed 2022/05/26

³ <https://www.insider.com/samsung-phones-default-beauty-mode-camera-airbrush-2016-6?amp>, last accessed 2022/05/26

with low self-esteem or negative body image use filters more and engage in more self-promotion online (Lee et al., 2014). Similarly, in a Korean study, Chae (2017) found that individuals use filters on their selfies because of their desire for a more ideal online self-presentation. However, discrepancies between the appearance of the digitally beautified self and the actual self can lower self-esteem and exacerbate body dysmorphia (Kleemans et al., 2018). Plastic surgeons report that increasing numbers of young people of both sexes are requesting surgeries to alter their physical face to look like a filtered face from their social media feed.⁴ A Chinese study found that young women's selfie-filtering behavior increased their willingness to consider cosmetic surgery when mediated by body surveillance and body shame (Lyu et al., 2021). Girls and young women are especially susceptible to societal pressures to look attractive; they are also the most active users of beauty filters (Dhir et al., 2016). Most of the studies cited here were conducted in Asia, where filter use is more widespread than in the West (Madan et al., 2018). In the present study, we interviewed filter users from Asia and the West, and our analysis controls for interviewee gender.

2.3 Online Deception

Lying is a common behavior in daily life and is especially ubiquitous online. Technology and online settings allow for more deception (Keyes, 2004) as a result of the interplay between users and the technical affordances of the medium in which they are communicating, the meticulous control they have over their self-presentation strategies (Walther, 1996), and the psychological rewards that might lead users to lie to present themselves more positively (Hancock & Guillory, 2015). Of particular interest here are private interactions over video-mediated communication, including those in the context of online dating (Drouin et al., 2016; Toma & Hancock, 2010). Dating and friendship-seeking sites are typically oriented to the possibility of meeting offline, so in theory, online daters should present themselves truthfully. Nevertheless, in one study, 51% of online daters admitted to misrepresenting themselves (Whitty, 2008).

Women are more likely to lie about their physical characteristics (Lo et al., 2013), while men most often lie about their relationship status, goals, and height (Schmitz et al., 2013). Relatedly, less attractive individuals are more likely to engage in deception in their online dating profiles (i.e., via enhanced photographs and physical characteristics), suggesting a strong connection between self-presentation strategies, perceived attractiveness, and deception in the online realm (Toma & Hancock, 2010).

One study found that nine out of 16 female college students interviewed had experienced catfishing, and four out of 16 had been targets of online impersonation involving fake or heavily photoshopped pictures (Simmons & Lee, 2020). Another study found that 16% of European online daters had encountered scammers (Buchanan & Whitty, 2014). "Catfishing," an activity involving the creation of a fake online profile for deceptive purposes (Harris, 2013), is common in online dating contexts. Women are more likely to be targets and men are more likely to perpetrate this form of deception (Mosley

⁴ <https://people.com/health/snapchat-dysmorphia-plastic-surgery-trend/>, last accessed 2022/05/26

et al., 2020). The motives behind catfishing may not always be entirely criminal or malicious, however. For instance, people might misrepresent themselves due to loneliness, dissatisfaction with physical traits, or for sexual identity exploration (Santi, 2019). Nonetheless, victims of catfishing can be exposed to severe physical, psychological, and financial harm (Koch, 2017; Santi, 2019). Public awareness of these behaviors in the United States has been raised through the popular TV show “Catfishing.” However, there are no laws tailored to criminalize catfishing in the US (Santi, 2019). Catfishing victims can also be found in other countries, including China (Huang et al., 2015), India (Kaur & Iyer, 2021), and Korea (Kim, 2015).

Increasingly, online dating scammers are using deepfake videos to fool their victims into believing that they are interacting with a desirable potential sexual or romantic partner.⁵ The effects of deepfake videos overlap with those of face-transforming filters, particularly filters where the user swaps their face with that of another person (Westerlund, 2019). Filters can also be used to perpetuate deception and fraud, as illustrated by the case of a beautiful young Chinese vlogger who was revealed to be an unattractive 58-year-old woman when her filter failed due to a technical glitch. Outrage followed this revelation, particularly since the “beautiful” vlogger had been soliciting cash gifts from her followers.⁶ Thus online self-presentation strategies can have real-life consequences if the online self does not match the users’ actual self, resulting, for example, in a breakdown of trust (Whitty & Buchanan, 2016).

2.4 Trust

Schoorman et al. (2007) define trust as someone’s perceptions about an individual’s ability, benevolence, and integrity. In their model of trust behavior, ability refers to the individual’s perceived skills and knowledge, benevolence accounts for how much someone perceives others’ intentions to be good-natured, and integrity denotes others’ personal and moral principles. While these three categories can be strengthened in face-to-face interaction via behavioral and emotional cues, they can be undermined by deceptive behaviors, as described above. There can also be real-life consequences for trusting others, especially when those individuals prove to be untrustworthy. Social psychology research has identified gender differences regarding trust behavior: Women are less likely to lose trust and also more likely to restore trust after being the target of a transgression than men are (Haselhuhn et al., 2015). Moreover, there is a relationship between trust and attractiveness. Women are more likely to perceive attractive men as more trustworthy based on their profile picture, while men tend to find attractive women’s profile pictures less trustworthy (McGloin & Denes, 2018). In this study we focus on the relationship between gender and perceived trustworthiness in a hypothetical scenario in which attractiveness is presumed to be enhanced through the use of video beauty filters.

⁵ <https://slate.com/technology/2021/09/deepfake-video-scams.html>; <https://www.freepressjournal.in/mumbai/sextortion-25-year-old-man-from-mumbai-gets-blackmailed-over-a-fake-video-clip>, last accessed 2022/05/26

⁶ <https://www.bbc.com/news/blogs-trending-49151042>, last accessed 2022/05/26

3 Methods

3.1 Data Collection

Data for this study were collected through two means, an online screener survey and one-on-one interviews. The online survey was designed to screen and recruit participants who use video filters; it also asked about use of different social media platforms and use of different types of filters. A recruitment message containing a link to the screener was sent out via several electronic mailing lists at a large university in the midwestern United States and posted on the researchers' social media accounts. To qualify for participation in the interview study, individuals had to be 18 or older, able to participate in a spoken interview in English, and be from China, India, South Korea, Spain, or the US. We focused on these countries for three main reasons. First, they represent broad cultural differences between the East and the West, and filter differences along this axis can be expected, since filters are more widespread in Asia than the West. Second, the selection allows us to explore differences within the two broad cultures. India, although part of Asia, differs culturally from East Asia, and there are cultural differences between Spain and the US. Finally, the members of our research team have first-hand knowledge of these cultures. This study received Institutional Review Board approval from the authors' university.

We conducted one-on-one, semi-structured interviews with survey respondents who met our requirements, all of whom were familiar with video filter use on social media and who were video filter users themselves. Forty-four participants were interviewed from five different cultural backgrounds, as summarized in Table 1. All but five of the interviewees were residing in the US at the time of the interviews (3 were in South Korea, and 2 were in Spain), and most were students. The interviewees ranged in age from 19 to 38; 64% are female, 32% are male, one identified as non-binary, and one interviewee declined to provide their gender. The interviewees reported using video filters on Instagram, Snapchat, TikTok, Zoom, Kakao Talk, and other social media platforms.

Table 1. Gender and cultural background of interview participants.

	China	India	South Korea	Spain	USA	Total
Female	8	4	5	4	7	28
Male	1	5	2	1	5	14
Other	1	0	0	0	1	2
Total	10	9	7	5	13	44

The semi-structured interviews lasted 45-60 minutes and were conducted and recorded over Zoom. The 45-item interview protocol included questions about participants' use of filters, the effects of filter use on their self-concept, and their perceptions of others' filter usage in social interaction. This study focuses on responses to a subset of the interview questions involving beauty filters, including comments made about motivations for beauty filter use, their effects on perceived authenticity and deception, and perceptions of frequency of filter use by others. These responses are described in

aggregate as background for an in-depth, quantitative analysis of interviewee responses to the following two-part hypothetical question:

- a. If you were to meet someone in person whom you previously chatted with exclusively over video online, and you found out that their face looked very different from their videos, how would you feel about them? Why?
- b. Would you feel differently if it was a potential romantic partner? Why?

3.2 Data Analysis

The interviews were transcribed by Zoom and manually checked for accuracy. A thematic content analysis approach was used to code the transcripts (Vaismoradi et al., 2013). Some themes were informed by the wording of the hypothetical question (i.e., *emotional reaction* and *action taken in response*). Other themes emerged from the data using a grounded theory approach (Glaser, 2002) and an iterative coding process. This involved discussion among the four authors until consensus on the themes was reached. We ended up with seven themes, which we coded independently for each part of the hypothetical question. In addition, we coded for whether the interviewee said they would respond differently to a potential romantic partner. The resulting codebook (variables and values) can be seen in Tables 2 and 3 in section 4.2. After training, each author coded part of the data, then all four authors together checked a random sample of more than half the data to confirm the consistency of the coding.

3.3 Research Questions

In this study we address two research questions:

- RQ1: How does the use of beauty filters affect trust in social interaction with a stranger?
- RQ2: Do participants' answers to this question vary according to their gender and cultural background?

This is an exploratory study; no specific hypotheses are advanced. However, based on the literature reviewed in the previous sections, we expected to find gender differences in trust and in attitudes toward beauty filter use.

4 Findings

4.1 Beauty Filter Use

Most participants, except for some males and the nonbinary participant, reported that they use video beauty filters, explaining that they employ such filters to subtly enhance their natural appearance. The interviewees perceived beauty filter use to be prevalent in video mediated communication, estimating that a majority of social media users, 60% to 90% on average, use filters to enhance their appearance. Young people and females were generally thought to use beauty filters the most.

4.2 Responses to the Hypothetical Question

We asked each interviewee, “How would you feel if you found that someone you’d been video chatting with online was using a filter all along and looked very different when you met them in person?” The interviewees overwhelmingly interpreted this question as referring to beauty enhancement filters. Responses to the hypothetical scenario varied according to gender and culture. The gender breakdown is presented quantitatively and discussed qualitatively, and the cultural differences are discussed qualitatively, due to the small numbers of participants in some of the culture groups.

Gender Differences. The results for each thematic variable broken down by gender are presented in Tables 2 and 3.

Overall, women more often than men expressed disappointment, compassion, understanding, respect, and speculated about (lack of) self-confidence, and more women (initially) said they would proceed normally. Men attributed dishonesty as the motive more, were more distrustful, and expressed indifference and mixed judgments more than women did. These overall patterns are illustrated in the following quotes.

- 1) S_06(F): If I’m meeting for a date or something, I would feel *disappointed*. I would say, like, that person’s a bit superficial, but, if it’s ... for friendship or something, I would say, “okay”. It wouldn’t upset me if it’s like, if that person doesn’t like to post their authentic face because *maybe something’s going on with their self-esteem, and it has to be treated, and we cannot judge that, because it’s like, the more you judge, the more they’re gonna hide themselves from society. So I wouldn’t judge them.*
- 2) I_16(M): [I’d be] *angry*. If they don’t have a good reason for using the filter, then I would end it there. I don’t know what a good reason would be for using the filter with a potential romantic partner, and not letting him or her know that you are using a filter. So that’s basically *deceiving*. You are practically *lying* to me even before we have met. So why would I continue the relationship if they can lie about such small things.

Most participants said their response would be different if the other (henceforth, O) was a potential romantic partner. Women said ‘slightly’ different more, and men said ‘strongly’ different more (Figure 1).

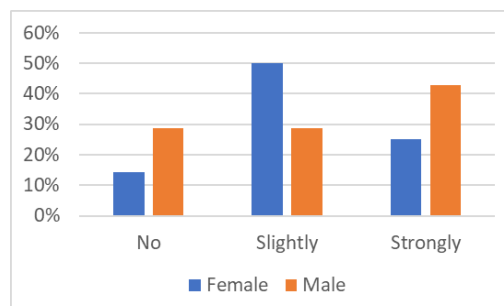


Fig. 1. Would your answer be different if it was a potential romantic partner?

Table 2. Responses to initial question (Part a.).

Variable	Values	Females N = 28	Males N = 14	Other N = 2	All N = 44
Motive for O's presumably deceptive use of filters	1) Dishonesty	7 (23)	7 (39)	0 (0)	14 (27)
	2) Self-confidence	11 (35)	4 (22)	1 (50)	16 (31)
	3) Other	5 (16)	4 (22)	0 (0)	9 (18)
	4) Didn't mention	8 (26)	3 (17)	1 (50)	12 (24)
	Total	31 (100)	18 (100)	2 (100)	51 (100)
Reaction to O presumably being less attractive in real life	1) Disappointed	5 (18)	0 (0)	0 (0)	5 (11)
	2) Don't care	2 (7)	0 (0)	1 (50)	3 (11)
	3) It's expected	2 (7)	0 (0)	0 (2)	2 (5)
	4) Didn't mention	19 (68)	14 (100)	1 (50)	34 (77)
	Total	28 (100)	14 (100)	2 (100)	44 (100)
Respectful of O's choice to use filters	1) Yes	10 (36)	4 (29)	1 (50)	15 (34)
	2) No	6 (21)	2 (14)	0 (0)	8 (18)
	3) Neutral/mixed	5 (18)	6 (43)	0 (0)	11 (25)
	4) Didn't mention	7 (25)	2 (14)	1 (50)	10 (23)
	Total	28 (100)	14 (100)	2 (100)	44 (100)
Emotional reaction to O's presumably deceptive self-presentation	1) Cheated/betrayed	8 (29)	2 (13)	0 (0)	10 (22)
	2) Disappointed	3 (11)	1 (7)	0 (0)	4 (9)
	3) Sad	2 (7)	0 (0)	0 (0)	2 (4)
	4) Surprised	3 (11)	3 (20)	0 (0)	6 (13)
	5) Indifferent	3 (11)	2 (13)	0 (0)	5 (11)
	6) Other	6 (21)	7 (47)	1 (50)	14 (31)
	7) Didn't mention	3 (11)	0 (0)	1 (50)	4 (9)
	Total	28 (100)	15 (100)	2 (100)	45 (100)
Attitude toward O	1) Compassionate	4 (13)	0 (0)	0 (0)	4 (9)
	2) Understanding	6 (20)	2 (13)	1 (50)	9 (19)
	3) Indifferent	4 (13)	3 (20)	0 (0)	7 (15)
	4) Mixed	2 (7)	5 (33)	0 (0)	7 (15)
	5) Judgmental	13 (43)	5 (33)	1 (50)	19 (40)
	6) Didn't mention	1 (3)	0 (0)	0 (0)	1 (2)
	Total	30 (100)	15 (100)	2 (100)	47 (100)
Would you still trust O?	1) Yes	4 (14)	3 (21)	0 (0)	7 (16)
	2) No	4 (14)	4 (19)	1 (50)	9 (20)
	3) Mixed/neutral	0 (0)	1 (7)	0 (0)	1 (2)
	4) Didn't mention	20 (71)	6 (43)	1 (50)	27 (61)
	Total	28 (100)	14 (100)	2 (100)	44 (100)

Action taken in response	1) Proceed normally	5 (18)	1 (7)	0 (0)	6 (14)
	2) Say something	3 (11)	1 (7)	1 (50)	5 (11)
	3) Cut off contact	2 (7)	1 (7)	0 (0)	3 (7)
	4) Seek other qualities	0 (0)	1 (7)	0 (0)	1 (2)
	6) Didn't mention	18 (64)	10 (71)	1 (50)	29 (66)
Total		28 (100)	14 (100)	2 (100)	44 (100)
Response dif- fers if a po- tential roman- tic partner	1) No	4 (14)	4 (29)	1 (50)	9 (20)
	2) Yes, slightly	14 (50)	4 (29)	1 (50)	19 (43)
	3) Yes, strongly	7 (50)	6 (43)	0 (0)	13 (30)
	4) Didn't mention	3 (11)	0 (0)	0 (0)	3 (7)
Total		28 (100)	14 (100)	2 (100)	44 (100)

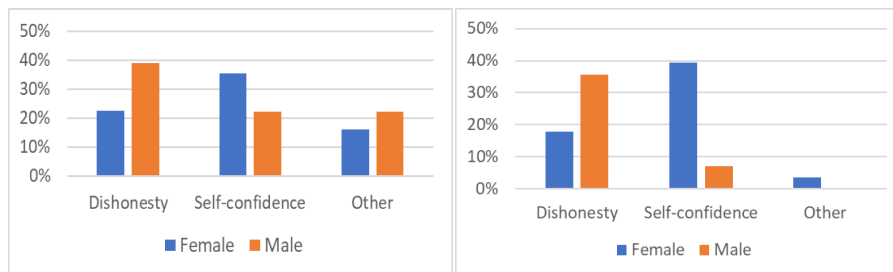
Table 3. Responses to question about potential romantic partner (Part b.).

Variable	Values	Females N = 28	Males N = 14	Other N = 2	All N = 44
Motive for O's presuma- bly deceptive use of filters	1) Dishonesty	5 (18)	5 (36)	0 (0)	10 (23)
	2) Self-confidence	11 (36)	1 (7)	1 (50)	13 (30)
	3) Other	1 (4)	0 (0)	0 (0)	1 (2)
	4) Didn't mention	11 (39)	8 (57)	1 (50)	20 (45)
Total		28 (100)	14 (100)	2 (100)	44 (100)
Reaction to O presumably being less at- tractive in real life	1) Disappointed	12 (43)	4 (29)	0 (0)	16 (36)
	2) Don't care	2 (7)	0 (0)	0 (0)	2 (5)
	3) It's expected	1 (4)	0 (0)	0 (0)	2 (2)
	4) Didn't mention	13 (46)	10 (71)	2 (100)	25 (57)
Total		28 (100)	14 (100)	2 (100)	44 (100)
Respectful of O's choice to use filters	1) Yes	3 (11)	1 (7)	0 (0)	4 (9)
	2) No	5 (18)	6 (43)	0 (0)	11 (25)
	3) Neutral/mixed	10 (36)	1 (7)	0 (0)	11 (25)
	4) Didn't mention	10 (36)	6 (43)	2 (100)	18 (41)
Total		28 (100)	14 (100)	2 (100)	44 (100)
Emotional re- action to O's presumably deceptive self- presentation	1) Cheated/betrayed	7 (24)	6 (43)	0 (0)	13 (29)
	2) Disappointed	7 (24)	2 (14)	0 (0)	9 (20)
	3) Sad	2 (7)	0 (0)	0 (0)	2 (4)
	4) Disrespected	0 (0)	2 (14)	0 (0)	2 (4)
	5) Other	8 (28)	2 (14)	0 (0)	10 (22)
	6) Didn't mention	5 (17)	2 (14)	2 (100)	9 (20)
Total		29 (100)	14 (100)	2 (100)	44 (100)

Attitude toward O	1) Compassionate	2 (7)	1 (7)	0 (0)	3 (6)
	2) Understanding	10 (33)	0 (0)	1 (50)	11 (23)
	3) Indifferent	1 (3)	1 (7)	0 (0)	2 (4)
	4) Mixed	2 (7)	2 (13)	0 (0)	4 (9)
	5) Judgmental	14 (47)	10 (67)	0 (0)	24 (51)
	6) Didn't mention	1 (3)	1 (7)	1 (50)	3 (6)
Total		30 (100)	15 (100)	2 (100)	47 (100)
Would you still trust O?	1) Yes	4 (14)	1 (7)	0 (0)	5 (11)
	2) No	4 (14)	3 (21)	0 (0)	7 (16)
	3) Didn't mention	20 (71)	10 (71)	2 (100)	32 (72)
Total		28 (100)	14 (100)	2 (100)	44 (100)
Action taken in response	1) Proceed normally	0 (0)	0 (0)	0 (0)	0 (0)
	2) Say something	1 (4)	1 (7)	0 (0)	2 (5)
	3) Cut off contact	9 (32)	3 (21)	0 (0)	12 (27)
	4) Seek other qualities	2 (7)	1 (7)	0 (0)	3 (7)
	6) Didn't mention	16 (57)	9 (64)	2 (100)	27 (61)
Total		28 (100)	14 (100)	2 (100)	44 (100)

In comparing the responses to the first and the second parts of the question (Tables 2 and 3), it can be seen that men's responses differ more than women's. Men mention (lack of) self-confidence as a motive less, a lack of respect more, feeling cheated more (and other feelings less), more negative judgment, and less trust. Women are more disappointed and more understanding, but say they will cut off contact more often (for reasons of character/lack of self-esteem) when thinking of a potential romantic partner. These differences are highlighted in what follows for each theme that emerged from the content analysis. Quotes by the interviewees are provided, as well as figures that compare the responses of females and males to the two parts of the hypothetical question. (The figures exclude the two 'other' gender participants and the N/A responses for the sake of clarity.)

Figs. 2a and 2b show the distribution of responses by females and males relating to the theme of **motivation**.



Figs. 2a & 2b. Motivation in initial and potential romantic partner responses.

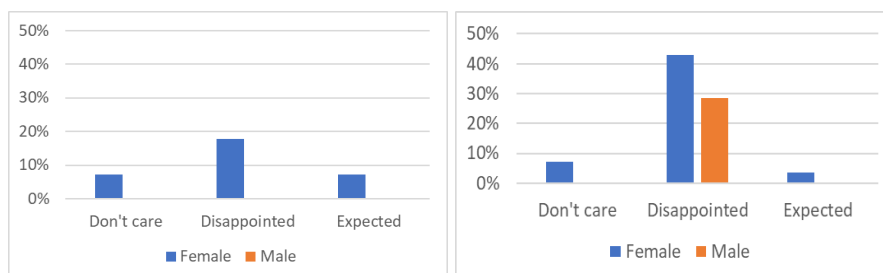
When speculating about the hypothetical other's (O's) motivations, a surprisingly high number of participants attribute a lack of self-confidence to O. This is especially true for women, as illustrated in quote 1) above and quotes 3) and 4) below.

- 3) C_02(F): I may be curious why they are using filters at the moment even though they use it before. And I may feel, *she is not very confident or he's not very confident*, and that he wants the filter at the moment.
- 4) SK_06(F): I understand them. I think I could think that *they are not confident enough in their appearances* and that's why they use video filters. I do it sometimes.

Mentions of self-confidence and seeking to understand O's motives decrease for males when considering a potential romantic interest. Both genders, especially men, also characterize O's behavior as dishonest or deceptive, as in quote 2). Thinking that O is deceptive is not incompatible with seeking to understand what personal or psychological issues might motivate O's behavior, however, as illustrated in quote 5):

- 5) US_31(F): That's *catfishing*, and I'm against that. I mean I probably wouldn't just like, rule them out as a person, like if I really like their personality and this was like coming out of something that was like *a deep insecurity*? I mean I don't love being *lied to*, and I do think that's *lying*. But you know I don't think that's like the worst thing in the world.

The second theme, **attractiveness**, is mentioned more by women and mentioned more overall for romantic relationships, as shown in Figs. 3a and 3b.



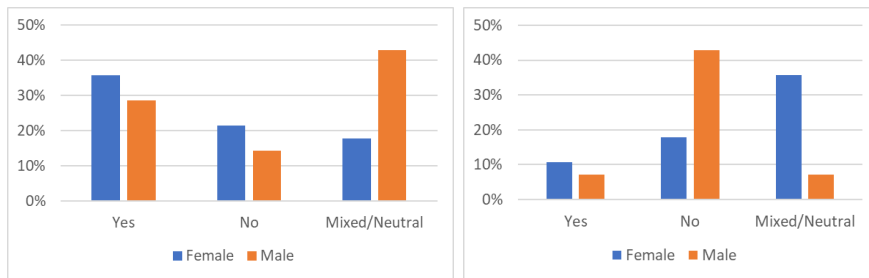
Figs. 3a & 3b. Attractiveness in initial and potential romantic partner responses.

Disappointment is expressed most often in mentions of attractiveness, as illustrated in quotes 6) - 8) below.

- 6) SK_12(F): I think agreeing to meet an online person in the real world is all about physical appearances. His looks are the only real thing I know about him. The online relationship is built on physical appearances. The chats we have online is meaningless until we actually meet in person and experience things together. So I would communicate with someone online for his looks but if he doesn't look like that? it's a lie? a fake? Then I wouldn't want to meet him. I would run away.
- 7) C_04(F): So, normally, they will look very good online. And then, like in real life would not be that attractive, so I will be disappointed. If that's the case, I'll be disappointed. Yeah.

- 8) SK_13(M): I would be a little disappointed, but that wouldn't affect the relationship much. There are a lot of other things besides the looks. I would be a little disappointed at first, as long as she doesn't look totally different, I would still try to get to know her better.
 Interviewer: What if she looks VERY different?
 SK_13(M): Then I will be VERY disappointed.

Responses relating to the theme of **respect** are shown in Figs. 4a and 4b.



Figs. 4a & 4b. Respect in initial and potential romantic partner responses.

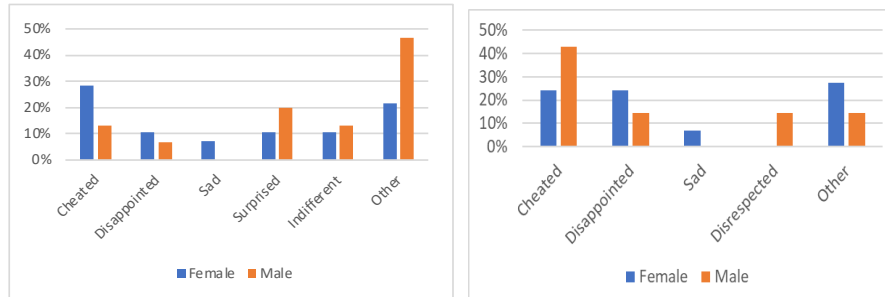
A surprising number of interviewees indicated that they would respect O's choice to use a filter (quote 10), despite many of them finding the behavior deceptive, as in 9).

- 9) SK_02(F): In my personal opinion, I respect their intention to use video filter, but I do not want to follow them. Because even on social media we try to show our real life not a fake one.
- 10) C_09(M): I mean, I'm, I'm okay with that. Because, I mean, I know that if you're having, like, a ... since it's not in person, you want yourself to look good, that's totally fine. So, I'm totally cool with that.

But respect decreases sharply when considering a potential romantic partner, especially for men, as illustrated in quote 2) above and quote 11).

- 11) US_27(M): I think if it was a potential romantic partner I'd feel kind of *led on*, or like *cat-fished*, in a sense, *lied to*. Obviously, maybe if it was such a drastic difference, I was expecting this person to look like this, but then they actually looked like this, I don't think I'd want to pursue a relationship with them anymore. It just shows that like they're, to me, like they're not confident in themselves and that I was just *deceived*.

Our hypothetical scenario asked the interviewees specifically about their **emotional reactions** (How would you feel?). Women's and men's responses are summarized in Figs. 5a and 5b.

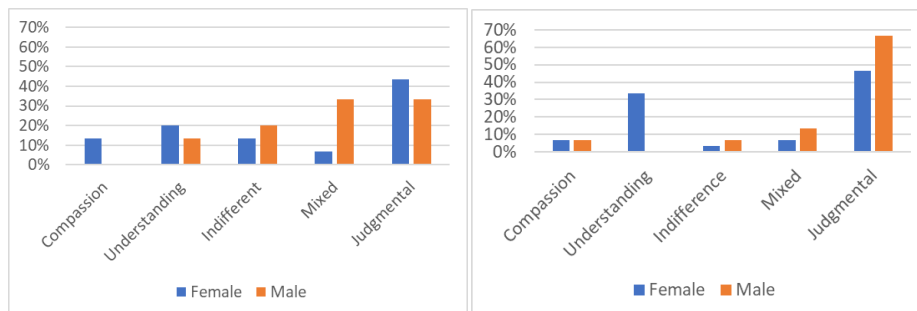


Figs. 5a & 5b. Emotional reactions in initial and potential romantic partner responses.

Both genders say they would feel cheated or deceived, while more women than men would feel disappointed, especially in a romantic relationship (see quote 6). Men feel much more cheated in romantic relationships and mention fewer other feelings, as in quotes 2), 11), and 13).

- 12) US_18(F): I would feel catfished. I would definitely feel cheated, which I think is one of the reasons why it makes it so that I don't want to use misleading filters like that.
- 13) SK_11(M): I wouldn't be too emotionally invested if it was a guy, but if it is a romantic partner, I would feel betrayed. It will be difficult to maintain the relationship.

Regarding **attitude** toward O, as Figs. 6a and 6b show, the women were more compassionate and understanding than the men; see quotes 1), 4), and 14), although the women were also more judgmental initially (quote 15). But men were more judgmental if it concerned a potential romantic relationship (quotes 2 and 11).

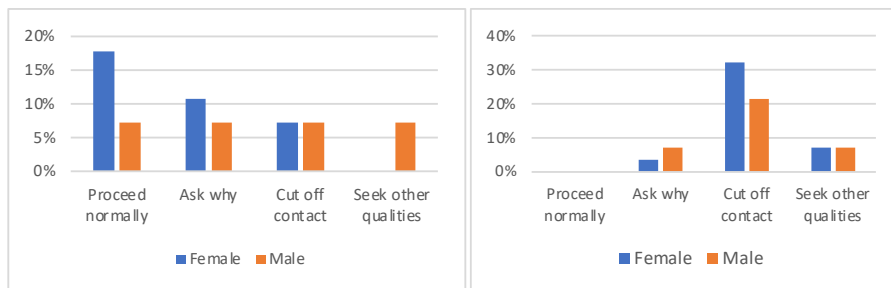


Figs. 6a & 6b. Attitude and judgment in initial and potential romantic partner responses.

- 14) C_10(F): Depends on my relationship with them. If I'm not dating any of them, I don't really care. Cuz like I understand it's kind of a common thing right now. And in my mind, I'll go like "oh poor girl, another one who doesn't feel comfortable with her face". But only in my head, I won't say anything. I just wish them the best. Hope they can make peace of [sic] their facial appearance in the future.
- 15) I_10(F): I would definitely feel deceived, and, because, they, like, put in something that was completely different, especially if it was like a romantic relationship, and the

relationship kind of matters for looks too, then I would feel completely deceived of them, like, kind of, they are trying to portray someone they are not, and I might ... not go follow with it, not because of how they look, but because they won't, like, comfortably show their true selves.

Some interviewees mentioned **actions** they would take in response to the hypothetical scenario; see Figs. 7a and 7b.



Figs. 7a & 7b. Actions taken in response in initial and potential romantic partner responses.

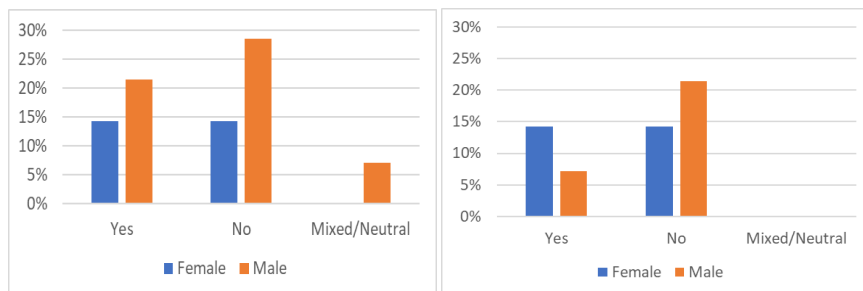
Women were initially more likely to carry on as if nothing was amiss (quote 16).

16) SK_14(F): I would think oh that was all photoshop. I would feel a little bit cheated but *I wouldn't tell her that you look completely different from your video.*

No one mentioned doing this with a potential romantic partner, though. Both genders indicated they would 'cut off contact' more in that case, especially women (quotes 13, 15, 17).

17) I_20(F): If it's somebody that I would like to date then I would feel much more betrayed than somebody I just want to be friends with. And I would just consider that oh, they feel more confident and happy by using a filter so it's not my space to comment on them, but I would be taken aback, I would be thrown off. *I would likely not engage with them more.*

Finally, the gender breakdown for **trust** is given in Figs. 8a and 8b.



Figs. 8a & 8b. Trust in initial and potential romantic partner responses.

Although we expected the hypothetical scenario to elicit reactions of distrust, mentions of trust or distrust were relatively infrequent overall. Some occurred only when prompted by a question about trust from the interviewer. Overall, male interviewees more often expressed distrust of O (quotes 18 and 19), and positive expressions of trust decreased sharply, from 21% to 7%, for men in a potential romantic relationship.

- 18) US_16(M): If it was someone I had, like, a personal connection with in any way, it might be kind of a red flag just because it's like, "oh, well, they're not showing who they really are", like it kind of gives that idea of, like, "is everything they're talking about truthful?", and it goes down that rabbit hole.
- 19) SK_11(M): I would be skeptical if I were to have a business relationship with that person because he might be deceiving.
Interviewer: So you would distrust that person
SK_11(M): Yes. I wouldn't take it for granted that he is a distrustful person, but I will be suspicious.

Surprisingly, however, several interviewees of both genders said that they would still trust O, despite finding O's behavior deceptive (as reported above).

- 20) Interviewer: Would you trust them?
S_08(F): Yeah. I would trust them, I would just think that this thing ... makes them feel more reassured, and that's the reason why they use the video filter, but I would trust them. I judge honesty by different standards.
- 21) SK_14(F): yeah [I would trust him], its just about his appearance that he lied about so I don't think that relates to his actual personality, yeah.
- 22) Interviewer: What does it mean to be disappointed? Does it mean you can't trust her anymore?
SK_13(M): I won't distrust her. Like I said, the looks are not everything I am looking for in a relationship. I would be disappointed in her looks but that doesn't mean I will be disappointed in her personality. So I'll try to get to know her.

Cultural Differences. Some patterns in the interviewee responses appear to be related to culture. For example, Indian and US participants most often attributed the **motivation** for O's deceptive video filter use to dishonesty (quotes 2 & 11). In contrast, South Korean and Spanish interviewees, who were mostly females, more often associated O's use of video filters with a lack of self-confidence (quotes 2 & 4). Chinese participants often did not offer any motivation for O's video filter use.

A very different picture emerges in the case of a potential romantic partner. South Korean participants mentioned dishonesty as a possible motivation more than self-confidence (quote 6), while US participants attributed O's behavior to a lack of self-confidence more. The Chinese and Indian interviewees often did not mention potential motivations for using filters in this hypothetical scenario.

Participants were more likely to be disappointed about O's **attractiveness** if O was a potential romantic partner, with Koreans and Chinese emphasizing the importance of attractiveness more and expressing the most disappointment (quotes 6-8).

Regarding **respect** for O's choice to use a beauty filter, South Korean (quote 9) and Spanish (quote 1) participants expressed the most respect, whereas Indians expressed the least respect (quote 2). US participants had mixed responses, and Chinese interviewees often did not mention respect. However, in the hypothetical situation involving a potential romantic partner, none of the participants from China, India, or the US, and only a few of the Spanish interviewees, mentioned respecting O's choice.

The **emotional reactions** mentioned by Indian and US participants mostly related to feeling cheated and betrayed (quotes 2 & 12), while Spanish interviewees mostly showed disappointment (quote 1). For a potential romantic partner, negative feelings associated with dishonesty (i.e., feeling betrayed) were most common overall, and Chinese, South Korean, and American participants also mentioned being more disappointed (quotes 7 & 8).

Overall, in terms of **attitudes** towards O's use of a video filter, participants were more judgmental than compassionate. Americans and Indians were the most judgmental (quotes 11 & 15), followed by South Korean, Chinese, and Spanish interviewees. Chinese participants were the most understanding overall (quote 14), while Spanish interviewees were the most compassionate (quote 1); both cultural groups were predominantly females. When responding to a potential romantic partner, US and Chinese participants were the most understanding, although Americans also expressed judgmental attitudes in this hypothetical situation (quotes 11 & 12).

In terms of **actions** they would take in response to O's deceptive use of a video filter, US and Spanish participants mentioned confronting O about the discrepancy between O's filtered and real self, while Chinese and South Korean participants would say nothing and proceed normally. We hypothesize that this could be due to a cultural difference regarding conflict avoidance in Asian cultures (Leung et al., 2002).

Finally, mentions of **trust** were infrequent overall. Interestingly, US participants mentioned a lack of trust more often initially (quote 18), but if O was a potential romantic partner, they expressed less distrust. Overall, however, participants were less trusting of O when O was a potential romantic interest.

5 Discussion

In this study we asked, "How does the use of beauty filters affect trust in social interaction with a stranger?" and "Do participants' answers to this question vary according to their gender and cultural background?"

With respect to the first question, the responses to our hypothetical scenario suggest that beauty filter use does not strongly affect assessments of untrustworthiness among young video filter users, and that it is likely to be interpreted in different – and perhaps naïve – ways, especially by women. The female participants sought to understand – and sometimes sympathized with – the "deceptive" filter users' motivations in terms of body image and self-esteem. This finding is consistent with past research findings that women have more compassion than men toward others (López et al., 2018). When women expressed a negative emotional reaction, it was often disappointment that O

lacked self-confidence, was not honest, or was not as physically attractive as initially advertised, especially if O was a potential romantic partner.

In contrast, male participants more often responded with negative judgment and distrust, especially toward potential romantic partners. However, only a few interviewees, most of them men (e.g., quote 18), indicated that they would be suspicious of O's motives more generally. Other men said that they might be surprised but would not care, because "it's [O's] choice" to present themselves as they please.

The scenario presented in our interview study is not only hypothetical. Our findings contrast with evidence of actual risk of deception in online environments, particularly risk to women, who are more often targets of online dating deception (Mosley et al., 2020). Romance fraud, for example, is on the rise since the Covid-19 lockdown in 2020. Fraudsters sometimes groom their victims for months to establish trust, and they increasingly make use of video communication to do so.⁷ Victims have been defrauded of money, and some have experienced online sexual abuse (Buchanan & Whitty, 2014). The risks associated with online dating can transition to the offline world, including rape and murder (Santi, 2019). Males, too, have been victimized by sextortion schemes that involve scammers misrepresenting themselves as attractive women online. In one tragic recent case in the US, a 17-year old boy who was being blackmailed after sharing nude photos with a "young girl" committed suicide rather than face the shame of exposure.⁸ As filter use becomes increasingly normalized, awareness of these risks needs to be raised among young internet users.

With respect to the second question, some cultural differences were observed, although these should be interpreted with caution, because the numbers of interviewees in some of the culture groups are small. The Indian and US interviewees were most likely to attribute O's behavior to dishonesty, although the US participants also attributed it to lack of self-confidence if O was a potential romantic partner. Overall, Spanish and South Korean participants were the most respectful towards O's beauty filter use, whereas Indian and US interviewees were the most judgmental. There is likely an interaction here with the gender findings, in that the Spanish and South Korean participants are mostly females, while about half of the Indian and US interviewees are males. Finally, the female Spanish participants were the most compassionate of all the demographic groups, although the Chinese interviewees showed the most understanding of O's possible reasons for using a deceptive video filter with a potential romantic partner. These findings suggest that the broad distinction between East and West is insufficient to explain cultural differences in perceptions of trustworthiness in video filter use; rather, the histories of individual cultures (e.g., their experiences with cases of deceptive filter and deepfake use) and the demographics of individuals within those cultures should also be taken into account.

⁷ <https://www.which.co.uk/news/article/online-dating-fraud-up-40-through-pandemic-aKHlv5M09iYX>, accessed 2022/05/27

⁸ <https://www.weau.com/2022/05/24/high-school-senior-dies-by-suicide-after-falling-victim-online-sextortion-family-says/>, accessed 2022/05/27

6 Conclusion

We have presented what we believe is the first cross-cultural study of video filter use. Our findings indicate that beauty filters are considered normal and acceptable by young social media users, and most users say that they employ such filters only to subtly enhance their natural appearance. However, in response to a hypothetical situation where a stranger used enhancement filters that significantly changed their appearance in private video interactions, participants attributed the discrepancy between the filtered and the real appearance to motives such as dishonesty and lack of self-confidence. Gender and culture-related differences were evident in their responses.

Despite many of our participants explicitly mentioning deception as the motivation for O's use of video filters and judging it negatively, most participants did not seem inclined to distrust O. This suggests that for young social media users, beauty filter use does not strongly affect assessments of trustworthiness in private interactions over video. However, this lack of distrust could render young adults susceptible to becoming targets of online fraud or abuse, a possibility that is especially concerning as regards women, who tend to trust more than men and to trust attractive men more than less attractive men (Haselhuhn et al., 2015; McGloin & Denes, 2018).

A limitation of this study is that we only interviewed active video filters users; the perceptions of non-users, who might also be vulnerable to deception, were not taken into account. Also, we did not interview children or teenagers, the most vulnerable age groups. A second limitation is that this exploratory study reports on a sample of participants that is relatively small and culturally unbalanced, especially for Spain and Korea. Moreover, some of the latter participants were located in their home countries rather than the US. While their responses might be more culturally authentic, their experiences differ from those of the majority of interviewees. Finally, there is an imbalance in the ratio of male to female participants in some culture groups. A larger, more balanced interview population is desirable for future studies.

Meanwhile, the future outlook for filter use is increasing tolerance and acceptance of beauty filters. As beauty filter use becomes more expected and normalized, more people will use them to enhance their appearance online. The filters themselves will also become more sophisticated; some, such as the Chinese app Meitu, already have the ability to virtually modify the bone structure of the user's face. Finally, we foresee a time when social media filters and deepfake apps, already similar in some of their effects, will converge and become functionally indistinguishable, raising further challenges for identifying deception in online interaction.

References

1. Buchanan, T., Whitty, M. T.: The online dating romance scam: causes and consequences of victimhood. *Psychology, Crime & Law* 20(3), 261-283 (2014).
2. Chae, J.: Virtual makeover: Selfie-taking and social media use increase selfie-editing frequency through social comparison. *Computers in Human Behavior* 66, 370–376 (2017).

3. Chua, T. H. H., Chang, L.: Follow me and like my beautiful selfies: Singapore teenage girls' engagement in self-presentation and peer comparison on social media. *Computers in Human Behavior* 55, 190-197 (2016).
4. Dhir, A., Pallesen, S., Torsheim, T., Andreassen, C. S.: Do age and gender differences exist in selfie-related behaviours? *Computers in Human Behavior* 63, 549–555 (2016).
5. Drouin, M., Miller, D., Wehle, S. M., Hernandez, E.: Why do people lie online? “Because everyone lies on the internet”. *Computers in Human Behavior* 64, 134-142 (2016).
6. Fribourg, R., Peillard, E., McDonnell, R.: Mirror, Mirror on my phone: Investigating dimensions of self-face perception induced by augmented reality filters. In: Marchal, M., Ventura, J., Olivier, A. H., Wang, L., Radkowski, R. (eds) 2021 IEEE INTERNATIONAL SYMPOSIUM ON MIXED AND AUGMENTED REALITY (ISMAR), pp. 470-478. IEEE Computer Society, Los Alamitos, CA (2021).
7. Glaser, B. G.: Conceptualization: On theory and theorizing using grounded theory. *International Journal of Qualitative Methods* 1(2), 23-38 (2002).
8. Gregory, S.: Cameras everywhere revisited: How digital technologies and social media aid and inhibit human rights documentation and advocacy. *Journal of Human Rights Practice* 11(2), 373–392 (2019).
9. Hancock, J. T., Guillory, J.: Deception with technology. In: Sundar, S. (ed.) *The handbook of the psychology of communication technology*, 270-289 (2015).
10. Harris, A.: Who coined the term catfish? *Slate*, 18 January (2013). <https://slate.com/culture/2013/01/catfish-meaning-and-definition-term-for-online-hoaxes-has-a-surprisingly-long-history.html>, last accessed 2022/05/27
11. Haselhuhn, M. P., Kennedy, J. A., Kray, L. J., Van Zant, A. B., Schweitzer, M. E.: Gender differences in trust dynamics: Women trust more than men following a trust violation. *Journal of Experimental Social Psychology* 56, 104-109 (2015).
12. Huang, J., Stringhini, G., Yong, P.: Quit playing games with my heart: Understanding online dating scams. In: *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, pp. 216-236. Springer (2015).
13. Javornik, A., Marder, B., Brannon Barhorst, J., McLean, G., Rogers, Y., Marshall, P., Warlop, L.: ‘What lies behind the filter?’ Uncovering the motivations for using augmented reality (AR) face filters on social media and their effect on well-being. *Computers in Human Behavior* 128, 107-126 (2022).
14. Kaur, G., Iyer, S.: Digital crimes on Indian online dating platforms during Covid-19: Impact on women. *International Journal of Law Management & Humanities* 4(4), 1277-1291 (2021).
15. Keyes, R.: *The post-truth era: Dishonesty and deception in contemporary life*. St. Martin's Publishing Group, New York (2004).
16. Kim, J. H.: Technique for identifying cyber crime using clue. *Journal of the Korea Institute of Information Security & Cryptology* 25(4), 767-780 (2015).
17. Kleemans, M., Daalmans, S., Carbaat, I., Anschütz, D.: Picture perfect: The direct effect of manipulated Instagram photos on body image in adolescent girls. *Media Psychology* 21(1), 93-110 (2018).
18. Koch, C. M.: To catch a catfish: A statutory solution for victims of online impersonation. *University of Colorado Law Review* 88(1), 233-280 (2017).
19. Lee, H. R., Lee, H. E., Choi, J., Kim, J. H., Han, H. L.: Social media use, body image, and psychological well-being: A cross-cultural comparison of Korea and the United States. *Journal of Health Communication* 19(12), 1343-1358 (2014).
20. Leung, K., Koch, P. T., Lu, L.: A dualistic model of harmony and its implications for conflict management in Asia. *Asia Pacific Journal of Management* 19(2), 201-220 (2002).

21. López, A., Sanderman, R., Ranchor, A. V., Schroevers, M. J.: Compassion for others and self-compassion: Levels, correlates, and relationship with psychological well-being. *Mindfulness* 9(1), 325-331 (2018).
22. Li, Y.: Research on user behavior of beauty filters in Tik Tok app: Taking college students as an example. Master's Thesis, Southwest University, China (2021).
23. Lo, S. K., Hsieh, A. Y., Chiu, Y. P.: Contradictory deceptive behavior in online dating. *Computers in Human Behavior* 29(4), 1755-1762 (2013).
24. Lyu, Z., Jiao, Y., Zheng, P., Zhong, J.: Why do selfies increase young women's willingness to consider cosmetic surgery in China? The mediating roles of body surveillance and body shame. *Journal of Health Psychology* 27(5), 1205-1217 (2021).
25. Madan, S., Basu, S., Ng, S., Ching Lim, E. A.: Impact of culture on the pursuit of beauty: Evidence from five countries. *Journal of International Marketing* 26(4), 54-68 (2018).
26. McGloin, R., Denes, A.: Too hot to trust: Examining the relationship between attractiveness, trustworthiness, and desire to date in online dating. *New Media & Society* 20(3), 919-936 (2018).
27. Mosley, M. A., Lancaster, M., Parker, M. L., Campbell, K.: Adult attachment and online dating deception: A theory modernized. *Sexual and Relationship Therapy* 35(2), (2020).
28. Santi, A.: "Catfishing": A comparative analysis of US v. Canadian catfishing laws & their limitations. S. Ill. ULJ 44, 75-104 (2019).
29. Schmitz, A., Zillmann, D., Blossfeld, H. P.: Do women pick up lies before men? The association between gender, deception patterns, and detection modes in online dating. *Online Journal of Communication and Media Technologies* 3(3), 52 (2013).
30. Schoorman, F. D., Mayer, R. C., Davis, J. H. An integrative model of organizational trust: Past, present, and future. *Academy of Management Review* 32(2), 344-354 (2007).
31. Simmons, M., Lee, J. S.: Catfishing: A look into online dating and impersonation. In: *International Conference on Human-Computer Interaction*, pp. 349-358. Springer (2020).
32. Toma, C. L., Hancock, J. T.: Looks and lies: The role of physical attractiveness in online dating self-presentation and deception. *Communication Research* 37(3), 335-351 (2010).
33. Vaismoradi, M., Turunen, H., Bondas, T.: Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences* 15(3), 398-405 (2013).
34. Walther, J. B.: Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research* 23(1), 3-43 (1996).
35. Westerlund, M.: The emergence of deepfake technology: A review. *Technology Innovation Management Review* 9(11), 39-52 (2019).
36. Whitty, M. T.: Revealing the 'real' me, searching for the 'actual' you: Presentations of self on an internet dating site. *Computers in Human Behavior* 24(4) 1707-1723 (2008).
37. Whitty, M. T., Buchanan, T.: The online dating romance scam: The psychological impact on victims—both financial and non-financial. *Criminology & Criminal Justice* 16(2), 176-194 (2016).
38. Xue, L., Parker, C. J., Hart, C. A.: How to design effective AR retail apps. In: tom Dieck, M. C., Jung, T. H., Loureiro, S. M. C. (eds.) *Augmented reality and virtual: New trends in immersive technology*, pp. 3-16. Springer (2021).
39. Yim, M. Y.-C., Park, S.-Y.: "I am not satisfied with my body, so I like augmented reality (AR)": Consumer responses to AR-based product presentations. *Journal of Business Research* 100, 581-589 (2019).
40. Yu, H., Lee, M.: Effects of the virtual makeup using beauty makeup applications on mood, body satisfaction, and self-esteem among female university students. *Journal of the Korean Society of Clothing and Textiles* 44(4), 727-738 (2020).