2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) | 979-8-3503-2838-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMAR59233.2023.00019

Using Identification with AR Face Filters to Predict Explicit & Implicit Gender Bias

Marie A. Jarrell *
Lab-STICC, IMT Atlantique
Brest, France

Etienne Peillard†
Lab-STICC, IMT Atlantique
Brest France

Abstract

Augmented Reality (AR) filters, such as those used by social media platforms like Snapchat and Instagram, are perhaps the most commonly used AR technology. As with fully immersive Virtual Reality (VR) systems, individuals can use AR to embody different people. This experience in VR has been able to influence real world biases such as sexism. However, there is little to no comparative research on AR embodiment's impact on societal biases. This study aims to set groundwork by examining possible connections between using gender changing Snapchat AR face filters and a person's predicted implicit and explicit gender biases. We discovered that participants who experienced identification with gender manipulated versions of themselves showed both greater and lesser amounts of bias against men and women. These results depended the user's gender, the filter applied, and the level of identification users reported with their AR manipulated selves. The results were similar to past VR findings but offered unique AR observations that could be useful for future bias intervention efforts.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies

1 Introduction

AR technology is incredibly varied including; wearable displays, handheld apps and games, and even GPS mapping systems. However, by far the most common and most engaged with use of AR is the AR photo filter. Snapchat, a popular platform known for its AR filters, has 332 million daily active users worldwide [15]. One report found 59% of US Snapchat users access the app on a daily basis for an average of 26 minutes [14]. These filters allow users to virtually play with their appearance and grants them the chance to explore their identities in bold new ways [27, 50]. Unfortunately, there may be negative consequences. There is growing public concern that AR beauty filters could contribute to sexist stereotypes and increase mental distress over one's appearance leading to drastic measures like plastic surgery [4, 11, 16, 49]. Others worry that the filters promote sexist or transphobic messages on how people of various genders should look [53]. The filter's have even brought up ethical questions about what identities are and are not appropriate to virtually replicate. FaceApp, a popular app that also changes people's faces using filters like Snapchat, had to remove a "race changing" filter after public outcry that it was contributing to racism and racist practices like black face [8]. Many of these concerns come from the belief that AR filter use can impact real world perceptions of ourselves and others and this belief may have merit.

We know identification with VR avatars can change how users perceive themselves to reflect the avatar they are embodying [37]. This correlates with increased empathy for individuals similar to the avatar which results in significant changes to audiences' perceptions of other societal groups [13, 18]. Studies have used "embodied identification" to combat biases of all kinds including gender [9, 28], race [6, 36, 48], and age [7, 46]. However, there is little research reproducing these studies with AR technology. AR is different as it can not only replace a person with a VR avatar but digitally change them. Thus our study seeks to replicate fully immersive VR identification research on gender swapping and gender bias. We use AR filters to digitally "change" participants' genders and see if there are correlations between identification with them and participants implicit and explicit gender biases using the following research questions:

RQ1: Is there a connection between the use of an AR filter to "change" a person's gender and that person's perceived level of gender bias whether implicit or explicit?

RQ2: Does a person's level of identification with an AR filter manipulation modify the predicted level of bias related to said filter?

2 Background

2.1 Identification and Bias Reduction

In interactive media identification is a phenomenon that merges user and avatar, forming a new user identity [37]. Identification with virtual avatars via embodiment can impact a wide variety of cognitive behaviors and perceptions. Gabbiadini et al. found identification with characters in hyper-masculine violent-sexist games increased participants endorsement of masculine beliefs and reduced empathy towards female victims of domestic violence [23]. While, Gonzalez-Liencres et al. found identification with a virtual victim of abuse could increase empathy for victims and be used to rehabilitate abusers [28]. Kishore et al. found identification via embodiment with a virtual African American victim of racial abuse from a virtual police officer led to police officer participants being more helpful toward said victim than those who witnessed the scene via third person perspective [36]. It seems embodied identification can have a wide range of effects on societal biases. Groom et al. found embodied identification with black applicants in racially priming situations like job interviews could increase racial bias [30]. Lopez et al. found that participants embodying female avatars during a virtual Tai Chi lesson had significantly higher implicit gender bias than those embodying male avatars [42]. Lopez et al. suggested that similar to how interview settings can trigger racial biases due to negative stereotypes about black individuals being unqualified [30], athletic scenarios could trigger gender based biases due to negative stereotypes about athletic activity and women [42]. We can not assume all types of virtual embodiment are the same and should evaluate how different virtual embodiment experiences, like those unique to AR, impact users.

2.2 AR Embodiment & Face Filters

While a longstanding concept in VR, AR embodiment research and knowledge is extremely limited. A large amount of writing on mixed-reality features technological papers about different systems [44]. When human centric research is conducted it often deals

*e-mail: marie.jarrell@imt-atlantique.fr.
†e-mail: etienne.peillard@imt-atlantique.fr.

DOI 10.1109/ISMAR59233.2023.00019

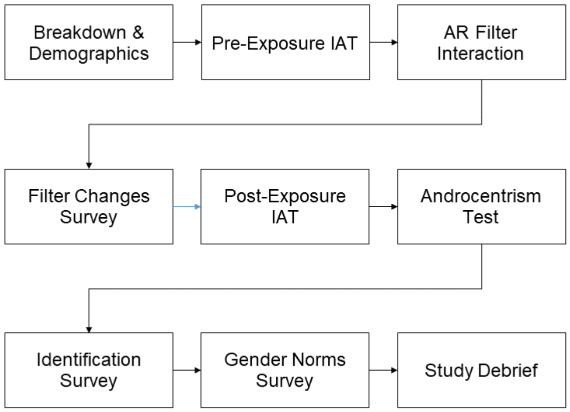

Figure 1: Workflow diagram of the procedure.

with basic mental functions such as; sensory perceptions, usability, emotional responses, cognitive load, learnability, attention, and motion sickness [44]. Research on embodiment is further split as different systems must first demonstrate that perceived presence in the system is possible [44]. Furthermore, when presence is established and embodiment is the key focus the effects may be explored in ways unique to the system [26]. Luckily, Genay et al. conducted a systematic review of AR research and embodiment and found that a sense of embodiment could be achieved across a wide variety of completely different AR systems as long as certain stimuli and induction conditions were met [26].

Of the many different AR systems the one most available to the public today are AR photo filters. Due to their ease of use and accessibility through smart phones, AR apps are both widely available and extremely impactful [3,32]. Findings suggest most of these users are women who, according to one reporting by the Dove Self-Esteem Project, show extremely high filter use [32]. The report showed 52% of surveyed girls reported daily filter use and 80% of girls reported to have used an app to change their appearance before the age of thirteen [32]. However, it is important to remember that these platforms have universal appeal. While Snapchat's audience predominantly consists of female users at 56.4% there is still an incredibly large male audience at 43.1% [14]. Simply put, AR filters are a huge part of the daily lives of millions of young people across the world and this influence is only expected to grow. Which is concerning as these filters may not only change projections of a user's face, but they could be capable of causing significant psychological changes in users as well. Recently, face filters, specifically beauty filters, have been accused of negatively impacting users' mental health and self confidence [4]. These concerns are relatively new so there is not adequate scientific evidence to verify the phenomenon.

Still, as mentioned there is evidence that embodying VR avatars can impact people's behaviors and mental attitudes. Beyond personal worries and accounts, there is evidence AR filters can impact perception. A study by Fribourg et al. found manipulating certain facial traits with filters impacted self perception in line with previous research on perception of others [21]. Meaning there is a possibility, however slim, that embodying and identifying with AR filters could encourage cognitive changes in users. However, there were unique effects in Fribourg et al.'s study on self-perception which suggested filters could also decrease participants' credence towards their own image [21]. Basically, there is a noted "specificity" in AR that needs to be studied. Our research can be seen as an attempt to remedy

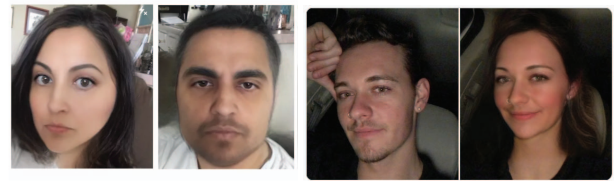

Figure 2: Examples of the Snapchat "My Twin" filters' real-time manipulations [33]. (Left: female to male, Right: male to female)

our lack of understanding on AR's influence. We want to see if identification with an AR filter can predict a person's feelings and actions. Specifically, if identification with a gender AR manipulated version of oneself correlates with different perceived levels of various gender biases. In doing so, we could provide empirical evidence showing if identification with AR filters does in some way connect with users' real-world gender biases.

3 Methods

This study examines connections between identification with an AR face filter which augmented the perceived gender of the user and their perceived gender bias. To do so we utilized the existing "My Twin" filters on Snapchat, Figure 2. Using Implicit Association Tests (IATs) [29], an androcentrism test [5], and survey questions on gender norms [24,40,43,54] we looked at how interacting with a "My Twin" filter that "changed" a user's gender impacted implicit and explicit gender biases. Now the concept of gender is incredibly broad and how we define and construct biases based on gender varies based on multiple factors such as an individual's culture, race, age, etc [1,52,56]. For our study we categorize participants based on their self identified gender and use the Snapchat filters parameters for deciding if a filter is a "female gender filter" or "male gender filter" since they were designed and largely adopted by users for these specific services [33]. We looked at not only the filter gender presentation but examined if participants' self-perceived levels of embodied identification with their AR manipulated selves (measured via a pre-validated questionnaire) significantly moderated the filters' correlated levels of perceived biases. A visual representation of our prediction model can be found in the appendix, Figure 9. These predicted interactions are represented by the following hypotheses:

Gender Embodiment hypothesis (H1): That people of different gender groups using AR filters that enhance (man x man filter) or swap (man x woman filter) their genders will display significantly different levels of perceived implicit and/or explicit gender bias. Specifically, that participants who use a gender filter that presents them as a gender other than their own (male -> female or female -> male) will exhibit less gender bias. Either in implicit associations between themselves and the other gender group via the IAT (H1a), the gender they default to in the androcentrism test (H1b), and/or their acceptance of gender norms (H1c).

Identification moderates Embodiment hypothesis (H2): There is an interaction between embodied identification and AR filter group. High levels of identification with an AR filter will strengthen predicted bias differences with regards to implicit associations (H2a), androcentrism (H2b), and/or acceptance of gender norms (H2c).

3.1 Demographics

The study was distributed globally online, and in person via recruitment at IMT Atlantique. The study was available in english and french with translations reviewed by native french and english speakers. The online survey was shared via email, social media links on LinkedIn and WhatsApp, and online survey sharing sites including SurveySwap, SurveyCircle, and PollPool. We recruited

Group	Items
Male	Man, Son, Father, Boy, Uncle, Grandpa, Husband
Female	Woman, Daughter, Mother, Girl, Aunt, Grandma, Wife
Same	Similar, Identical, Alike, Related, Equal, Match, Kindred
Other	Contrasting, Opposed, Different, Unalike, Contrary Separate, Contradictory

Table 1: Implicit Association Test (IAT) categories and items

246 participants who agreed to partake in the study after the full study was completed. After removing inattentive participants using attention checks, a final sample group of 122 participants was used for analysis. Of this sample there were 53 male participants, 66 female participants, 1 non-binary participant, and 2 participants who preferred not to disclose their gender. Since the last two groups were too small to examine we limited the study to the 119 male and female identifying participants. Members from both groups were evenly randomly split into each filter group.

The study was global but a majority our respondents listed the USA 16%, United Kingdom 16%, France 15%, India 8% or Germany 4% for one of their nationalities. Of our participants, 55% were aged 18-24, 26% were aged 25-30, 9% were aged 31-34, 7% were aged 35-40, and 3% were 41 and above. Other demographic metrics including race, language fluency, and education level were collected to compare our sample to the larger population. Of our sample the largest racial demographics were 67.2% White / Caucasian – Non Hispanic, 11.8% Asian, 08.4% Hispanic / Latino, and 07.6% Indian. Participants listed their most common fluent language which was largely english at 60.5%, then french 16%, german 07.6%, and spanish 04.2%. The education level was mostly Bachelor's Degree at 46.2%, and Master's Degree at 31.9 %, with 87.4 % having a college level degree (associates or higher). We asked participants about their political social and economic leanings as there have been correlations between these factors and perceptions of gender appropriate behavior. Socially the sample was, 03.4% Conservative, 08.4% Slightly Conservative, 14.3% Neutral, 10.1% Slightly Liberal, 31.9% Liberal, 28.6% Very Liberal, and 03.4% Preferred not to say. Economically the breakdown was, 05.9% Conservative, 08.4% Slightly Conservative, 21.8% Neutral, 17.6% Slightly Liberal, 31.1% Liberal, 10.9% Very Liberal, and 04.2% Preferred not to say. Information was collected on how much time individuals spent on image sharing social media platforms like SnapChat or Instagram. Overall use was 15.1% saying they I never go on such platforms, 21.8% 1 – 2 hours per week, 15.1% 3-5 hours, 28.6% 6-10 hours, 12.6% 11-20 hours, and 06.7% 21-40 hours. We also asked how often participants used image filters or photo manipulation applications to modify their pictures. In our sample 30.3% of participants reported they had never or basically never used them, 42.0% had used them a couple of times but not frequently, 15.1% used them about half of the time, 06.7% used them frequently but not all the time, and 05.9% used them on nearly all their photos. The study was reviewed and approved by the IMT Atlantique Ethics Board.

3.2 Procedure

3.2.1 Initial Survey & IAT

Figure 1 outlines the full study workflow. In the beginning of the study participants were taken to the Qualtrics page where the study was hosted. There they viewed an informed consent form telling them they would be using Snapchat and answering questions on the experience. Participants were initially told the study would be examining how AR filter interactions affected cognitive abilities such as creativity or mental logical task performance. Participants were informed that there would be no penalties for refusing to take part in the study and that they could exit out of the study at any time. Individuals were informed that participation was completely voluntary and there were no financial benefits for taking part in the study. Before being assigned a Snapchat filter, participants filled out several demographics questions as well as two questions on their current social media and filter use; "On average how much time do you spend on image sharing social media platforms ex SnapChat or Instagram?" and "Do you use image filters or image manipulation applications to modify your pictures". After the demographics survey, participants completed the pre-exposure IAT that tested their implicit connections to either the male or female gender. In this gender IAT, participants had to quickly categorize certain attributes (e.g., otherness) with pairs of words (e.g., female – same vs. male – other). Past IATs have measured implicit bias by whether or not participants unconsciously connect men or women with leadership positions [51] or with the concepts of good and bad [28]. A new IAT was used since we did not expect filter use to impact perceived leadership capability or morality. We constructed a new IAT to measure unconscious bias via association based on the concepts of "Inclusion of Other in the Self" [31] and ingroup / outgroup bias [22, 34]. A positive IAT score corresponded to a preference for pairing words connected to sameness with female coded words meaning a stronger connection to women. Whereas a negative IAT score corresponded to a preference for pairing sameness and male coded words providing evidence of a stronger connection to men. Table: 1 breaks down the specific attributes and words used in the IAT. The IAT was based on former IATs and used the same gendered terms even though the same vs other terms were new [6, 7, 20, 30, 34, 35, 42]. Participants were told this was a logic based test to examine their matching abilities so they were not primed to think of the concepts involved before the test.

Embodied Iden. [41] AVE: .844 Largest Correlatio: .218	When using the filter, it felt as if I was the filtered person.	.736
	I felt like the filter was a part of me when using the filter.	.833
	When using the filter, it was as if I became one with the filter.	.924
	When using the filter, I was transported into the filtered person.	.854
	When using the filter, it felt as if the filtered body became my own.	.900
	When using the filter, it was as if I acted directly through the filtered person.	.814

Table 2: Embodied Identification questions

3.2.2 Filter Interaction

After taking the first IAT participants were randomly assigned to the following Snapchat set-ups: 1) No Filter, to create a baseline for predicted bias when no filters are present. 2) Male Filter, for bias predictions when interacting with a male identified gender AR manipulation. 3) Female Filter, for bias prediction when interacting with a female identified gender AR manipulation. Due to software and phone capabilities limiting some filters to a partial manipulation, that was only fully applied when users took a photo of themselves, participants were led through a test to determine if the filters were fully applied in real time. (Only those with a fully applied real time filter, 122 total, were used in data analysis) Participants were given a simple task to induce identification with their AR augmented selves. They were asked to make faces of themselves displaying five emotions; joy, anger, sadness, disgust, and fear [19]. This task would make them aware of the connection between their facial movements and the filtered faces in order to increase identification similar to full body movement identification tasks [6]. The emotion task was cho-

Appearance: Male Norms [40,43] AVE: .867 Largest Correlation: .858	I think a young man should try to be physically tough, even if he's not big. It is important for men to look physically attractive in public. Men should not wear make-up, cover-up, or bronzer. Men should be concerned with their bodies and try to gain muscle.	.936 .743 .905 .883
Appearance: Female Norms [40,43] AVE: .822 Largest Correlation: .880	I think a young woman should try to be physically attractive, even if she's not thin. Women should always try to wear make-up. Women should be concerned with their bodies and try to lose weight.	.852 .828 .785
Behavior: Male Norms [24,40,54] AVE: .888 Largest Correlation: .880	Boys should prefer to play with trucks rather than dolls. Boys should excel at contact sports. A boy should prefer watching action movies to reading romantic novels. It is more important to encourage boys than girls to participate in athletics. Men should be the leader in any group. A man should always be the major provider in their household. A man should provide discipline in the family. Men make better bosses than women. If both husband and wife have jobs, the husband should work to make more money rather than the mother.	.944 .859 .916 .856 .873 .930 .877 .829 .907
Behavior: Female Norms [24,40,54] AVE: .838 Largest Correlation: .781	Girls should prefer to play with dolls rather than trucks. It is all right for a girl to want to play football or build toy trains. Girls should excel in domestic skills. A girl should prefer watching YA (Young Adult) dramas instead of watching football games. It is more important to encourage girls than boys to take home economics classes. Women should take a support role in any group. A woman should always be in charge of the cleaning, cooking, & decorating in their household. If both husband and wife have jobs and their child gets sick at school, the school should call the mother rather than the father.	.939 -.706 .875 .896 .834 .772 .888 .793

Table 3: Gender norm related factors and questions

sen to keep participants from taking part in gender priming activities such as making "sexy poses" and replicated the short play sessions users of Snapchat engage in while making photos and videos. Participants were told to make a face for each emotion for 6 seconds taking a total of 30 seconds overall (6 seconds x 5 emotions). Then keeping Snapchat open with the filter applied participants answered several questions about how the app and filter changed their face. Questions asked included such items as "Did the app change the shape of your face?","Did the app change your skin?","Did the app change features on you face?"(i.e. eyes, nose, and/or lip size) "Did the app change your facial hair" (added, removed, no change) and "Overall do you think the app made your face more masculine or more feminine?". We also included an open answer question for participants to explain how they perceived the changes and how using the filter made them feel. This allowed us to ensure participants both used the filter they were assigned and further increased participants' awareness of the changes the filters made.

3.2.3 Final IAT & Survey

After interacting with the filter participants completed a second IAT test. By comparing the pre and post IATs the filter interactions' impact on gender group connection could be observed. Next participants did a "creativity test" by coming up with the name and hobby of "a typical person". Once done they were taken to the next page and answered demographic questions about the person to test androcentrism. Androcentrism is a term used to described the phenomenon of centering men and the male perspective, and associating men with the default identity [2]. By asking for a name and hobby first participants were unconsciously led to consider the person's gender, as names are often gendered, without explicit coaching. The question about the character's gender was included in a list of other demographic questions such as age, nationality, and information on their family life to minimize its importance. By seeing the created characters' gender distributions among the filter groups we could see if participants were more likely to view men as a "typical person" despite nearly equal actual distribution of men and women [5].

Then, we asked participants two more questions about their past experience with the filters used in the study. Including "Have you used the filter featured in this study before?" and "If Yes how often have you used the filter". Afterwards, participants answered questions on their identification with the AR manipulation of themselves created by the face filter. We used two pre-validated scales on embodied identification one based on embodied presence [55] and one based on body ownership [10]. Lastly, participants answered questions on various gender norms facing both men and women. The norms included appearance norms, hobby norms, and societal role norms. Each of these constructs was measured with several 7-point scale items, mostly ranging from 1-Completely Disagree to 7-Completely Agree that were adapted from existing validated scales from similar past research [6,10,24,25,40,41,43,47,54,58].

Finally we re-confirmed the validity of the measurement scales with a Confirmatory Factor Analysis (CFA). We discovered Identification via Body Ownership was not facilitated by our AR interactions and thus not applicable. We also found that our items for examining biases regarding societal roles and hobbies were not significantly distinct from each other and were instead related to a larger encompassing factor. This new factor was dubbed Behavior norms and was a combination of elements from both Hobby and Role norms. Table 2 shows the questions for Identification via Embodied Presence or Embodied Identification. Table 3 shows the two main gender norm attitudes (appearance and behavior) used in the experiment. All the factors show convergent validity (AVE

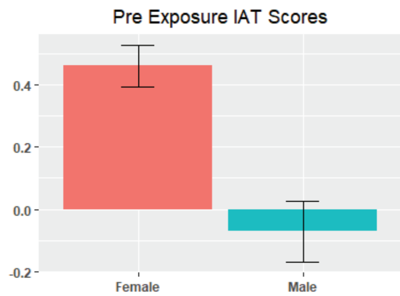

Figure 3: Male and female IAT scores pre Snapchat filter interactions.

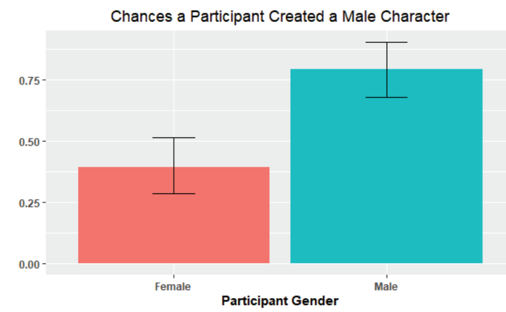

Figure 5: Percent chances (0-1 = 0-100%) a male character was created when asked to imagine an average person.

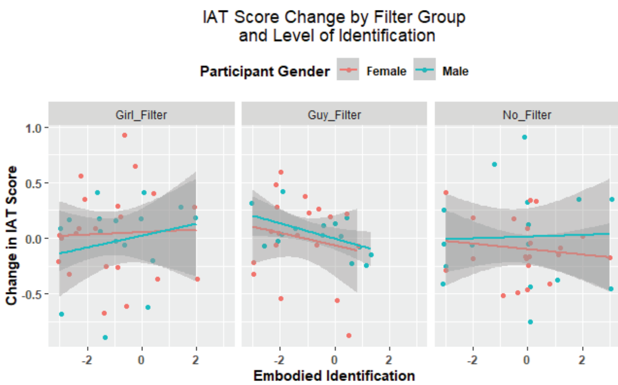

Figure 4: Male and female participants' changes in IAT (ΔIAT > 0 = greater female x same association, ΔIAT < 0 = greater male x same association) moderated by embodied identification

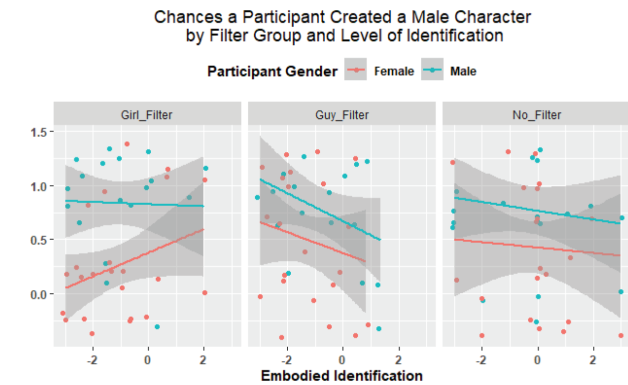

Figure 6: Chances of creating a male character (androcentrism) moderated by embodied identification.

> 0.50). For discriminant validity we did not compare the male norms to the female norms as they were supposed to measure the same thing for different populations not different factors altogether. So while covariance between say male appearance norms could be compared to female behavior norms we did not compare male appearance norms to female appearance norms. With this technique in place each factor showed discriminant validity ($\sqrt{\text{AVE}} >$ largest correlation). Analysis of the CFA suggested we had a great model fit: $\chi 2$ (395) = 628.807, p $<$ 0.001; and our robust RMSEA was not greater than .05, as the result of the P-value RMSEA <= 0.05 test was = 0.001. Finally, both our CFI and TLI scores were above satisfactory levels, CFI = 0.978, TLI = 0.975 [38].

4 Results

4.1 IAT Data

We constructed and evaluated the IAT scores using *iatgen.org* [12]. As a reminder an IAT or implicit association test gives a numerical value that illustrates unconscious associations between two groups of concepts. In our study these concepts were men/women vs same/other. A positive score noted an unconscious association of women-same:men-other and an negative score noted an unconscious association of men-same:women-other. The larger the value the stronger the association. Using a one-way ANalysis Of VAriance (ANOVA) we found that pre-exposure (before using a Snapchat filter) female participants had significantly more positive IAT scores than male participants: F(1,106) = 79.74, $p < 0.001$, $\eta 2 = 0.43$) ie female participants more implicitly associated women with similarity and men to otherness (M = 0.47, SD = 0.27). Whereas and male participants associated men with similarity and women with otherness (M = -0.09, SD = 0.38), Figure 3. We computed the change between the pre and post IATs (ΔIAT) by taking the post-interaction with a gender filter IAT value and subtracting the pre-interaction with a gender filter IAT value, as done in past IAT studies [6, 42, 48]. This is because by using (ΔIAT) we can legitimately attribute any differences in IAT scores between the filter groups to interaction with the filter and not the unique characteristics of the sample groups.

We used a three-way linear regression model with between–subject factors; Filter Type (Girl, Guy, None), Participant Gender (Male, Female), and Level of Embodied Identification. The last factor (Level of Embodied Identification) was obtained by turning the "identification via embodied presence" survey answers (from hereon, "embodied identification") into a weighted sum score to allow for interactions with the independent variables. That is all answers to the 6 questions were added together and weighted by subtracting the overall average score. So a person with a positive score was showing an above average level of identification in relation to their peers. We found neither interacting with the girl filter (SE = 0.02, t = 0.09, $p = 0.93$) nor the guy filter (SE = 0.02, t = -0.36, $p = 0.72$) produced a significant interaction. Further analysis found regardless of the participants' gender, filter used, or level of identification ΔIAT scores were not significantly impacted. This is illustrated in Figure 4 which shows IAT scores compared to averaged identification weighted sum scores. Meaning the sum of all six questions was subtracted by the average sum identification score and divided by 6. This converted embodied identification to

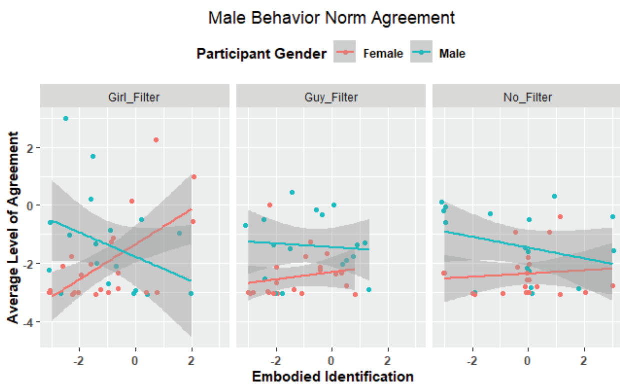

Figure 7: Participants' level of agreement with norms about male behavior for each group; no filter (baseline), MyTwin guy filter, and MyTwin girl filter, moderated by embodied identification.

Condition	Mean	SD
No Filter : Female Participant : Low-Ave	-2.733	0.365
No Filter : Female Participant : +Ave	-2.236	0.855
No Filter : Male Participant : Low-Ave	-0.667	1.167
No Filter : Male Participant : +Ave	-1.750	1.130
Girl Filter : Female Participant : Low-Ave	-2.489	0.685
Girl Filter : Female Participant : +Ave	-1.056	2.040
Girl Filter : Male Participant : Low-Ave	-0.798	1.837
Girl Filter : Male Participant : +Ave	-2.206	1.074
Guy Filter : Female Participant : Low-Ave	-2.603	0.841
Guy Filter : Female Participant : +Ave	-2.208	0.568
Guy Filter : Male Participant : Low-Ave	-1.500	1.267
Guy Filter : Male Participant : +Ave	-1.296	0.977

Table 4: Average Male Behavior Norm Agreement split by low to average, and above average level of identification.

a 7 point range to reflect the 7-point scale the items used, where -3 = Extremely Low Identification, 0 = Neutral, and 3 = Extremely High Identification. This weighted sum average representation of identification is used in all identification comparison graphs.

4.2 Androcentrism Data

For the androcentrism study we examined if the gender of the character participants created was male. This showed if participants were more likely to think of an average person as being a man. Meaning female, non-binary, or non-disclosed genders were group together. However, only 3% (4 of the 119) participants listed the character's gender as something other than male or female. We found evidence of androcentrism as 68 of 119 participants created a male "typical person" or roughly 57%. Using a one-way ANOVA we discovered that the participant's gender significantly impacted this prediction, F(1,117) = 22.32, $p < 0.001$, $\eta 2 = 0.16$. Male participants (M = 0.79, SD = 0.41) were far more likely to create a male character than female participants (M = 0.39, SD = 0.49), Figure 5. As with ΔIAT a three-way linear regression model with between–subject factors Filter Type (Girl, Guy, None), Participant Gender (Male, Female), and Level of Embodied Identification was used to predict androcentrism. We found that regardless of the filter; girl (SE = 0.02, t = -0.77, $p = 0.44$) or guy (SE = 0.02, t = -0.13, $p = 0.90$), there were no significant interactions, Figure 6.

4.3 Gendered Norms Data

To examine level of explicit bias via agreement with various gendered norms we used a Structural Equation Model (SEM) instead of a traditional t-test to keep the independent variables as a standardized latent factor. A three-way interaction between the Participant's Gender, the Filter Group assigned, and the Level of Embodied Identification reported was used to predict agreement with the SEM factors regarding stereotypes about men and women, Figure 3. This three-way interaction was not found to significantly predict norm agreement levels for three of the gendered norms; Male and Female Appearance norms, and the Female Behavior norms. However, Male Behavior norm agreement could be significantly predicted by the three-way interaction (SE = 0.139, z = 0.060, $p = 0.020$). This was validated by an omnibus test, $\chi 2(2) = 6.2651874$, $p = 0.044$. Figure 7 illustrates the differences in levels of agreement with stereotypes on Male Behavior across the different gender and filter groups based on reported level of embodied identification with oneself while using the filter. Thus verifying our hypotheses H1c and H2c as high levels of identification with a unique filter did predict different levels of bias for men and women. Male participants who identified with the female filter were predicted to display less gender bias. Male participants reporting high levels of embodied identification (EI*18, with 18 as the max level of sum identification on our scale 6 questions * +3 points) with the female filter were predicted to be 1.405 standard deviations lower in male behavior norm agreement than their base male peers (non-filter, average identifying). So they were more likely to disagree with stereotypes that policed how men and young boys behaved. Table 6 in the appendix further breaks down the SEM statistical results for the male behavior norm.

However, lower bias was only predicted for men using the girl filter. In contrast, women who identified with the girl filter were predicted to have the highest levels of explicit bias against men that did not follow male behavior stereotypes. Female participants with high levels of embodied identification (EI*18) with the female filter were predicted to be 2.192 standard deviations higher in their male behavior norm agreement rating than their base (non-filter average identifying) female peers. So they were more likely to agree with stereotypes that policed how men behaved. To put this in perspective when identification was split into two groups (1: participants reporting average levels of identification or less, and 2: participants reporting above average levels of identification) above average female girl-filter users' average level of agreement (M = -1.056, SD = 2.04) was 1.433 points higher on a seven point scale than average-below female girl-filter users (M = -2.489, SD = 0.685). Meanwhile, above average male girl-filter users' average level of agreement (M = -2.206, SD = 1.074) was 1.408 points lower than average-below male girl-filter users (M = -0.798, SD = 1.837). Table 4 breaks down the results, most groups somewhat disagreed with the stereotypes M < 0 but remember a single 1 point change can be a jump from a Disagree to a Somewhat Disagree. While this three way interaction was only significant for the male behavior norms the same trends appeared across all explicit bias tests.

It seems identification had an overall impact on all the predicted gender norm agreement results. For the three remaining norms Identification x Participant Gender was a significant two-way interaction predictor for agreement. However as before, when predicting how much participants would agree with stereotypes policing how men and women should look and how women should act, experiencing identification with oneself while using Snapchat impacted the predictions differently based on one's identified gender group. Female participants who identified with themselves displayed higher levels of bias by having higher agreement scores with every norm. Meanwhile and male participants displayed lower levels of bias as observed from their lower overall agreement scores. Female participants with high levels of identification displayed 0.918 standard deviations greater male appearance norm agreement (SE = 0.025, z

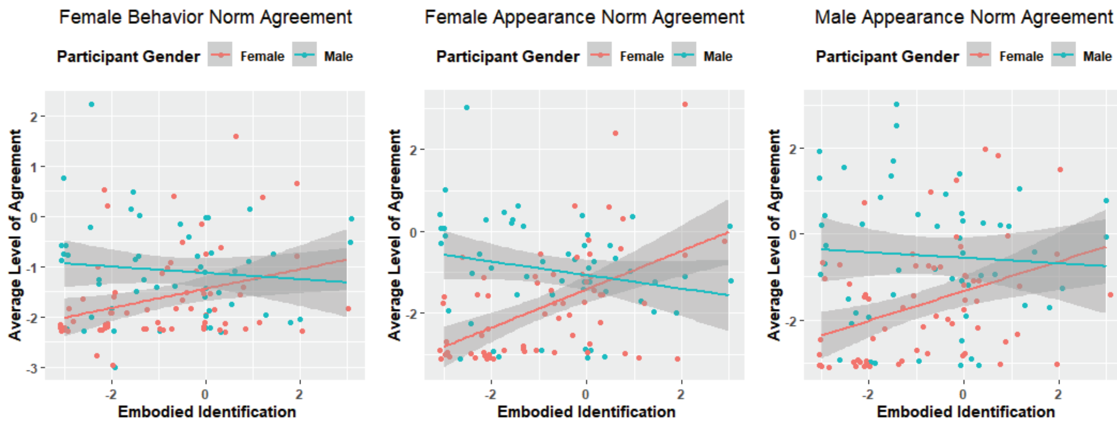

Figure 8: [Participants' level of agreement with stereotypical norms regarding female behavior and male and female appearance by participants' gender group, moderated by embodied identification.

Condition	Mean	SD
Female Behavior Norm Agreement		
Female Participant : Low-Ave	-1.853	0.732
Female Participant : +Ave	-1.344	1.007
Male Participant : Low-Ave	-0.980	1.152
Male Participant : +Ave	-1.174	0.778
Female Appearance Norm Agreement		
Female Participant : Low-Ave	-2.412	0.808
Female Participant : +Ave	-1.198	1.471
Male Participant : Low-Ave	-0.653	1.480
Male Participant : +Ave	-1.226	0.965
Male Appearance Norm Agreement		
Female Participant : Low-Ave	-2.140	1.027
Female Participant : +Ave	-1.086	1.409
Male Participant : Low-Ave	-0.320	1.826
Male Participant : +Ave	-0.688	1.272

Table 5: Each filter and gender groups' average amount of Female Behavior, Female Appearance, and Male Appearance Norm Agreement for low-average and above average level of identification groups.

= 2.426, $p=0.015$), 1.494 standard deviations greater female appearance norm agreement (SE = 0.028, z = 3.679, $p<0.001$), and 0.81 standard deviations greater female behavior norm agreement (SE = 0.025, z = 2.157, $p=0.031$) than female participants who reported a base level of identification with their filters. Male participants with high levels of identification displayed 0.162 standard deviations less male appearance norm agreement, 0.378 standard deviations less female appearance norm agreement, and 0.144 standard deviations less female behavior norm agreement than their base peers, see Table 7 in the appendix. As seen, an above average identifying female girl-filter users' average level of agreement with female behavior norms (M = -1.344, SD = 1.007) was 0.509 points higher than average-below identifying female girl-filter users (M = -1.853, SD = 0.732). They were 1.214 points higher (M = -1.198, SD = 1.471) than average-below users identifying users (M = -2.412, SD = 0.808) for female appearance norms. Finally, they were 1.054 points higher (M = -1.086, SD = 1.409) than average-below identifying users (M = -2.140, SD = 1.027) for male appearance norms. For male users level of agreement with female behavior norms for above average identification users (M = -1.174, SD = 0.778) was .194 points lower than average-below identifying users (M = -0.980, SD = 1.152). The predicted level of agreement was 0.573 points lower between high (M = -1.226, SD = 0.965) and average to low identifying users (M = -0.653, SD = 1.480) for female appearance norms. Lastly, predicted agreement for male appearance norms was 0.368 points lower between above average identifying (M = -0.688, SD = 1.272) and average-below identifying (M = -0.320, SD = 1.826) users.

5 Discussion

Our study examined the relationship between embodied identification with different gender based AR manipulations of oneself and a person's perceived gender biases. We looked at biases a person may have about both women and men that were both implicit and explicit. The key finding of our research was that level of identification was a significant predictor when if came to explicit gender biases. Proving our hypothesis that there is some sort of connection between one's ability to identify with an AR gender manipulated self and one's probability of holding biased beliefs. Higher level's of identification with an AR manipulate self along with a participant's gender and sometimes the type of filter they used correlated with both higher stereotype agreement scores for women and lower scores for men. This was not always reliant on which filter was applied nor if a filter was applied at all. Though for predicting male behavior norm agreement a significant three way interaction between Participant Gender, Filter Type, and level of Embodied Identification was observed. This three way interaction was not significant for the other there norms but the trends in bias predictions were nearly identical. This might suggest that for women connecting with oneself on Snapchat correlates with greater amounts of biased thinking. Whether this is because identification increases bias or that women who are more able to connect with themselves on the app are more likely to agree with gender biases we can't say.

Whatever the case this seems to be flipped for men who largely showed reduced levels of perceived bias when reporting high levels of identification. This does reflect past research which suggests for many simply using filters can be seen as a feminine act [39]. One study by Lavrence and Cambre found that some men perceived using a filter as "unspeakably feminine" and "emasculating" with boys reporting being policed by their peers and discouraged from the using them all together [39]. So male participants that are willing to engage and identify with AR filtered versions of themselves, especially female versions of themselves, may be more likely to reject stereotypes about what is appropriate for men and women as they are engaging in something seen as feminine. Meanwhile women identifying with filters are engaging in acts that are seen as feminine and thus promoting ideas about what is appropriate for their gender group. While, we can not say if this connection is corollary or causal the existence of this connection opens up all new possibilities for AR developers and researchers to understand and manipulate societal biases. If this relationship is causal this would mean promoting identification could be useful for decreasing bias in the male population but promoting dis-identification would be the goal for the female population. Unfortunately, this goes against both the design goals of most filters designed for women [50]and what users seek when using them [57]. We used pre-existing AR face filters from Snapchat to allow our findings to be applicable to the effects of real world use of the technology. Our findings suggest current filters could at best be used to identify groups with biases perspectives and at worst be actively promoting the adoption of these ideas. Further analysis on the nature of the relationship and if it can also be causal as with VR identification would be needed.

The best lesson developers can take from our work is that there is likely no one perfect intervention. As mentioned our analysis revealed that identification with an AR manipulated self can predict a user's explicit biases. However, these predictions were not homogeneous. Female participants who identified with themselves using the girl MyTwin filter showed the highest levels of explicit bias regarding male behavior norms while male participants showed the lowest. While the filter type was not significant for the other explicit norms the identification mismatch remained. High identification predicted greater bias levels for women but lower levels for men. Plus, identification only significantly predicted explicit biases. Meanwhile the filter group a participant was in was only significant for the explicit male behavior norm prediction. So, the same interaction with one filter that correlates with lower levels of various biases amongst men may correlate with higher levels of those same biases amongst woman. Furthermore, none of the filters had any significant relationship with any of the implicit biases regardless of identification level. So if an interaction could have a causal impact on agreement with stereotypes that would not automatically mean there would be any change to implicit biases like unconscious preferences for a gender group or making gendered assumptions about a person's identity. Both of which can still have significant impacts on people's lives as explained by Ah et al. on how the "male as default" assumption in medicine has had adverse affects on women's health [2]. However, with VR past studies suggested identification could impact implicit bias [6, 7, 30, 34, 42]. There are multiple differences between those studies and ours including interaction time, the use of full body avatars, and inclusion of more mentally involved tasks. Perhaps future studies could examine each of these areas to discern if AR has similar implicit bias manipulating abilities like VR or if there are unique contexts for each technology. Either way, it does not seem that a developer could create one application that could address all of these biases and be universally applied to every population and produce the same results. Instead of one universal bias intervention developers will likely need to create unique experiences based on the gender group and bias they want to target.

6 Limitations and Future Work

There are several limitations to our work. For one, while the gender swap app we used was marketed and accepted by the wider online community as providing an "authentic" change in gender [33] it does make use of a generalized and stereotypical perception of gender. It does not consider how gender is expressed differently among various groups and cultures [1, 52, 56]. For example, many African women have short or shaved hairstyles which would be seen as masculine on European women. Our bias indicator questions may also reflect a WEIRD (Western/White, Educated, Industrialised, Rich, Democratic) bias. We also treated all the biases and the participants' results as neutral, despite the fact that historically different societies have pushed for women to be extra aware of "proper" norms for both genders [1, 17]. It is possible the extreme differences in opinions women expressed compared to men are in part a result of their different upbringings in their cultures. In fact, work by Obremski et al. which explored racial bias showed that when a multicultural lens was applied to racial bias research it significantly impacted participants perceived levels of explicit racial bias [45].

We did strive to recruit participants from a wide range of cultural backgrounds and we made the study internationally available to insure we were not only looking at one population. The apps only reflected one extremely limited idea of how a person from the two gender groups should look. In fact, they excluded or changed the features of many individuals assigned to the gender group they identified as, such as removing some women's squarer jaws. This algorithmic stereotyping is harmful as it promotes a hierarchy of gender representation, placing some individuals as less manly/womanly than others and does not account for gender groups outside of the male/female binary. However, all these limitations do reflect the ongoing reality of how people react when made to conform to a society's stereotypical outlook on gender. After all, the apps were designed by a global billion-dollar corporation and were promoted and shared by users because of how "well" they transformed the diverse faces of their user base into the genders they claimed to represent. The apps reflect the imperfect and limited view of gender that societal standards place on people and allow us to research how engagement with those standards enforces, combats, and/or predicts stereotypical thinking. Using our study as a starting point, further research could investigate culturally specific filters and norms.

7 Conclusion

Our study starts conversation on if identification with an AR emulation of a societal group related to attitudes about that group. Specifically, if there was a connection between a person's ability to identify with themselves after having their gender presentation manipulated in AR and said person's perceived levels of implicit and/or explicit gender biases. We found that based on a participant's gender group, identifying with an AR gendered face filter could correlate with either lower or higher levels of perceived explicit bias, such as one's amount of agreement with gendered stereotypes, though not implicit biases. These findings suggest many difficult realities for researchers studying AR and societal biases and for developers hoping to use AR to influence societal biases. The biggest takeaway from our research being developers should not seek to create a universal intervention for bias, but instead separate interventions based on the group targeted. Our findings suggest it may be beneficial to encourage more men to use and identify with themselves on Snapchat. However for women it is better to encourage the rejection of extremely feminized versions of themselves as for them identification predicted more bias, not less.

8 Open Access to Research Data

In accordance with the Horizon 2020 Pilot on Open Research Data free online access to all used data from this study can be accessed via the following link: (https://cloud.imt-atlantique.fr/index.php/s/cBxEfbcZsApgKMR). Access will also be granted on the lead researchers website: https://mariejarrell.wordpress.com/ . Interested individuals can reach out to the lead researcher at marie.jarrell@imt-atlantique.fr if they require any further aide.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and programme under the Marie Skłodowska-Curie grant agreement No 899546.

References

[1] J. Acker. The problem with patriarchy. *Sociology*, 23(2):235–240, 1989. 2, 8

[2] B. Ah, L. M, and D. Jf. Is man the measure of all things? a social cognitive account of androcentrism. *Pers Soc Psychol Rev*, 23(4):307–331, 2019. Publisher: Pers Soc Psychol Rev. doi: 10.1177/1088868318782848 4, 8

[3] A. O. Alkhamisi and M. M. Monowar. Rise of Augmented Reality: Current and Future Application Areas. *International Journal of Internet and Distributed Systems*, 1(4):25–34, Nov. 2013. Number: 4 Publisher: Scientific Research Publishing. doi: 10.4236/ijids.2013.14005 2

[4] R. M. Alsaggaf. The Impact of Snapchat Beautifying Filters on Beauty Standards and Self-image: A Self-Discrepancy Approach. In *The European Conference on Arts & Humanities 2021: Official Conference Proceedings*, pp. 47–58. ECAH2021, UCL, London, UK, Sept. 2021. doi: 10.22492/issn.2188-1111.2021.4 1, 2

[5] A. H. Bailey and M. LaFrance. Anonymously male: Social media avatar icons are implicitly male and resistant to change. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 10(4), Dec. 2016. Number: 4. doi: 10.5817/CP2016-4-8 2, 4

[6] D. Banakou, P. D. Hanumanthu, and M. Slater. 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, Nov. 2016. doi: 10.3389/fnhum.2016.00601 1, 3, 4, 5, 8

[7] D. Banakou, S. Kishore, and M. Slater. Virtually being einstein results in an improvement in cognitive task performance and a decrease in age bias. *Frontiers in psychology*, 9:917, 2018. Publisher: Frontiers. doi: 10.3389/fpsyg.2018.00917 1, 3, 8

[8] K. Bell. FaceApp removes racist selfie filters, Aug. 2017. Section: Tech. 1

[9] K. Beltran, C. Rowland, N. Hashemi, A. Nguyen, L. Harrison, S. Engle, and B. F. Yuksel. Reducing implicit gender bias using a virtual workplace environment. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, number 277 in CHI EA '21, pp. 1–7. Association for Computing Machinery, New York, NY, USA, May 2021. 1

[10] E. Bolt, J. T. Ho, M. R. Lesur, A. Soutschek, P. N. Tobler, and B. Lenggenhager. Effects of a Virtual Gender Swap on Social and Temporal Decision-Making. Technical report, PsyArXiv, Mar. 2021. type: article. doi: 10.31234/osf.io/84v9n 4

[11] J. Brucculieri. 'Snapchat Dysmorphia' Points To A Troubling New Trend In Plastic Surgery, Feb. 2018. 1

[12] T. P. Carpenter, R. Pogacar, C. Pullig, M. Kouril, S. Aguilar, J. LaBouff, N. Isenberg, and A. Chakroff. Survey-software implicit association tests: A methodological and empirical analysis. *Behavior research methods*, 51:2194–2208, 2019. 5

[13] M. Christofi and D. Michael-Grigoriou. Virtual reality for inducing empathy and reducing prejudice towards stigmatized groups: A survey, Oct. 2017. ISSN: 2474-1485. doi: 10.1109/VSMM.2017.8346252 1

[14] B. Dean. Snapchat Demographic Stats: How Many People Use Snapchat in 2022?, Jan. 2022. 1, 2

[15] S. Dixon. Snapchat: daily active users worldwide, May 2022. 1

[16] J. Eshiet. *"REAL ME VERSUS SOCIAL MEDIA ME:" FILTERS, SNAPCHAT DYSMORPHIA, AND BEAUTY PERCEPTIONS AMONG YOUNG WOMEN*. PhD thesis, San Bernardino, California State University, June 2020. 1

[17] S. M. Evans. Sons, daughters, and patriarchy: Gender and the 1968 generation. *The American Historical Review*, 114(2):331–347, 2009. 8

[18] X. Fang, S. Singh, and R. Ahluwalia. An Examination of Different Explanations for the Mere Exposure Effect. *Journal of Consumer Research*, 34(1):97–103, June 2007. doi: 10.1086/513050 1

[19] C. for Clinicial Psychology. Basic emotions | learn more, 2022. 3

[20] G. Freedman, M. C. Green, M. Seidman, and M. Flanagan. The Effect of Embodying a Woman Scientist in Virtual Reality on Men's Gender Biases. *Technology, Mind, and Behavior*, 2(4), Nov. 2021. doi: 10.1037/tmb0000046 3

[21] R. Fribourg, E. Peillard, and R. Mcdonnell. Mirror, Mirror on My Phone: Investigating Dimensions of Self-Face Perception Induced by Augmented Reality Filters. In IEEE, ed., *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 470–478. IEEE Computer Society, Bari, Italy, Oct. 2021. doi: 10.1109/ISMAR52148.2021.00064 2

[22] I. Fritsche, Eva Jonas, and Thomas Fankhänel. The role of control motivation in mortality salience effects on ingroup support and defense. *Journal of Personality and Social Psychology*, 95(3):524, Sept. 2008. doi: 10.1037/a0012666 3

[23] A. Gabbiadini, P. Riva, L. Andrighetto, C. Volpato, and B. J. Bushman. Acting like a Tough Guy: Violent-Sexist Video Games, Identification with Game Characters, Masculine Beliefs, & Empathy for Female Violence Victims. *PLOS ONE*, 11(4):e0152121, Apr. 2016. doi: 10.1371/journal.pone.0152121 1

[24] N. L. Galambos, A. C. Petersen, M. Richards, and I. B. Gitelson. The Attitudes Toward Women Scale for Adolescents (AWSA): A study of reliability and validity. *Sex Roles*, 13(5):343–356, Sept. 1985. doi: 10.1007/BF00288090 2, 4

[25] N. L. Galambos, A. C. Petersen, M. Richards, and I. B. Gitelson. The attitudes toward women scale for adolescents (awsa): A study of reliability and validity. *Sex roles*, 13(5-6):343–356, 1985. 4

[26] A. C. S. Genay, A. Lecuyer, and M. Hachet. Being an Avatar "for Real": a Survey on Virtual Embodiment in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, Early Access:1–1, 2021. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi: 10.1109/TVCG.2021.3099290 2

[27] T. Goetz. Swapping Gender is a Snap(chat): Limitations of (Trans) Gendered Legibility within Binary Digital and Human Filters. *Catalyst: Feminism, Theory, Technoscience*, 7(2), Oct. 2021. Number: 2. doi: 10.28968/cftt.v7i2.34839 1

[28] C. Gonzalez-Liencres, L. E. Zapata, G. Iruretagoyena, S. Seinfeld, L. Perez-Mendez, J. Arroyo-Palacios, D. Borland, M. Slater, and M. V. Sanchez-Vives. Being the victim of intimate partner violence in virtual reality: First- versus third-person perspective. *Frontiers in psychology*, 11:820, 2020. Publisher: Frontiers. doi: 10.3389/fpsyg.2020.00820 1, 3

[29] A. G. Greenwald, D. E. McGhee, and J. L. K. Schwartz. Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6):1464–1480, 1998. Place: US Publisher: American Psychological Association. doi: 10.1037/0022-3514.74.6.1464 2

[30] V. Groom, J. N. Bailenson, and C. Nass. The influence of racial embodiment on racial bias in immersive virtual environments. *Social Influence*, 4(3):231–248, July 2009. doi: 10.1080/15534510802643750 1, 3, 8

[31] S. Gächter, C. Starmer, and F. Tufano. Measuring the closeness of relationships: A comprehensive evaluation of the 'inclusion of the other in the self' scale. *PLOS ONE*, 10:e0129478, June 2015. doi: 10.1371/journal.pone.0129478 3

[32] A. Haines. From 'Instagram Face' To 'Snapchat Dysmorphia': How Beauty Filters Are Changing The Way We See Ourselves, Apr. 2021. Section: Style & Beauty. 2

[33] C. Harbison. Snapchat's New Gender Swap Filter Will Make You Question Your Identity: How to Get the Male to Female Filter, May 2019. 2, 8

[34] B. S. Hasler, B. Spanlang, and M. Slater. Virtual race transformation reverses racial in-group bias. *PLOS ONE*, 12(4):e0174965, Apr. 2017. doi: 10.1371/journal.pone.0174965 3, 8

[35] P. Implicit. Select a test, 2011. 3

[36] S. Kishore, B. Spanlang, G. Iruretagoyena, S. Halan, D. Szostak, and M. Slater. A virtual reality embodiment technique to enhance helping behavior of police toward a victim of police racial aggression. *PRESENCE: Virtual and Augmented Reality*, 28:5–27, Oct. 2021. doi: 10.1162/pres_a_00339 1

[37] C. Klimmt, D. Hefner, P. Vorderer, C. Roth, and C. Blake. Identification With Video Game Characters as Automatic Shift of Self-Perceptions. *Media Psychology*, 13(4):323–338, Nov. 2010. doi: 10.1080/15213269.2010.524911 1

[38] R. B. Kline. *Principles and practice of structural equation modeling*. Guilford publications, New York, NY, USA, 2015. 5

[39] C. Lavrence and C. Cambre. "Do I Look Like My Selfie?": Filters and the Digital-Forensic Gaze. *Social Media + Society*,

6(4):2056305120955182, Oct. 2020. Publisher: SAGE Publications Ltd. doi: 10.1177/2056305120955182 7

[40] R. F. Levant, T. J. Rankin, C. M. Williams, N. T. Hasan, and K. B. Smalley. Evaluation of the factor structure and construct validity of scores on the Male Role Norms Inventory—Revised (MRNI-R). *Psychology of Men & Masculinity*, 11(1):25–37, 2010. Place: US Publisher: Educational Publishing Foundation. doi: 10.1037/a0017637 2, 4

[41] J. V. Looy, C. Courtois, M. D. Vocht, and L. D. Marez. Player Identification in Online Games: Validation of a Scale for Measuring Identification in MMOGs. *Media Psychology*, 15(2):197–221, May 2012. doi: 10.1080/15213269.2012.674917 3, 4

[42] S. Lopez, Y. Yang, K. Beltran, S. J. Kim, J. Cruz Hernandez, C. Simran, B. Yang, and B. F. Yuksel. Investigating Implicit Gender Bias and Embodiment of White Males in Virtual Reality with Full Body Visuomotor Synchrony. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. ACM, Glasgow Scotland Uk, May 2019. doi: 10.1145/3290605.3300787 1, 3, 5, 8

[43] J. R. Mahalik, E. B. Morray, A. Coonerty-Femiano, L. H. Ludlow, S. M. Slattery, and A. Smiler. Development of the Conformity to Feminine Norms Inventory. *Sex Roles*, 52(7):417–435, Apr. 2005. doi: 10.1007/s11199-005-3709-7 2, 4

[44] L. Merino, M. Schwarzl, M. Kraus, M. Sedlmair, D. Schmalstieg, and D. Weiskopf. Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009-2019). In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 438–451. IEEE, Porto de Galinhas, Brazil, Nov. 2020. ISSN: 1554-7868. doi: 10.1109/ISMAR50242.2020.00069 1, 2

[45] D. Obremski, O. B. Akuffo, L. Lücke, M. Semineth, S. Tomiczek, H.-F. Weichert, and B. Lugrin. Reducing racial bias by interacting with virtual agents: An intervention in virtual reality. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, pp. 747–755, 2023. 8

[46] S. Y. Oh, J. Bailenson, E. Weisz, and J. Zaki. Virtually old: Embodied perspective taking and the reduction of ageism under threat. *Computers in Human Behavior*, 60:398–410, 2016. doi: 10.1016/j.chb.2016.02.007 1

[47] T. C. Peck, M. Doan, K. A. Bourne, and J. J. Good. The Effect of Gender Body-Swap Illusions on Working Memory and Stereotype Threat. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1604–1612, Apr. 2018. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi: 10.1109/TVCG.2018.2793598 4

[48] T. C. Peck, S. Seinfeld, S. M. Aglioti, and M. Slater. Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, 22(3):779–787, Sept. 2013. doi: 10.1016/j.concog.2013.04.016 1, 5

[49] K. Ramphul and S. G. Mejias. Is "Snapchat Dysmorphia" a Real Issue? *Cureus*, 10(3):e2263, Mar. 2018. doi: 10.7759/cureus.2263 1

[50] H. Rex. the augmented reality artists behind your favourite instagram filters, May 2019. 1, 8

[51] S. Schulze, T. Pence, N. Irvine, and C. Guinn. The Effects of Embodiment in Virtual Reality on Implicit Gender Bias. In J. Y. Chen and G. Fragomeni, eds., *Virtual, Augmented and Mixed Reality. Multimodal Interaction*, Lecture Notes in Computer Science, pp. 361–374. Springer International Publishing, Cham, 2019. doi: 10.1007/978-3-030-21607-8_28 3

[52] J. W. Scott. Gender: A useful category of historical analysis. *The American historical review*, 91(5):1053–1075, 1986. 2, 8

[53] S. Sonoma. Snapchat's gender-swapping photo filters can result in real transphobia. 1

[54] J. T. Spence, R. Helmreich, and J. Stapp. A short version of the Attitudes toward Women Scale (AWS). *Bulletin of the Psychonomic Society*, 2(4):219–220, Oct. 1973. doi: 10.3758/BF03329252 2, 4

[55] J. Van Looy, C. Courtois, M. De Vocht, and L. De Marez. Player Identification in Online Games: Validation of a Scale for Measuring Identification in MMOGs. *Media Psychology*, 15(2):197–221, May 2012. Publisher: Routledge _eprint: https://doi.org/10.1080/15213269.2012.674917. doi: 10.1080/15213269.2012.674917 4

[56] J. Wallach Scott. Gender: still a useful category of analysis? *Diogenes*, 57(1):7–14, 2010. 2, 8

[57] A. Youn. What Is the Ideal Instagram Filter? *Aesthetic Surgery Journal Open Forum*, 1(2), Apr. 2019. doi: 10.1093/asjof/ojz019 8

[58] J. L. Zaichkowsky. The personal involvement inventory: Reduction, revision, and application to advertising. *Journal of advertising*, 23(4):59–70, 1994. 4