Research Article **ISSN 2639-944X**

Journal of Medical - Clinical Research & Reviews

Effects of Facial Photo Filters on Judgments of Attractiveness and Trustworthiness: An Online Experiment with AI-Generated Faces

Darius Krutzek*

Hochschule Fresenius University of Applied Sciences, Frankfurt am Main, Germany.

***Correspondence:**
Darius Krutzek, Hochschule Fresenius University of Applied Sciences, Frankfurt am Main, Germany.

Received: 11 Sep 2026; **Accepted:** 24 Oct 2026; **Published:** 01 Nov 2026

Citation: Darius Krutzek. Effects of Facial Photo Filters on Judgments of Attractiveness and Trustworthiness: An Online Experiment with AI-Generated Faces. J Med - Clin Res & Rev. 2026; 10(11): 1-8.

ABSTRACT

Facial photo filters are widely used when sharing selfies, but how others perceive filtered faces has rarely been examined experimentally. In this online experiment, I examined whether a typical facial photo filter from Instagram affects judgments of the attractiveness and trustworthiness of female faces. To rule out familiarity effects, 40 AI-generated (StyleGAN) faces of women (39 analyzed) were presented either unfiltered or modified with a filter that combines soft-focus, contrast, and exposure adjustments. Participants (N = 69, 39 women and 30 men, aged 18–63 years) rated each face on 10-point scales. The four questionnaire versions were randomly allocated to recruitment channels rather than to individual participants, so condition was confounded with viewer gender (filter: 28 men and 6 women, no filter: 33 women and 2 men). Filtered faces received slightly lower mean ratings of attractiveness (Hedges' g = −0.32, 95% CI [−0.79, 0.15]) and trustworthiness (g = −0.32, 95% CI [−0.79, 0.15]), but neither difference was statistically significant. Mixed-effects models with crossed random effects for participants and faces led to the same conclusion. Adjusting for viewer gender reduced the attractiveness difference to nearly zero but changed the trustworthiness difference only slightly. Gender differences could not be tested. About half of the women and a quarter of the men reported editing photos of themselves. The extent of editing was not significantly correlated with age or education. The findings do not support the assumption that such filters lower these judgments. Adequately powered studies with randomization at the participant level are needed.

Keywords

Facial photo filters, Instagram, facial attractiveness, trustworthiness, AI-generated faces.

Introduction

The use of photo filters is a widespread trend on social media platforms such as Instagram and Snapchat [1]. Photo filters are software functions that change the appearance of a photograph before it is shared, for example by altering exposure, color, or contrast. Filters designed specifically for editing faces are particularly popular. These filters, commonly called beauty or makeup filters and referred to in this article as *facial photo filters*, typically combine soft-focus and smoothing effects with changes in brightness and contrast. With a single tap, they can lighten or darken the skin tone and individual facial areas and remove visible skin blemishes. Ninety percent of young women report using filters before posting photographs of their own face, so-called selfies [2].

Photo manipulation itself is not new. Professional image editing software such as Photoshop has existed since the late 1980s [3], but its complexity and price limited its use to a small group of users. Smartphone apps such as Facetune, VSCO, and FaceApp and filters built into social media platforms have since made the editing of faces available to everyone, in some cases with AI-based changes to facial features that are intended to match Western beauty ideals [4]. This development has raised clinical and societal concerns. The term "Snapchat dysmorphia", for example, refers to the wish to look like one's filtered selfies, which may lead people to seek cosmetic surgery [1] and has been discussed in relation to body dysmorphic disorder [5]. In Norway, influencers have been required to label retouched bodies and faces in advertising photographs since summer 2022 [6].

Attractiveness and Trustworthiness of Selfies

In 2022, Instagram had more than 1 billion users worldwide,

about 70% of whom were between 13 and 34 years old [7]. Young people share photos on Instagram to preserve memories and share them with friends and family, but also to present and develop their own identity [8] and posting selfies has been linked to impression management and self-esteem [9]. Likes act as a social reward. Social media engagement follows the principles of reward learning [10] and in adolescents, photos with many likes elicit stronger activation of the nucleus accumbens than photos with few likes [1]. Day and Carelli [12] and Klawonn and Malenka [13] describe this region as a key structure of the reward system. In the study by Sherman et al. [11], the effect was found in both hemispheres when participants viewed their own photos. Because attractive selfies are more likely to be liked, users have a strong incentive to optimize their appearance in photos, and facial photo filters are a convenient tool for doing so.

Selfies also play a central role in online dating. On Tinder, which had about 75 million monthly active users in 2022 [14] and on similar apps, users decide whether they are interested in another person mainly on the basis of profile photos, supplemented by only a few social cues [15]. Such decisions are often made in less than a second [16], so physical appearance plays a decisive role. At the same time, filtered photos may raise doubts about a person's honesty. In focus groups conducted by Lavrence and Cambre [17], young adults, and men in particular, reacted negatively to conspicuously filtered selfies. They suspected that the depicted person looked quite different in reality or had something to hide, whereas subtle retouching of blemishes was considered acceptable. Consistent with this, Hong et al. [18] found in a content analysis that Instagram selfies with filters received fewer likes than selfies without filters.

These findings suggest that facial photo filters may lower rather than raise how attractive and trustworthy a person appears. At the time the present study was conducted, however, experimental evidence on this question was lacking. It was also unclear whether women and men judge filtered faces differently. Young women use filters far more often [2] than adults in general [19], yet women and men in the study by Lavrence and Cambre [17] evaluated conspicuous filtering in a similarly negative way.

The Present Study

In the present experiment, I examined how a typical facial photo filter available on Instagram affects judgments of the attractiveness and trustworthiness of female faces. AI-generated faces were used to rule out familiarity with the depicted persons. Based on the qualitative findings described above, filtered faces were expected to be judged as less attractive and less trustworthy than unfiltered faces, and the effect on attractiveness was expected to occur among male but not among female viewers. These predictions were specified before data collection. In addition, I explored how participants' own editing of photos was related to their age and education.

Method

Design

The online experiment was implemented in SoSci Survey [20] and used a between-subjects design with two conditions (filter vs. no filter). In both conditions, participants were to rate the same 40 faces, either unfiltered or filtered (for deviations, see Stimuli). Each condition was implemented in two questionnaire versions that presented the faces in a randomized base sequence or in the reversed sequence to counteract serial position effects. The four questionnaire versions were randomly allocated to four recruitment channels, and each channel received only the link to its version. Randomization therefore took place at the level of recruitment channels rather than individual participants. With only two channels per condition, this procedure could not ensure that the groups were comparable in participant characteristics.

Participants

Participants were recruited between April 11 and May 5, 2022, via WhatsApp groups, personal contact networks, and social media. They were invited to take part in a psychological experiment on the "psychological effects of different facial presentations." Participation was voluntary. Participants had to be at least 18 years old, because the women shown in the photographs were estimated to be between 18 and 45 years old, and to have normal vision. No a priori power analysis was conducted. Based on feasibility, a sample of 50 to 70 participants was planned.

Of the 98 persons who consented and began the survey, 74 (75.5%) completed it. Most of those who discontinued did so during the rating phase. Data quality was screened with two indicators of meaningless responding [21]. Two participants were excluded because their relative speed index exceeded 2.0, which means that they completed the survey more than twice as fast as the typical participant. Three participants were excluded because they gave identical ratings to all 40 faces on at least one of the two scales. None of the remaining participants' mean ratings met the outlier criterion defined before data collection (values more than three interquartile ranges beyond the quartiles). The final sample comprised $N = 69$ participants (39 women and 30 men, none of whom selected the option "other") aged 18 to 63 years ($M = 28.8$, $SD = 12.9$, $Mdn = 23$). Their highest educational qualification was a general higher education entrance qualification (Abitur, 42.0%), an intermediate secondary school certificate (20.3%), a university of applied sciences entrance qualification (18.8%), a university degree (14.5%), a doctorate (1.4%), a lower secondary school certificate (1.4%), or none of these (1.4%). Most participants used Instagram (79.7%) and Snapchat (58.0%), 40.6% used TikTok, and 20.3% used at least one dating app.

Because randomization took place at the level of recruitment channels, the two conditions differed markedly in their composition (see Table 1). The filter condition ($n = 34$) consisted predominantly of men (28 men, 6 women), whereas the no-filter condition ($n = 35$) consisted predominantly of women (33 women, 2 men), Fisher's exact test, $p < .001$. Participants in the filter condition were also older ($M = 34.2$, $SD = 15.9$) than those in the no-filter condition (M

= 23.5, *SD* = 5.4), $t(40.22) = 3.71$, $p < .001$.

Table 1: Descriptive Statistics by Condition and Viewer Gender.

		Age (years)		Attractiveness		Trustworthiness	
Condition and gender	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*
No filter							
Women	33	23.3	5.4	5.52	1.27	5.78	1.11
Men	2	26.5	4.9	6.55	2.12	6.90	1.45
Total	35	23.5	5.4	5.58	1.31	5.84	1.13
Filter							
Women	6	27.0	17.7	5.97	1.17	5.90	0.95
Men	28	35.7	15.4	5.00	1.16	5.36	1.34
Total	34	34.2	15.9	5.17	1.20	5.45	1.28
Total sample	69	28.8	12.9	5.38	1.26	5.65	1.22

Note. Attractiveness and trustworthiness are participants' mean ratings of 39 AI-generated female faces on scales from 1 to 10, with higher values indicating higher perceived attractiveness or trustworthiness. Because the four questionnaire versions were randomly allocated to recruitment channels rather than to individual participants, condition and viewer gender were confounded.

Stimuli

The stimuli were 40 portraits of women generated with the generative adversarial network StyleGAN [22,23]. The portraits therefore did not depict real persons. Synthetic faces were used because photographs from sources such as stock photo archives may already be familiar to participants, and prior exposure increases liking [24]. By 2022, faces generated with StyleGAN could no longer be reliably distinguished from photographs of real people [25], as also illustrated by the website whichfaceisreal.com [26]. When generating the faces, only a White appearance and an apparent age of 18 to 45 years were specified, so the faces varied in other features such as hair color, eye color, and head posture. In each square image, the head occupied approximately 90% of the image area, and the eyes were always at the same position. Only images without typical StyleGAN artifacts, such as asymmetric earrings or eyeglasses, distorted teeth, or irregular backgrounds, were selected.

In the filter condition, the faces were modified with the facial photo filter *No Filter*, a typical facial photo filter that was available on Instagram as a one-click application [27]. The filter applies strong soft-focus blurring across the face, increases global contrast, and modifies exposure across the entire image. It was applied to all photographs in a standardized manner. Figure 1 shows an example.

A post hoc pixel-wise comparison of the images embedded in the questionnaires revealed two deviations. One image in the no-filter versions was the filtered version of a face that also appeared unfiltered in the same questionnaire, and one face in the filter versions had no unfiltered counterpart. These items were excluded, so all analyses are based on 39 faces that were presented unfiltered in the no-filter condition and filtered in the filter condition. The assignment of images to items was verified for the two versions with the base sequence and assumed for the versions with the reversed sequence. This assumption was supported for the reversed no-filter version, whose face-level mean attractiveness ratings correlated highly with those of the base version ($r = .94$), but less clearly for the reversed filter version ($r = .51$). The filtered images also had a slightly higher resolution (approximately 550 × 550 pixels) than the unfiltered images (480 × 480 pixels).



Figure 1: Example of a Face Presented Without and With the Facial Photo Filter.

Note. The face was generated with StyleGAN and does not depict a real person. Left: unfiltered version (no-filter condition). Right: version modified with the facial photo filter "No Filter" (filter condition).

Procedure and Measures

After information on the study and on data protection, participants gave their informed consent by clicking "Next." According to the screening questionnaire of the Ethics Committee of Hochschule Fresenius, the study involved no vulnerable groups, no deception, no collection of personally identifying data, and no risks beyond those of everyday life, so a full ethics application was not required. Participants then rated each face on two scales from 1 to 10. Attractiveness was assessed with the item "Please rate the attractiveness of the woman shown in the picture" (1 = *I do not like the face at all*, 10 = *I like the face very much*). Trustworthiness was assessed with the item "How trustworthy would you consider this otherwise unknown person to be if you saw the photo on her social media profile?" (1 = *I would not trust this person at all*, 10 = *I would trust this person completely*).

Afterward, participants reported their gender, age, highest educational qualification, and use of social media and dating apps. Finally, they indicated whether they edited photos of themselves (*yes* or *no*) and how strongly they did so before uploading them, on a scale from 1 to 5 without verbal labels. The second question was presented to all participants, regardless of their answer to the first. All materials were presented in German, and the wordings reported here are translations. The median completion time was 6.4 min (interquartile range: 5.2–8.2 min).

Data Analysis

Data were prepared in Python 3.11 (pandas) and analyzed in R 4.3.3 [28] with the packages lme4 1.1-35.1 [29], lmerTest 3.1-3 [30] and effectsize 0.8.6 [31]. For each participant, mean attractiveness and

mean trustworthiness ratings were computed across the 39 faces. Within conditions, the internal consistency of these mean scores was high (Cronbach's α = .95–.96 for attractiveness and .93–.96 for trustworthiness).

The two conditions were compared with Welch's *t* tests, which do not assume equal variances [32]. Although directional effects were expected, two-tailed *p* values are reported (α = .05), and one-tailed values are given for completeness. Effect sizes are reported as Hedges' *g* with 95% confidence intervals (CIs), following Lakens [33]. Because participants responded to a sample of faces, I also fitted linear mixed-effects models with crossed random effects for participants and faces to the individual ratings [34]. The models contained the fixed effect of condition (coded −0.5 = no filter, 0.5 = filter), random intercepts for participants and faces, and random slopes of condition for faces. This is the maximal random-effects structure justified by the design [35], because participants were nested within conditions and random participant slopes were therefore not possible. The models were estimated with restricted maximum likelihood. Degrees of freedom were approximated with the Satterthwaite method [30] and 95% CIs of the fixed effects were based on the corresponding *t* distribution. A standardized effect size *d* was computed by dividing the fixed effect by the square root of the summed variance components, with the variance of the face slopes weighted by 0.25, the squared contrast code [36].

Sensitivity analyses repeated the main analyses (a) including all 74 participants who completed the survey, (b) excluding the filter version with the reversed sequence, for which the assignment of images to items was least certain, (c) restricting the analyses to the two versions with the base sequence, and (d) adding viewer gender, and in a further model also age and sequence, as covariates to the mixed-effects models. Because each condition comprised only two recruitment channels, the clustering of participants within channels could not be modeled. Because condition and viewer gender were almost completely confounded, gender differences in the effect of the filter could not be tested. Gender-specific results are therefore reported descriptively and are supplemented by an exploratory comparison among women. Exploratory analyses examined participants' own photo editing. Gender differences in whether participants edited photos of themselves were tested with Fisher's exact test, and gender differences in the self-reported extent of editing with a Mann–Whitney *U* test. Associations of the extent of editing with age and education were tested with Spearman correlations, excluding the education option "none of these". The normality of the participant means was checked with Shapiro–Wilk tests, and the association between mean attractiveness and mean trustworthiness ratings was assessed with a Pearson correlation. With about 35 participants per condition, the Welch tests had a power of 80% to detect effects of $d \geq 0.68$ (two-tailed, α = .05).

Results

Descriptive Statistics

Table 1 presents the descriptive statistics by condition and viewer gender, and Figure 2 displays the participants' mean ratings. Mean ratings of attractiveness and trustworthiness were strongly correlated, $r = .56$, 95% CI [.38, .71]. The distributions of the participant means did not deviate markedly from normality (Shapiro–Wilk *p*s > .05).

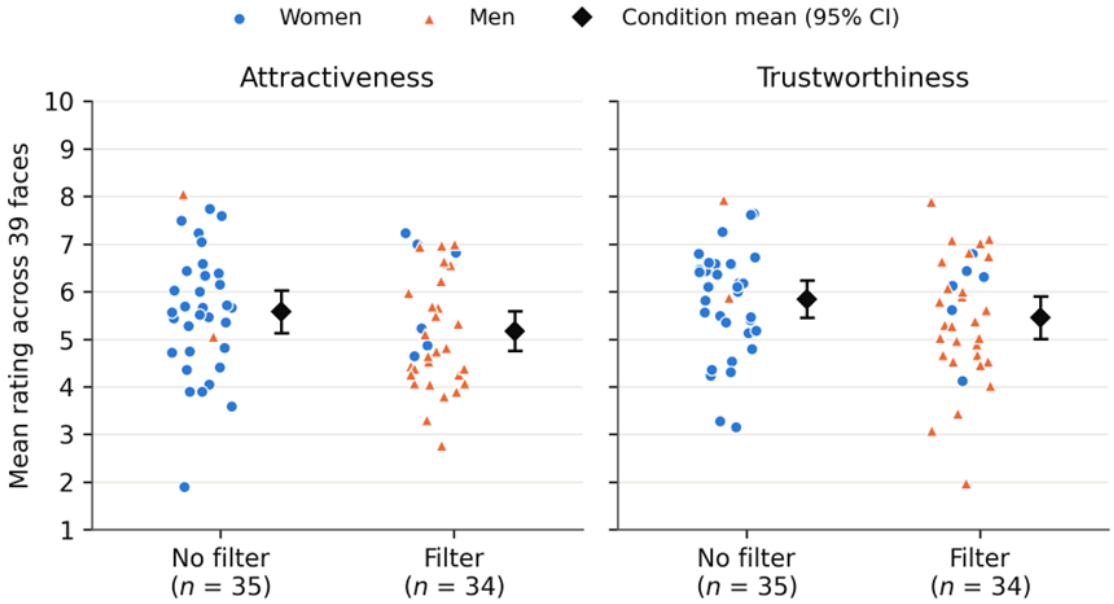


Figure 2: Participants' Mean Ratings of Attractiveness and Trustworthiness by Condition and Viewer Gender.
Note. Points represent individual participants' mean ratings across 39 faces (circles: women, triangles: men, horizontally jittered). Black diamonds and error bars represent condition means and 95% confidence intervals. The filter condition consisted of 28 men and 6 women, the no-filter condition of 33 women and 2 men.

Effects of the Filter on Attractiveness and Trustworthiness

Filtered faces were rated as slightly less attractive ($M = 5.17$, $SD = 1.20$) than unfiltered faces ($M = 5.58$, $SD = 1.31$). The difference of −0.41 points, 95% CI [−1.01, 0.19], was not statistically significant, $t(66.82) = -1.36$, $p = .180$ (one-tailed $p = .090$), $g = -0.32$, 95% CI [−0.79, 0.15]. Filtered faces were also rated as slightly less trustworthy ($M = 5.45$, $SD = 1.28$) than unfiltered faces ($M = 5.84$, $SD = 1.13$). This difference of −0.39 points, 95% CI [−0.97, 0.19], was not statistically significant either, $t(65.52) = -1.34$, $p = .186$ (one-tailed $p = .093$), $g = -0.32$, 95% CI [−0.79, 0.15].

The linear mixed-effects models led to the same conclusions (see Table 2). The effect of condition was $b = -0.41$, 95% CI [−1.04, 0.22], $t(77.93) = -1.29$, $p = .200$, $d = -0.17$, for attractiveness and $b = -0.39$, 95% CI [−1.00, 0.22], $t(78.18) = -1.27$, $p = .206$, $d = -0.18$, for trustworthiness. The models also showed that the faces differed considerably in how attractive they were judged to be (*SD* of face intercepts = 1.27) and that the effect of the filter varied across faces (*SD* of face slopes = 0.59 for attractiveness and 0.58 for trustworthiness). The expected negative effects of the filter were therefore not confirmed.

Sensitivity Analyses

Including all 74 participants who completed the survey yielded slightly larger but still nonsignificant differences for attractiveness, $t(70.20) = -1.75$, $p = .085$, $g = -0.40$, 95% CI [−0.86, 0.06], and trustworthiness, $t(69.84) = -1.66$, $p = .102$, $g = -0.38$, 95% CI [−0.84, 0.08]. The corresponding mixed models yielded $b = -0.57$, $t(81.89) = -1.68$, $p = .097$, for attractiveness and $b = -0.47$, $t(83.77) = -1.58$, $p = .118$, for trustworthiness. Excluding the filter version with the reversed sequence did not change the conclusions, with $b = -0.45$, $t(58.23) = -1.23$, $p = .224$, for attractiveness and $b = -0.34$, $t(62.95) = -1.05$, $p = .299$, for trustworthiness. Restricting the analyses to the two versions with the base sequence ($n = 36$) led to the same result, with $b = -0.49$, $t(36.86) = -1.15$, $p = .257$, for attractiveness and $b = -0.70$, $t(39.29) = -1.85$, $p = .072$, for trustworthiness. When viewer gender was added to the mixed-effects models, the effect of condition decreased to $b = -0.09$, 95% CI [−1.06, 0.88], $t(70.87) = -0.19$, $p = .850$, for attractiveness but changed only slightly for trustworthiness, $b = -0.32$, 95% CI [−1.26, 0.62], $t(71.05) = -0.69$, $p = .494$. With age and sequence as additional covariates, the estimates were $b = -0.20$, 95% CI [−1.25, 0.84], $t(68.12) = -0.39$, $p = .696$, for attractiveness and $b = -0.69$, 95% CI [−1.68, 0.29], $t(68.48) = -1.40$, $p = .166$, for trustworthiness. Given the near-complete confounding of condition, viewer gender, and age, these adjusted estimates are very imprecise.

Viewer Gender

Because only two men participated in the no-filter condition and only six women in the filter condition, gender differences in the effect of the filter could not be tested. Descriptively, men in the filter condition gave lower ratings than women in the filter condition (5.00 vs. 5.97 for attractiveness and 5.36 vs. 5.90 for trustworthiness, see Table 1). Women rated filtered faces as slightly more attractive ($M = 5.97$, $SD = 1.17$) than unfiltered faces ($M = 5.52$, $SD = 1.27$). An exploratory comparison among women showed no significant effect of the filter on attractiveness, $t(7.29) = 0.84$, $p = .428$, $g = 0.35$, 95% CI [−0.51, 1.20], or trustworthiness, $t(7.70) = 0.28$, $p = .784$, $g = 0.11$, 95% CI [−0.74, 0.96].

Exploratory Analyses of Participants' Own Photo Editing

Twenty-six participants (37.7%) reported editing photos in which they themselves could be seen. This proportion was higher among

Table 2: Linear Mixed-Effects Models of Attractiveness and Trustworthiness Ratings.

			95% CI				
Parameter	**Estimate**	***SE***	***LL***	***UL***	***t***	***df***	***p***
Attractiveness							
Intercept	5.38	0.25	4.87	5.88	21.22	74.94	< .001
Condition (filter vs. no filter)	−0.41	0.32	−1.04	0.22	−1.29	77.93	.200
SD participant intercepts	1.23						
SD face intercepts	1.27						
SD face slopes (condition)	0.59						
Face intercept–slope *r*	−.16						
SD residual	1.66						
Trustworthiness							
Intercept	5.65	0.18	5.29	6.01	31.19	98.04	< .001
Condition (filter vs. no filter)	−0.39	0.31	−1.00	0.22	−1.27	78.18	.206
SD participant intercepts	1.18						
SD face intercepts	0.67						
SD face slopes (condition)	0.58						
Face intercept–slope *r*	−.54						
SD residual	1.74						

Note. Each model is based on 2,691 ratings of 39 faces by 69 participants. Condition was coded −0.5 (no filter) and 0.5 (filter), so the estimate for condition represents the difference between the conditions. Models were estimated with restricted maximum likelihood. Degrees of freedom were approximated with the Satterthwaite method, and confidence intervals are based on the corresponding *t* distribution. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

women (19 of 39, 48.7%) than among men (7 of 30, 23.3%), Fisher's exact test, $p = .045$, $OR = 3.07$, 95% CI [0.98, 10.55]. Because the confidence interval of the odds ratio, which is computed with a different method than the p value, included 1, this difference should be interpreted with caution. Among the 26 participants who edited their photos, most rated the extent of their editing as low, with 84.6% choosing 1 or 2 on the 5-point scale. Across all participants, the extent of editing did not differ between women ($M = 1.59$, $SD = 0.75$) and men ($M = 1.47$, $SD = 0.90$), Mann–Whitney $U = 500.5$, $p = .240$, and was not significantly correlated with age, $r_s = -.15$, $p = .224$ ($n = 69$), or with education, $r_s = .04$, $p = .775$ ($n = 68$). Nine of the 43 participants who reported not editing their photos nonetheless chose a value of 2 or higher on the extent scale.

Discussion

In this study, I examined whether a typical facial photo filter lowers judgments of the attractiveness and trustworthiness of AI-generated female faces. Faces presented with the filter received slightly lower mean ratings on both dimensions, by about 0.4 points on a 10-point scale. Neither difference was statistically significant, and the mixed-effects models, which also account for the sampling of faces, led to the same conclusion. The expected effects were therefore not confirmed. The confidence intervals, however, were wide and included zero as well as effects of moderate size. The data thus do not show that the filter has no effect. They are inconclusive with regard to the direction and size of a potential effect.

The interpretation of the small negative difference is further limited by the composition of the groups. Because the questionnaire versions were randomly allocated to recruitment channels rather than to individual participants, the filter condition consisted mostly of men and was older on average, whereas the no-filter condition consisted almost exclusively of young women. Descriptively, men gave lower ratings than women within the filter condition, and women rated filtered faces as slightly more attractive than unfiltered faces. When viewer gender was taken into account, the difference in attractiveness ratings was close to zero, whereas the difference in trustworthiness ratings changed only slightly. All adjusted estimates, however, were very imprecise. At least for attractiveness, the small overall difference may therefore reflect differences between the groups of viewers rather than an effect of the filter. For the same reason, gender differences in the effect of the filter could not be tested, and it remains open whether men judge filtered faces more harshly than women do.

The present findings neither confirm nor rule out the negative evaluations of conspicuously filtered selfies described in qualitative research [17] or the lower number of likes received by filtered selfies on Instagram [18]. In the focus groups of Lavrence and Cambre [17], male participants reacted negatively to conspicuously filtered selfies of women. One of them stated: "When I'd see a girl that has like, first picture's a filter picture, I'd swipe left, like 'nah I'm good'" (p. 10). Female participants, in contrast, regarded restrained, selective retouching as acceptable: "Just like little things that don't really change how you look, they're just changing one little thing that was off in that moment" (p. 7). Such evaluations may depend on whether a filter is recognized as such, on its intensity, and on the context in which a photo is seen. In the present study, each participant saw only one version of each face and could not compare versions directly. Participants were also not told that some images had been edited, so the filter may have gone unnoticed by many of them. In addition, faces synthesized with StyleGAN tend to be rated as more trustworthy than real faces [25], which may have left limited room for further changes in perceived trustworthiness.

About half of the women and a quarter of the men reported editing photos of themselves, and most of them rated the extent of their editing as low. These proportions are markedly lower than the 90% of young women reported by Gill [2]. Gill's figure, however, refers to any use of filters when posting selfies, including effects that do not beautify, whereas the present item asked about editing photos of oneself. In this sample, the extent of editing was not significantly correlated with age or education.

Limitations and Future Directions

The study has several limitations. First, randomization took place at the level of four recruitment channels rather than individual participants. With only two channels per condition, randomization could not balance participant characteristics, and condition was confounded with viewer gender and age. This limits causal conclusions about the effect of the filter. Because the clustering of participants within channels could not be modeled, the reported p values may also be somewhat too liberal. Second, the sample was small. With about 35 participants per condition, only effects of $d \geq 0.68$ could be detected with adequate power, whereas the effects observed here were considerably smaller. Third, a post hoc check of the stimuli revealed that one image in the no-filter condition had been filtered and that one face had no unfiltered counterpart. Both items were excluded, and the assignment of images to items in the versions with reversed order was assumed rather than verified. In addition, filtered and unfiltered images differed slightly in resolution, so filtered images may have been displayed slightly larger on some screens. Fourth, only one specific filter with strong global smoothing was examined. Because Meta discontinued third-party augmented reality effects on Instagram in January 2025 [37], this filter is no longer available, although the filtered stimuli themselves have been preserved. Fifth, the stimuli were AI-generated faces of young to middle-aged White women. Whether the findings generalize to real selfies, to other groups of depicted persons, and to authentic social media or dating contexts remains open. Finally, sexual orientation was not assessed, so questions concerning attraction in dating contexts could not be addressed, and photo editing was measured with single items without verbal scale labels.

Future studies should randomize individual participants to conditions, for example, via a single link with automatic random assignment, and should determine sample sizes a priori. Detecting an effect of $d = 0.30$ with 80% power in a between-subjects design would require about 176 participants per condition (two-tailed, α

= .05). Within-subjects designs, in which each participant rates filtered and unfiltered versions of different faces in counterbalanced order, would be considerably more efficient. Filter intensity could also be manipulated dimensionally, that is, along a continuum from subtle retouching to strong smoothing, to identify the degree of editing at which judgments of attractiveness and trustworthiness change. Such studies should further assess viewers' gender, sexual orientation, and own use of filters, as well as whether viewers recognize a photo as filtered.

Conclusion

In this online experiment, a typical facial photo filter did not significantly lower judgments of the attractiveness or trustworthiness of AI-generated female faces. The small negative tendency observed could not be separated from differences in viewer gender and age between the groups. The assumption that facial photo filters make faces appear less attractive and less trustworthy is therefore not supported by the present data, and whether men judge filtered faces more harshly than women do could not be tested. Both questions should be examined in adequately powered studies with randomization at the participant level before conclusions are drawn for social media or online dating.

References

1. Ramphul K, Mejias SG. Is Snapchat dysmorphia a real issue. Cureus. 2018; 10: e2263.
2. Gill R. Changing the perfect picture Smartphones social media and appearance pressures. City University of London. 2021. https://www.citystgeorges.ac.uk/__data/assets/pdf_file/0005/597209/Parliament-Report-web.pdf
3. Encyclopaedia Britannica. Adobe Photoshop. In Encyclopaedia Britannica. 2022. https://www.britannica.com/technology/Adobe-Photoshop
4. Stark H. Introducing FaceApp The year of the weird selfies. Forbes. 2017. https://www.forbes.com/sites/haroldstark/2017/04/25/introducing-faceapp-the-year-of-the-weird-selfies/
5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders 5th ed. 2013.
6. Norwegen bekämpft Schönheitswahn Influencer müssen Fake-Fotos kennzeichnen. Norway fights the beauty craze Influencers must label fake photos. 2021. https://www.n-tv.de/panorama/Influencer-muessen-Fake-Fotos-kennzeichnen-article22618520.html
7. Statista. Distribution of Instagram users worldwide as of April 2022, by age group Graph. 2022. https://www.statista.com/statistics/325587/instagram-global-age-group/
8. Chen H. College-aged young consumers' perceptions of social media marketing The story of Instagram. Journal of Current Issues & Research in Advertising. 2018; 39: 22-36.
9. Pounders K, Kowalczyk CM, Stowers K. Insight into the motivation of selfie postings Impression management and self-esteem. European Journal of Marketing. 2016; 50: 1879-1892.
10. Lindström B, Bellander M, Schultner DT, et al. A computational reward learning account of social media engagement. Nature Communications. 2021; 12: 1311.
11. Sherman LE, Payton AA, Hernandez LM, et al. The power of the like in adolescence Effects of peer influence on neural and behavioral responses to social media. Psychol Sci. 2016; 27: 1027-1035.
12. Day JJ, Carelli RM. The nucleus accumbens and Pavlovian reward learning. Neuroscientist. 2007; 13: 148-159.
13. Klawonn AM, Malenka RC. Nucleus accumbens modulation in reward and aversion. Cold Spring Harbor Symposia on Quantitative Biology. 2018; 83: 119-129.
14. Iqbal M. Tinder revenue and usage statistics. Business of Apps. 2022. https://www.businessofapps.com/data/tinder-statistics/
15. Walther JB, Parks MR. Knapp ML, et al. Cues filtered out, cues filtered in Computer-mediated communication and relationships. Handbook of interpersonal communication. Sage. 2022; 529-563.
16. Chopik WJ, Johnson DJ. Modeling dating decisions in a mock swiping paradigm An examination of participant and target characteristics. Journal of Research in Personality. 2021; 92: 104076.
17. Lavrence C, Cambre C. Do I look like my selfie Filters and the digital-forensic gaze. Social Media + Society. 2020; 6: 2056305120955182.
18. Hong S, Jahng MR, Lee N, et al. Do you filter who you are Excessive self-presentation social cues and user evaluations of Instagram selfies. Computers in Human Behavior. 2020; 104: 106159.
19. Santanachote P. Majority of Americans who know of social media beauty filters find them troubling CR survey shows. Consumer Reports. 2021. https://www.consumerreports.org/electronics/social-media/majority-of-americans-who-know-of-social-media-beauty-filter-a2740948053/
20. Leiner DJ. SoSci Survey Version 3.1.06. SoSci Survey GmbH. 2019a. https://www.soscisurvey.de
21. Leiner DJ. Too fast too straight too weird Non-reactive indicators for meaningless data in internet surveys. Survey Research Methods. 2019b; 13: 229-248.
22. Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR. 2019; 4396-4405.
23. Karras T, Laine S, Aittala M, et al. Analyzing and improving the image quality of StyleGAN. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR. 2020; 8107-8116.
24. Moreland RL, Zajonc RB. Exposure effects in person perception Familiarity similarity and attraction. Journal of Experimental Social Psychology. 1982; 18: 395-415.
25. Nightingale SJ, Farid H. AI-synthesized faces are

indistinguishable from real faces and more trustworthy. Proc Natl Acad Sci USA. 2022; 119: e2120481119.

26. West J, Bergstrom C. Which face is real. 2019. https://www.whichfaceisreal.com/
27. Alexandra rmh. IG filters & inspo. Instagram. Retrieved July 14, 2022. https://www.instagram.com/alexandra_rmh/channel/
28. Core Team R. RA language and environment for statistical computing Version 4.3.3. R Foundation for Statistical Computing. 2024. https://www.R-project.org/
29. Bates D, Mächler M, Bolker B, et al. Fitting linear mixed-effects models using lme4. Journal of Statistical Software. 2015; 67: 1-48.
30. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package Tests in linear mixed effects models. Journal of Statistical Software. 2017; 82: 1-26.
31. Ben-Shachar MS, Lüdecke D, Makowski D. effectsize Estimation of effect size indices and standardized parameters. Journal of Open Source Software. 2020; 5: 2815.
32. Delacre M, Lakens D, Leys C. Why psychologists should by default use Welch's t-test instead of Student's t-test. International Review of Social Psychology. 2017; 30: 92-101.
33. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science A practical primer for t-tests and ANOVAs. Front Psychol. 2013; 4: 863.
34. Judd CM, Westfall J, Kenny DA. Treating stimuli as a random factor in social psychology A new and comprehensive solution to a pervasive but largely ignored problem. Journal of Personality and Social Psychology. 2012; 103: 54-69.
35. Barr DJ, Levy R, Scheepers C, et al. Random effects structure for confirmatory hypothesis testing Keep it maximal. J Mem Lan. 2013; 68: 255-278.
36. Westfall J, Kenny DA, Judd CM. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. J Exp Psychol Gen. 2014; 143: 2020-2045.
37. Meta Spark. A Meta Spark update. Meta Spark Blog. 2024. https://spark.meta.com/blog/meta-spark-announcement/

© 2026 Darius Krutzek. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License