

Spark AR cosmetic filters reversal: Safety briefings

Plan

On Thursday, August 6th at 8am PT, we will revert an interim policy decision we made towards the end of last year to disallow AR effects associated with cosmetic surgery. The plan is to pre-brief safety or policy partners whom we expect to have a public negative reaction to this announcement, who may benefit from hearing from us beforehand. While we initially wanted to bundle this with a briefing on private by default to soften the impact, the pushback of that launch requires a standalone briefing.

Key messages

- At the end of 2019, we temporarily removed some face altering AR effects from Instagram, following a debate around their potential impact on mental health.
- After receiving significant feedback from creators and evaluating the way in which our community uses face altering effects, we have decided to once again allow them.
- However, we also recognise that there may be potential unintended consequences from face altering effects, so we will not allow these effects to be discovered or amplified in our Effects Gallery.
- We will continue to uphold our ban on face-altering effects that directly promote or encourage potentially dangerous cosmetic surgery procedures

Additional context for key messages

At the end of 2019, we temporarily removed some face altering AR effects from Instagram, following a debate around their potential impact on mental health. It was always meant to be an interim measure, giving us time to reach out to our creators and users, and collect a range of perspectives. Since then, we have received significant feedback - both from creators and users - that

1. People enjoy creating and using face altering effects, and miss being able to create artistic or fantastical effects like they can on other services.

We found that creators of these effects predominantly used face alteration and feature augmentation to create artistic, surreal, fantasy effects that many enjoy, and that these effects are widely available across other platforms.

2. Face filters allow people to express themselves in a greater range of ways, particularly to subvert typical gender norms

"Face filters emphasize the face as part of an image, and the realtime interaction naturally draws us closer to the camera. Without knowing, women who use face filters automatically discard the long-standing norm of face-ism. An exception to this pattern are the male filtered selfies, where men seemingly willingly experiment with feminine aesthetics and gender display. Unfortunately, image-based research gives limited insights in their motivations and experiences, but the typically feminine aesthetics of filters at least enable this experimental practice and I can imagine that transgender people find this empowering... In the most stereotypical female filtered selfies, the filters seem to make the image even more stereotypically feminine. However, the filters that convey stereotypical norms of gender are not necessarily used in conventional ways. Some men subvert their meaning by using them as a way of identity experimentation. The feminine aesthetics that are inherent in most face filters are actually useful for them to play with."

[[HYPERLINK](#)

"https://dspace.library.uu.nl/bitstream/handle/1874/380003/AWholeNewWaytoSeeYourselfie_MSchipper_NMDC_thesis_5944635.pdf?sequence=2&isAllowed=y" \h] (Utrecht University, 2018)

3. Filters allow people to express themselves and have greater agency around how they share parts of their identities on social media

"In the private realm, digital self-portraits are seemingly controlled by the user/photographer on almost every level. We get to choose how to frame and capture our images, how to crop and filter them and where to post them. As such, selfies represent a moment of complete agency and self-expression, which is almost certainly pleasurable and meaningful, otherwise why would so many people do it?"

[[HYPERLINK "https://www.academia.edu/17789374/_selfie_Digital_self-portraits_as_commodity_form_and_consumption_practice"](https://www.academia.edu/17789374/_selfie_Digital_self-portraits_as_commodity_form_and_consumption_practice) \h] (2016, Consumption Markets & Culture)

4. This is still an evolving space with limited longitudinal research available on the impact of face altering effects, and we're committed to understanding it better

We welcome opportunities to collaborate and learn more about this space from third party experts and academics, as the current landscape of longitudinal research is limited.

Following this analysis and feedback gathering exercise, we have decided to once again allow face altering effects. However, we also recognise that there may be potential unintended consequences from face altering effects, so we will not allow these effects to be discovered or amplified in our Effects Gallery and Effects Tray (which is how many people find them).

We will continue to uphold our ban on face-altering effects that directly promote or encourage potentially dangerous cosmetic surgery procedures - for example, effects that offer the sale of these procedures, effects that depict cosmetic procedures through surgery lines, or effects that offer instructions for such procedures.

Creators who've previously submitted effects that have been held in review do not need to take action. Effects will be automatically re-reviewed for publishing.

Additional resources

- [HYPERLINK "https://docs.google.com/document/d/1czbQ-3fiZQNf9EG4oI_VS8YoL4j7qCLml3KBc8AxpM4/edit" \h]
- [HYPERLINK "https://docs.google.com/document/d/1cr5gOmw1y9RU2Dai0HJBhk2O_FlsQqjmoDXZ8_uRbyE/edit" \h]

Outreach tracker

Region	Organisation / Individual	Briefing POC	Status	Feedback
NorAm	NEDA / Claire Mysko (CEO) and Chevese Turner (Chief Policy Officer)	Dayna Geldwert	Done	Neutral - respect the nuance/complexity, but concerned about impacts. "Will you continue to be vigilant about how these features are being used?" "On the surface, I can understand why you're making this decision, but we know that there are things going on underneath the surface that aren't captured in research... young people don't always know how these things are impacting them." Wants to know if we can share more details on our internal findings and the external research we used to inform our decision. Also said greater specificity around which filters are/aren't allowed are important to landing this publicly. Were glad to hear we are adding friction to the discoverability of these filters.
EMEA	Young Minds / Fiona and Emily	KWOC		Wasn't able to schedule a call, but emailed with update will let you know what they come back with

EMEA	Diana Award	KWOC	Done	Neutral to negative - but worried about how we came to the decision and would have liked to been consulted
EMEA	Internet Matters	KWOC	Done	Neutral Understood why we made change but worried about what this means for our pressure to be perfect work as it contradicts
EMEA	The Mix	KWOC	Done	Neutral to negative - but worried about how we came to the decision and would have liked to been consulted
LatAm	SaferNet	[REDACTED]	done	Neutral - hesitant on the fine lines between cosmetic filters and filters that promote surgery. They asked for more information on how to report filters.
US	Northeastern / Rachel Rodgers	Vaishnavi / Sarah	done	Appreciate the thoughtful point regarding the creative expression, experimentation with identity, and opportunity to transgress stereotypical norms of presentation. Also difficult to draw a line in an increasingly blurry "beauty market" of permanent makeup, and "minimally invasive" procedures that may progressively be normalized and not viewed as harmful or substantial, and yet contribute to a slowly rising minimal beauty standard. E.g. wedding industry now provides beauty packets of what would previously have been considered cosmetic surgery but now is just a means of looking "fresh" on your special day). Agree that longitudinal data in this space is lacking.
US	University of Washington / Katie Davis	Vaishnavi / Sarah	done	It sounds like your team has been very thoughtful in how you've approached this decision. It will be interesting to see how different groups (particularly youth) make use of these effects over the long-term. I imagine there will be quite a range, with most of them harmless and potentially even beneficial with respect to identity exploration.
US	Harvard / Bryn Austin	Vaishnavi / Sarah	done	It sounds like you were able to keep in some protections for younger or more

				vulnerable users on Instagram. All of these technologies evolve so quickly that keeping the platform safe will be a moving target over time, I would bet... I'm always happy to help and have been really impressed with your approach to developing safety measures for your users around these issues.
APAC	Dr. Jamie Chiu	[REDACTED]	done	"Thank you for the update. I am glad to have been able to help."
US	Cornell / Janice Whitlock	Vaishnavi / Sarah	done	I think your approach is sound -- it seems to blend a number of key considerations into something workable, at least for now.

Email for stakeholders we previously sought feedback from [final]

Dear [REDACTED],

I hope you're doing well.

Late last year, we connected on the topic of AR cosmetic surgery effects on Instagram, and you shared insights that were very helpful to us in considering a thoughtful policy approach to this challenging topic. Thank you once again for taking the time to speak with us. I wanted to circle back and let you know where we landed and how expert input like yours helped us reach this result.

You may recall that last year we temporarily removed some face altering AR effects from Instagram, following questions raised about the potential impact of these effects on mental health. Removing these effects was an interim measure that provided us time to reach out and collect a range of perspectives from creators, users, and experts in psychology, research, body image activism, and augmented reality.

Starting today, we'll once again allow people to create and share face-altering effects on Instagram, however we will not allow these effects to be discovered or promoted in our Effects Gallery (which is how many people find effects). We will continue to uphold our ban on effects that directly promote or encourage potentially dangerous cosmetic surgery procedures - for example, effects that offer the sale of these procedures, effects that depict cosmetic procedures through surgery lines, or effects that offer instructions for such procedures.

As we dug into this issue, we found that creators of these effects predominantly used face alteration and feature augmentation to create artistic, surreal, and fantasy effects. Surrealistic and fantastical effects were seen by experts as of much less concern than hyper-realistic face alteration, though many of the experts we consulted shared the sentiment that drawing a firm line between the realistic and the unrealistic is quite hard to do. Research we consulted also indicated that face filters can allow people to express themselves in a greater range of ways - particularly to subvert typical gender norms - and to have greater agency around how they share parts of their identities on social media. This is validated by the fact that many users told us they enjoyed both creating and using face altering effects, and were unhappy with our interim decision to remove them entirely.

This is still an evolving space with limited longitudinal research available on the impact of face altering effects, and we're committed to understanding it better. We welcome further opportunities to collaborate and learn more about this space from experts like you. Thank you again for making time to speak with us and please let us know if you have any questions.

Regards,

yyy