Cosmetic Surgery Effects (a/c priv)

A. Meeting Goal

Decide on whether to continue, modify, or lift the ban on Cosmetic Surgery AR Effects that we instituted in October 2019 in response to expert and press concerns about potential negative wellbeing impacts (especially on minors). Cosmetic Surgery Effects (as referenced in this doc) are AR / camera effects that change people's facial structure in a way that must be achieved by plastic surgery (such as enlarging eyes or lips, changing the position of eyes or changing the shape of nose or chin) and cannot be achieved by ordinary makeup techniques, *unless* the result is an exaggerated or humorous transformation (e.g., rapid aging, babyface, character transformation). Continuing to ban Cosmetic Surgery AR Effects negatively impacts creator voice & innovation, and may impact long-term growth (although we lack definitive data).

B. Decision Needed & Options

Should we continue, modify, or lift the ban on Cosmetic Surgery AR Effects that we instituted in October 2019 in response to expert and press concerns about potential negative wellbeing impacts (especially on minors)?

- **Option 1: Continue the temporary ban with current Policy and re-evaluate when there is more definitive data on wellbeing and product impacts [Policy + Comms, Responsible Innovation Team recommendation]**
    a. Pros: Mitigates wellbeing concerns; no PR/regulatory risk
    b. Cons: Limits growth; limits voice and innovation; arguably western-centric
- **Option 2: Lift the ban but remove these effects from IG recommendation surfaces [FB App + Instagram Recommendation / Spark AR + MSGR + i18n open to a version of this]**
    a. Pros: Fosters voice and innovation; lower impact to growth
    b. Cons: Still notable wellbeing risk (since recs are marginal source of virality); possible bad PR
- **Option 3: Lift the ban [Boz/Spark AR, MSGR + i18n recommendation]**
    a. Pros: Fosters voice and innovation; lowest impact to growth; supports regional perspectives
    b. Cons: Highest wellbeing risk; possible bad PR

CAVEAT: Data on product impact is limited, and research into wellbeing impacts of AR effect is too new and not yet robust. Meaningful research may take years. Therefore, we must make a **value-based** decision here -- essentially, safety vs. voice.

C. Key Considerations

Wellbeing & Safety Concerns: There is not yet robust causal evidence about the relationship between Cosmetic Surgery AR Effects and wellbeing. However, experts from around the world -- APAC included -- generally agree that these effects are cause for concern for mental health and wellbeing, especially amongst vulnerable populations (females, youth, those with a history of mental health concerns, etc.) Additionally, academic research on related topics (e.g., wellbeing impacts of seeing "ideal" images in media) gives us cause for concern. See appendix for additional information.

PR/Regulatory: The wellbeing concerns have resulted in press and regulator criticism of platforms allowing such effects, and our October ban received positive press coverage. Reversing it risks a PR / regulatory backlash. Instagram is the primary platform for AR usage and would likely face the brunt of the criticism if we reverse.

Competitiveness / Growth: We do see adoption from both users and creators, though we don't have conclusive product impact given the ban was not tested. We face a competitive landscape of sharing effects, and it is likely that we will see a negative growth impact, simply because any restriction is likely to reduce engagement if people go elsewhere. Long-term, limiting our platform could handicap us as we look to bootstrap our network of Creators to build for our future devices.

Voice, Creativity, and Innovation: By banning these Effects, we may limit people's voice in expressing themselves through developing certain Effects or presenting themselves as they choose. This can limit innovation in AR Effects creation on our Apps and limit our ability to foster cultural moments and viral trends.

Asia Perspective: AR Effect usage and preferences vary across regions. It is likely that this ban disproportionately impacts Effects preferred by people in the APAC region. Those not persuaded by the safety/wellbeing warnings of experts may argue that this ban is Western-centric and about exporting values.

D. History & Current Policy

History:

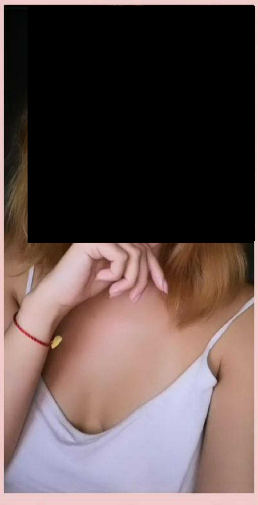
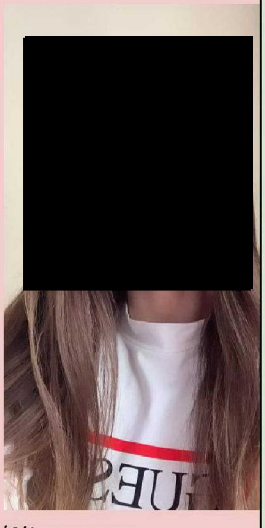
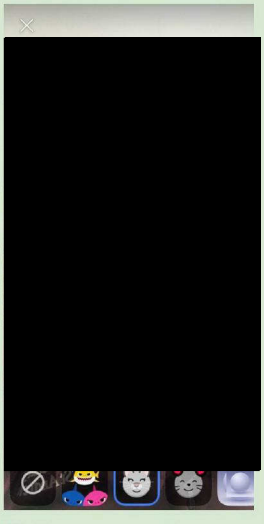
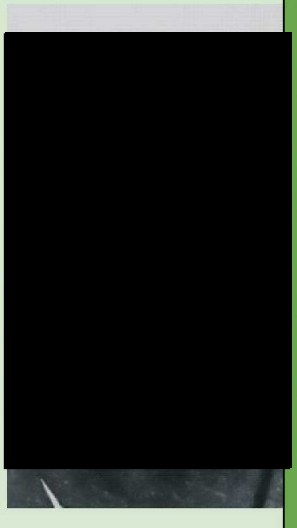
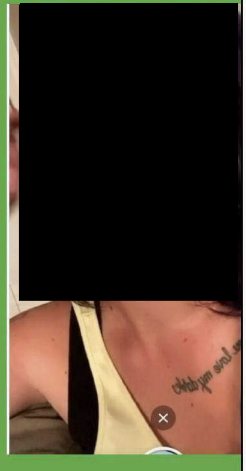
- May 2017 -- IG began offering first-party AR Effects (but never any 1st party Cosmetic Surgery Effects)
- Aug 2019 -- We launch third-party (developer) AR Effects to 100% of users. (Cosmetic Surgery Effects not yet prohibited.)
- Sept 2019 -- Public complaints and regulatory questions about Cosmetic Surgery Effects.
- Oct 2019 -- Temporary ban on Cosmetic Surgery Effects while we research safety risks.
- Dec 2019 -- After consulting 21 independent experts around the globe, we decide to continue the ban, pending leadership review.

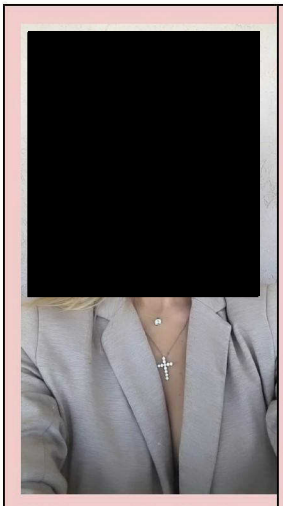
Current Policy (since October 2019):

- No AR effects that change people's facial structure in a way that must be achieved by plastic surgery (such as enlarging eyes or lips, changing the position of eyes or changing the shape of nose or chin) and cannot be achieved by ordinary makeup techniques unless in service of creating a character transformation.
- No AR effects that explicitly promote plastic surgery (e.g. nip/tuck lines on the face)
- No AR effects that target skin tone to lighten or promote racist stereotypes
- We permit:
    a. Exaggerated or humorous transformations (e.g., rapid aging, babyface)
    b. Fantastical effects that change a person's facial structure for the purpose of turning the person into a character or animal
    c. Make-up effects (e.g., darkening eyelashes, or adding lip color) or skin smoothing
    d. Uploading, via other photo manipulation services, images that mimic cosmetic surgery

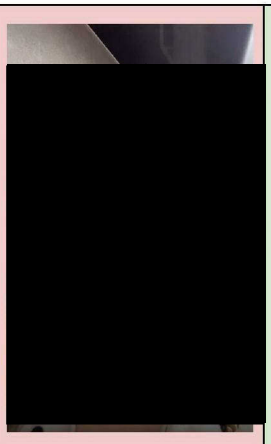
E. Examples of AR Effects

Red = currently not allowed under interim policy
Light green = currently allowed under interim policy
Dark green = always allowed

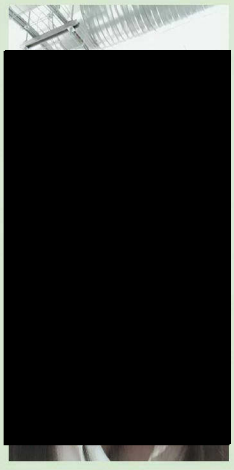
1. Cosmetic Surgery Effects	2. Cosmetic Surgery Effects + Decorations	3. Cosmetic Surgery Effects + Character Transformation	4. Cosmetic Surgery Effects + Exaggerated or Humorous Transformation	5. Makeup effects
Changes size or shape of facial features only	Changes size or shape of facial features PLUS adds one or more decorations	Changes size or shape of facial features PLUS adds two or more features that are integral to a character	Drastically changes size or shape of facial features in a way that is obviously exaggerated or humorous and not for beauty purposes	Adds makeup to the user and does NOT change the size or shape of facial features
	Ex: flower crown, freckles, glasses, bow, hat, sparkles	Ex: cat ears, whiskers	Ex: witch nose, exaggerated eyebrows, bug eyes, face-aging	Ex: lipstick, eyelashes, blush, eyeshadow, skin smoothing, blush, etc.
 (Enlarges the eyes)	 (Slims nose, enlarges lips)	 (Enlarges eyes)		



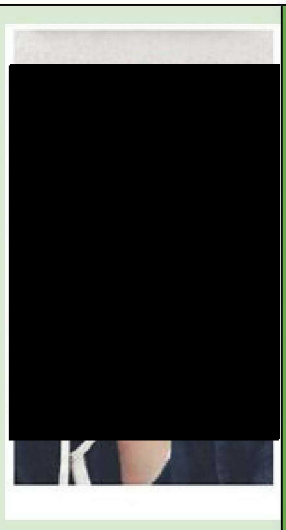
(Enlarges the lips)

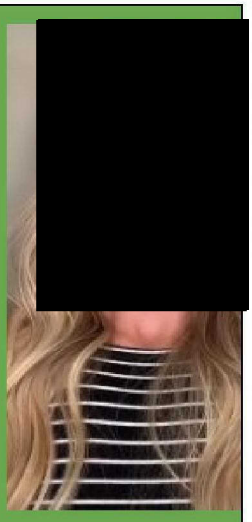


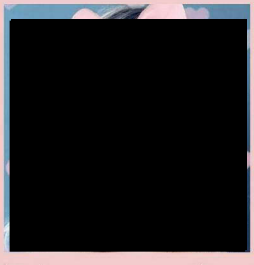
(Slims nose, enlarges lips)



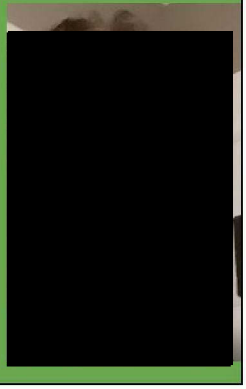
(Slims chin)



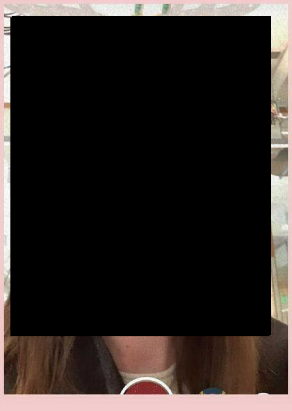




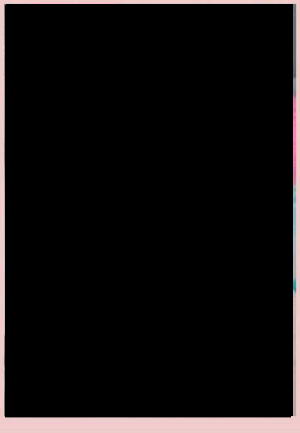
(Enlarges eyes, slims chin)



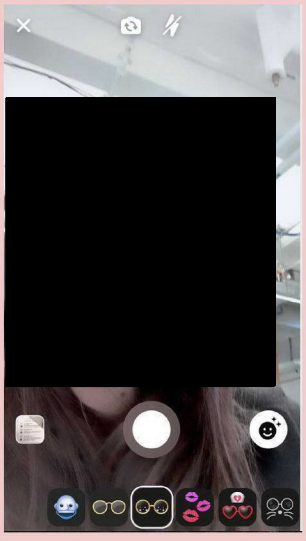
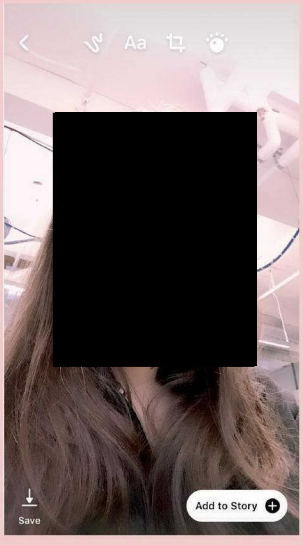
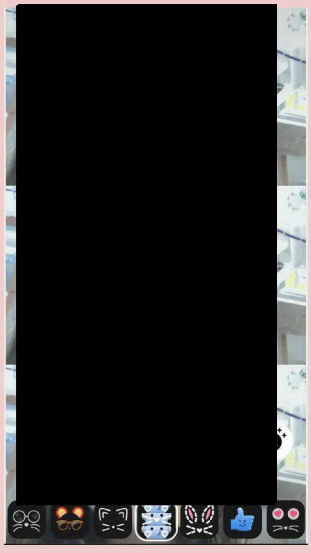
However, the grey area across the spectrum #1 through #5 above is big. Here are some examples of borderline effects.

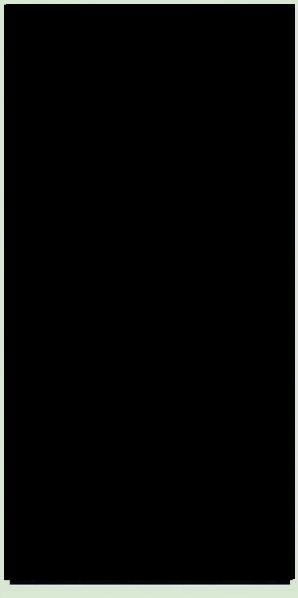
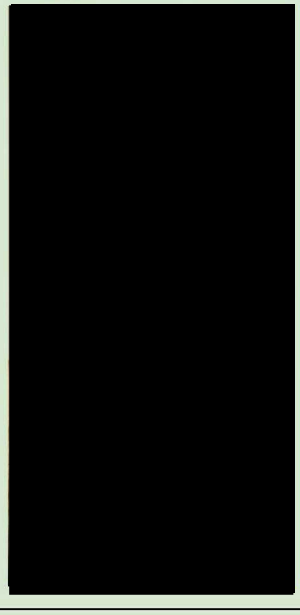


Reject: The character is Rudolph but it only has two character elements (antlers and nose). Effect makes eyes bigger and changes the shape of the forehead.



Reject: Not really a "character" and only has two elements (hair clips + face art)

		
Reject: No character transformation, adds 2 decorations. Eye enlarging, chin slimming.	*Reject:* No character transformation. Eye enlarging, chin slimming.	*Reject:* No character transformation, adds two decorations. Chin slimming.

	
Approve: Not a recognizable "character" but has three elements (face jewels, face paw prints, ears)	*Approve:* Character looks like pikachu, has three elements.

F. Appendix

Operations / Review of AR Effects

AR Effects are proactively reviewed by FB, and users can also directly report them to us. We also use automated detection to identify possible violations and route them for human review. In general, when we identify a violating AR Effect, we reject it and notify the AR creator. However, since the October temporary ban on Cosmetic Surgery Effects, we have not been rejecting cosmetic surgery effects but rather placing them in a "holding queue." There is now a backlog of 33K effects (c. 8x daily new effect volume) impacted by the temporary ban. Once we confirm (or lift) the ban, we will notify these creators accordingly.

History & Policy -- additional information

In October 2019, there was a series of articles skewering Instagram about third party Cosmetic Surgery effects. The articles focused on the negative impact these effects could have on mental health and body dysmorphia issues for teens, especially girls. Leads from the Policy, Instagram, and Spark product teams agreed to a temporary policy to stop approving Cosmetic Surgery AR Effects. While the press response was overwhelmingly positive, they did not take away the temporary nature of the policy, and we foresee significant comms fallout if the interim policy is reversed. Company spokespeople, including IG Head of Product Vishal Shah, have spoken publicly about the policy as part of IG's well-being efforts. (See coverage here).

In November, the Content Policy team aimed to finalize the policy for Cosmetic Surgery Effects through the Product Policy Forum -- the company's regular policymaking process. During that process, stakeholders from Content Policy, IG public policy, Spark AR product, IG product, Comms, Operations, Research, and Product Marketing generally supported continuing the ban on Cosmetic Surgery AR Effects. This policy development process explored various options, including age-gating Cosmetic Surgery AR Effects and removing Cosmetic Surgery AR Effects from recommendation surfaces. Ultimately, the ban was provisionally passed, pending completion of further APAC-focused data analysis, additional outreach to experts, and additional literature review. Following this additional work, we decided to escalate the issue to App Leads to ensure alignment with the recommendation to continue the ban.

Wellbeing - additional info

We are not aware of any robust research on the impact of AR Effects on wellbeing. There has been research on adjacent topics that may help inform our thinking:

- Exposure to "ideal" images throughout media (often thin, attractive women or muscular men; images digitally edited and/or with flattering poses and lighting) can lead to increased body dissatisfaction (see 1, 2), though the effects are small and strongest for women with underlying vulnerability (e.g., pre-existing body dissatisfaction).
- The few experiments focusing on edited images involve *briefly* exposing participants to edited images and measuring *short term* changes related to well-being (e.g., body dissatisfaction), generally finding that edited images result in more body or face dissatisfaction than unedited

images; there is little information on the effects of either longer term exposure to edited images, nor the durability of any associated well-being effects (3, 4, 5, 6).

- Disclaimers that an image has been edited (e.g., icons, text, warnings) are not consistently protective against negative effects on body satisfaction; in fact, several studies found potentially harmful effects (harmful: 6, 7, 8; mixed: 3, 9).
- Most of the research has been on the impact of viewing others' edited images -- there is very little research on impact of selfie editing behavior to the individual editing their own photos; correlationally and longitudinally it is associated with greater body dissatisfaction (10, 11, 12), and experimental assignment to edit a selfie did not improve anxiety associated with posting a selfie (13).
- Anecdotal evidence from the plastic surgery community suggests viewing and editing one's own selfie images could activate desire for cosmetic surgery ('Snapchat dysmorphia', see 14, 15); rates of requesting plastic surgery to improve appearance in selfies has gone up from 42% to 55% from 2015 to 2017 (15). Some hypothesized mechanisms where edited images could lead to harmful effects include internalization of an unattainable beauty standard (15) or increased self-discrepancy (a contrast between one's actual vs. ideal self), which can be associated with dysphoria and/or attempts to reduce the discrepancy (16, see also 17).
- Information about potential positive effects is lacking and generally unclear; one small cross-sectional study from Korea found that the negative impacts of thin ideals and media pressures on body satisfaction were lessened for frequent photo editors (18). However, in general, the arguments that editing serves to improve self-esteem or control over self-presentation are in the context of a narrative that is generally negative about well-being (e.g., reinforcing negative processes like self-sexualization (19), narcissistic self-promotion (20), or addictive-like dependence on peer feedback (21)).

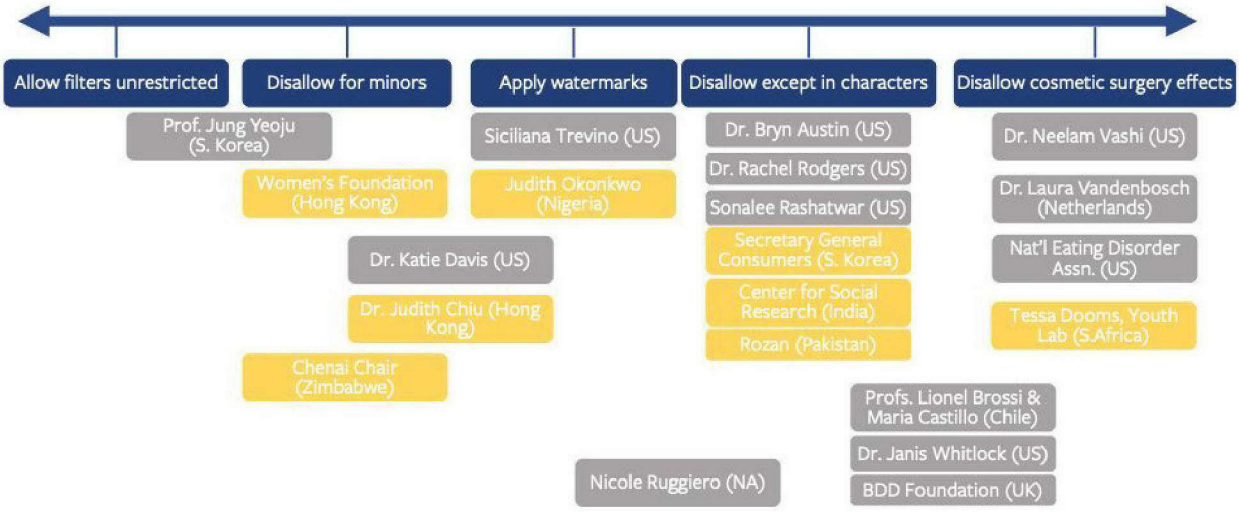
To augment this review of research, we consulted 21 independent experts around the world. Here is a summary of our findings:

- These extreme cosmetic surgery effects can have severe impacts on both the individuals using the effects and those viewing the images.
- Children are particularly vulnerable, however many others are vulnerable as well: those with a history of mental health challenges, eating disorders, etc.
- Generally, stakeholders support making a distinction between effects achieved via makeup versus those that can only be achieved via cosmetic surgery, though no line will be perfect.

(Note: After the Policy Development Process in November, an additional round of outreach was done with 9 experts from APAC and MENA and these stakeholders reiterated the dangers these filters have in advancing unrealistic beauty standards and impacting mental health and body image and their feedback was consistent with prior stakeholder feedback.)

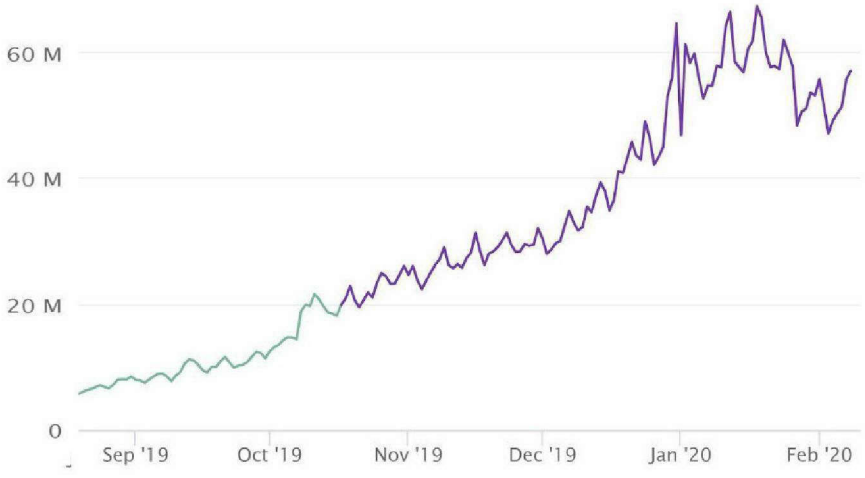
[See: Full stakeholder engagement summary]

This map represents the positions preferred by the stakeholders we consulted to develop this proposal:



Competitiveness / Growth - additional information

Adoption of AR effects have grown significantly since we launched 3P effects last half. This chart shows DAP for Spark AR Effects on IG before (green) and after the ban (purple):



We do not have conclusive product impact given the ban was not tested. AR appearance effects remain one of the main drivers for AR effect shares and make up 33% of total AR effect shares. Beauty effects (a superset of Cosmetic Surgery AR Effects) have a higher share+save rate than 3P non-beauty effects (15% vs 13%). Some users are using beauty AR extensively in their day to day lives to make Stories more interesting, to save time on make-up, and for self-esteem. Where we have inconclusive evidence is with regards to policy driving a downturn in the number of clients and project cancellations. Long-term,

limiting our platform can handicap us as we look to bootstrap our network of Creators to build for our future devices.

Voice, Creativity, and Innovation -- additional information

Creators are important because creators create content that goes viral. Creator feedback on the Cosmetic Surgery AR Effects ban is mixed:

- Some creators contend that this ban limits their artistic expression and feel uneasy about the line that feels like it's being drawn by IG/FB about beauty standards.
  - *"Effects can be gender affirming and body affirming, a company like Facebook deciding on this makes me feel really uncomfortable and I'm not happy about it."*
  - *"It's strange that Instagram has the power to police how we represent ourselves"*
- Some, after hearing the rationale behind the interim policy (user well being), were accepting of the policy.
  - After a one-on-one session with top creator Holy Mariia to explain the policy background and rationale, that creator then went on to create other popular (non-cosmetic surgery) effects.
  - Another popular creator we worked one-on-one with, @paigepiskin, posted publicly: *As you can see on my page, the LIL icey eyes filter had to be changed according to policy guidelines. Social platforms are concerned with wellbeing and mental health and how filters that alter the face effect them. I think its very noble that they are trying to consider how to allow creative freedom while also considering well being protection, it must be a bit challenging for them to navigate this filter boom. [sic]*

Virality: Organic Use Trumps Recommendation Surfaces. There are two third party AR Effect recommendation surfaces on Instagram, the Effects Gallery and Effect Tray. On average, 60% of Effect shares come from the camera tray (a partial recommendation surface of first and (soon to add) third party, plus effects the user has previously interacted with). Only 2% of effect shares come from the Effects Gallery (a recommendation surface of both first and third party effects). 16% of effect shares come from organic story attributions from IG users. However, if an effect goes viral, story attribution becomes the dominant surfaces people discover and share effects from. For example, during the first few days that the "Which Disney" effect went viral, 61% of shares came from story attribution.

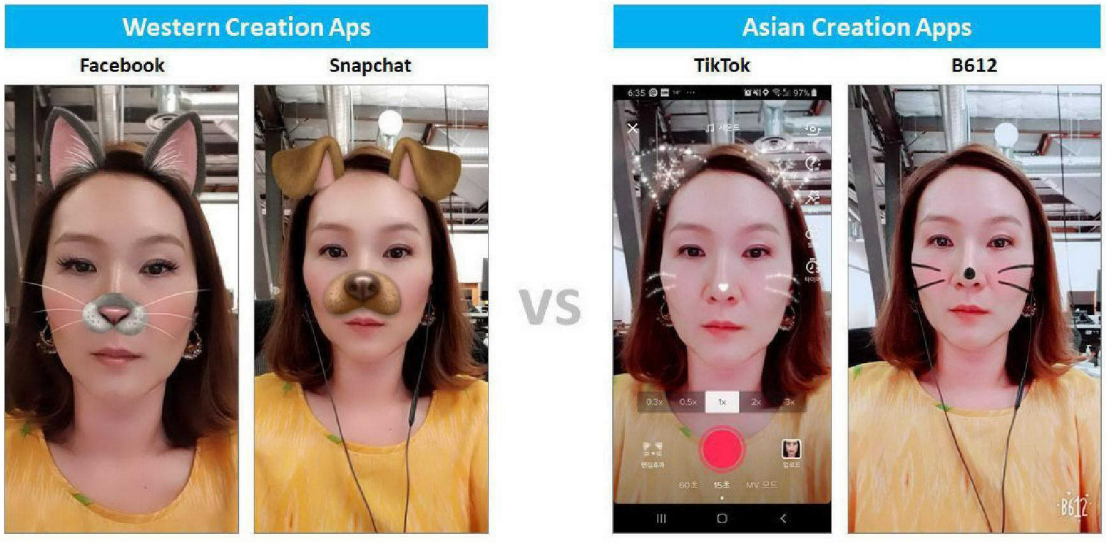
Asia Perspective

While social apps are widely used across Asia, the way Asians interact on these platforms differs from the West. This is due to the Asian tradition of humility that shuns self-promotion and encourages indirect and subtle ways of expression. The same principle applies to Asian users' expectations from AR effects and therefore, local creation apps that allow subtle facial feature modification are preferred over Western styled creation apps such as Snapchat (data).

A regional gap in the preference of AR effects can be clearly seen below where Asian apps apply more subtle effects compared to Western apps. These subtle effects can be achieved by makeup only. While they can result in a very slight facial distortion, i18n highly recommends to - at a minimum - enable this in consideration of users in Asia. To be clear, the Asian Creation App examples below would be banned under the current policy. When the i18n team partnered with the Messenger team to test the impact of

such subtle AR effects specifically designed for the Asian audience, a holdout test confirmed an 8% increase in camera shares in Asia compared to the global AR effects.

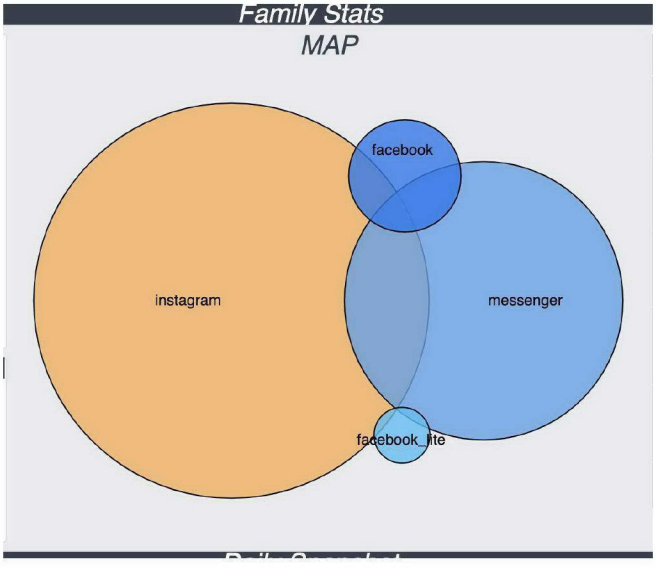
Since the current ban prohibits all types of facial distortion regardless of the intensity of changes (unless it is a character transformation), this blocks an opportunity for the company to empower Asians to freely express themselves in **their** ways.



Additional Data & Resources

AR Usage Overview (MAP):

- Instagram: 328.99 million
- Facebook: 13.92 million
- Messenger: 139.78 million



Note: This shows the usage of all AR effects (both first party and third party) across the Facebook, Inc. platform. All effects in Messenger are first party (Messenger-created) effects and do not include third party effects. Dashboard

Effect Captures by Platform

Average daily overview by platform:

Instagram
- 1st Party: 54M avg daily captures by 410 effects
- 3rd Party: 393M avg daily captures by 643K effects

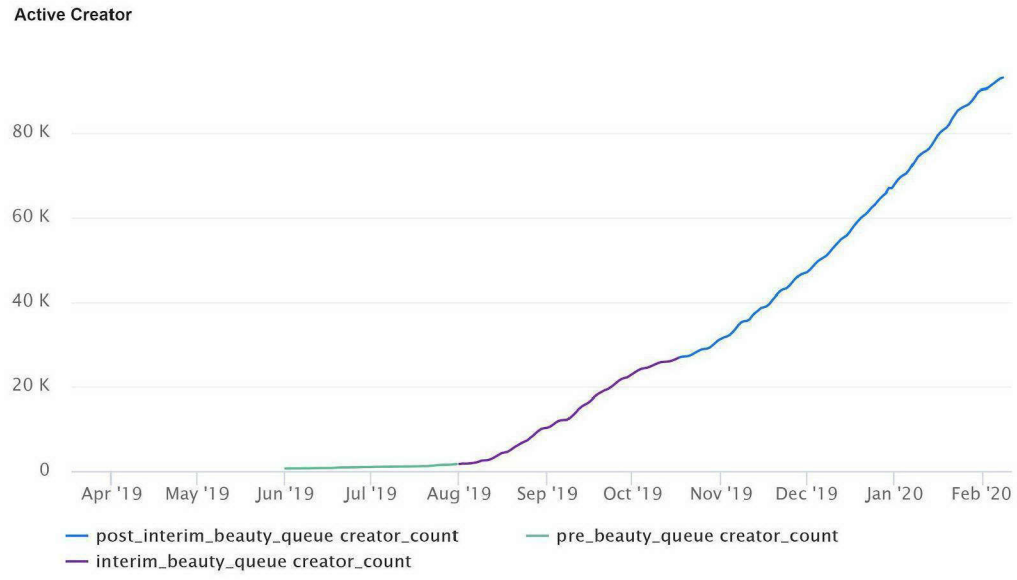
Facebook
- 1st Party: 2M avg daily captures by 664 effects
- 3rd Party: 2M avg daily captures by 115K effects

(Query - Screenshot)

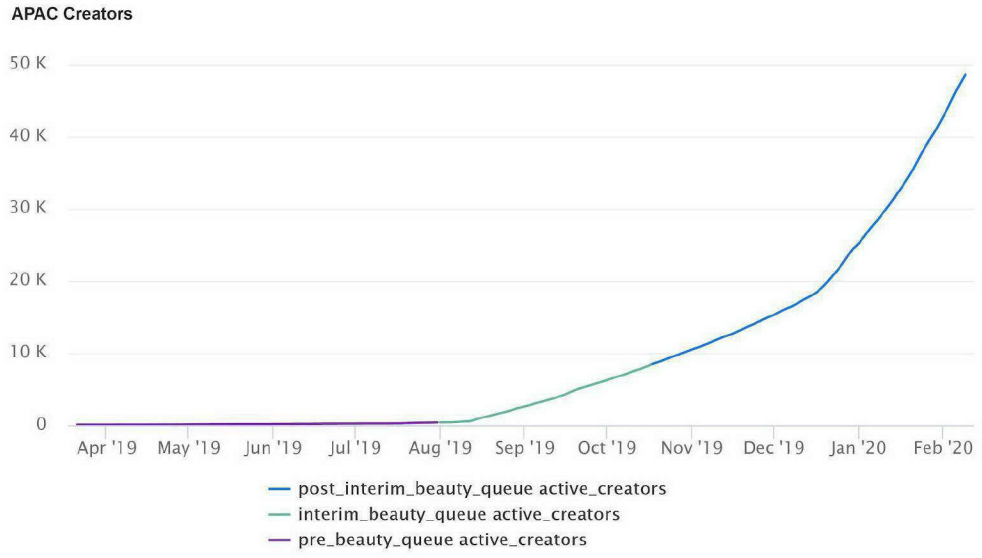
FaceWarp (any effect that distorts the face that includes, but is not limited to, Cosmetic Surgery Effects) effects are responsible for about 25% of third-party Effect Try-ons and 35% of Stories Impressions.
- 5.5% of total third-party Effect Try-ons currently come from cosmetic surgery effects which are later placed in 'Beauty' queue following user/employee reports.
- All faceWarp live effects, at least 20% of try-ons come from cosmetic surgery effects and mainly have been published before policy introduction.

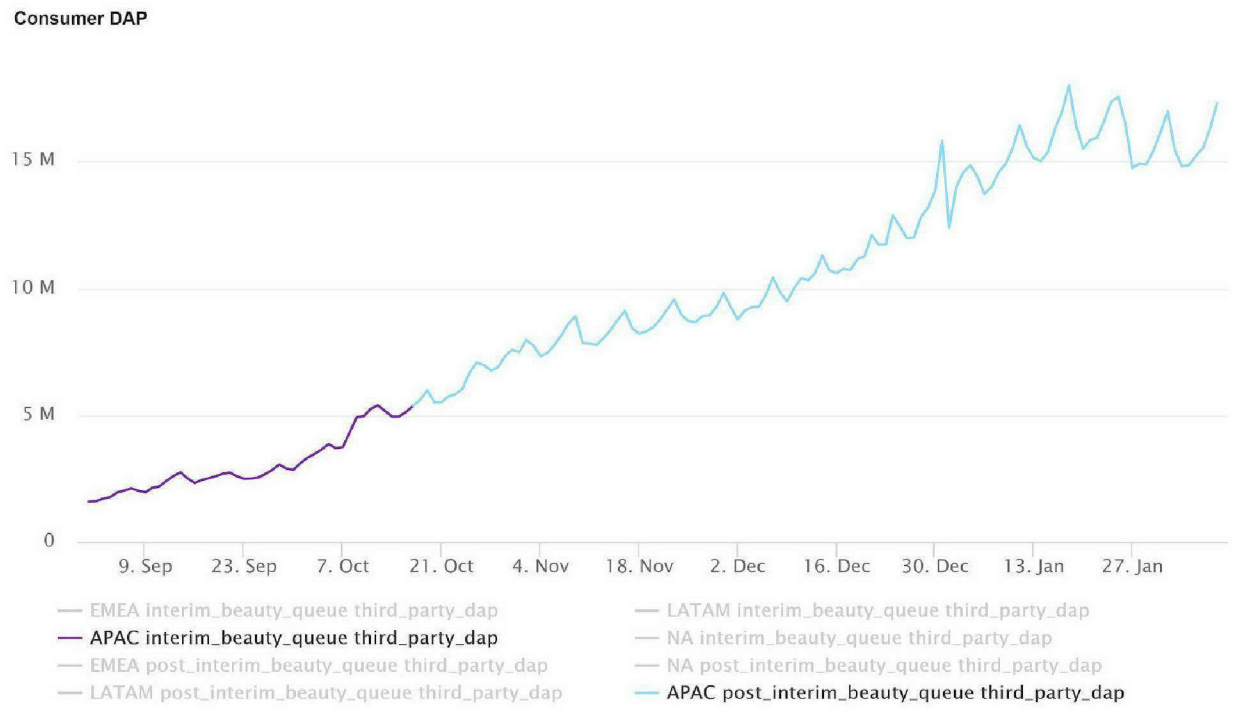
Facewarp Usage as a % of AR Usage: Query

Creator Growth pre/post policy: Query



APAC Creator Growth pre/post policy: Query



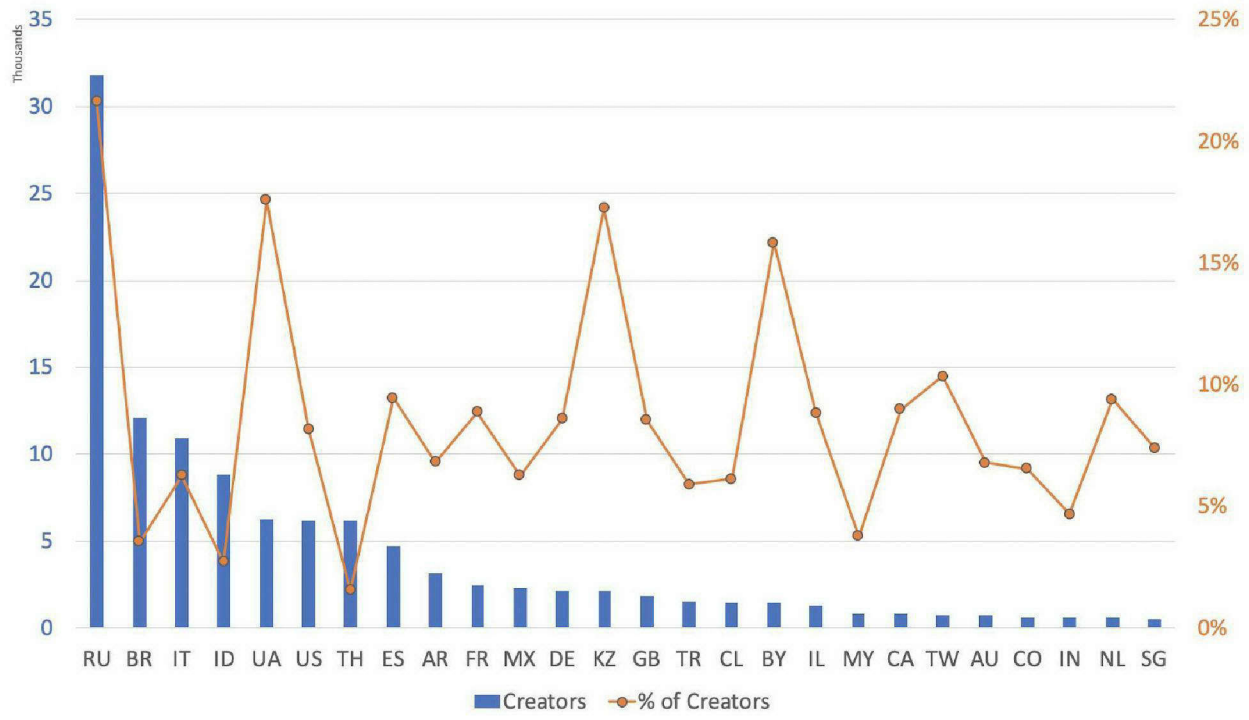


Top Effects: Until the interim policy launched in October, 0% of Cosmetic Surgery AR Effects were ranked in the Top 100 effects by try-ons.

Additional Effect Creation:
The data shows that creators who created Cosmetic Surgery AR Effects and who did not stop creating did so on average more often than their counterparts who did not create Cosmetic Surgery Effects.

- 42% of creators using "face warp" (any effect that distorts the face that includes, but is not limited to, Cosmetic Surgery Effects) ended up publishing other effects, which is in line with overall creator retention behaviour, so there is no reason to believe that creators might churn at a higher rate because of Cosmetic Surgery policy limitations.

Effects Usage: On average Cosmetic Surgery Effects have a slightly higher share+save rate than third party non-Cosmetic Surgery Effects (15% vs 13%).



US Accounts for 5% of the entire creator ecosystem. Russia @ 26%. About 12% of creators from the US have ever created an effect broadly categorized as *Beauty*. 30% of effect creators in the #1 market has made an effect categorized as *Beauty*.

Artists who built Beauty effects are primarily female (68% vs 43% for overall population), and come from our top creator markets: 41% of beauty creators are from Russia.

APAC Specific Camera Analysis

Historical analysis from the Local APAC team about specific camera trends.

- Asian beauty & makeup analysis - insights on Asian beauty/makeup as well as competitive insights.
- Asian camera effects trend - Asian camera trend in 4 focus areas in 2017
- Seoul Art team
- APAC Camera Effects Hold Out Test - 2018: A month-long holdout test resulted in a top line 8% increase in camera shares with masks