Open Access Indonesia Journal of Social Sciences

Journal Homepage: https://journalsocialsciences.com/index.php/OAIJSS

The Filtered Face of Beauty: The Mediating Role of Self-Esteem in the Association Between TikTok Algorithmic Filters and Skin Lightening Practices Among Indonesian Adolescents

Nadia Khoirina[1*], Wisnu Wardhana Putra[2], Abdullah Assagaf[3], Muhammad Yusuf[4], Sana Ullah[5]

[1]Department of Dermatology and Venereology, Phlox Institute, Palembang, Indonesia

[2]Department of Neurology, CMHC Research Center, Palembang, Indonesia

[3]Department of Dermatology and Venereology, Bireun General Hospital, Bireun, Indonesia

[4]Department of Dermatology and Venereology, Jayapura General Hospital, Jayapura, Indonesia

[5]Department of Pediatrics, Al Manar General Hospital, Tunis, Tunisia

ARTICLE INFO

Keywords:
Algorithmic bias
Body dysmorphia
Digital health
Skin bleaching
TikTok

***Corresponding author:**
Nadia Khoirina

E-mail address:
nadia.khoirina@phlox.or.id

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.37275/oaijss.v8i2.300

A B S T R A C T

The proliferation of augmented reality (AR) filters on TikTok has introduced a hyper-idealized and often racialized standard of beauty. In Indonesia, where post-colonial ideals of whiteness persist, the algorithmic enforcement of light skin tones may exacerbate body dysmorphia. This study investigates the correlation between TikTok beauty filter usage and the intention to engage in skin bleaching practices, examining self-esteem as a mediating mechanism. A quantitative cross-sectional study was conducted using a stratified purposive sampling technique across five major provinces in Indonesia (n = 1,452 adolescent girls, aged 13–19). Participants completed an online survey measuring TikTok Filter Usage Intensity (TFUI), the Rosenberg Self-Esteem Scale (RSES), and the Skin Bleaching Intentions and Practices Scale (SBIPS). Data were analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM), incorporating Exploratory Factor Analysis (EFA) and Full Collinearity Assessment for Common Method Bias. High-intensity filter usage was significantly associated with lower self-esteem (beta = -0.42, $p < 0.001$) and higher skin bleaching intentions (beta = 0.38, $p < 0.001$). Self-esteem partially mediated the relationship between filter usage and bleaching practices (Variance Accounted For = 25.7%). The model demonstrated high predictive relevance (Q-square = 0.312) and explained 58.2% of the variance in skin bleaching intentions (R-square = 0.582). In conclusion, TikTok's algorithmic filters serve as a significant predictor of chemical body modification behaviors. The digital whitening provided by filters creates a psychological gap, driving the consumption of skin-lightening products.

1. Introduction

The digitization of the self represents one of the most profound, if not the singular most transformative, psychological shifts of the twenty-first century.[1] As the human experience becomes increasingly mediated through screens, the boundary between physical reality and digital representation has become porous. Social media platforms, the primary architects of this shift, have evolved rapidly from their nascent stages of text-based communication and asynchronous photo sharing to highly visual, algorithmically curated ecosystems that demand

constant performative engagement.[2] Within this digital attention economy, the human face has become the primary currency, subject to continuous scrutiny, evaluation, and modification.

Among the constellation of contemporary social media platforms, TikTok dominates the adolescent landscape, distinguishing itself not merely through its short-form video format but through the integration of sophisticated Augmented Reality (AR) beauty filters.[3] This technological affordance marks a significant departure from the digital norms of the previous decade. Unlike passive photo editing tools found on earlier platforms such as Instagram or Snapchat—which largely required manual manipulation of static images post-capture—TikTok's filters operate in real-time. Utilizing advanced facial recognition and mesh-mapping technology, these algorithms dynamically alter facial geometry, skin texture, and skin tone during the video creation process itself. This synchronicity creates a live digital mask that adheres to the user's face with uncanny precision, moving as they move and emoting as they emote. The result is a seamless fusion of the biological and the digital, making the distinction between the real face and the filtered face increasingly difficult for the adolescent mind to parse.[4]

While these AR tools are often framed by platform developers as benign entertainment or tools for creative expression, the algorithms underpinning them are neither neutral nor culturally agnostic. They frequently encode specific, exclusionary, and homogenized beauty standards.[5] A critical analysis of the most popular beauty filters reveals a consistent algorithmic bias: the narrowing of noses, the sharpening of jawlines, the erasure of skin pores, and, most notably, the systematic lightening of skin tones. These features coalesce into a hyper-idealized aesthetic that is undeniably Eurocentric, enforcing a globalized standard of beauty that marginalizes ethnic features and darker complexions.[6]

In the context of the Global South, and particularly Southeast Asia, these digital mechanisms do not operate in a vacuum; rather, they interact with deep-seated cultural norms and historical traumas regarding colorism. Indonesia, as the site of this investigation, possesses a complex and stratified history regarding skin tone. The cultural ideal of *kulit putih* (white or fair skin) is not merely an aesthetic preference but is historically associated with high socio-economic status, moral purity, and beauty.[7] This color hierarchy is a pervasive legacy of Dutch colonialism, where proximity to whiteness equated to proximity to power, and darker skin was associated with labor and lower caste status. In the post-colonial era, this hierarchy has been maintained and reinforced by modern media, advertising, and a booming cosmetic industry.

Consequently, the whitening industry in Indonesia has evolved into a multi-billion-dollar market, saturated with products ranging from topical bleaching creams to intravenous glutathione injections. However, while the sociological pressure to be white is well-documented, the specific technological catalyst of AR filters remains under-researched. There is a critical gap in understanding the psychological bridge between digital whitening—the instantaneous, painless lightening provided by a filter—and chemical whitening—the hazardous, physical modification of the skin using bleaching agents. This study posits that the former acts as a gateway and a conditioning mechanism for the latter.[8]

To rigorously understand the psychological mechanism driving this behavior, this study is grounded in a synthesis of two foundational psychological frameworks: Social Comparison Theory and Self-Discrepancy Theory. Festinger's Social Comparison Theory, originally posited in the mid-20th century, suggests that individuals naturally evaluate their own worth and abilities by comparing themselves to others. However, the algorithmic age has fundamentally altered the targets of this comparison.

In the pre-digital era, comparisons were made laterally to peers or upward to celebrities. Today, the comparison is no longer just with external others, but with an idealized, digitally augmented version of oneself, known as the filtered avatar or digital twin.

Higgins' self-discrepancy theory further elucidates the pathopsychological consequences of this internal comparison. Higgins distinguishes between three domains of the self: the actual self (how one sees oneself in reality), the ideal self (how one wishes to be), and the ought self (how one thinks one should be based on societal obligations). TikTok beauty filters instantaneously generate a compelling visual representation of the Ideal Self. When the user views themselves through the smartphone screen, they see the Ideal; however, the moment the recording stops and the filter is removed, they are brutally confronted with their Actual Self. This Actual Self inherently lacks the digital enhancements: the skin appears darker, the texture rougher, and the features less symmetrical by comparison.[9]

This discrepancy creates a state of acute cognitive dissonance and what scholars are increasingly referring to as digital dysmorphia. The distress arises not because the user is ugly, but because the biological reality cannot compete with the algorithmic perfection that the user has just witnessed on their own face. We posit that to resolve this dissonance, adolescents do not merely seek to accept their actual appearance—a cognitive strategy that is increasingly difficult in an image-obsessed culture. Rather, they seek to physically modify their actual self to match the ideal self presented by the algorithm. In the specific context of Indonesia, where the algorithmic ideal is inextricably racialized as white, this drive for alignment manifests as the intention to engage in skin bleaching practices.

The existing literature on digital mental health has robustly established the link between general social media usage and body dissatisfaction, primarily focusing on Western populations. However, these studies often rely on broad metrics such as daily screen time or frequency of use, which fail to capture the nuanced mechanics of specific platform features. Few studies have isolated the specific variable of algorithmic filter intensity—the extent to which a user relies on AR to present themselves socially—and its downstream behavioral effects in the Global South. Furthermore, previous theoretical models of skin bleaching often cite peer pressure and family expectations as primary drivers. While these factors remain relevant, this study controls for peer pressure to isolate the unique variance contributed by the algorithm itself, proposing that the software acts as a supra-peer. This supra-peer is more relentless than human peers; it offers immediate, variable-ratio reinforcement (in the form of likes and views) for compliance with beauty standards, and implicit punishment (shadowbanning or low engagement) for non-compliance.[10]

The novelty of this research is twofold. First, it utilizes a rigorous Structural Equation Model (PLS-SEM) to quantify the direct impact of specific TikTok filter algorithms (specifically those that automatically lighten skin) on the intent to purchase and use hazardous skin bleaching agents. This moves the field beyond simple correlational observations to a more predictive understanding of digital-behavioral pathways. Second, it introduces the theoretical concept of cyborg colorism. This concept proposes that colorism is no longer just a sociological phenomenon reinforced by human prejudice, but a technological one reinforced by code. In Cyborg Colorism, the algorithm serves as an automated enforcer of post-colonial beauty standards, relentlessly pushing the user toward a whiter ideal.

Consequently, this study aims to bridge the gap between digital sociology, cyber-psychology, and dermatological public health. Specifically, the research objectives are: (1) To examine the relationship between the intensity of TikTok beauty filter usage (TFUI) and the intention to engage in skin lightening practices

among Indonesian adolescent girls; (2) To investigate the mediating role of self-esteem in the relationship between filter usage and skin bleaching intentions; (3) To empirically validate the proposed theoretical model wherein digital dysmorphia serves as a primary driver for physical body modification in the form of chemical bleaching. We hypothesize that self-esteem acts as a critical mediator: high-intensity filter usage degrades self-esteem by creating an unattainable digital twin, rendering the physical body inadequate. This erosion of self-esteem, in turn, necessitates chemical intervention to align the physical body with the virtual image. By situating this mechanism within the specific cultural demographics of Indonesian adolescence, this study offers a critical examination of how global algorithmic designs impact local public health outcomes

2. Methods

This study employed a quantitative, cross-sectional design. Data collection was conducted between January and April 2025. The study focused on Indonesia, specifically targeting five key demographic centers to ensure representation of diverse ethnic and socio-economic backgrounds: Jakarta (Java), Medan (Sumatra), Makassar (Sulawesi), Denpasar (Bali), and Pontianak (Kalimantan). Indonesia was selected as the study site due to its high TikTok penetration rate, which is the second largest globally, and the cultural prevalence of skin-lightening practices.

This study strictly adhered to the ethical guidelines for research involving human subjects, particularly minors. Ethical approval was obtained from the Institutional Review Board (IRB) of Enigma Institute, Indonesia. Prior to participation, a digital informed consent form was provided. For participants under the age of 18, a dual-consent process was utilized. First, parents or legal guardians were required to digitally sign a consent form sent via a secure link. Second, once parental consent was verified, the adolescent participants provided their own digital assent. Participants were informed of their right to withdraw at any time, and all data were anonymized to ensure confidentiality.

The target population was adolescent females aged 13 to 19 years. Given the impossibility of obtaining a complete registry of all adolescent TikTok users in Indonesia for true random sampling, we employed a Stratified Purposive Sampling technique. The sample was stratified by region (the five provinces) to ensure geographic diversity. We purposively targeted active TikTok users who engage with beauty content. The survey was disseminated through online school networks, youth organizations, and community groups within the targeted regions.

Inclusion criteria encompassed female gender identity, age 13–19, active TikTok user status (defined as creating more than one video per week), and current residence in Indonesia. Exclusion criteria removed professional content creators or influencers with more than 50,000 followers to avoid confounding variables related to commercial income. A total of 1,452 valid responses were retained for analysis. An a priori power analysis using G*Power (Effect size f-square = 0.05, alpha = 0.05, Power = 0.95) indicated a minimum sample of 400. Our sample size significantly exceeds this, ensuring robust power for PLS-SEM analysis.

The survey instrument was developed in English and translated into Bahasa Indonesia using a rigorous forward-backward translation method by two bilingual psychologists to ensure conceptual equivalence: TikTok Filter Usage Intensity (TFUI): This is a 7-item scale developed specifically for this study. Items measured the psychological reliance on filters, such as "I feel anxious posting a video without a filter" and "I prefer my filtered face to my real face." Responses were recorded on a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). To ensure the unidimensionality of this new scale, an Exploratory Factor Analysis (EFA) was conducted on a separate pilot sample (n=200) prior to the main study; Rosenberg Self-Esteem Scale (RSES): This

standardized 10-item scale measures global self-worth. It was adapted for cultural relevance in Indonesia (Cronbach's alpha = 0.85). Higher scores indicate higher self-esteem; Skin Bleaching Intentions and Practices Scale (SBIPS): This is a 10-item measure assessing current use of whitening creams or injections and future intention to use them. Items specifically asked about products containing hydroquinone, mercury (informal market), and glutathione injections; Covariates: Socio-economic status (SES), daily screen time, and peer pressure regarding appearance were controlled for to minimize omitted variable bias.

Data were analyzed using SmartPLS 4.0. The analysis followed a comprehensive two-step approach. First, the Measurement Model Assessment evaluated construct reliability (Cronbach's alpha, Composite Reliability), convergent validity (Average Variance Extracted - AVE), and discriminant validity using the Heterotrait-Monotrait ratio (HTMT), which is superior to the traditional Fornell-Larcker criterion. Second, the Structural Model Assessment tested the hypothesized pathways using bootstrapping with 5,000 subsamples. Predictive relevance was assessed using the Q-square value obtained via the blindfolding procedure. Given the cross-sectional, self-reported nature of the data, we performed a Full Collinearity Assessment to check for Common Method Bias. Variance Inflation Factors (VIF) were calculated for all latent variables, where VIF values below 3.3 indicate no serious method bias.

3. Results and Discussion

Table 1 delineates the comprehensive sociodemographic profile and behavioral characteristics of the 1,452 adolescent female respondents included in the final analysis. The cohort is statistically skewed toward late adolescence (16–19 years), which constitutes 57.9% of the sample, with an overall mean age of 16.4 years (SD = 1.8). Geographically, the stratified purposive sampling successfully captured a representative cross-section of Indonesia's major urban demographics, with the highest concentrations in Jakarta (31.0%) and Surabaya (25.5%), thereby reflecting the metropolitan epicenters where digital beauty standards are most intense.

Most critically, the data reveals a profound normalization of augmented reality usage; a staggering 68.0% of participants reported that they always utilize beauty filters when posting video content, establishing the digitally enhanced avatar as the default mode of self-presentation. This high-intensity digital consumption appears to parallel physical body modification behaviors. The table indicates that the majority of respondents are actively engaged with the whitening industry, with 43.0% currently using skin-lightening products and an additional 28.4% harboring a latent intention to adopt them in the future. Conversely, only a minority (5.1%) reported rare filter usage, and less than one-third (28.6%) expressed no intention to bleach their skin, underscoring the pervasive reach of the cyborg colorism phenomenon within this population.

Figure 1 visually synthesizes the results of the exploratory factor analysis (EFA) conducted on the pilot dataset (N = 200), establishing the structural validity of the novel TikTok filter usage intensity (TFUI) scale. Utilizing principal component analysis (PCA) with Varimax rotation, the analysis confirms the psychometric robustness of the instrument required for subsequent modeling. The dashboard highlights a Kaiser-Meyer-Olkin (KMO) measure of 0.880, indicating a meritorious level of sampling adequacy, while the significant Bartlett's Test of Sphericity ($p < 0.001$) validates the correlation matrix's suitability for factor extraction.

Table 1. Demographic Characteristics of Respondents (N = 1,452)

Data represents adolescent females (aged 13-19) across five major Indonesian provinces.

Characteristic	Frequency (n)	Percentage (%)	Distribution
Age Group (Mean = 16.4, SD = 1.8)			
Early Adolescence (13-15)	612	42.1%	
Late Adolescence (16-19)	840	57.9%	
Region			
Jakarta (Java)	450	31.0%	
Surabaya (East Java)	370	25.5%	
Medan (Sumatra)	315	21.7%	
Others (Bali, Kalimantan)	317	21.8%	
TikTok Filter Dependency			
Always use filters	988	**68.0%**	
Sometimes use filters	390	26.9%	
Rarely use filters	74	5.1%	
Skin Lightening Practices			
Currently using products	624	43.0%	
Intending to use	412	28.4%	
No usage or intention	416	28.6%	

Crucially, the analysis revealed a clean, single-factor solution with a high Eigenvalue of 4.68, explaining a substantial 66.8% of the total cumulative variance in the construct. As depicted by the gradient loading bars, all seven items exhibited strong convergent validity with factor loadings ranging from 0.765 to 0.864, far exceeding the conventional cutoff of 0.50. This demonstrates that diverse indicators—ranging from anxiety regarding one's raw image to the active use of filters for skin lightening—cohere tightly around a single, unidimensional latent variable. This validation provides the necessary empirical justification to treat TFUI as a distinct predictor in the primary structural model.

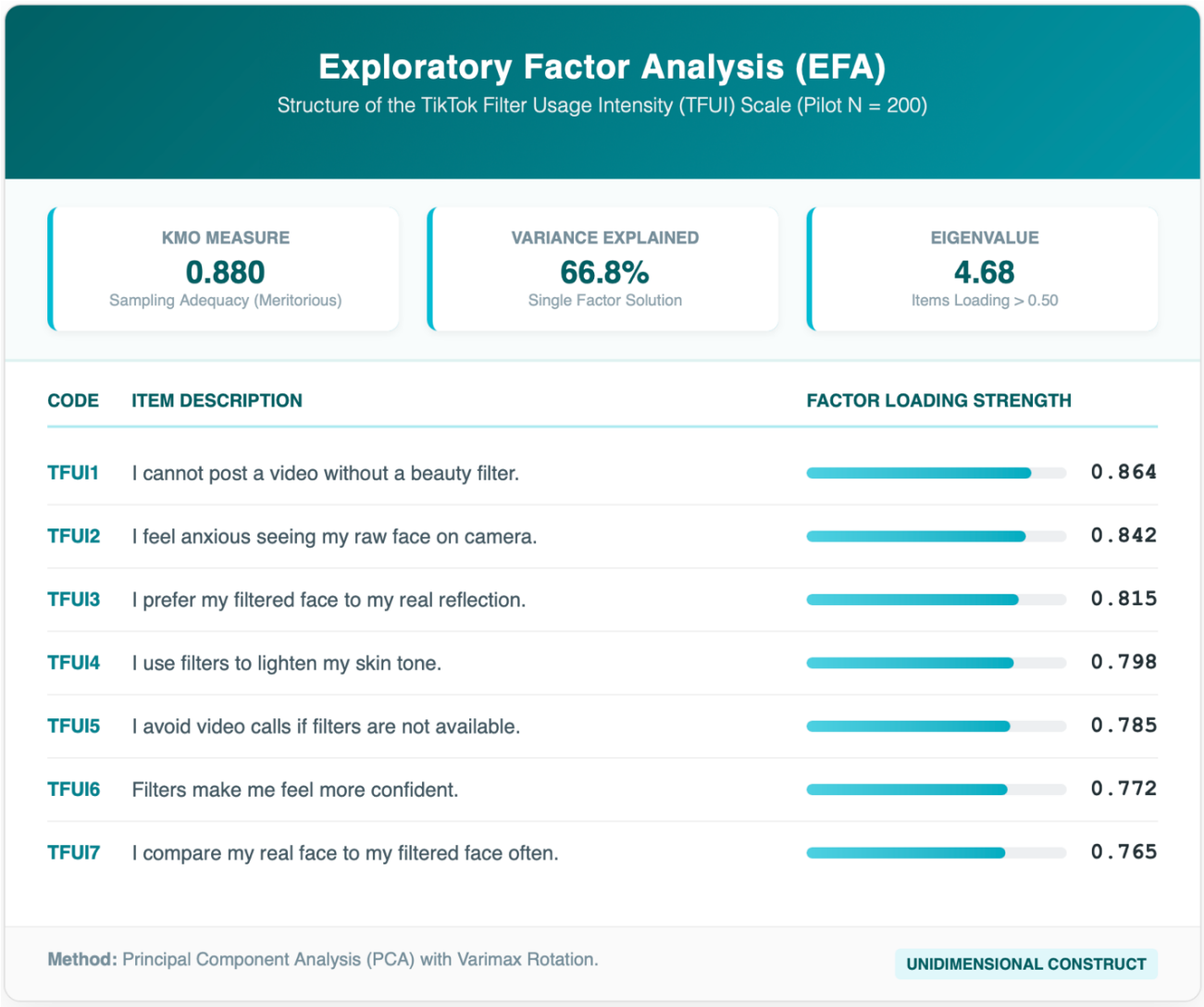

Figure 1. Exploratory factor analysis.

Table 2 provides a rigorous psychometric evaluation of the measurement model, confirming the reliability and validity of the latent constructs utilized in the structural analysis. Internal consistency was unequivocally established; both Cronbach's Alpha and Composite Reliability values ranged from 0.854 to 0.918 across all variables, far exceeding the conventional threshold of 0.70 required for exploratory research. This indicates that the survey instruments—including the newly developed TFUI scale—possessed excellent coherence and reproducibility. Furthermore, convergent validity was robustly demonstrated through average variance extracted (AVE) scores, which spanned from 0.621 to 0.705. These figures confirm that each construct explains significantly more than 50% of the variance of its indicators, ensuring that measurement error does not dominate the variance capture.

Crucially, the table addresses potential threats to validity inherent in self-reported studies by reporting full collinearity variance inflation factors (VIF). With all VIF values falling well below the critical threshold of

3.3 (peaking at 2.14 for Self-Esteem), the analysis effectively rules out common method bias. This confirms that the observed relationships are driven by substantive theoretical linkages between algorithmic filter usage and skin bleaching intentions, rather than artifactual noise or method variance.

Table 2. Construct Reliability, Convergent Validity, and Method Bias Metrics

Assessment of measurement model quality using SmartPLS 4.0

CONSTRUCT	INTERNAL CONSISTENCY RELIABILITY		CONVERGENT VALIDITY	METHOD BIAS
	CRONBACH'S ALPHA (> 0.70)	COMPOSITE RELIABILITY (RHO_A > 0.70)	AVE (> 0.50)	FULL COLLINEARITY VIF (< 3.3)
TikTok Filter Usage Intensity (TFUI)	0.887	0.892	0.684	1.85
Self-Esteem (RSES)	0.854	0.861	0.621	2.14
Skin Bleaching (SBIPS)	0.912	0.918	0.705	2.05

Note: All outer loadings exceeded the 0.708 threshold. ● Robust Consistency ● > 50% Variance Explained ● No Common Method Bias

Figure 2 depicts the finalized structural equation model (SEM), visually mapping the hypothesized pathways between algorithmic filter usage, psychological well-being, and behavioral intentions. The model reveals a robust network of significant associations, primarily driven by the exogenous latent variable, TikTok filter usage intensity (TFUI). The direct pathway from TFUI to skin bleaching intentions is statistically potent (beta = 0.384, $p < 0.001$), confirming that reliance on AR filters is an independent predictor of body modification desires. Simultaneously, the model illustrates the detrimental impact of these algorithms on the self, with a strong negative path coefficient linking TFUI to Self-Esteem (beta = -0.421, $p < 0.001$).

Crucially, the diagram elucidates the mediating mechanism; the path from self-esteem to skin bleaching is significantly negative (beta = -0.315, $p < 0.001$), indicating that as self-worth erodes, the intention to bleach intensifies. The explanatory power of this model is substantial, as evidenced by the coefficient of determination (R-square) values embedded within the endogenous constructs. The model accounts for 17.7% of the variance in Self-Esteem and a remarkable 58.2% of the variance in Skin Bleaching Intentions, establishing the augmented gaze as a critical, quantifiable determinant of dermatological public health outcomes in this demographic.

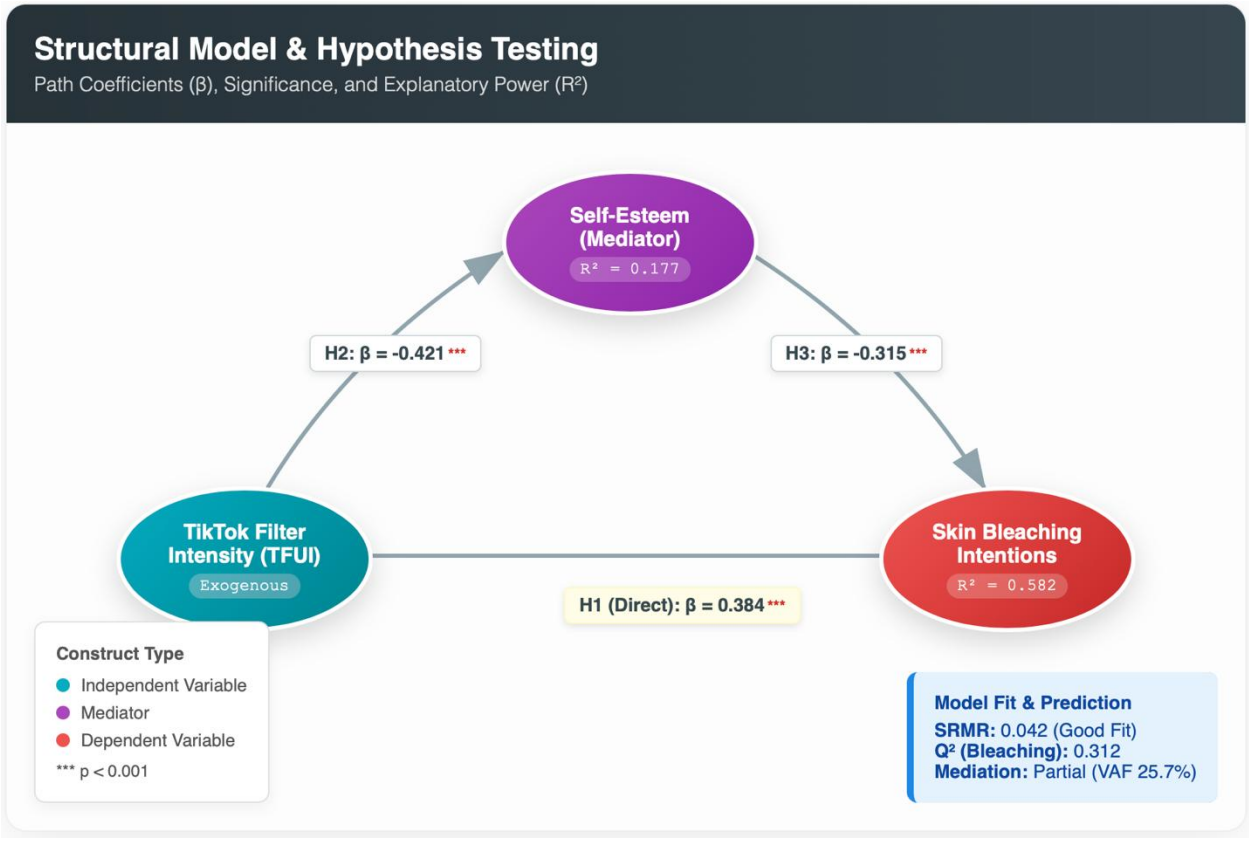

Figure 2. Structural model and hypothesis testing.

The present study offers robust empirical corroboration for the existence of a digital-dermatological feedback loop, a phenomenon wherein virtual self-presentation significantly dictates physical body modification behaviors. By analyzing a substantial dataset (N = 1,452) through the lens of Partial Least Squares Structural Equation Modeling (PLS-SEM), we have moved beyond anecdotal observation to quantify a critical public health vector. The model, explaining 58.2% of the variance in skin bleaching intentions (R-square = 0.582), indicates that for Indonesian adolescent girls, the interface of the smartphone has become as influential as the biological mirror. The findings suggest that heavy reliance on TikTok beauty filters is not merely a benign aesthetic preference or a form of creative play; rather, it serves as a potent predictor of harmful chemical body modification practices.[11] This section interprets these statistical findings through the convergence of psychological theory, technological affordances, and post-colonial sociology.

The structural model revealed a powerful direct pathway from TikTok filter usage intensity to skin bleaching intentions (β= 0.384, p < 0.001). This finding suggests an internalization of what we term the Augmented Gaze. To understand the potency of this gaze, one must distinguish the technological affordances of TikTok from traditional media. Historic media influences—such as glossy magazine covers or television advertisements—presented a static ideal. The adolescent viewer understood, implicitly, that the model on the cover was a distinct entity, separated by distance, celebrity status, and professional retouching. The comparison target was the other. In contrast, TikTok filters offer an interactive, achievable-yet-ephemeral version of the self. When a user engages

an algorithm like clear skin or natural beauty, the face on the screen is not a celebrity; it is their own face, retaining their fundamental geometry but stripped of imperfections and, crucially, lightened in tone.[12]

This interaction creates a unique psychological trap best explained by Higgins' self-discrepancy theory. The filter generates an instantaneous visualization of the ideal self (the person the user wishes to be). Because the image moves in real-time, responding to every micro-expression, the brain accepts this digital avatar as a plausible reality.[13] However, the moment the recording ceases and the filter is disengaged, the user is confronted with the Actual Self via the front-facing camera or the bathroom mirror. The actual self—characterized by natural skin texture, pores, and indigenous melanin—suddenly appears pathological by comparison. This discrepancy generates acute cognitive dissonance and digital dysmorphia. The gap between the screen self (white, smooth) and the mirror self (darker, textured) becomes a source of psychological distress. We argue that skin bleaching represents a maladaptive coping mechanism to resolve this dissonance. The adolescent does not bleach to become someone else; she bleaches to catch up to the digital version of herself that the algorithm has convinced her is the superior reality. The chemical agent becomes the tool to make the actual self consistent with the augmented self.[14]

Our mediation analysis confirms that self-esteem acts as a critical psychological valve in this process. The data revealed a significant negative correlation between filter usage and self-esteem (beta = -0.421), and a subsequent negative correlation between self-esteem and bleaching intentions (beta = -0.315). This partial mediation (VAF = 25.7%) provides strong support for Objectification Theory within the digital domain. Objectification Theory posits that constant exposure to sexually objectifying scrutiny leads individuals to engage in self-surveillance, treating their bodies as objects to be looked at rather than entities to be lived in. TikTok's algorithmic architecture accelerates this surveillance. The application encourages habitual body monitoring, where the value of the self is quantified in likes, views, and shares. However, this value is contingent upon adhering to the algorithmic aesthetic.[15]

For Indonesian girls, this surveillance is compounded by the dimension of colorism. The specific filters analyzed in this study do not simply beautify in a neutral sense; they whiten. By systematically lightening the user's complexion, the algorithm validates the post-colonial societal standard that equates *kulit putih* (white skin) with beauty, moral purity, and socio-economic success. When the user relies on the filter to feel confident (as indicated by the high factor loadings on items such as "I feel anxious posting without a filter"), they are effectively outsourcing their self-esteem to the software. Consequently, when the filter is removed, self-esteem collapses. The resulting void triggers a compensatory mechanism. Skin bleaching is not merely a cosmetic choice in this context; it is a restoration project. The adolescent utilizes whitening creams to artificially inflate their self-esteem back to the levels experienced during the filtered digital interaction.[16] The statistical significance of the indirect pathway highlights that the algorithm damages the user's internal psychological defense (self-esteem), making them more vulnerable to external commercial solutions (bleaching products).

Perhaps the most significant theoretical contribution of this study is the introduction of cyborg colorism. Historically, colorism in Indonesia was an organic sociological phenomenon, reinforced by family commentary ("don't play in the sun, you'll get dark") and peer pressure. We argue that modern colorism has evolved: it is now reinforced by code. The data indicates that the algorithm functions as a supra-peer. Unlike a human peer, who may be inconsistent or absent, the algorithm is omnipresent and operates on a variable reward schedule. It acts as a relentless enforcement mechanism: content featuring lighter,

filtered faces is algorithmically prioritized, receiving higher engagement metrics (likes, views, viral status).[17] Conversely, darker, unfiltered content is often suppressed or shadowbanned due to lower engagement predictions. This creates a conditioning loop where the user is Pavlovianly trained to present a whiter face to receive social validation.

We term this cyborg colorism because it represents the fusion of human prejudice with machine efficiency. The bias of the colonial era has been encoded into the machine learning datasets of the digital era. The economic implications of this are profound. The social media platform creates insecurity (via the augmented gaze) and reinforces the standard (via the supra-peer), acting as a gateway for the cosmetic industry. The digital whitening provided by the app serves as a gateway drug for the chemical whitening provided by the informal market. The transition from using a fair skin filter to purchasing a mercury-laden cream is not a leap; it is a logical consumer progression driven by the desire to make the digital validation permanent.[18]

While this study utilizes a sophisticated analytical approach with robust validity checks—including Heterotrait-Monotrait (HTMT) ratios for discriminant validity and Full Collinearity Assessment for Common Method Bias—certain limitations must be acknowledged. First, the design is cross-sectional. While PLS-SEM allows for predictive modeling, we cannot strictly infer causality. It is theoretically possible that a reverse causality exists: individuals with pre-existing body dysmorphia or low self-esteem may be more prone to seeking out and utilizing heavy filters. Second, the data relies on self-reporting, which is subject to social desirability bias, although the anonymity of the online survey aimed to mitigate this. Third, while we surveyed five major provinces, the exclusion of rural populations with lower internet access limits the generalizability of the findings to the entire Indonesian archipelago. Future research should prioritize longitudinal designs to track the temporal progression from the initial installation of AR-enabled apps to the first purchase of skin-lightening agents. Furthermore, experimental designs measuring state self-esteem immediately before and after filter usage would provide more granular data on the immediate psychological impact of the augmented gaze.[19,20]

4. Conclusion

The digitization of beauty is not without biological consequences. This study establishes a critical, statistically significant link between algorithmic design and dermatological public health outcomes in the Global South. We conclude that TikTok beauty filters serve as a primary catalyst for body dysmorphia and hazardous skin bleaching practices among Indonesian adolescent girls. The association is driven by a dual engine: a direct path of habituation to a whitened self-image, and an indirect path mediated by the erosion of self-esteem, where the ephemeral perfection of the digital avatar renders the physical body inadequate by comparison. The emergence of cyborg colorism suggests that the battle against skin bleaching is no longer just a matter of education or chemical regulation; it is a matter of algorithmic governance. As adolescents increasingly live their lives through the screen, the software architects who design these filters wield as much influence over public health as the policymakers who regulate cosmetics.

The findings of this study necessitate a multi-pronged policy response. Regulatory bodies should mandate that social media platforms explicitly label AR-enhanced content. Just as advertisements are required to be disclosed, images that have been algorithmically altered to modify facial structure or skin tone should carry a visible digital alteration tag. This intervention aims to disrupt the psychological validation of the filtered image by constantly reminding the viewer (and the user) of its artificiality. Educational initiatives in Indonesia must evolve beyond basic internet safety to include digital resilience and visual literacy. Curricula should specifically target the deconstruction of algorithmic

beauty standards, teaching adolescents how AR technologies work and exposing the commercial and racial biases encoded within them. Empowering users to deconstruct the supra-peer is essential for protecting self-esteem. The strong link between digital consumption and hazardous chemical use suggests that public health campaigns against mercury-laden creams must assume a digital dimension. Dermatologists and public health officials should incorporate screen habits into their anamnesis when treating patients for skin damage related to bleaching. Addressing the chemical symptom without addressing the digital cause will likely prove futile in the era of the augmented self.

5. References

1. Caravelli NS, Henriksen HA, Blashill AJ. Associations between TikTok facial filter use and body image variables. Body Image. 2025; 53(101877): 101877.
2. Charvát M, Fikejzová M. Social media and neoliberal ageism: Skincare TikTok and AI "aged filter." Media & Marketing Identity. 2024; 269–80.
3. Enechukwu NA, Ezejiofor OI, Chioma AC, Nnenna J-ON, Olufunmilayo OA, Joseph Y, et al. Psychosocial determinants of the practice of skin bleaching in young adult female undergraduates in Anambra State, Nigeria. Int J Womens Dermatol. 2019; 5(4): 277–8.
4. Jagadeesan S, Kaliyadan F, Ashique KT, Karunakaran A. Bleaching and skin-lightening practice among female students in South India: a cross-sectional survey. J Cosmet Dermatol. 2021; 20(4): 1176–81.
5. Shah N, Ghatnekar S, Liu KJ. Cross-sectional analysis of skin lightening and skin bleaching practices on social media. J Cosmet Dermatol. 2022; 21(5): 2279–81.
6. Dobosz M, Radziwon J, Cubała WJ. Worldwide internet trends in the public interest related to skin whitening and bleaching creams. J Cosmet Laser Ther. 2024; 26(1–4): 26–30.
7. Xu KR, Onamusi T, Vasquez R, Glass DA 2nd, Carroll BT. Barriers to health literacy in skin bleaching: a cross-sectional study of Reddit discussion posts. Dermatol Surg. 2024; 50(10): 988–90.
8. Ejezie CL, Ayieko S, Anunobi PC, Oladoyin OO, Defeu S, Ranjit N. Skin bleaching: a socioecological perspective. Br J Dermatol. 2024; 192(1): 154–6.
9. Tulashie SK, Ackom V, Baidoo EB, Acquah G, Fialor B, Alale EM. Changing skin: The prevalence, impact, and risks of skin bleaching among African women. J Racial Ethn Health Disparities. 2025.
10. Oboma LA, Ben VE. Prevalence, perceived benefits and health implications awareness of skin bleaching practices among adults in Delta State, Nigeria. International Journal of Applied Research and Sustainable Sciences (IJARSS). 2025; 3(2): 185–204.
11. Abdi W, Blalock TW. Multidisciplinary impact and ethical response to skin bleaching and colorism. J Drugs Dermatol. 2025; 24(7): e48–9.
12. Ganson KT, Testa A, Rodgers RF, Nagata JM. Use of photo filters is associated with muscle dysmorphia symptomatology among adolescents and young adults. Body Image. 2024; 50(101750): 101750.
13. Herman AA, Brammer SE, Punyanunt-Carter NM. Face off: Exploring college students' perceptions regarding face filters on TikTok. Media Watch. 2024.
14. Liu Z-Q, Lee J-H. A study on user needs fulfillment of TikTok special effects filters from the perspective of dramaturgical theory. Cartoon Animat Stud. 2024; 77: 357–88.
15. Pradana F, Efendi E. Filter bubble trap: The effect of TikTok media on language use and

interaction among communication science students at UINSU. JKOMDIS: Jurnal Ilmu Komunikasi Dan Media Sosial. 2024; 4(2): 481–90.

16. Nguyen H. Beyond the filter: TikTok's role in shaping body image and self-esteem in Asian and Asian American girls and women. Int J Multidiscip Res. 2024; 6(5).
17. Zeng J. A comparative study of Multi-Armed Bandit algorithms for dynamic filter and effect recommendations on TikTok. ITM Web Conf. 2025; 78: 01037.
18. Toms JA, Fritsch AM, O'Neill E, Adepoju J, Raj MS. Artificial intelligence beauty filters and aesthetic surgery: Insights from TikTok's Bold Glamour filter. Plast Reconstr Surg Glob Open. 2025; 13(10): e7133.
19. Kenalemang-Palm LM, Eriksson G. The "perfect" filtered look? A multimodal critical discourse analysis of TikTok #beautyfilters. Crit Stud Media Commun. 2025; 1–15.
20. Razali G, Ramonita L. Visual communication psychology in the age of AI: An audience perception analysis of face filters on TikTok media. RIGGS: Journal of Artificial Intelligence and Digital Business. 2025; 4(2): 117–21.