



SAFEGUARDING ONLINE SPACES (SOS) STUDY



IN COLLABORATION WITH GLOBAL SHAPERS COMMUNITY - SINGAPORE HUB
COPYRIGHT © 2024 SG HER EMPOWERMENT LTD. ALL RIGHTS RESERVED.
NOT TO BE REPRODUCED WITHOUT PRIOR PERMISSION FROM SHE.

contactus@she.org.sg

TABLE OF CONTENTS

Topline Findings	5
About This Study	6
Key Findings	
Key Concerns and Attitudes	7
Key Safety Gaps	9
Algorithm-Driven Content	11
Cancelling	15
Generative Artificial Intelligence	21
Youth Perspectives on Solutions	24
Recommendations	26

FOREWORD BY SHE CEO

Everyone – whether female or male – is entitled to safety and respect online.

SHE's work in this space is guided by our commitment to engaging with the community to learn about their experiences, concerns, and attitudes, in order to identify areas where positive change can be implemented.

As our Study findings show, whilst Singapore youth are keen to explore the promise of new technologies, they are also concerned with the potential downsides.

Many harbour fears surrounding the impact of emerging online trends and technologies; for example, they are afraid of being cancelled for expressing their personal views, concerned about the effect of 'echo chambers' on their social media, and worried that generative AI may be misused.

Female youth are particularly concerned about sexual online harms, including image-based sexual abuse (IBSA). Perhaps related to this, is the finding from our Study that unhealthy, gender-biased attitudes seem to have taken root amongst our male youth. This is concerning and certainly warrants further investigation, as it may impact the occurrence of online harms against females in Singapore.

That said, the Internet isn't 'all bad' and can offer a multitude of positive experiences and opportunities. We must therefore find a way to ensure that safety and opportunity co-exist.

Finding the right balance will not be easy, but we hope that this Study will spark discussions on how the community, industry, and Government stakeholders can all work together to co-create solutions.

On our part, SHE will continue working with partners to engage the public and empower them with information on how to keep themselves safe online. It is our belief that Internet users, especially young ones, should have at their disposal the right education, toolkits, and resources that enable them to not only navigate the online realm safely, but to also seek support and recourse if they encounter harmful online conduct.

We are thankful to the Global Shapers Singapore Hub for working with us on this insightful Study.



Simran Toor
CEO
SG Her Empowerment (SHE)

MESSAGE FROM GLOBAL SHAPERS COMMUNITY

In our ever-evolving digital landscape, the importance of understanding and addressing online safety concerns cannot be overstated. The Safeguarding Online Spaces (SOS) Study addresses these challenges head-on. By employing the acronym SOS — a universally recognised signal for distress and urgent help — the SOS Study's name serves a dual purpose. It highlights our commitment to enhancing the security and well-being of users in digital spaces, while drawing attention to the pressing need for action.

The foundation of our research lies in the invaluable insights provided by over 500 youth respondents from our survey and focus group discussions. Their candid responses have painted a vivid picture of the online world as seen through the eyes of our youth. We have listened to their experiences and concerns, and now it is our duty to translate these voices into meaningful action. As we embark on this journey, we look forward to further engagement and collaboration with government agencies, Big Tech companies, and NGOs working tirelessly to champion user safety. Understanding the concerns of youth, as well as their preferences for community education and support, can serve as a catalyst for informed policy decisions and industry practices that prioritise the safety and well-being of diverse online communities.

I have had the privilege to embark upon this Study alongside a dedicated team of youth deeply committed to advancing the cause of online safety for our youth. The active participation of youth in leading, developing and executing the SOS Study is a resounding testament to the power of youth voices in shaping policies and fostering change. They are not just the subjects of our Study; they are our partners in this endeavour.

We are especially grateful to SG Her Empowerment, for their support and commitment in making this Study a reality. Their dedication to empowering women and addressing critical issues facing our society is commendable, and it's an honour to collaborate with them in making the online world a safer place for all.

The fight to safeguard online spaces demands collective action, across communities, industry and government. I invite all like-minded organisations, individuals, and institutions to join us in exploring actionable solutions and constructive dialogue to ensure that our online spaces are truly safe, inclusive, and empowering for all. Together, we can build a digital world that encourages connection, community and diversity — the time to act is now.



Calissa Man
Vice Curator

Global Shapers Community – Singapore Hub

TOPLINE FINDINGS

#1: The top three online harms of concern to youth were image-based sexual abuse (IBSA), cyberbullying/harassment, and doxxing

Close to half of youth polled reported being concerned with IBSA (48 percent), cyberbullying/harassment (47 percent) and doxxing (42 percent).

Gen Zs were more concerned about IBSA (Gen Z: 54 percent, Millennials: 40 percent). Millennials were more concerned about impersonation/identity theft (Millennials: 41 percent, Gen Z: 29 percent).

Female Gen Zs were twice as likely to be concerned about IBSA than male Millennials (F Gen Z: 62 percent, M Millennials: 31 percent).

#2: Six in 10 youth (63 percent) reported being exposed to sensitive content without searching for it, and suffered sustained emotional impact from the exposure

Such content included graphic violence, nudity and sexual activity, unhealthy eating behaviour, and gender-based hateful or derogatory content.

Seven in 10 (68 percent) reported being upset after viewing the content, of which half (52 percent) reported feeling this way for at least a few hours.

#3: Youth cancelled others, yet feared being cancelled themselves

Four in 10 (39 percent) had cancelled others, primarily because they felt that what the cancelled person did was wrong.

But cancelling has a clear downside. Seven in 10 (73 percent) said that the fear of being cancelled shapes their behaviour online, including by causing them to refrain from sharing personal views.

#4: Youth were open to generative AI, but wanted clear rules of engagement

Half of youth polled (52 percent) have used AI-based tools with images.

However, focus group participants expressed the need for clear rules on what is harmful in this area, citing regulation as an important remedy. Many were aware of the misrepresentation of persons (55 percent) and the sexualisation or objectification of women (54 percent) as negative effects of AI generated or altered images.

#5: Youth lacked awareness of safety tools and recourse avenues

Four in 10 (38 percent) reported low awareness of self-help tools for online harms, such as in-app reporting.

Five in 10 (49 percent) reported low awareness of legal recourse options, with females more likely to indicate this (F: 54 percent, M: 44 percent).

ABOUT THIS STUDY

Why investigate youth concerns on online harms?

In September 2023, SHE released a research report¹ which revealed that:

- Online harms were prevalent in Singapore, particularly among the younger segment of the population aged 15 to 44
- Female youth were twice as likely to experience sexual harassment compared to male youth

Given that younger Singaporeans are a group especially at risk of online harms, SHE and Global Shapers Singapore Hub partnered on this Study to find out more about what our youth are experiencing online, how they feel about certain emerging technologies and trends, and how they perceive and deal with online harms.

Through the Study, we hoped to gain specific information on:

- Youth's concerns and attitudes on generative AI, algorithm-recommended content, and cancelling
- How online harms impact youth well-being
- Youth's help-seeking behaviours
- What youth believe must be done to tackle online harms

A list of key terms defined in this Study can be found at **Annex A**.

Who participated in the Study?

Between 11 March to 10 June 2023, 500 Singapore youth² (Singapore Citizens and PRs) aged 16 to 35 completed an online survey. This sample was representative of our local resident population by age, gender and ethnicity.

To provide a deeper insight, 44 youth aged 15 to 35 also participated in focus group discussions.³

¹ Visit <https://tinyurl.com/SHEOHS23> for more information.

² Sampled from Milieu Insight's online panel platform, based on quotas referenced from the Department of Statistics (DOS). A soft quota was applied for education. For survey respondent profile breakdowns refer to Annex B.

³ Three sets of virtual focus group discussions were conducted based on different topics with a total of 44 unique individuals. The session on Generative AI had n=20 participants. The session on content recommendations had n=17 participants. The session on cancel culture had n=14 participants. Quotations featured in this report have been edited for readability.

KEY CONCERNS & ATTITUDES

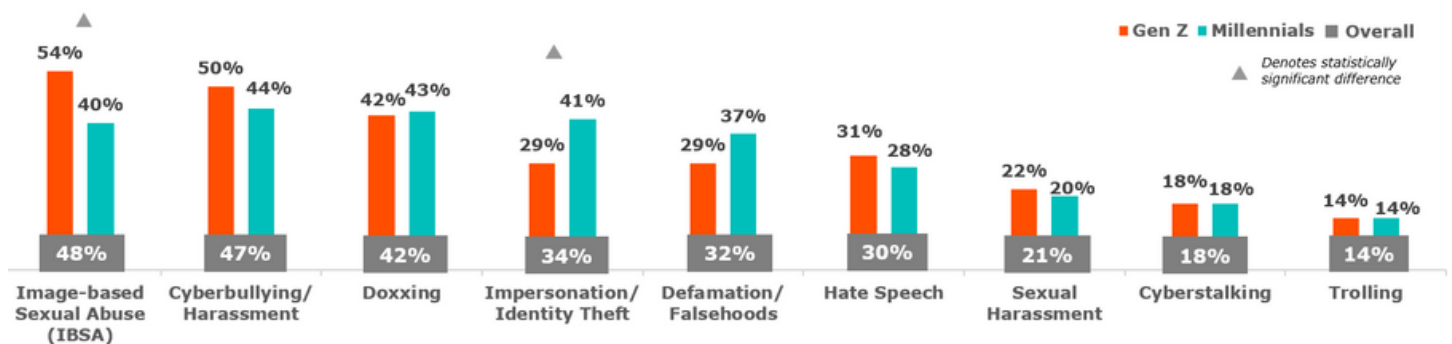
What were the top online harms of concern to youth?

Youth were most concerned about violations of dignity and privacy, such as IBSA, cyberbullying/harassment, and doxxing

As shown in [Figure 1](#) below, close to half were most concerned about:

- IBSA (48 percent)
- Cyberbullying/harassment (47 percent)
- Doxxing (42 percent)

Figure 1: Top Three Online Harms of Concern, by Age



Q: Please select the top three online issues which concern you the most.
Base: All Respondents, n=500

How did these concerns differ by age and gender?

Millennials were more concerned about impersonation/identity theft, while Gen Zs were more concerned about IBSA

With reference to [Figure 1](#), Millennials (those born between 1996 – 1988 in this sample) were significantly more concerned about impersonation/identity theft (Gen Z: 29 percent, Millennials: 41 percent).

Gen Zs (those born after 1996 in this sample) were significantly more concerned about IBSA (Gen Z: 54 percent, Millennials: 40 percent).

The differences may be a result of the types of online activities that each generation tends to engage in more frequently.

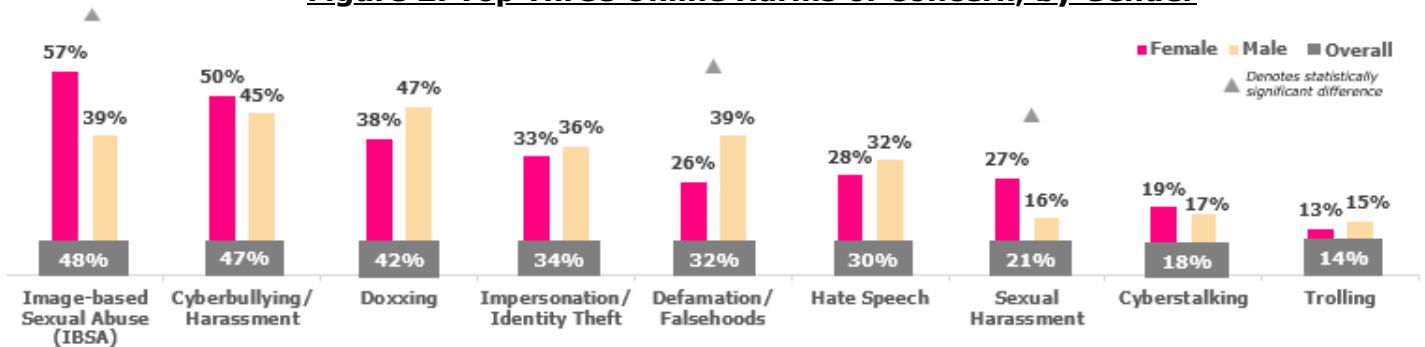
IBSA was the biggest concern for females, with the most concern arising from Gen Z females

As shown in [Figure 2](#), male youth were significantly more concerned about defamation/falsehoods (F: 26 percent, M: 39 percent).

Comparatively, female youth were significantly more concerned about IBSA (F: 57 percent, M: 39 percent) and sexual harassment (F: 27 percent, M: 16 percent). Gen Z females were the group most concerned about IBSA (F Gen Z: 62 percent, F Millennials: 50 percent, M Gen Z: 46 percent, M Millennials: 31 percent).

The greater concern by females over IBSA may be due to the perception that they are more at risk of being victimised this way, and/or of the potential severity of the negative impact that they may suffer.

Figure 2: Top Three Online Harms of Concern, by Gender



Q: Please select the top three online issues which concern you the most.
Base: All Respondents, n=500



KEY SAFETY GAPS

What digital safety gaps did the Study reveal?

Low awareness of safety and recourse avenues for online harms

Singapore youth appear apathetic towards learning about available self-help tools or recourse avenues for online harms:

- Four in 10 (38 percent) reported low awareness of self-help tools such as reporting on social media apps
- Five in 10 (49 percent) reported low awareness of legal recourse avenues
- Five in 10 (45 percent) reported low awareness of non-legal sources of help such as counselling

How did this differ by age and gender?

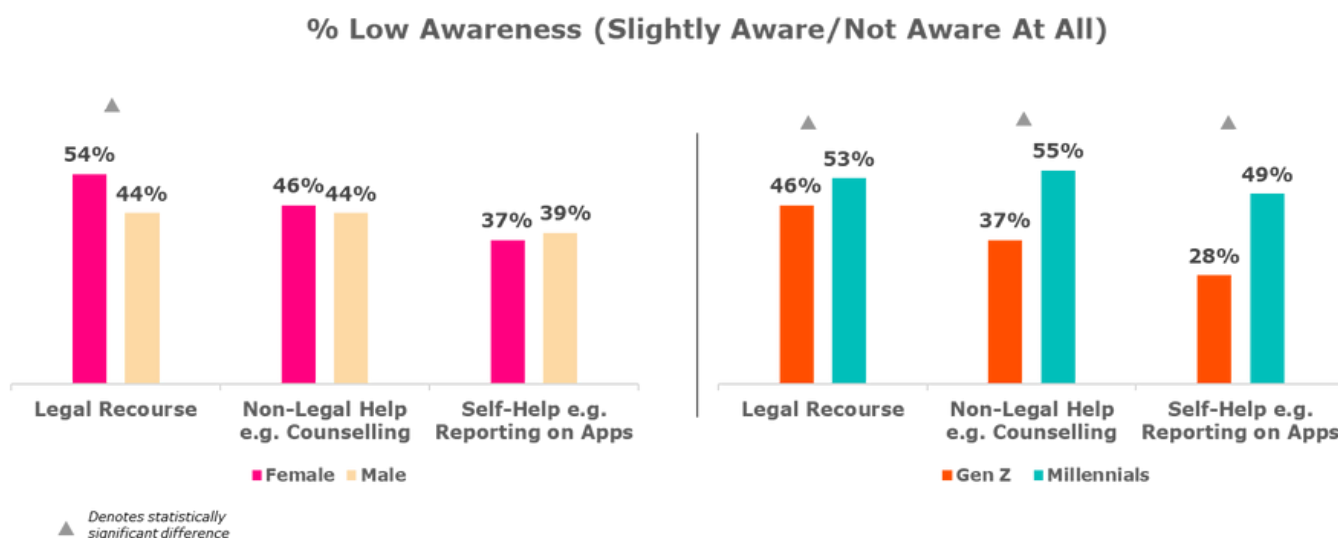
Millennials were generally less aware of safety and recourse avenues

As shown in [Figure 3](#), Millennials were less aware than Gen Zs, especially in relation to non-legal help avenues (Gen Z: 37 percent, Millennials: 55 percent) and on self-help tools (Gen Z: 28 percent, Millennials: 49 percent).

Female youth were less aware of legal recourse avenues

Concerningly, female youth were significantly less aware of legal recourse avenues (F: 54 percent, M: 44 percent).

Figure 3: Low Awareness of Help/Recourse Avenues, by Gender and Age



Q: To what extent are you aware of legal actions you may take if you experience any online harms?

Q: ...aware of non-legal sources of help (e.g., counselling) if you experience any online harms?

Q: ...aware of self-help tools/processes (e.g., reporting on social media apps) if you experience any online harms?

Base: All respondents, n=500

Unhealthy, gender-biased mindsets towards females becoming normalised amongst younger males

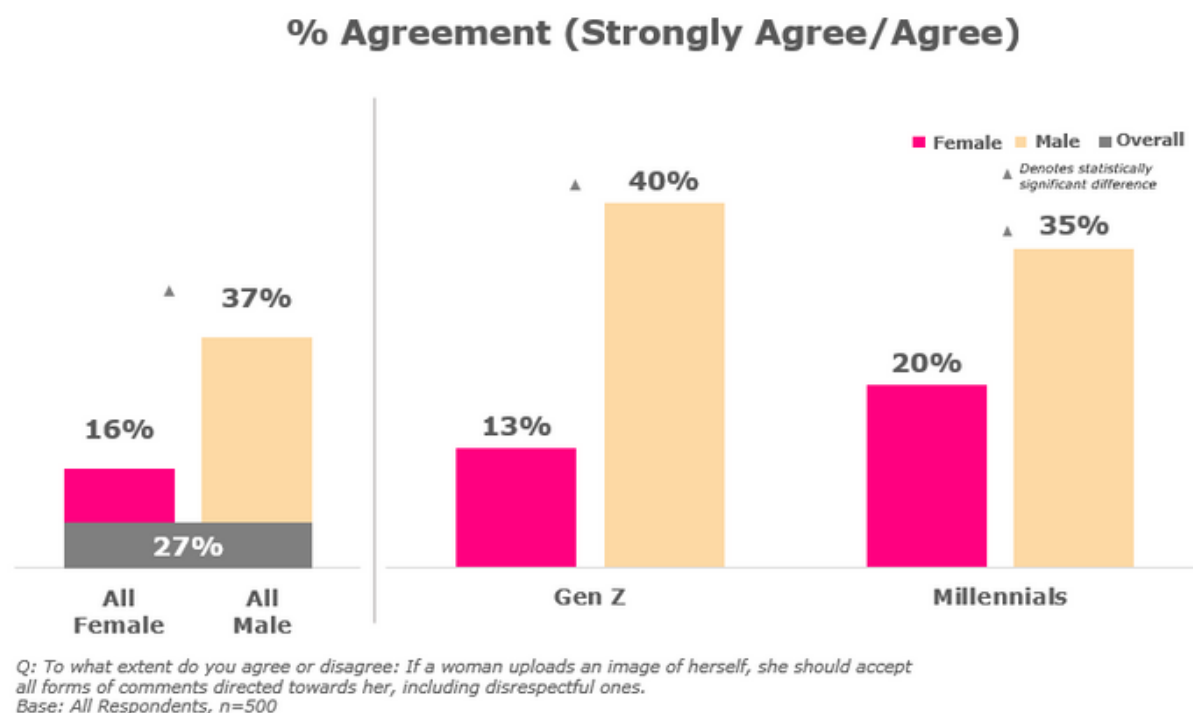
As shown in [Figure 4](#), one in four youth (27 percent) felt that if a woman uploads an image of herself, she should accept all forms of comments directed towards her, including disrespectful ones.

Abusive or disrespectful comments are never warranted, and it is important that such harmful misconceptions are addressed among our youth. If they are not, female targets of online harms could be deterred from seeking help, believing they deserve to be treated this way.

A significantly higher proportion of male youth - four in 10 (37 percent) - harboured this misconception, with Gen Z males being more likely than Millennial males to do so (M Gen Z: 40 percent, M Millennials: 35 percent).

This suggests that unhealthy, gender-biased attitudes may be becoming normalised amongst younger generations of males. This is a concerning trend. If left unchecked, such negative mindsets could proliferate, resulting in harmful online (or even offline) behaviours towards females, such as verbal abuse or victim blaming.

Figure 4: Unhealthy Gender-Biased Attitudes, by Gender and Age



ALGORITHM-DRIVEN CONTENT

Algorithms play a critical role in shaping our online experience. In the online world, an algorithm is a set of rules, signals and data that determines how content is filtered, ranked, selected, and recommended to users on Internet platforms such as social media sites. Algorithms also analyse user behaviour, interactions, and interests to understand content preferences and deliver personalised recommendations.⁴

What were youth's attitudes towards algorithm-driven content?

Youth recognise the value of personalised content recommendations, which can bring convenience and help them connect with like-minded people.

However, they believe this must be weighed against risks such as a lack of privacy, or of falling into 'echo chambers'⁵ and down 'rabbit holes.'⁶

"Track[ing] user behaviour [allows content] to be tailored to our needs. It brings convenience, but if **we want convenience**, then we must **sacrifice our privacy** as well." – Female, 30-34

"Sometimes I see **a bit too much** of a certain side of politics. I realised I had to learn about the other side as the personalised content won't show it to me." – Male, 20-24

"When it is very personalised content based on your interests...it **boosts your self-esteem** because **you find a community** that has similar interests. **It's a safe bubble...It can be good or bad.** [For example], if you're watching terrorist content repeatedly, you get **pulled into the ideologies.**" – Female, 15-19

⁴ See also our definition of 'recommendation algorithms' at Annex A.

⁵ See Annex A for our definition.

⁶ See Annex A for our definition.

What were the safety gaps uncovered?

A majority reported being recommended unsolicited sensitive content

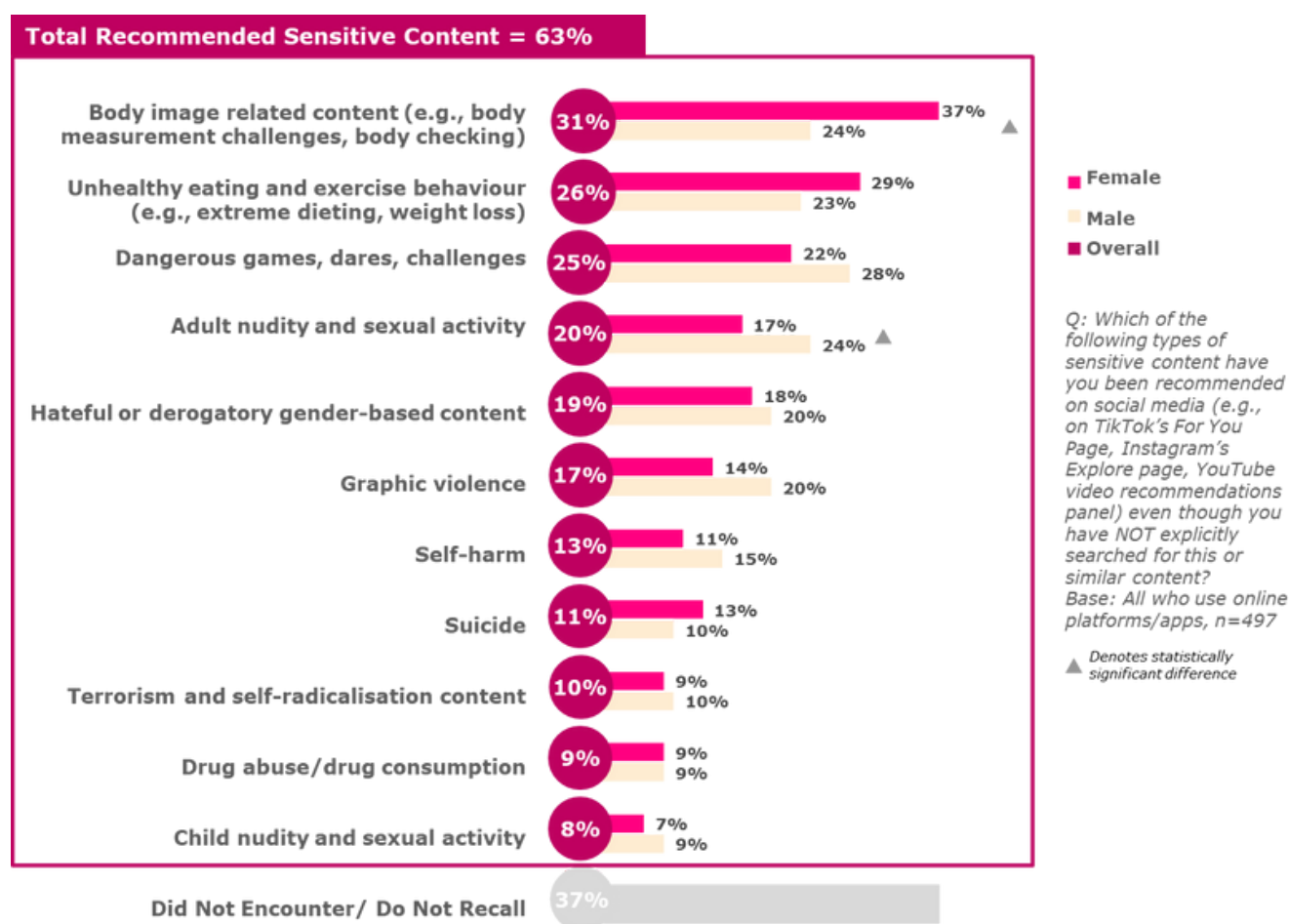
As seen in [Figure 5](#) below, six in 10 youth⁷ polled (63 percent) reported being recommended certain types of sensitive content even though they did not search for that content.

The top three forms of sensitive content most frequently encountered in this manner pertained to:

- Body image (31 percent)
- Unhealthy eating and exercise behaviour (26 percent)
- Dangerous games, dares and challenges (25 percent)

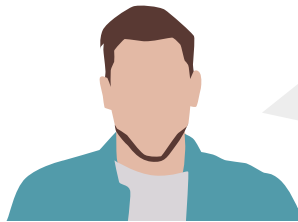
A noticeable number of youth reported being recommended content on adult nudity and sexual activity (20 percent), gender-based hateful or derogatory content (19 percent), and child nudity and sexual activity (8 percent). Again, this content was not searched for. This is concerning as such material can amount to breaches of online platform community standards, or even criminal conduct, and yet is nevertheless recommended as content to users.

Figure 5: Types of Sensitive Content Recommended



⁷ This Study adopted a selected list of sensitive content as stated in [Figure 5](#).

A majority of those who encountered unsolicited sensitive content reported sustained emotional impact



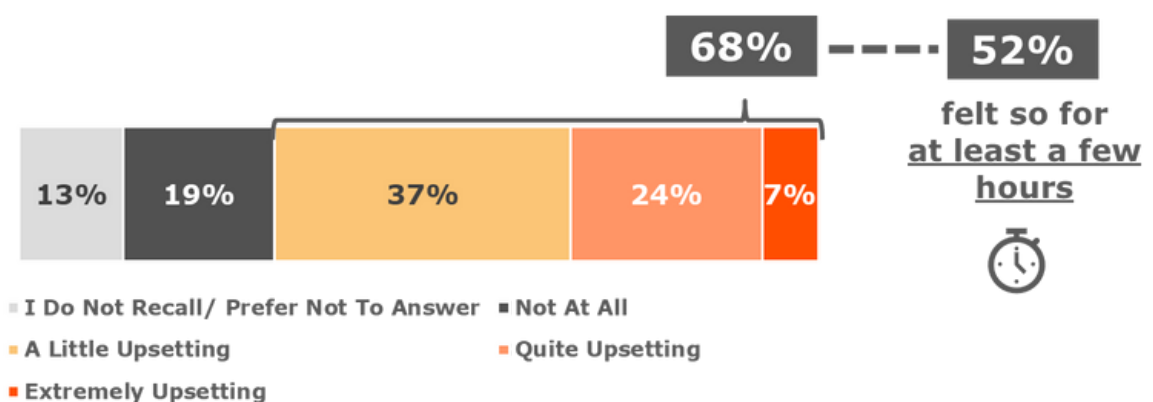
*"I was **quite disturbed** by this video of two boys in China doing a live stream. Suddenly an unknown man came from behind and hacked their arms off. The video went viral for a few days on Facebook, Twitter and TikTok. I don't watch violent content on these platforms since I stick to funny, PG topics most of the time. When this video was recommended in my feed, **it was quite horrifying for me.**" - Male, 25-30*

As shown in [Figure 6](#), seven in 10 (68 percent) of those recommended sensitive content without searching for it reported feeling upset and affected. A significant number said these feelings lasted for some time after exposure:

- One in three (31 percent) were quite or extremely upset
- Half of those (52 percent) who were upset felt so for more than at least a few hours

These findings demonstrate that the impact of sensitive content exposure is not momentary and will not simply disappear once our content feed refreshes. They underscore the need for us to recognise the potential emotional and psychological impact that such exposure can cause, and to develop appropriate and timely strategies to prevent or address the negative impact.

Figure 6: Impact of Unsolicited Sensitive Content and Duration



Q: Generally, what impact did encountering such sensitive content have on you?

Base: All who encountered sensitive content, n=314

Q: Thinking about the last time you encountered such sensitive content on your feed, how long did you feel upset about it?

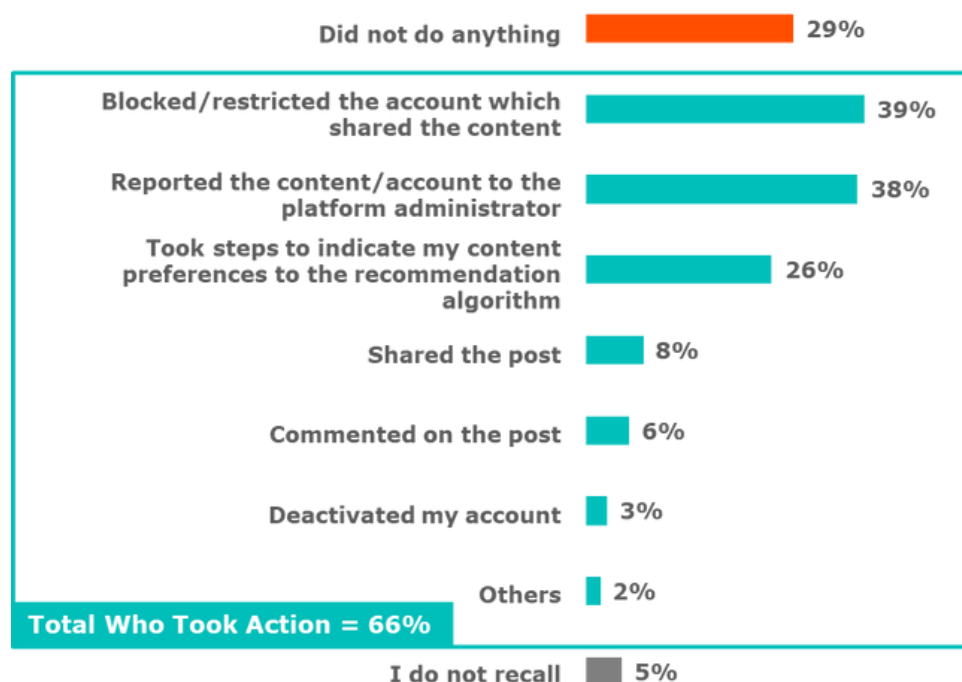
Base: All who were at least a little upset, n=216

Many did not take steps after encountering sensitive content

As shown in Figure 7, after encountering unsolicited sensitive content, three in 10 (29 percent) did not take any form of action to address what they saw, such as by blocking or restricting the account or reporting the content.

Those who took action (66 percent) primarily did so by blocking or restricting the account that shared the content (39 percent), or by reporting the content/account to the platform concerned (38 percent).

Figure 7: Action Taken in Response to Sensitive Content Recommendations



Q: If you encountered such content recommended on your feed, which of the following actions did you take?
Base: All who encountered sensitive content, n=314

Algorithms recommended different content to females and males respectively, reinforcing unhealthy stereotypes



*It's hard not to see women as just bodies when that's all the algorithm shows me...It's easy to forget there's a real person behind those pictures. It's **hard not to internalise those beliefs.**" - Male, 20-24*

As seen in Figure 5, female youth were significantly more likely to report being recommended body image-related content, as compared to male youth (F: 37 percent, M: 24 percent). Both male and female group participants shared how this content affected their self-esteem.

Comparatively, male youth were significantly more likely to report being recommended content on adult nudity and sexual activity (F: 17 percent, M: 24 percent). Male focus group participants shared how such content encouraged them to view women as sexual objects.

CANCELLING

Cancelling online refers to calling for others to shun, boycott, ostracise and/or express their disapproval towards individuals, public figures and companies after they have done or said something considered objectionable or offensive.

It is observed in 'cancel campaigns' or 'cancel culture', where people participate in and encourages others to call for and create pressure to punish a person for what he/she may have said or done (e.g., by shaming them or pressuring their employer to fire them).⁸

What were youth's attitudes towards cancelling?

Cancelling seen as an accountability tool that can be weaponised

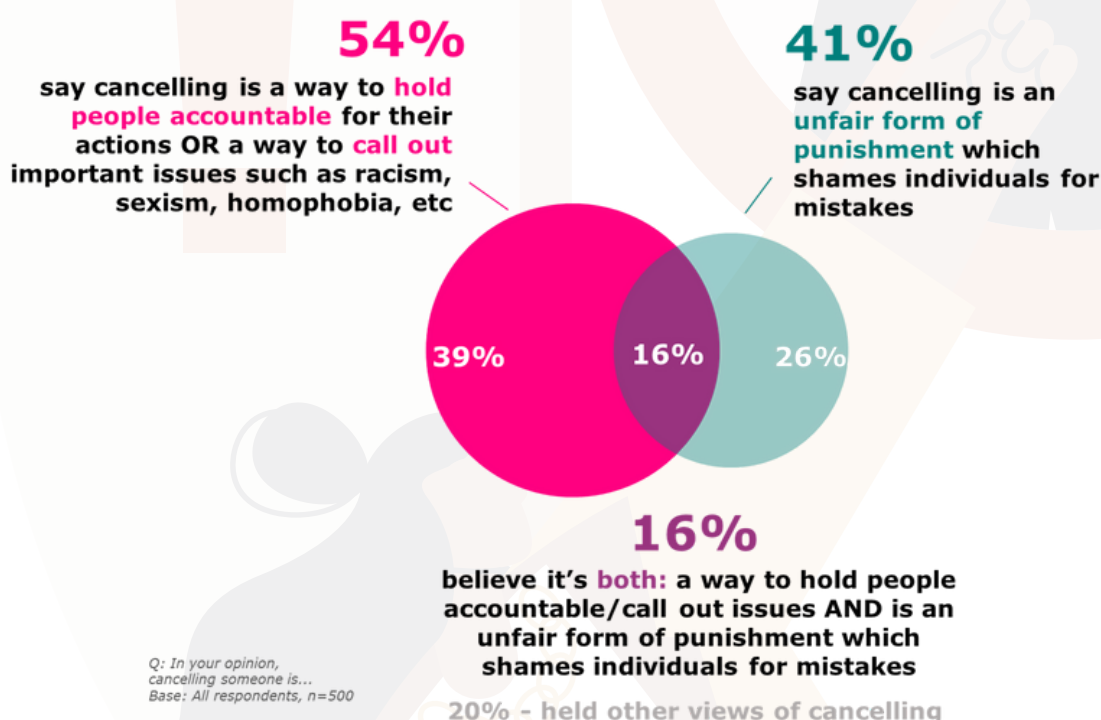


*"Social media can be **weaponised to harm** someone because of motives or agendas...if I really hate someone [then] I can [say] untruthful things about the person to bring him down so it is easy to create a bad perception...with cancel culture he might lose his job and reputation easily". – Female, 25-29*

As shown in [Figure 8](#), one in two youth (54 percent) perceive cancelling as either a way to hold others accountable for their actions or as a means to call out important issues.

Conversely, 41 percent viewed cancelling as an unfair form of punishment that shames individuals for mistakes.

Figure 8: Perceptions of Cancelling



⁸ See also our definition of 'cancelling' at Annex A.

A majority believed there were reasonable ways to cancel someone

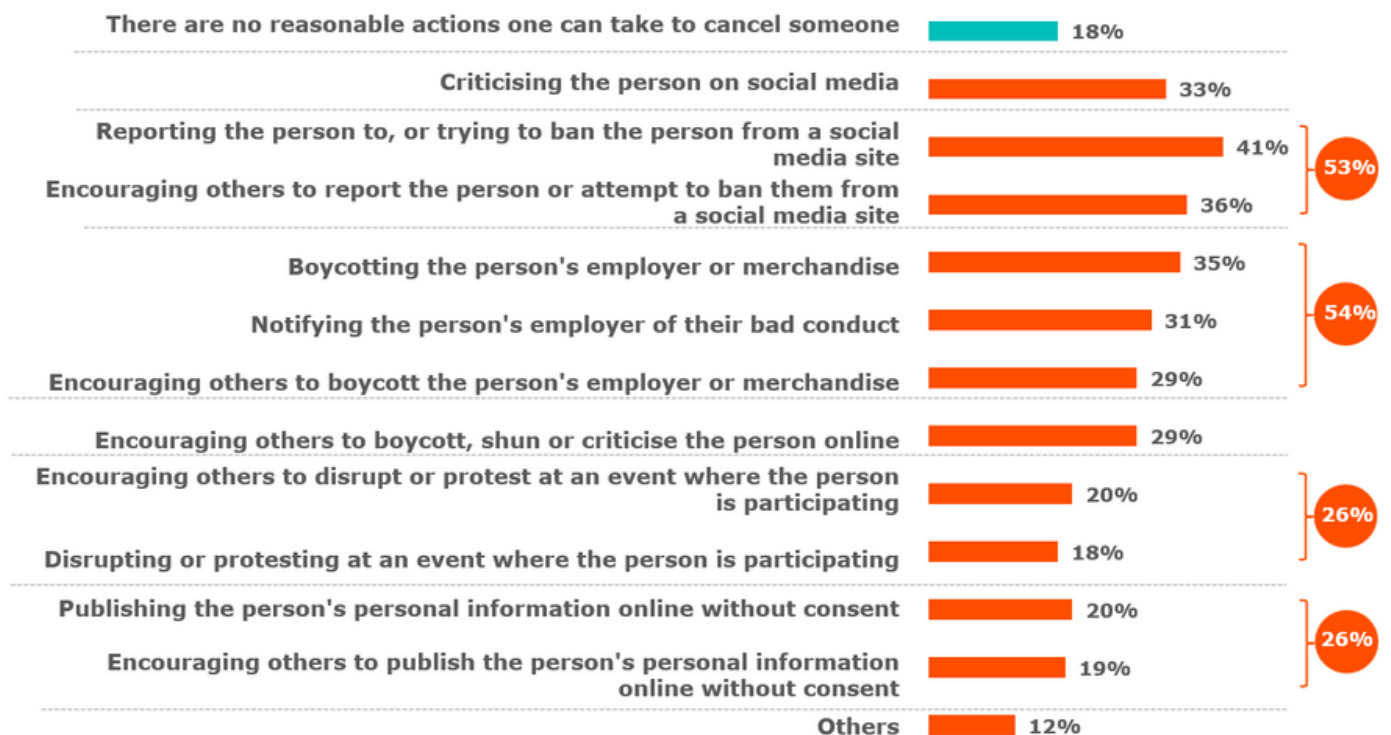
82 percent of youth polled thought there were reasonable means of cancelling someone, reflecting how normalised cancelling has become.

As shown in [Figure 9](#), seeking to ban an individual from a social media platform was the top cited reasonable method of cancelling someone (41 percent), followed by encouraging others to cancel that person (36 percent).

Many felt it was reasonable to publish a person's personal information online without consent (20 percent), or to encourage others to do so (19 percent). This is even though doing so may amount to the criminal offence of doxxing.

Only one in five (18 percent) felt that no act of cancelling was justifiable. Millennial respondents were more likely than their Gen Z counterparts to adopt this view (Gen Z: 15 percent, Millennials: 22 percent), suggesting that cancelling is normalised to a greater degree among younger segments of the population.

Figure 9: Reasonable Forms of Cancelling



Q: In your view, which of the following, if any, are reasonable actions taken to 'cancel' a person?
Base: All Respondents, n=500

Many had actively participated in cancelling to call out behaviour they deemed wrongful

As shown in [Figure 10](#), around four in 10 (39 percent) had actively participated in cancelling. Most did so by:

- Reporting the cancelled person to, or trying to ban the cancelled person from a social media site (17 percent)
- Boycotting the person's employer or merchandise (13 percent)
- Criticising the person on social media (12 percent)

[Figure 11](#) further shows that among those who had cancelled someone, most did so because they:

- Thought that what the person did was wrong (65 percent)
- Did not like the person they were cancelling (31 percent)
- Thought cancelling was a way of asserting their beliefs (24 percent)

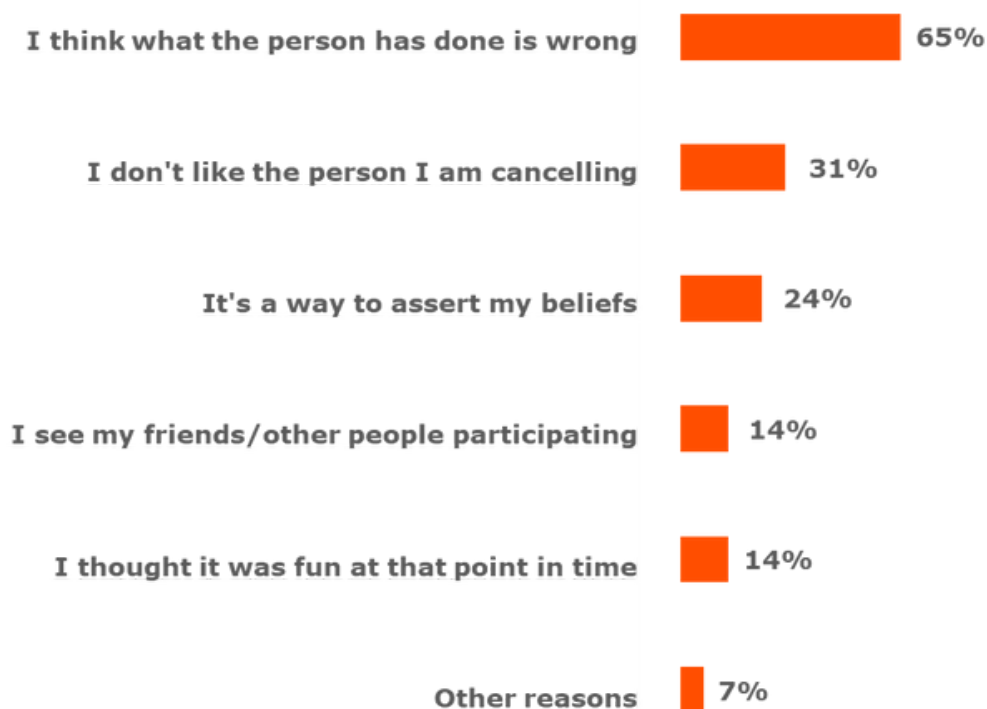
Figure 10: Actions Taken to Cancel Others



Q: Which of the following actions, if any, have you personally taken to 'cancel' someone?

Base: All Respondents, n=500

Figure 11: Reasons for Cancelling Others



*Q: Think about the last time you participated in cancelling someone.
Why did you do so?
Base: All who have cancelled someone, n=195*

What were the safety gaps uncovered?

Although many had engaged in cancelling, a majority agreed that it could cause serious consequences to the cancelled individual

As shown in [Figure 12](#), more than half of youth polled agreed that cancelling someone could lead to that person suffering the following negative effects:

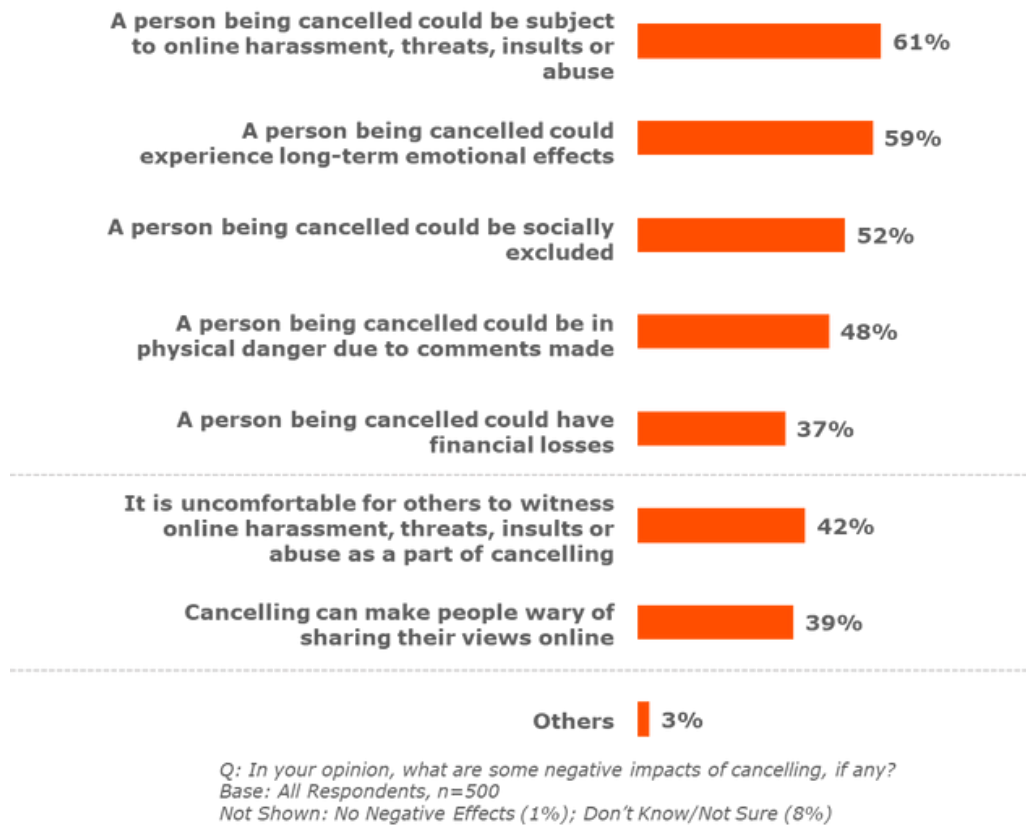
- Online harassment, threats, insults, or abuse (61 percent)
- Long term emotional impact (59 percent)
- Social exclusion (52 percent)

48 percent also agreed that cancelling someone could place that person in physical danger.

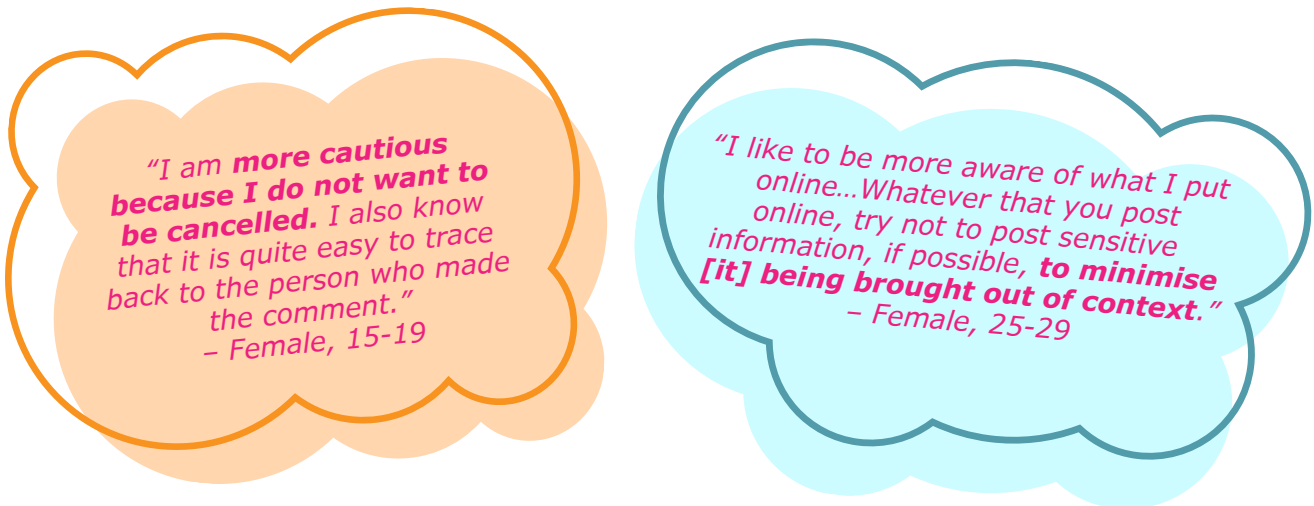
Clearly, Singapore youth appreciate how serious and sustained the negative impact of cancelling can be on the target of the cancelling.

It is concerning that while most youth agree that cancelling can lead to negative effects, four in 10 (39 percent) of our participants had actively participated in cancelling (as seen in [Figure 10](#)).

Figure 12: Negative Effects of Cancelling



A majority said that the fear of being cancelled affected how they behaved and communicated online

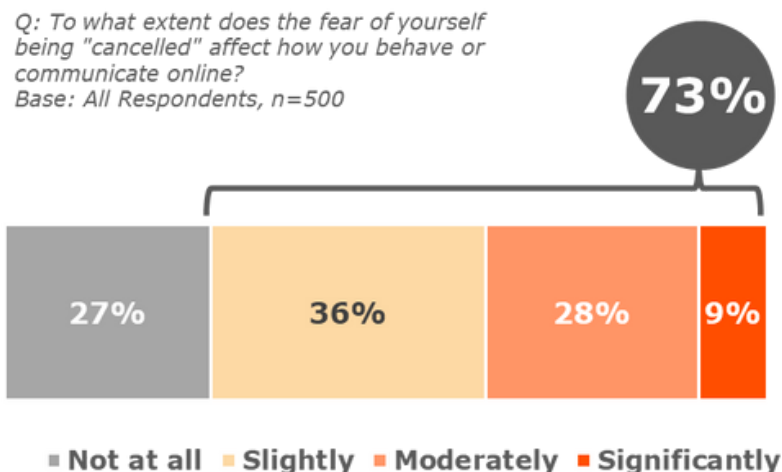


As shown in [Figure 13](#), 73 percent of youth polled said that the fear of being cancelled affects how they behave and communicate online.

To avoid being cancelled by others, focus group participants shared that they refrain from sharing their personal views online. This is consistent with the findings in [Figure 12](#), which showed that 39 percent of youth agreed that cancelling could make people wary of sharing their personal views online.

These findings suggest that cancelling causes a chilling effect on healthy online discourse, with many refraining from sharing their views for fear of being the target of a cancel campaign.

Figure 13: Fear of Cancellling Affects Behaviour



Most did not report others who instigated cancelling

As shown in Figure 14, 86 percent of youth who had witnessed someone encouraging others to cancel a person failed to report the instigator's actions.

Two in three of those who did not report this behaviour felt it was wrong to encourage others to cancel someone (69 percent).

One in three (31 percent) did not report this behaviour because they did not think it was wrong to encourage others to cancel someone.

These findings suggest that there is some reluctance or apathy towards reporting people who encourage others to participate in cancelling another. While more research is required to understand the reasons behind this, steps should be taken in the meantime to ensure that young people understand why instigating cancelling is wrong and may in fact amount to criminal conduct in certain circumstances.

Figure 14: Witnessing and Reporting Someone Instigating Cancelling



Q: If you have seen someone encouraging others to cancel a person (e.g., content encouraging others to boycott an individual), have you reported one or more of these instances?

Base: All Respondents, n=500

Base for breakdown of reporting: All who witnessed cancelling, n=315

GENERATIVE AI

Generative Artificial Intelligence (AI) is a type of AI that uses algorithms and machine learning to create new and original content, including images, text, videos, and other media. It learns from a large reference database and continually trains on more data. These models utilise neural networks to identify patterns in existing data and predict outcomes.⁹ AI image generators use trained artificial neural networks to use textual input to create high-quality, realistic images. In this Study, we focused on generative AI as it pertains to altering or creating images.

What were youth's attitudes towards generative AI?

Youth were keen to explore generative AI, with many having used AI-based tools to alter or create images



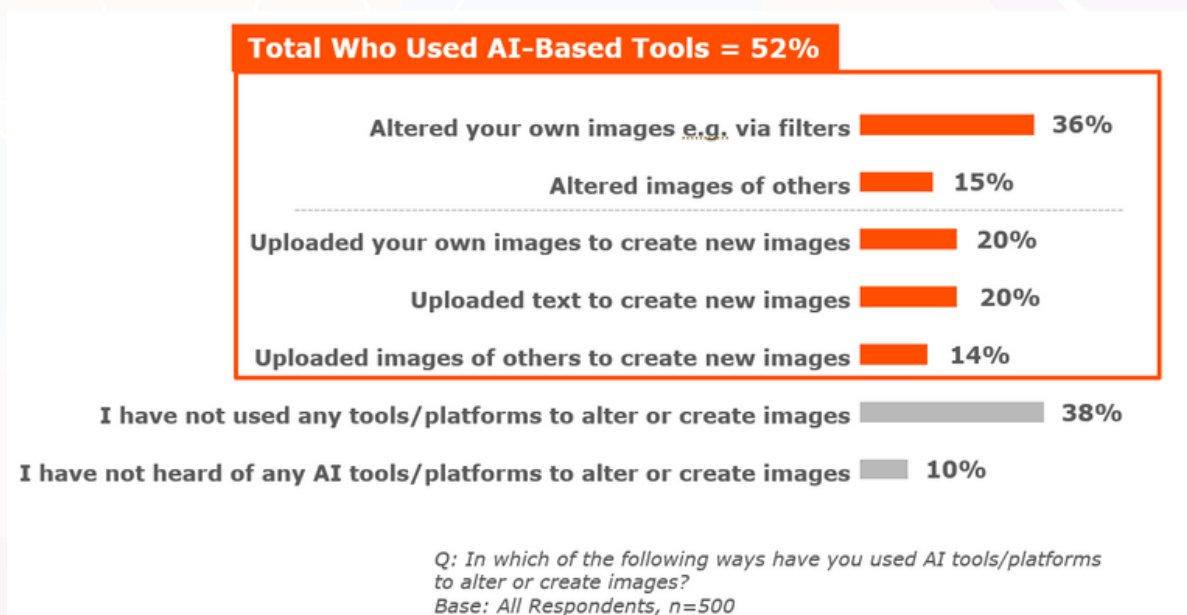
*"For example, you want a picture of like a monkey in space. Let's create that. **You can't take a picture of that.** The use of [these tools] comes in for such situations." – Female, 15-19*

As shown in [Figure 15](#), half of youth polled (52 percent) have used AI-based tools to alter or create images. The most common means of doing so included:

- Altering their own images via filters (36 percent)
- Uploading their own images to create new images (20 percent)
- Uploaded text to create new images (20 percent)

Focus group participants shared that generative AI tools are exciting opportunities for experimentation, bringing ideas or concepts into fruition in a novel way.

Figure 15: Use of AI-Based Tools to Alter or Create Images



⁹ See also our definition of 'generative AI' at Annex A.

What were the safety gaps uncovered?

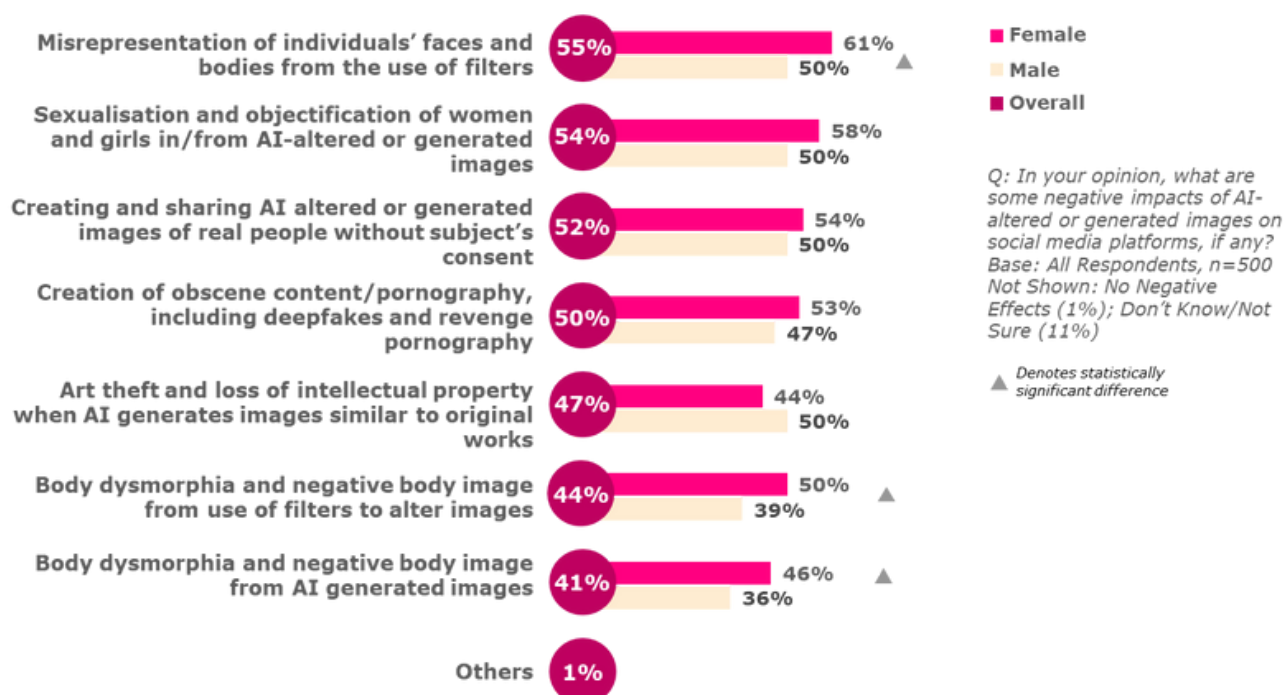
Misrepresentation of faces and bodies from the use of filters was the top-cited negative effect of AI-altered or generated images

As shown in [Figure 16](#), the most cited negative effects of AI-altered or generated images were:

- Misrepresentation of faces and bodies from the use of filters (55 percent)
- Sexualisation and objectification of girls and women (54 percent)
- Creation and sharing of AI-altered or generated images of a person without their consent (52 percent)

Given interest among youth in adopting AI-based tools, we must mitigate their exposure to these potential risks and misuses.

Figure 16: Potential Negative Effects of AI-Altered or Generated Images



Females youth were comparatively more likely to cite negative effects of AI-altered or generated images

As shown in [Figure 16](#), female youth were generally more likely than male youth to agree that negative effects of AI-altered or generated images exist. They were also significantly more likely to cite the following as negative effects:

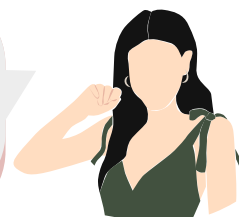
- Misrepresentation of faces and bodies from the use of filters (F: 61 percent, M: 50 percent)
- Body dysmorphia and negative body image from the use of filters to alter images (F: 50 percent, M: 39 percent)
- Body dysmorphia and negative body image from AI-generated images (F: 46 percent, M: 36 percent)

Several female focus group participants shared their concerns about the impact of AI-altered or generated images on body image norms:



*“The use of AI for filters **affects women mostly because the beauty standard for us is just way higher.** The standard for us is to wear makeup and look good already. But for men, that's not their standard.” – Female, 20-24*

*“Young people [may] start to think [those images] are real, but it's actually computer generated. **It may not be something that you can actually attain, no matter how much we try.** They may be continuously be frustrated about why they still do not look like their altered image.” – Female, 25-29*



Female focus group participants also shared their concerns over the featuring of women as sexual objects in malicious uses of generative AI. These findings suggest that young females are generally more concerned than their male counterparts about how AI-altered or generated images may be misused, play a role in exacerbating body image issues, and perpetuate the sexual objectification of women.

*“If the image or deep fake looks real enough and if people can't discern if it's fake, then **the harm is already done - it just destroys the girl's reputation.** There's nothing stopping them from just creating a horrible image. With just a face, you could just “deepfake” it onto someone else's body.” – Female, 20-24*

*“What worries me is how generative AI **seems to always lean towards a sexualised image when a woman or girl is the subject.** The more I see AI-generated imagery, the more I feel we're being pushed into a box of unrealistic and sexualised portrayals.” – Female, 25-29*

*“Usually **the victims [of deepfakes] I see are women and young girls.** I rarely see cases of guys or men.” – Female, 20-25*



YOUTH PERSPECTIVES ON SOLUTIONS

What do youth think should be done?

Algorithm-driven content

Close to half of youth polled cited the following as potential methods of addressing sensitive content recommendations; the top-cited method being an improvement in safety settings and tools on online platforms:

- Stronger online safety settings and tools on online platforms (56 percent)
- Stricter enforcement of laws prohibiting certain types of content online (52 percent)
- Public awareness campaigns and educational programmes (52 percent)
- Stricter laws prohibiting certain types of content online (48 percent)

Cancelling

Comparatively, when it comes to actions associated with cancelling, youth favoured an increase in regulatory and enforcement measures:

- Stricter enforcement of laws prohibiting certain actions associated with cancelling e.g., hateful comments, doxxing (52 percent)
- Stricter laws prohibiting certain actions associated with cancelling e.g., hateful comments, doxxing (51 percent)
- Stronger online safety settings and tools on online platforms (47 percent)
- Public awareness campaigns and programmes (42 percent)

Generative AI

As the use and misuse of generative AI remains a nascent issue, the Study sought to engage youth's views on problem-solving via our focus groups. The following themes emerged:

- Around half of youth in the focus groups agreed that further regulation is needed to set out clear limits on acceptable vs. wrongful use:

"In terms of laws, it should catch up. [For example] in intellectual property law...[if] you send something through an AI generated tool, it should be illegal to make profits out of it. Since you didn't own the original IP, [you should] not monetise that." – Male, 20-24.

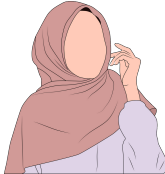
"For example, if someone takes your likeness and puts it in a video without your consent... the law should recognise that there's a transgression here... It's only when it's in a law that people will take it seriously." – Female, 25-29.

- Of the focus group participants who felt regulation was important, all agreed that this would only be part of the solution:



*"You can put in regulations on what you should and shouldn't do, but there'll be inevitably some people who still abuse it."
– Female, 25-29*

- Around half of the focus group participants felt that Internet platforms or AI developers should be proactive in tackling the potential harmful effects of generative AI, by using more innovative, safety-focused design tools:



*"[They could] create some filter to stop people from generating...the more harmful new images."
– Female, 20-24*

*"I feel that the tech companies who are creating the generative AI models in the first place should take the responsibility for it. If you leave it to the masses, you cannot control individuals who might misuse it..."
– Female, 30-35*



*"On top of being able to have an identifier feature on social media [to flag] when there's copyright content [so] you don't post, it could [flag] if [the user is] not that person [in the] photo and be restricted [from posting]."
– Female, 15-19*

Youth recognised that they play an important role

Slightly more than half of our focus group participants emphasised the importance of youth exercising their agency and individual responsibility to contribute to a safer online space. Whilst platform rules and regulations are necessary, these alone will not work if young users themselves do not adopt a proactive role to online safety.

Discussions centred around being mindful of what one posts and showing consideration towards others online, but also extended to reporting harmful content to benefit both themselves and the wider community.

*"Before we post a comment, do a double check...will it bring harm or good to the person?"
– Female, 30-35*



*"I see reporting unsuitable content not just as a personal duty but as a collective effort...to help protect other users."
– Male, 25-29*



*"The community should play a big part in the exposure and publicity of what is wrong, what is right, to use it in the good or bad way."
– Female, 30-35*



RECOMMENDATIONS

The Study findings have provided useful insight into some of the safety gaps that youth encounter online.

It is of paramount importance that these be addressed as a matter of priority, so that the next generation of Internet users can enjoy an online space that is safe, respectful, and empowering.

Our views on how this can be done are shared below.

1. Raise awareness of help, support, and recourse avenues

- The findings revealed that youth in Singapore had low awareness of self-help tools, legal recourse avenues, and non-legal resources such as counselling. Female youth, who are at risk of certain online harms, were significantly less aware of legal recourse avenues.
- Young people should know how to seek help or recourse if they encounter harms online. This is particularly important as the impact of online harms can be serious and long lasting, e.g., in the cases of cancelling or the misuse of one's images using generative AI.
- More public education can be undertaken to raise awareness of the available tools, support systems, and recourse avenues, with specific efforts targeted at female youth, who appear to have lower awareness.
- On its part, SHE will continue to raise awareness of and run the SHECARES@SCWO support centre, which offers holistic help to targets of online harms, via free counselling and legal clinics, and assistance with reporting.¹⁰

2. Develop efforts to tackle gender-based online harms

- The findings suggest that female youth were disproportionately targeted and were concerned about IBSA and sexual harassment. In addition, algorithm-recommended content plays a role in perpetuating unhealthy gender stereotypes among Internet users, particularly in the areas of body image and the sexual objectification of females.
- Our Study also revealed that male youth subscribed to gender-biased beliefs which could normalise online misconduct towards women.
- Targeted efforts are needed to tackle these issues. This may include stricter laws and enforcement to deter perpetrators of IBSA and sexual harassment, and efforts that counter the proliferation of unhealthy gender stereotypes online.
- This must be complemented by better public education to tackle unhealthy gender-biased beliefs, particularly among male youth, to ensure they do not develop attitudes that drive harmful or abusive conduct towards females online.

¹⁰ Visit <https://she.org.sg/shecares-scwo> for more information.

3. Increase support for 'safety by design' efforts

- A collaborative approach between users, stakeholders and industry is important, as it will allow digital products to be developed to better address safety gaps.
- Community organisations and advocacy groups that support online harm targets (such as SHE) can play an important role, as they can shed light on user experiences and emerging online harm trends. It would thus be important to ensure there are platforms or channels for such organisations to share ground-up data with developers, Internet platforms, and regulators alike.
- In addition, more can also be done to encourage females to play a role in product design and safety. An inclusive approach would strengthen and add robustness to new tools and technologies offered in the market, and more comprehensively tackle gender-based online harms against females.

4. Encourage better norms of online behaviour

- Youth harboured many concerns and fears around online harms, and with respect to major harms like cancelling, caused them to change their online behaviour, *e.g.*, by self-censoring their actual views or disengaging from online interactions.
- If youth begin to withdraw from online participation because they fear being cancelled, this will affect the quality of online discourse, their ability to build a sense of community online, and possibly cause them to lose out on beneficial experiences and opportunities.
- As a society, we should encourage better norms of online behaviour among users of the Internet. This can be done via public education campaigns, with a focus on younger Internet users who are still developing their online habits. To help build an online space that is both accessible and safe, we must encourage responsible, civil, and respectful online behaviour.

5. Provide platforms for co-creation with youth

- Young people are the group that spend the most time online, but they do not have the same experiences. As our Study findings showed, different generations of youth (Gen Zs and Millennials) had different views and concerns on online safety.
- To ensure that these multi-faceted issues are meaningfully understood and addressed, it is important to ensure that young people have opportunities to share information and ideas with the relevant stakeholders, and to explore how they can collaborate on solutions.
- For example, some Internet platforms have youth advisors or councils, which provide input and views about their community's needs and experiences. This approach can be adopted by more Internet platforms as a means to support trust-building and safety efforts.

ANNEX A: KEY TERMS IN SOS STUDY

Recommendation Algorithms: It refers to the set of computing instructions that determines what a user will be served based on many factors. This is done by applying machine learning techniques to the data held by online services, to identify user attributes and patterns and make recommendations to achieve particular goals. It is a key component of a content recommendation system, which prioritises content or makes personalised content suggestions to users of online services.¹¹

Generative AI: It refers to artificial intelligence (AI) models that use machine learning to produce novel content such as text, images, audio, video in response to prompts. Examples include text-based chatbots (e.g ChatGPT, Google Bard), image or video generators (e.g., DALL-E 2), and voice generators (e.g., Microsoft VALL-E).¹²

Cancelling: It refers to calling for others to shun, boycott, ostracise and/or express their disapproval towards individuals, public figures and companies after they have done or said something considered objectionable or offensive.¹³

Echo Chamber: An echo chamber is an environment where a person only encounters information or opinions that reflect and reinforce their own. Echo chambers can create misinformation and distort a person's perspective so they have difficulty considering opposing viewpoints and discussing complicated topics.¹⁴ Recommendation algorithms could reinforce echo chambers as they prioritise content based on users' past behaviours, preferences, and interactions. What is known as a filter bubble forms around individuals' online experiences.

Rabbit Holes: Rabbit holes can refer to time-consuming process of continually consuming content online. It can also refer to the process by which a user starts in a rich information environment and ends up viewing similar yet increasingly extreme content. Recommendation algorithms exacerbate rabbit holes by continually reinforcing a feedback loop of content to users.¹⁵

¹¹ eSafety Commissioner. (2022). *Recommender systems and algorithms – position statement*. eSafety Commissioner. URL: <https://www.esafety.gov.au/industry/tech-trends-and-challenges/recommender-systems-and-algorithms>

¹² Ibid.

¹³ Norris, P. (2023). Cancel culture: Myth or reality?. *Political Studies*, 71(1), 145-174.

¹⁴ Goodwill Community Foundation (GCF) Global Learning. (2023). *Digital Media Literacy – What is an Echo Chamber?*, GCF Global Learning. URL: <https://edu.gcfglobal.org/en/digital-media-literacy/what-is-an-echo-chamber/1/>

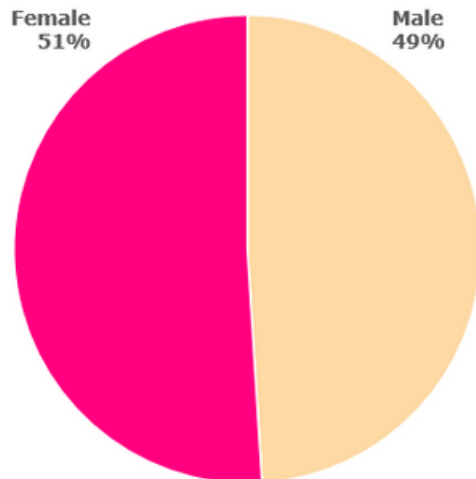
¹⁵ Brown et al. (2022). *Echo chambers, rabbit holes, and ideological bias: How YouTube recommends content to real users*. Brookings. URL: <https://www.brookings.edu/articles/echo-chambers-rabbit-holes-and-ideological-bias-how-youtube-recommends-content-to-real-users/>

This study also touched on nine different types of online harms, defined as follows:

Online Harm	Description
1. Sexual harassment	Unwanted messages of a sexual nature or obscene images
2. Cyberbullying/ Harassment	Online bullying, harassment, threats, insults, or abuse (including body shaming and hateful comments on the basis of race, religion, sexuality <i>etc.</i>)
3. Impersonation/ Identity Theft	Posing as someone else online (<i>e.g.</i> , setting up a social media account under another person's identity)
4. Defamation/ Falsehoods	False or misleading statements, images, or videos about a person
5. Image-based sexual abuse (IBSA)	Distribution of nude, intimate or sexually explicit images or videos (including altered images or videos) of a person without his/her consent
6. Cyberstalking	Online stalking (<i>e.g.</i> , by tracking one's whereabouts through his/her social media accounts)
7. Hate Speech	Being targeted with hate speech/hate comments because of the community/segment of society you belong to or identify with (<i>e.g.</i> , racist comments)
8. Doxxing	Doxxing (<i>i.e.</i> , publishing a person's identity information to cause harassment, alarm or distress, fear of violence, or to provoke or facilitate the use of violence)
9. Trolling	Intentionally instigating conflict, hostility, or arguments in an online platform

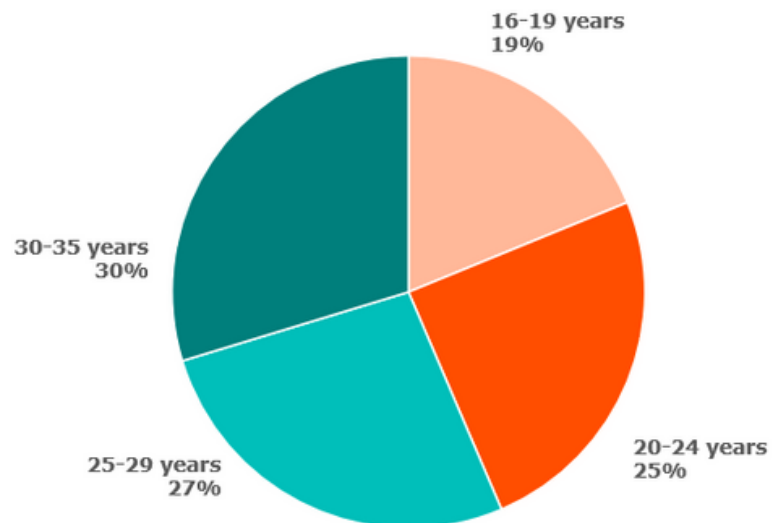
ANNEX B: RESPONDENT PROFILE

Gender



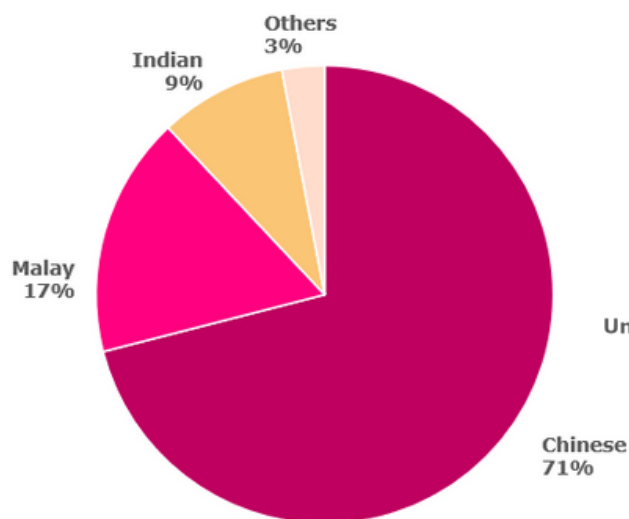
Q: Gender (based on Panel details)
Base: All respondents, n=500

Age



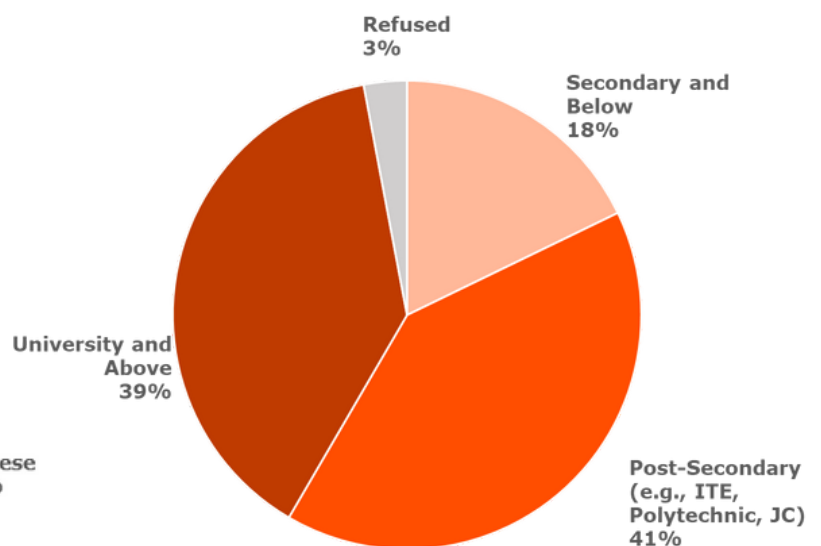
Q: What is your age?
Base: All respondents, n=500
Percentages may not total 100% due to rounding.

Ethnicity



Q: What is your race?
Base: All respondents, n=500

Education



Q: What is your highest education level attained?
Base: All respondents, n=500
Note: As the sample is of youth, many are still schooling.
Proportions reflect their highest qualification attained to date
Percentages may not total 100% due to rounding.

ACKNOWLEDGEMENTS

This report was a collaborative effort between SHE and the Global Shapers Community (GSC) - Singapore Hub. We thank individuals who contributed to the data collection, analysis, and writing for this study:

- Ms Annie Lem, Volunteer, SHE
- Ms Calissa Man, Vice Curator, GSC – Singapore Hub
- Mr Chan Shawn Kit, Volunteer - GSC – Singapore Hub
- Ms Cheryl Tan, Director, Programmes and Partnerships, SHE
- Mr Colin Goh, Director, Communications, SHE
- Mr Darryl Shya, Volunteer, GSC – Singapore Hub
- Ms Deborah Tan, Research Consultant, SHE
- Mr Ethan Teo, Nominee Shaper, GSC – Singapore Hub
- Ms Gan Xin Chen, Volunteer - GSC – Singapore Hub
- Ms Gigi Huang, Global Shaper, GSC – Singapore Hub
- Dr Han Ei Chew, Adjunct Senior Research Fellow, IPS; Board Member, SHE
- Ms Harleen Kaur, Programme and Research Executive, SHE
- Ms Harkiran Kaur, Programme and Research Executive, SHE
- Ms Ilyna Tan, Global Shaper, GSC – Singapore Hub
- Ms Jamie Yau, Volunteer, GSC – Singapore Hub
- Mr Justin Quek, Global Shaper, GSC – Singapore Hub
- Ms Natalie Chia, Research Manager, SHE
- Ms Nur Hazeem Abdul Nasser, Global Shaper, GSC – Singapore Hub
- Ms Rachel Kuo, Global Shaper, GSC – Singapore Hub
- Ms Rachel Wong, Volunteer, GSC – Singapore Hub
- Mr Sean Lim, Volunteer, GSC – Singapore Hub
- Ms Simran Toor, CEO, SHE
- Ms Suhaila Zainal Shah, Global Shaper, GSC – Singapore Hub
- Ms Urvashi Ghosh, Global Shaper, GSC – Singapore Hub

About SG Her Empowerment

SG Her Empowerment (“SHE”) is an independent non-profit organisation, with Institution of Public Character status, that strives to empower girls and women through community engagements and partnerships, and advocating for gender equality and equity in Singapore. SHE facilitates conversations, listens to the community, and gives voice to their views. We take a rigorous, research-based approach to defining solutions, forging partnerships with and between key interest groups, organisations and the government to advocate for change. In this way, we drive conversations and partnerships to effect meaningful change. Establishing equality. Building equity. Shaping a better Singapore for all.

About the Global Shapers Singapore Hub

Borne of the World Economic Forum, the Global Shapers Community is a network of inspiring youth leaders under the age of 30 working together to address local, regional and global challenges. The Singapore Hub of the Global Shapers Community comprises young professionals who are emerging and accomplished leaders in the public, private, and people sectors.

Brought to you by:



In collaboration with:

