Sunlight AfA

Detailed Sensing Poll Findings and Research Roadmap



Contents

ltem	Pages
Detailed sensing poll findings	3 – 33
Research Roadmap	34 – 44

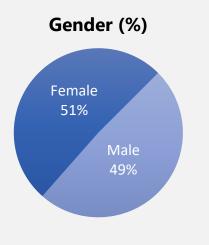
Sunlight AfA Detailed Sensing Poll Findings

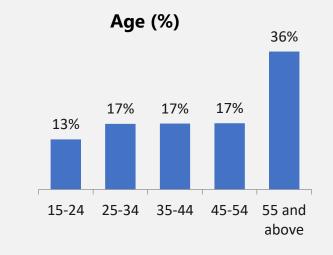


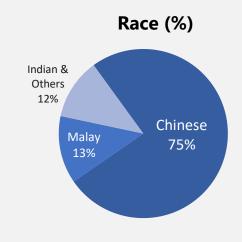


About the Poll

Data collection method	Target audience	Final sample size	Fieldwork period
Online survey by RySense Ltd	Singapore Citizens and PRs aged 15 and above	n=1,049	03 Jan 2022 to 09 Jan 2022







Data representative of Singapore resident population by Gender, Age, Race

Categories of Online Harms

Stalking & Impersonation	Doxxing, Threats & Harassment	Sexual-related Harms
 Befriend you using a fake identity Create a hoax social media account using your identity, Stalk you online 	 Bully, blackmail, harass or threaten you Ask you to commit a crime or join an illegal organization Post mean, humiliating, abusive, or offensive remarks, images or videos about you Reveal private information about you without your consent. 	 Distribute nude, intimate, or sexually explicit images or videos of you without your consent Exhibit unwanted sexual behaviour to you Create and/or share fake pornography featuring you or someone you care about, including "deep fakes" Send you unwelcomed and unwanted images.

For this poll, "online harms" is defined as the use of the Internet to engage in activities that result in harm or suffering to a person or a group of people online or offline, and "gender-based online harms" is defined as online harms due to one's gender.

Summary of Key Insights

Summary of Key Insights and Proposed Action Areas

Observations	Action Area
 While more males reported having experienced online harms, females were more likely to (i) feel unsafe online and (ii) targeted by gender-based online harms (GBOH) Females aged 25-35 were most likely to experience GBOH compared to males 	Consider needs of females in efforts to create a safer online space for all
 Overall, more than 40% of respondents know how to use platforms' safety and privacy tools (e.g., blocking, turn off location sharing) Awareness of help-seeking avenues for online harms is low, especially among females. More than 50% of respondents expressed concern over possible social stigma if they were to report online harms Other deterrents to help-seeking include (i) the perception that taking action would not make any difference and (ii) not knowing what to do 	Ensure availability and improve awareness of help-seeking avenues; address deterrents to help-seeking.
 Few respondents who experienced GBOH filed reports to technology companies and/or relevant authorities Respondents ranked (i) reporting systems for complaints and (ii) laws/legislations to address online harms as the top 2 measures that would facilitate action after individuals experience GBOH 	Ensure ease of using reporting channels, raise awareness of such channels and encourage reporting as an action to take after experiencing GBOH
 Respondents ranked technology companies as the stakeholder whose response to gender-based online harms had the greatest room for improvement Stricter enforcement of laws prohibiting GBOH and public awareness campaigns and programmes were perceived to be the most effective solutions to reducing GBOH 	Need for multiple stakeholders (i.e. technology companies, the Government and community organisations) to collectively address online harms

Perception

Perceptions of Online Harms and Digital Safety

Over 3 in 5 respondents believe online harms are common in Singapore.

Online harms are common in Singapore nowadays

Strongly Disagree/Disagree Neutral Strongly Agree/Agree

Fewer Singaporeans felt safe from online harms compared to walking alone at night.

night in Singapore? Generally, do you feel safe from online harms when you are online?

Generally, do you feel

safe walking alone at

92% 8%

Yes No

66% 34%

Notes:

5%

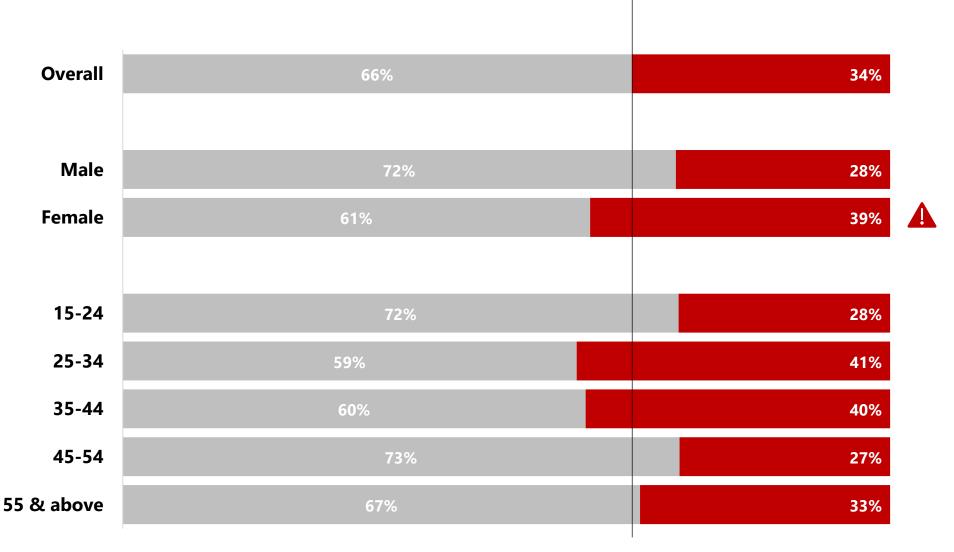
31%

- Data cell weighted by gender, age, race. n=966 (excluded "Don't know/Not Sure), n=1049.

64%

Proportion who do not feel safe from online harms when they are online

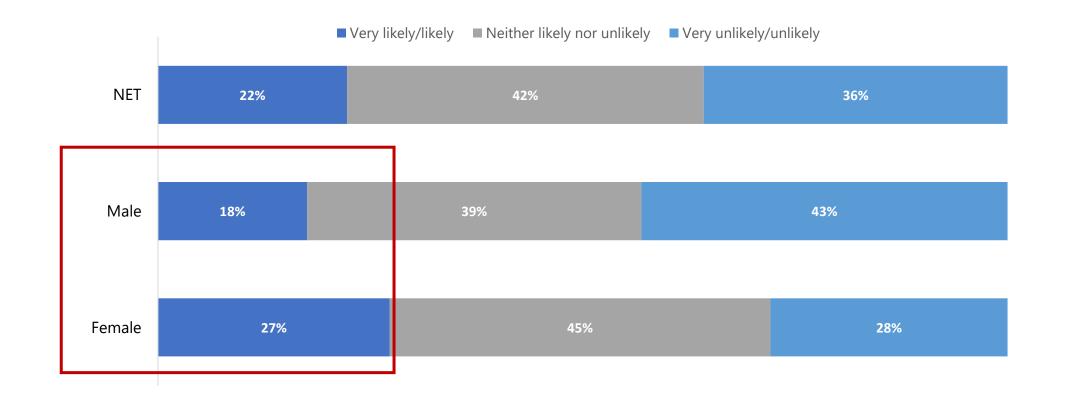
Women are less likely to feel safe from online harms



Generally, do you feel safe from online harms when you are online? [n=1049]

Perceptions of Gender-Based Online Harms

Around 1 in 5 respondents reported feeling targeted by gender-based online harms. Females were more likely than males to feel targeted by online harms because of their gender.

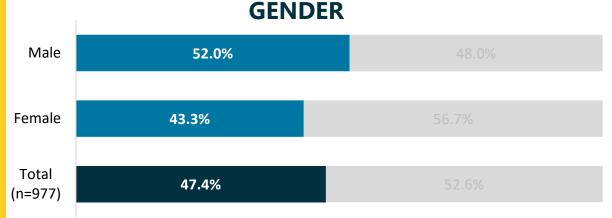


In your opinion, how likely are you to be targeted by online harms because of your gender? [n=1049]

Prevalence

Overall Prevalence of Online Harms

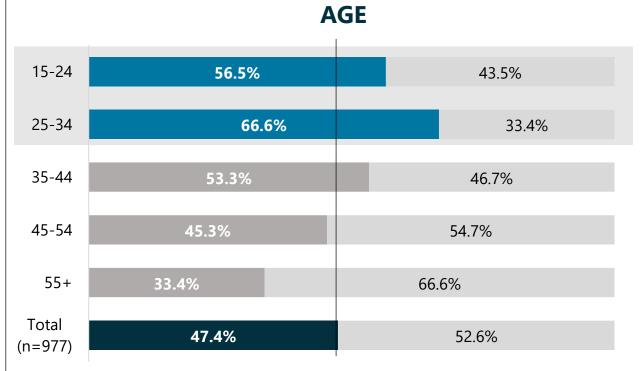
Nearly one in two Singaporeans have experienced online harms.



Personally affected by one or more online harms

Among those who preferred not to answer these questions relating to experiences with online harms (n=72), females outnumber males

More than half of youths aged 15 to 35 years old report experiencing online harms.



Personally affected by one or more online harms

Personal Experiences with Online Harms by Type of Online Harms

Top forms of online harms experienced were: contact with fake identities; receipt of unsolicited material; online stalking; and online bullying/blackmail/harassment.

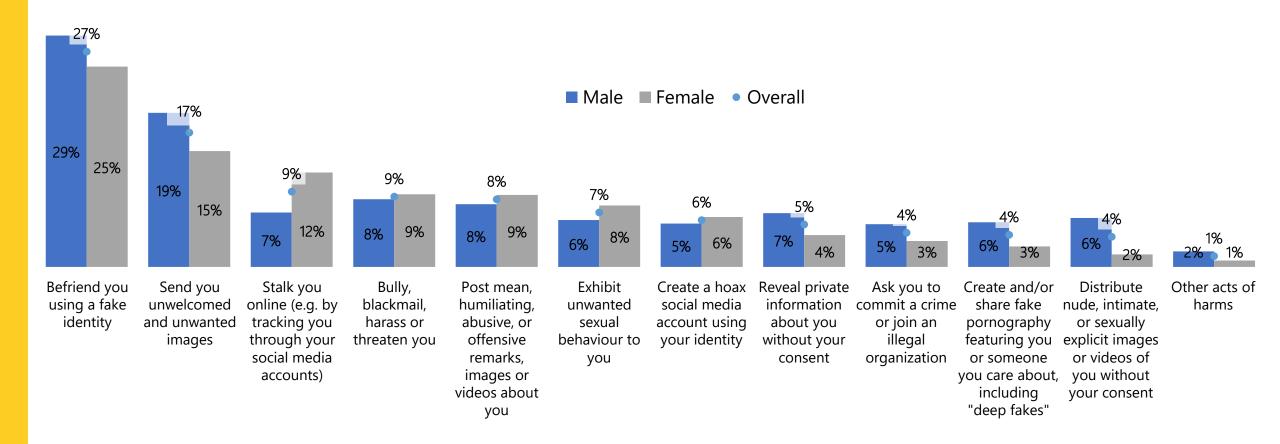
Voc

	Yes
Befriend you using a fake identity	27.1%
Send you unwelcomed and unwanted images	16.9%
Stalk you online (e.g. by tracking you through your social media accounts)	9.5%
Bully, blackmail, harass or threaten you	8.8%
Post mean, humiliating, abusive, or offensive remarks, images or videos about you (e.g. comments or messages on your posts/profile)	8.5%
Exhibit unwanted sexual behaviour to you	6.8%
Create a hoax social media account using your identity	5.9%
Reveal private information about you without your consent	5.3%
Ask you to commit a crime or join an illegal organization	4.3%
Create and/or share fake pornography featuring you or someone you care about, including "deep fakes"	4.0%
Distribute nude, intimate, or sexually explicit images or videos of you without your consent	3.8%
Other acts of harms on the Internet	1.3%

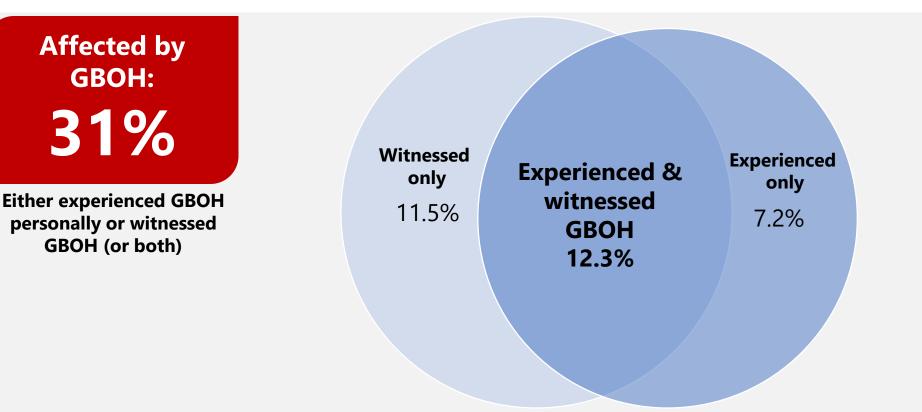
Has anyone ever used the Internet to...? [n=977]

Types of Online Harms Experienced by Gender

Women were more likely to have experienced online stalking and online bullying, while men were more likely to have received unwelcomed images.



Prevalence of Gender-based Online Harms (GBOH)



Unaffected by GBOH: 69%

19%

of all respondents reported they have been <u>personally</u> <u>affected</u> by gender-based online harms in the last 12 months. In the past year (last 12 months), how often have you been affected by online harm(s) due to your gender (i.e. gender-based online harms)?

202 / 1049 [or 19% of all respondents];

202 / 535 [or 38% of those who experienced online harms] [n=202; 35 preferred not to answer; 298 have not experienced GBOH; Total: 535]

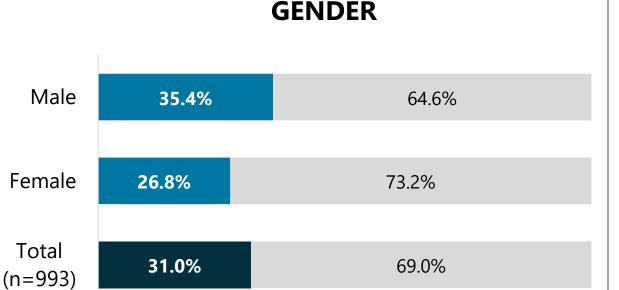
23%

of all respondents reported they have been <u>seen or heard</u> <u>about gender-based online</u> harms occurring in the last 12 months. In the past year (last 12 months), how often have you seen or heard about online harm(s) occurring to others whom you know personally due to their gender?

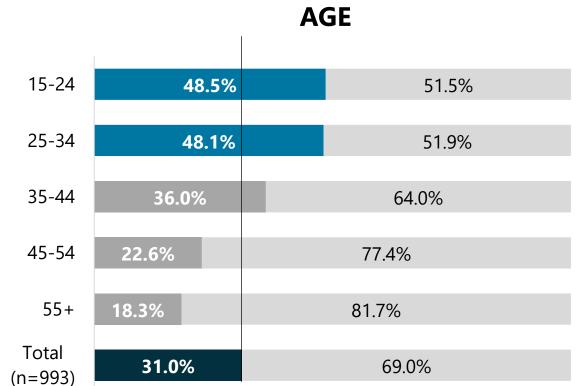
242 / 1049 [or 23% of all respondents]; 242 / 434 [or 56% of those who witnessed online harms] [n=242; 42 preferred not to answer; 150 have not witnessed GBOH; Total: 434]

Prevalence of GBOH by Gender and Age

The majority of those who experienced or witnessed GBOH are between ages 15 to 34.



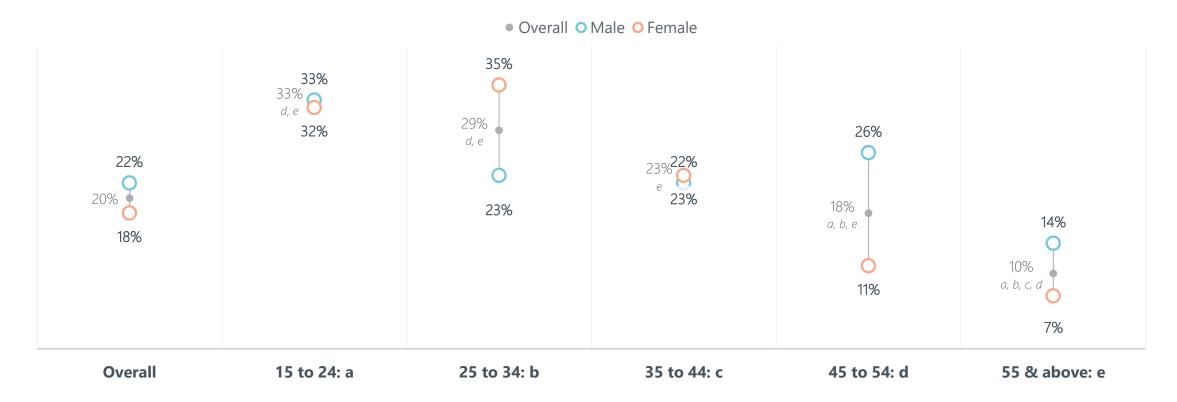
Either experienced GBOH personally or witnessed GBOH (or both)



Either experienced GBOH personally or witnessed GBOH (or both)

Breakdown of respondents who reported having experienced GBOH

Females aged 25 to 34 were most likely to have ever experienced at least one form of online harms listed due to their gender.

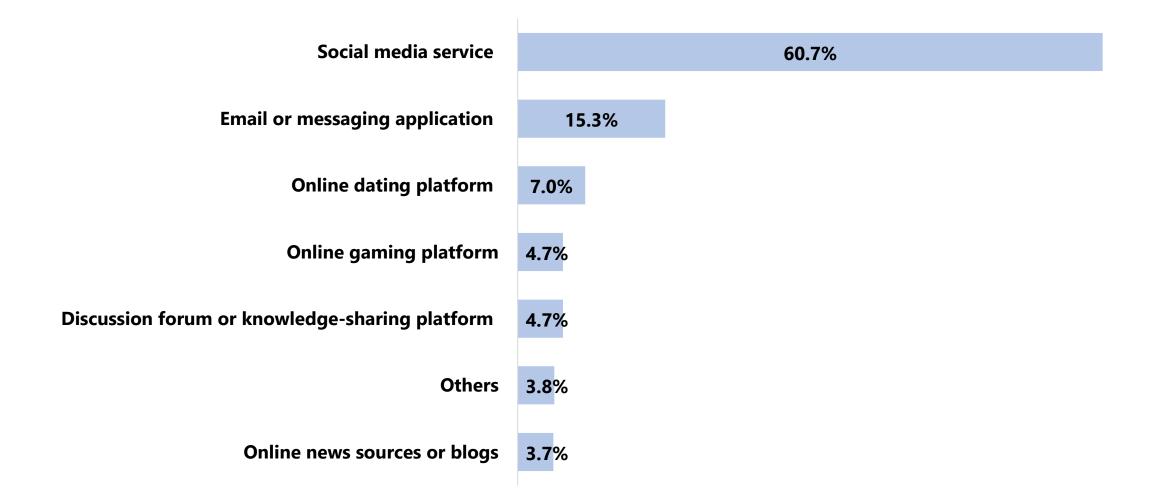


Notes:

- Data cell weighted by gender, age, race. n=1049.

- Alphabet denotes post-hoc significant difference(s) at a 95% level, amongst Age Groups for Overall figures..

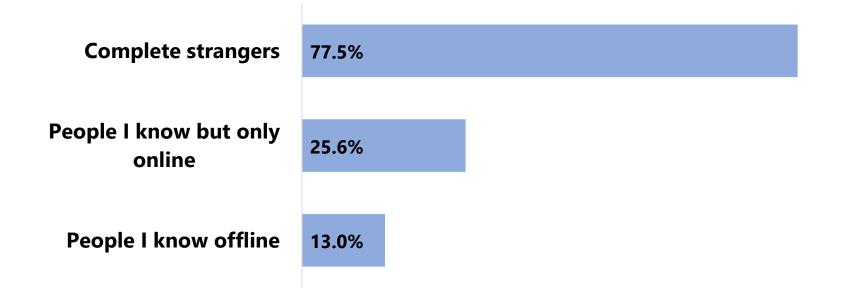
Platforms Where Respondents Experienced GBOH Most



On which platform have you mainly experienced gender-based online harms? [n=209]

Perpetrators of GBOH

Perpetrators of GBOH are overwhelmingly strangers.

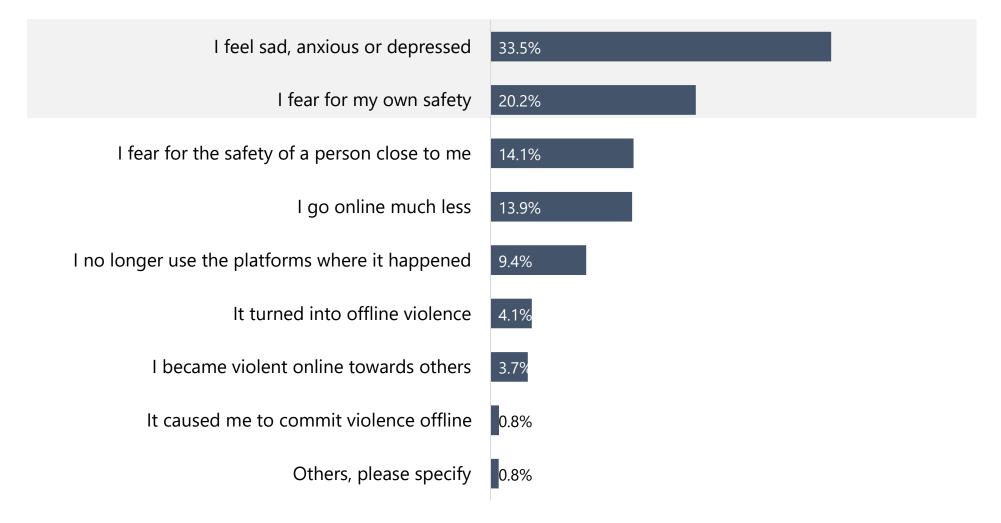


Who commits acts of gender-based online harms against you? [n=209]

Impact and responses to online harms

Main impact of GBOH

Those subjected to GBOH reported psychological impact such as fear and anxiety.



Which of the following is the main impact of gender-based online harms on you? [n=129 who reported at least one GBOH impact]

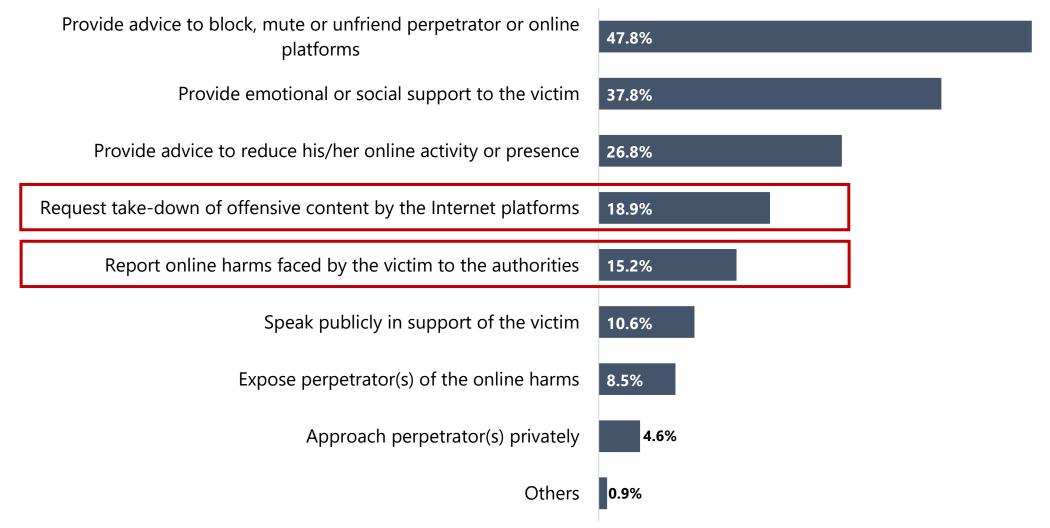
Action taken after experiencing GBOH

Blocking the perpetrator and changing privacy settings are the most common actions taken. Few respondents sought help or filed reports to platforms and/or relevant authorities.



Actions that witnesses took after seeing or hearing about GBOH

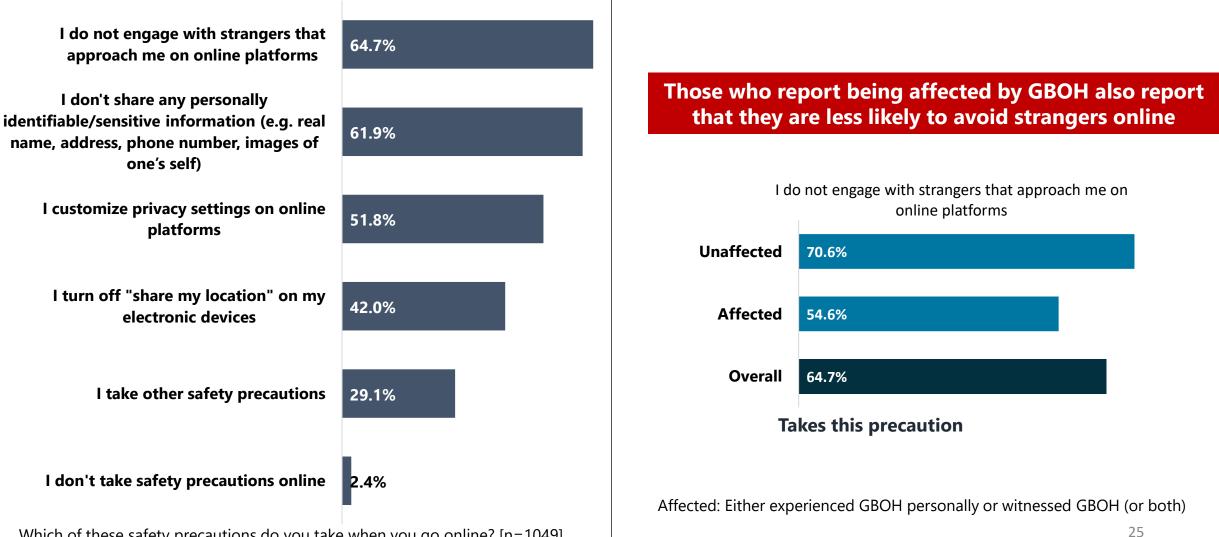
Few witnesses advised victims to report GBOH to platforms or relevant authorities



Which of the following actions did you take after seeing or hearing the incident(s)? [n=208, 34 did not do anything; Total n=242]

Safety precautions respondents take when going online

As a safety precaution, over 60% of respondents do not engage with strangers online and share personally identifiable information.

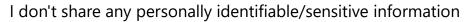


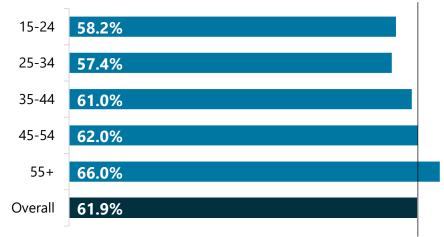
Which of these safety precautions do you take when you go online? [n=1049]

Safety precautions respondents take when going online by age

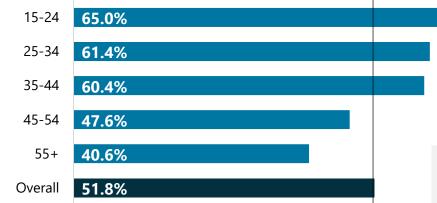


I do not engage with strangers that approach me on online platforms

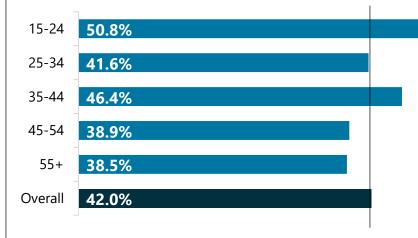




I customize privacy settings on online platforms



I turn off "share my location" on my electronic devices



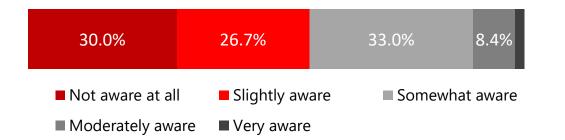
Younger respondents are significantly more likely to take safety precautions that require higher digital / technical literacy like customising privacy settings and turning off location sharing.

Which of these safety precautions do you take when you go online? [n=1049]

Help-seeking

Awareness of help-seeking avenues

More than half (57%) are unaware or only slightly aware of help-seeking avenues if they experience any online harms



To what extent are you aware of the help-seeking avenues available if you experience any online harms?

Male (n=510) 52.2% 47.8% Female (n=539) 60.9% 39.1% 15-24 (n=136) 48.9% 51.1% 25-34 (n=178) 55.9% 44.1% 35-44 (n=179) 58.5% 41.5% 45-54 (n=180) 55.8% 44.2% 55 + (n = 375)59.4% 40.6% Unaffected (n=685) 61.8% 38.2% Affected (n=308)46.4% 53.6% Overall (n=1049) 56.7% 43.3% Low awareness Aware

Affected: Either experienced GBOH personally or witnessed GBOH (or both)

Females are less aware of help-seeking avenues

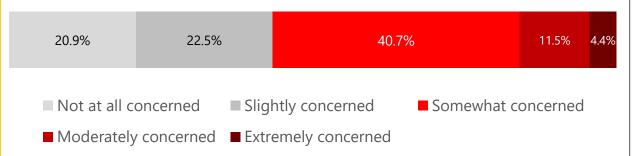
Measures that encourage action after GBOH **Deterrents of action after GBOH** Having reporting systems for complaints is the top The perception that taking action would not make that will encourage any difference, not knowing what to do and the respondents measure who anonymity of perpetrators are key deterrents to experience GBOH to take action. taking action. Reporting systems for complaints 53.0% Legislation or laws to address and prosecute Thinking it would not make any 49.4% 43.6% online harms difference Awareness of personal rights online 43.7% Not knowing what to do 43.4% Culture of reporting online harms without 40.9% fear or shame Not knowing the identity of the 43.2% abusive users Support group to seek help/support from 38.4% Agency/regulator to implement and enforce Shame or guilt 29.5% 37.2% online safety Online platforms to enforce acceptable 33.4% content and behaviour Fear of damaging reputation 27.3% Educational campaigns on threat of online 31.6% harms Others 1.5% Others

What would encourage you to take action if you are a victim of online harms? [n=1049]

What would discourage you from taking action if you are a victim \mathfrak{G} online harms? [n=1049]

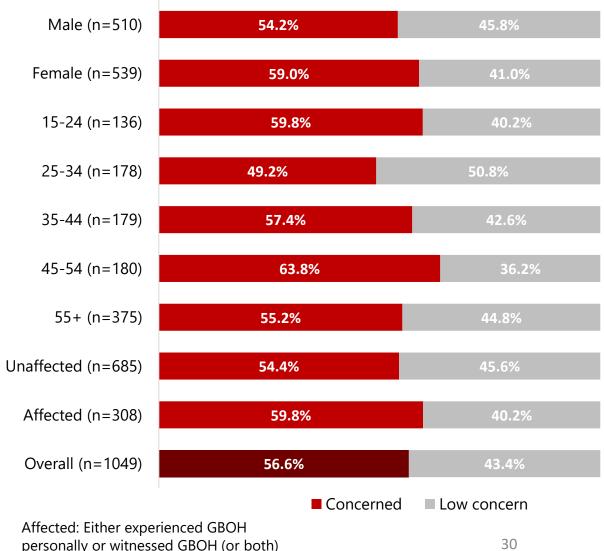
Concerns about social stigma

Slightly more than half (57%) would be personally concerned about social stigma if they had to report online harms.



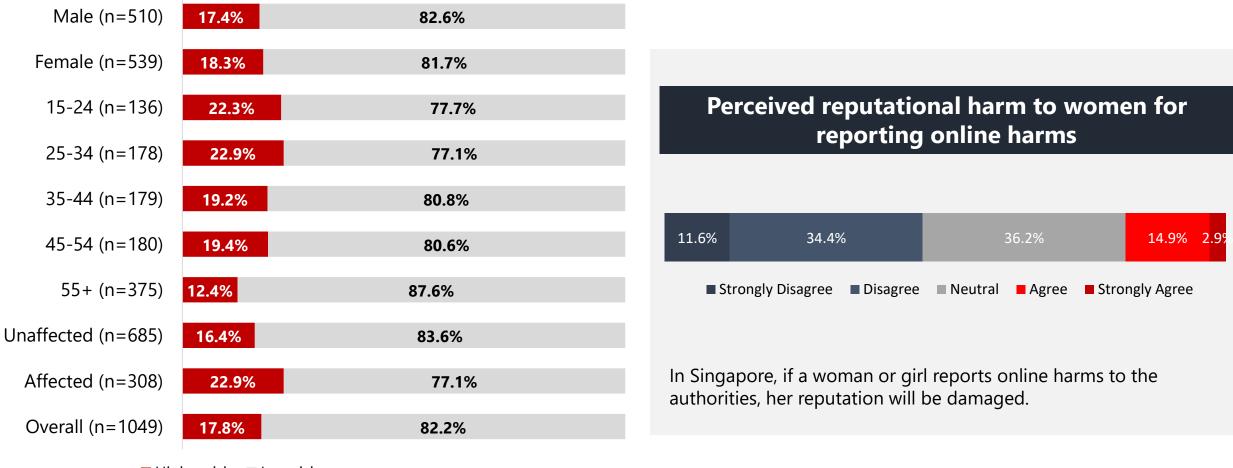
How concerned are you with social stigma if you were to report online harms?

Concerns about social stigma by gender, age and experiences with GBOH



Perceived reputational harm by gender, age and experiences with GBOH

The perceived reputational harm for women and girls who report online harms is assessed to be low, although younger respondents were more concerned.



■ Higher risk ■ Low risk

Higher risk: Agree or Strongly agree that a woman or girl's reputation will be damaged if she reports online harms

Affected: Either experienced GBOH personally or witnessed GBOH (or both)

Perceptions of most effective solutions to GBOH

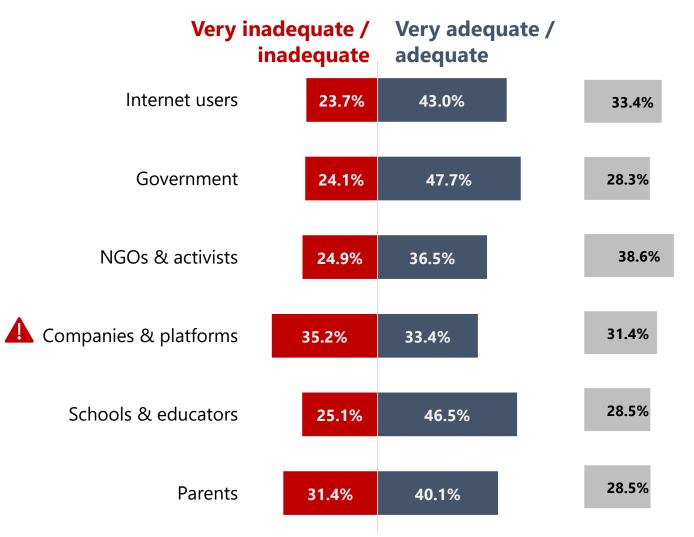
Stricter enforcement of relevant laws and public awareness campaigns and programmes were perceived to be the most effective solutions to reducing GBOH.

Stricter enforcement of laws prohibiting gender-based online harms	42.7%
Public awareness campaigns and programmes	34.0%
Stricter laws prohibiting gender-based online harms	31.5%
Stronger privacy settings and tools on online platforms	30.6%
Improved policies and practices among online service providers	18.7%
Penalizing online service providers	17.2%
Compensation for harm caused by gender-based online harms	8.3%
Others	0.8%

What would be most effective in reducing gender-based online harms? [n=1049; Choose up to 2]

Perceived effectiveness of stakeholders' current response

Companies and platforms (sites where online harms occur) have the most room for improvement in their response to GBOH.



How would you rate [stakeholder's] current response to gender-based online harms? [n=1049]

33

Sunlight AfA Detailed Research Roadmap

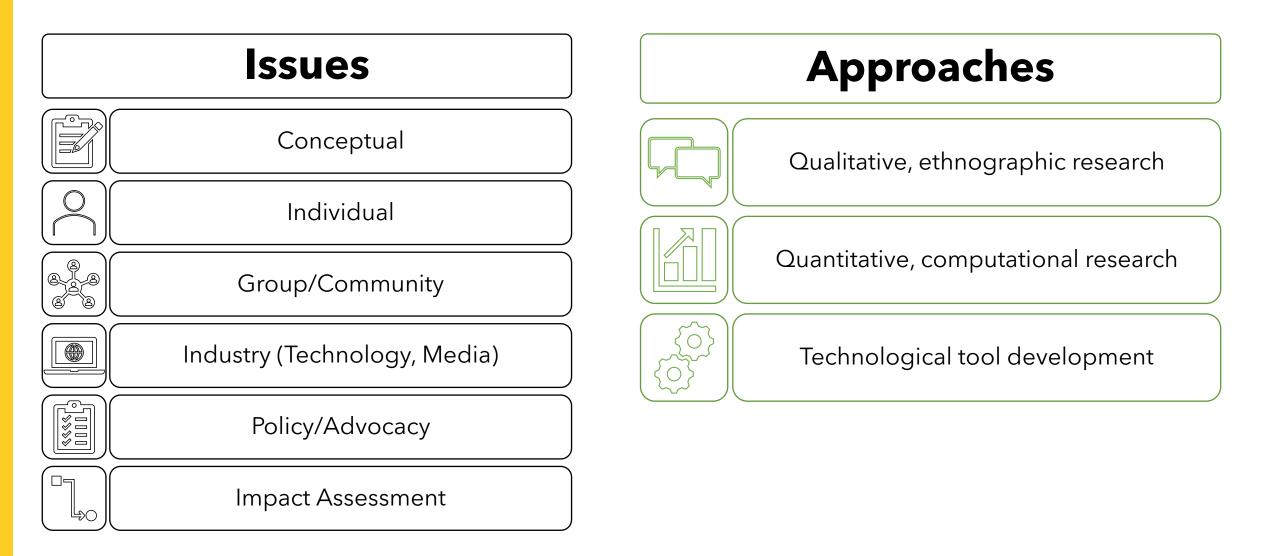


SUNLIGHT AFA RESEARCH ROADMAP

Building on the sensing poil findings, the AIA developed a radomap for Singapore's research community and engaged task-holders from the academic community as well as the people and public sectors to distill the key research thrusts to inform long-term efforts to tackle online harms. In doing so, the roadmap seeks to guide future research to depend on our understanding of online dynamic, card potential interventions and asses their effectiveness.



Overview of Research Roadmap



Conceptual



- Defining harms: Clarify/quantify impact and severity of harms, as well as evolving standards of the spectrum of harms across different platforms in different virtual realms, which helps calibrate proportionality of responses including reporting and responses by companies, criminal liability, and enforcement
- Nomenclature: Develop accurate and value-neutral terms to describe different types of online harms
- Develop scales of online harms based on (i) severity ("how awful") and (ii) legality ("how lawful"). Technology companies can retrieve harmful/objectionable content they have taken down and invite different stakeholders to "tag" the harms according to their severity and legality to establish the "community standards" and need for regulatory action
- Investigating (mis)assumptions about online harms e.g. men tend to be perpetrators and women are more likely to be subject to online harms, females would not perpetrate gender-based online harms on other females, what happens online stays online

Individuals

- Victims: understand feelings experienced; impact on physical and mental wellbeing; support required; victim-centric concepts of online harms
- Examine victim experiences during and after reporting incident to authorities (e.g. to police), barriers and concerns regarding reporting
- Bystanders: develop interventions for them to support victims; prevent diffusion of responsibility
- Perpetrators: understand perpetrator intent and psychology; high risk demographic profiles; rehabilitation process
- Accomplices: understand their "gatekeeping" role in aiding and/or abetting online harms and develop ways to intervene
- Understand and develop individual awareness and resilience in responding to online harms
- Understand impact of family structures and dynamics; class; household composition on individual resilience
- Understand risk factors behind individuals': (i) susceptibility to online harms; (ii) likelihood of perpetrating online harms

Group/Community

- Understand group norms in different online platforms, and reticence in calling out bad behaviour
- Develop social resilience: study support for victims from social groups and networks (e.g. direct assistance or other resources); deep-dive into issues of resource capacity and nature and scope of support
- Understand influence of religious and cultural groups (e.g. ethnic groups; youth groups; pop cultures): relationship across three aspects (i) social stigma; (ii) community and religious leadership and; (iii) discourses on online harms (comparative research across religions)
- Design community responses to online harms to help those who lack family support
- Understand subgroups of interest: families (guidance on parental mediation and bridging inter-generational gaps), young males (exposure to porn, normalisation of attitudes towards sexual relationships), vulnerable groups (e.g. domestic workers)

Industry (Technology, Media)



- Industry best practices: platforms for knowledge sharing and industry self-regulation including industry codes of practice
- Efficacy of safety features and reporting systems on media and tech platforms and the data required for evaluation
- Initiate industry-academia research collaborations
- Value of CSR initiatives
- Guidelines on privacy protection/codes of professional practice in reporting sexual incidents
- Role of app design and metrics in promoting hostile online environments e.g. dehumanisation, monetisation, use of avatars
- Develop ways to detect and report so that the burden of responsibility does not rely only on the individual

***** * * *

Policy/Advocacy

- Comparative analysis of global approaches e.g. legislation to enhance online safety
- Develop deeper understanding of regional and Asian policy frameworks and approaches, assess feasibility of regional synchronization
- Feasibility analysis of a dedicated enforcement agency for online harms
- Study role of mainstream education to raise youths' understanding of online harms and their responsibility and agency (up to and including tertiary education)
- Enhance public understanding and education about evolving forms of personal data, how they are used in detection/reporting

Impact Assessment

- Periodic assessment of effectiveness of public education, legislation, technology company responses
- Perceptions of different respondent profiles and trends over time
- Tracking impact of online harms in the real-world context
- Longitudinal analysis of impact of policies

Qualitative, ethnographic research

 Interviews, focus groups, discourse and conversational analysis, case-based research, online ethnography (rich contextualised findings on victim trauma, bystander attitudes and interventions etc.)

Quantitative, computational research



- Panel study to follow a pool of participants for assessing longitudinal effects
- Developing accurate survey measures for sensitive questions relating to online harms (tracking non-response rate to particular questions, by different respondent profiles etc.)
- Machine learning supported content analysis of videos and discussion forum threads with harmful content (trends in engagement with misogynistic content, common tropes etc.)
- Assess and analyze data from past police and formal security complaint reports

Technological Tool Development

- Detect misinformation and online harms with optimisation for multilingual context, including automatic and AI driven detection
- Collect evidence for claim verification
- Analyse social media content to ascertain if images have been repurposed or taken out of context
- Construct profiles of people to help verify identity of persons to potentially befriend
- Develop design nudges to guide or encourage prosocial behaviour
- Machine learning supported content analysis of videos and discussion forum threads with harmful content (trends in engagement with misogynistic content, common tropes etc.)
- Data sharing arrangement for tech companies and academia to collaborate on technological tool development

Please contact the Sunlight AfA secretariat (email: <u>MCL AfA_Secretariat@mci.gov.sg</u>) if you have any questions about the report or use of the findings.