Minorities, Advocates and Digital Harms: Evaluating Support and Guidance Materials

Rob Cover,[1] and Shiyang Zhu
RMIT University, Australia

Abstract: This paper reports findings from a study of the wellbeing, support and remedy guidance needs among minority communities subject to digital harms. The study developed and tested six guidance documents specifically for minority populations: online abuse and harassment; disinformation; scams; doxxing, deepfakes and how to report digital harms. Organizations which provide support services and/or advocate for minority groups—CALD and migrant communities, LGBTQ+, people with disabilities and older people—were invited to evaluate the draft guidelines. Participating organizations overwhelmingly felt that each of the issues were serious for their communities, but considered guidelines targeting minorities 'in general' rather than particular communities did not address specific needs or allow marginalized people to see themselves represented. Some reported that the format, language and readability may be inaccessible to some minority populations. The study found that generic guidelines to help protect users' wellbeing, increase knowledge of digital harms and manage remediation were deemed less useful, in favor of co-designed community-specific guidelines.
Keywords: Digital harms; Minorities; Guidance; Online abuse; Disinformation; Reporting.

Minority identity groups are subject to higher rates of online abuse, harassment and doxxing, and are more often the subject of disinformation and misleading synthetic media than other users in the digital ecology of the 2020s (Obermaier et al., 2023; Walther, 2022). This is particularly the case for migrant and ethnic group minorities, gender- and sexuality-diverse minorities (LGBTQ+), culturally and linguistically diverse populations living in wider nation-state communities (UN Special Rapporteur on Minority Issues. 2022), and people living with disabilities (Burch, 2018; Graaff, 2021). Despite nearly a quarter of a century of legislation and international initiatives to combat hate speech, it has flourished across digital platforms in ways which are harmful both to individual users and the security of the global community (UN Special Rapporteur on Minority Issues, 2022). At the same time, online content and behavior that falls short of hate speech but is nevertheless harmful to some users continues to increase—including insulting, offensive or misleading stereotypes (Poland, 2016), persistent trolling and other forms of harassment (Lupton, 2015; O'Connell et al., 2024), discriminatory synthetic media (Meikle, 2023; Tan, 2020), and attempts to defraud through targeted digital scams (ACCC, 2022).

---

[1] Corresponding author: Professor of Digital Communication and Director, Digital Ethnography Research Centre, RMIT University, La Trobe Street, Melbourne, Australia. E-mail: rob.cover@rmit.edu.au

Increasing rates of online hostility has sponsored inquiries, legislation and policy initiatives in a range of jurisdictions (Flew, 2021), concerns about wellbeing and mental health (Keighley, 2022), calls for more effective platform moderation (Gillespie, 2018), and more stringent state regulation of platforms (Christchurch Call, 2019). Arguably, platforms can no longer be relied upon to intervene and prevent online harms—both in general and in regard to high-target minorities—leaving the labor of remedy to victim-survivors, nation-state governments, and community support groups (Dogan & Karaosmanoglu, 2025). Calls for increased support for minorities in the form of better guidance on rights, available assistance and clearer definitions of the kinds of digital harms to which users may be subject have been made (e.g., Thomas et al., 2022). Some governments that have established digital regulatory bodies have attempted to provide guidance documents for victim-survivors of digital harms, including information on how to seek remedies and information on self-care practices. For example, the United Kingdom's OfCom, auspiced by the Online Safety Act 2023, provides guidance for consumers including how to report harmful online content (OfCom, 2023). Likewise, Australia's eSafety Commissioner has published guidance describing cyberbullying of adults and children, advice on distressing content, and tips on how to deal with fraud attempts such as sextortion (eSafety Commissioner, 2025). However, in most cases, the guidance is addressed to the *general user*, ignoring the specific needs of minority victim-survivors such as language, content accessibility and fear of secondary victimization, and specific, targeted forms of abuse (Waqas et al., 2019) Given the negative health and wellbeing impact on minority community members subject to online hostility and hate speech (Waqas et al., 2019), there is an urgent need to understand what kind of guidance supports minority community users when targeted, experiencing or witnessing digital harms, including scams, abuse, harassment and hate speech.

In this context, with limited regulatory initiatives, difficulties navigating the interjurisdictional setting (Vincent, 2017), unclear platform policies and limited intervention and moderation (Wachs and Wright 2018; Wilson 2019), more guidance on definitions, terminology, remedy, self-care and support is needed for minorities. This is a point emerging particularly in light of and, more alarming retraction by several major platforms from their already limited policies that ban hate speech against minorities and the supply of reactive moderation and intervention on digital platforms, particularly X (formerly Twitter) and the Meta platforms (Facebook, Instagram and Threads) (Frenkel et al., 2023; GLAAD, 2025; Isaac & Schleifer, 2025). More research is needed on (i) what kind of guidance for the cultural specificity of minority community members is needed, (ii) the form and accessibility of guidance documents and texts, (iii) inclusion and exclusions of information, and (iv) the range of digital harms most pertinent to specific minorities where increased knowledge may lead to remediation.

This article draws on research undertaken for a major project combatting digital harms to understand the forms of guidance needed and recommended for minority communities vulnerable to online abuse, harassment, misleading information, fraud and hate speech. To understand the guidance needs of minorities, we drew on analyses of platform policies, statutory law and survey data, ethnographies focused on care, protection and intervention among a range of users and victim-survivors of digital harms, and stakeholder engagement, and developed a set of six (early draft) two-page guidelines in the form of fact sheets to elicit clear community-specific needs. The fact sheets outlined key data, issues and remedies relevant to minority groups—broadly speaking—in Australia, and were circulated to key informants whose lived experience or role in support can help us understand guidance needs in more inclusive ways. They were designed on the theoretical premise that mutual care and a sense of interdependency is an important resilience factor in managing being targeted by digital harms (Cover, 2024).

This article outlines the key findings from the evaluation of elicit guidance material with the intent of providing new knowledge to those whose regulatory, support and statutory

responsibilities guide and educate members of the public about harmful digital content and behaviors and available remedies—government agencies, platforms, supra-national representation bodies, and service providers. In light of the scant literature on minority guidance documentations in relation to digital harms, and the wider paucity of evaluations of intervention measures for migrant minority communities and communities of people with a disability (Blaya, 2019) we begin immediately with background on how our elicitation material (draft fact sheets) was developed and the process of having minority community support and advocacy organizations evaluate them to inform community-specific needs vis-à-vis increasing rates of digital harms. We analyse their quantitative and qualitative responses to the seriousness of each of the topic issues raised by the fact sheets, their assessment of the usefulness, and the recommendations made for increasing utility and value for their communities.

Developing Guidance to Elicit Culturally-Specific Needs

It is now increasingly understood in scholarship, policy and community attitudes that problematic, offensive, hateful or disinformation/false content online can have serious harms for users and for the quality of the digital ecology. It is recognized that both governmental and platform regulatory measures are broadly ineffective and unsatisfactory in dealing with digital harms. And there is nascent knowledge that users are increasingly turning to self-care and mutual care to manage individually and collectively the impact of digital harms (Cover, 2022a, 2024), often resulting from the inconsistent and haphazard moderation and intervention practices are neither globally applied nor recognize the specific needs or injuries relevant to specific minorities (Jiang et al., 2021). The express needs of minority communities in regard to self-care, understanding and mutual support is unknown. The *Online Hostility in Australian Digital Cultures* project investigates a broad range of everyday experiences of online abuse and other digital harms, with one focus point on the needs of members of minority communities in Australia, particularly CALD, LGBTQ+, people living with disabilities and older Australians.

One mechanism to assess the needs of minorities has been to draw on project findings from ethnographies and stakeholder engagement, in combination with extant literature, to develop early drafts of guidelines documents on a range of areas, and have leads of organizations which support, care for or advocate for various minority groups assess their value and describe their concerns, any absences, special requirements for their organization, the format of the guidelines and the extent of their interest in sharing edited versions with the communities or clients, either in print (e.g., in a lobby) or digitally (e.g., on a website or electronic newsletter).

The fact sheets were on six topics that represent a range of what we refer to as the 'umbrella' of contemporary digital harms: (1) Online abuse and harassment; (2) Disinformation; (3) Online and Mobile Scams; (4) Doxxing and (5) Deepfakes. We added an additional guidance document to ensure that exposure to the above could be more quickly remediated: (6) Reporting online abuse. The inclusion of this explainer was in response to new knowledge that not enough people generally are aware of reporting options when encountering or experiencing digital harms (Australian Government, 2023). These topics were chosen as a starting point for eliciting needs among communities. We chose to include the cognate area of digital and mobile scam communication which, we argue, is a contemporary digital harm albeit one usually addressed through different regulatory mechanisms (e.g., criminalization of financial and identity fraud) due to the higher rate of victimization of those who belong to minority groups, including particularly older members of those groups. In our view, the use of digital communication devices to communicate scam messages is part of a wider problematic use of the digital, is known to be psychologically harmful even when scam attempts are not

successful (Alves & Wilson 2008; James et al., 2014) and is usefully seen as part of the picture of everyday toxification of digital communication.

Although primarily focused on abuse and harassment, we included the topics of doxxing (sharing another user's private details, workplace, family member names or contact details with the intent of encouraging others to harass) and deepfakes (synthetic media) due to their topicality in Australia during this study and, again, the recognition that both of these harmful digital formations have a significant negative impact on minorities (Meikle, 2023). A high-profile doxxing scandal emerged in Australia in early 2024 when the names and details of members of WhatsApp group of Jewish writers creative artists discussing privately their position on the Gaza War (Taylor, 2024). New privacy legislation was enacted in Australia to criminalize doxxing as a result, making the offense punishable by up to seven years' imprisonment under the Privacy and Other Legislation Amendment Act 2024. Deepfakes, or AI-generated videos in which a subject's face or body has been digitally altered to make them look like someone else (Cover, 2022b), have become increasingly topical over the past five years alongside other synthetic media such as digitally-generated images, particularly in relation to elections and the generation of disinformational scandal; some studies have indicated the very high propensity of everyday users to fail to recognize deepfakes and/or to require better guidance on how to recognize them as non-authentic (Meikle, 2023).

In the form of draft guidance documents differentiated from 'generic user' support information available online, our elicitation material drew on everyday user accounts through ethnographic work with victim-survivors in this project, and the authors' related studies. A fact sheet format (two pages per document) comprised of a short introduction, dot-point factual information under discrete sub-headings and links to further information or resources was chosen as the most appropriate format for the guidelines, given fact sheets have long been recognized as the most manageable design format for the strategic delivery of information quick, succinctly and to generate uptake and knowledge acquisition (Eggensperger & Salvatore, 2022). Readability was tested using Microsoft Word's readability statistics, with each fact sheet edited to have a Flesch-Kincaid grade level of no higher than 13—that is, at the level of accessible reading suitable to a high-school graduate and appropriate for a target audience of general adults (Williamson & Martin, 2010).

Methodology

Given that the goal of our study was to uncover the culturally-derived perspectives on the *value* of written guidance documentation for culturally- and linguistically-diverse minority communities, including gender- and sexuality-diverse communities and people living with a disability, we approached the project by drawing on approaches to *care* that recognize justice, resilience and liveability as achievable through frameworks of self-care and mutual care grounded in interdependency among communities and community specificity (Cover 2024). We therefore tested the draft guidance material as an elicitation exercise by gathering feedback in the form of Likert-scale questions and open-ended responses from a selection of community advocacy organizations serving as key community respondents and representatives who are central to support and care processes among their communities, and with key informant knowledge about community needs.

This study adopted a mixed-methods, participatory evaluation design to investigate the perceived usefulness, accessibility, and cultural adequacy of digital harms guidance materials for minority communities in Australia. The methodological approach was informed by scholarship on digital harms, care ethics, and co-design with marginalized communities, which emphasizes the importance of participatory knowledge production, cultural specificity, and attentiveness to differential vulnerability within digital ecologies (Ahmed 2021; Gillespie 2018; The Care Collective 2020).

Participant recruitment

Organizations providing support, advocacy, or services to minority populations in Australia were purposively sampled. These included organizations working with culturally and linguistically diverse (CALD) and migrant communities, LGBTQ+ communities, people living with disabilities, and older Australians. Australians. The aim was to gain key insights on what sort of guidance was perceived to work for a range of different minority community groups, attentive to intersectionality (Hancock, 2011), without risking the circulation of draft-only guidelines among at-risk minority community members themselves. This organizational-level recruitment strategy was adopted to avoid exposing individual community members to draft materials while still drawing on expert, practice-based knowledge of minority needs and vulnerabilities.

Ninety-four organizations were invited via email to participate in an online evaluation. Seventy organizations completed evaluations of at least three fact sheets, yielding a response rate sufficient for thematic saturation across qualitative items. Responding organizations represented disability advocacy and support (42%), migrant and CALD communities (26%), gender- and sexuality-diverse communities (17%), and older persons (15%), with several organizations supporting intersectional constituencies.

Data collection

Data were collected through an online survey instrument combining structured Likert-scale questions and open-ended qualitative prompts. Quantitative items assessed perceptions of the seriousness of each digital harm, the usefulness of each fact sheet for the organization's community, clarity of language, and interest in sharing revised versions. Qualitative questions invited detailed feedback on accessibility, cultural relevance, missing information, language requirements, format, and community-specific concerns.

This combination of quantitative and qualitative data enabled both descriptive comparison across harm categories and in-depth exploration of how guidance materials intersect with lived experience, organizational practice, and community trust. The approach aligns with mixed-methods strategies commonly used in digital harm and wellbeing research (Costello et al. 2019; Thomas et al. 2022).

Data analysis

Quantitative data were analysed descriptively to identify patterns in perceived seriousness, usefulness, and interest across the six topic areas and organizational types. Qualitative responses were analysed using reflexive thematic analysis, following Braun and Clarke's iterative coding approach. Initial codes were developed inductively, with attention to issues of accessibility, representation, language, cultural specificity, and fears of secondary victimization. Themes were refined through repeated comparison across organizational categories, enabling identification of both shared concerns and community-specific divergences.

Ethical considerations

The study received institutional ethics approval from RMIT University Human Research Ethics Committee (Approval number 26459). Participation was voluntary, organizational identities were anonymized, and no individual service users were involved. This design was chosen to minimize risk while still centring the knowledge of those with

responsibility for supporting communities disproportionately affected by digital harms. The methodological emphasis on elicitation and co-design reflects ethical commitments to care, mutual responsibility, and the avoidance of extractive research practices in work with marginalized populations (Ahmed 2021; Cover 2024).
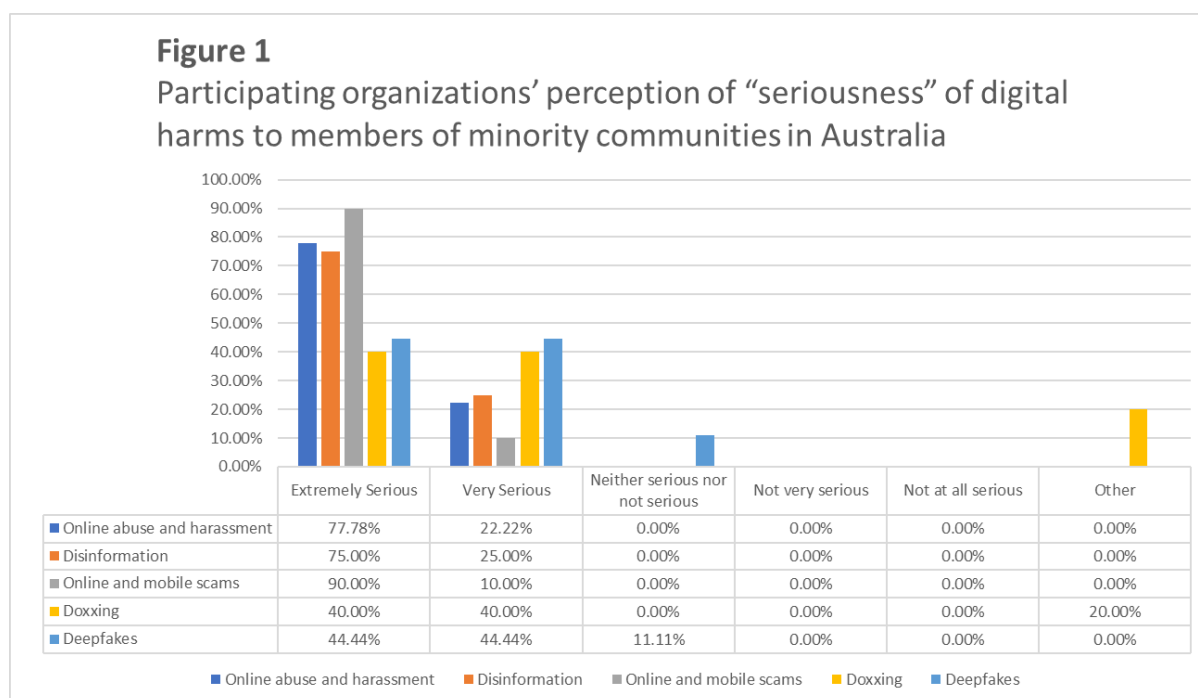
Findings

Through coding and analysis of responses given via a survey tool as both Likert-scale and qualitative answers, four emergent areas of significance were revealed related to digital harms, guidance for understanding contemporary problematic online content, and for managing wellbeing or reporting as relevant to minority communities: (1) that minority community support and advocacy organizations perceive these digital harms to be very serious, even though none of them have an active mission supporting their communities or clients in relation specifically to the experience of digital harms; (2) that guidance is seen as important and necessary for the wellbeing of their clients and communities; (3) that fact sheets and other textual documents for general population and/or targeting 'minorities' as a group may not be as inclusive and accessible as intended; and (4) that although producing guidance for minorities— in contrast to the general population—is important, tailoring guidance for individual communities and intersectional communities may be warranted. We outline key findings organized around two key headings: *seriousness* of digital harms to minority communities and *usefulness* of guidance documentation.

I: Seriousness of digital harms

Significant in the evaluations undertaken by the participating minority community support and advocacy organizations was that the five harm areas for which there was a fact sheet (online abuse and harassment; disinformation; scams; doxxing and deepfakes) were broadly recognized as being matters that were either "extremely serious" or "very serious" for their clients/communities *as* minority subjects (see Figure 1). Two participating organizations marked "other" in the case of the Doxxing and minorities fact sheet—both indicated in commentary that the terminology of doxxing was new to them and they were yet to fully assess its social impact. Only one organization found a digital harm topic to be less serious than 'extreme' or 'very'—in this outlier instance the organization perceived deepfakes to be "neither serious nor not serious". Although no reason for this choice was provided in the commentary, it can be reasonably extrapolated that they view the AI-generation of synthetic media as, in itself, not harmful except when used for disinformation or other malicious purposes (Cover, 2022b), and arguably beginning to be recognized as a tool for creative industries in providing cost-effective post-production enhancement (Aldredge, 2020; Lees et al. 2021).

Figure 1 Participating organizations' perception of "seriousness" of digital harms to members of minority communities in Australia

Figure 1
Participating organizations' perception of "seriousness" of digital harms to members of minority communities in Australia

	Extremely Serious	Very Serious	Neither serious nor not serious	Not very serious	Not at all serious	Other
Online abuse and harassment	77.78%	22.22%	0.00%	0.00%	0.00%	0.00%
Disinformation	75.00%	25.00%	0.00%	0.00%	0.00%	0.00%
Online and mobile scams	90.00%	10.00%	0.00%	0.00%	0.00%	0.00%
Doxxing	40.00%	40.00%	0.00%	0.00%	0.00%	20.00%
Deepfakes	44.44%	44.44%	11.11%	0.00%	0.00%	0.00%

Other than this case, the organizations were in broad agreement that these five fields or types of digital harm were matters of serious concern for minorities. Our analysis suggests that this may be a combination of broad community knowledge and the lived experience of their community members or clients. It has become better known over the past half-decade that minorities, including particularly migrant populations, people with disabilities and gender and sexuality diverse people are increasingly encountering the use of pejorative or discriminatory language in many settings, and witnessing or experiencing personal attacks based exclusively on identity factors, religion or race and gender identity (Butler, 2024; UN Special Rapporteur on Minority Issues, 2022).
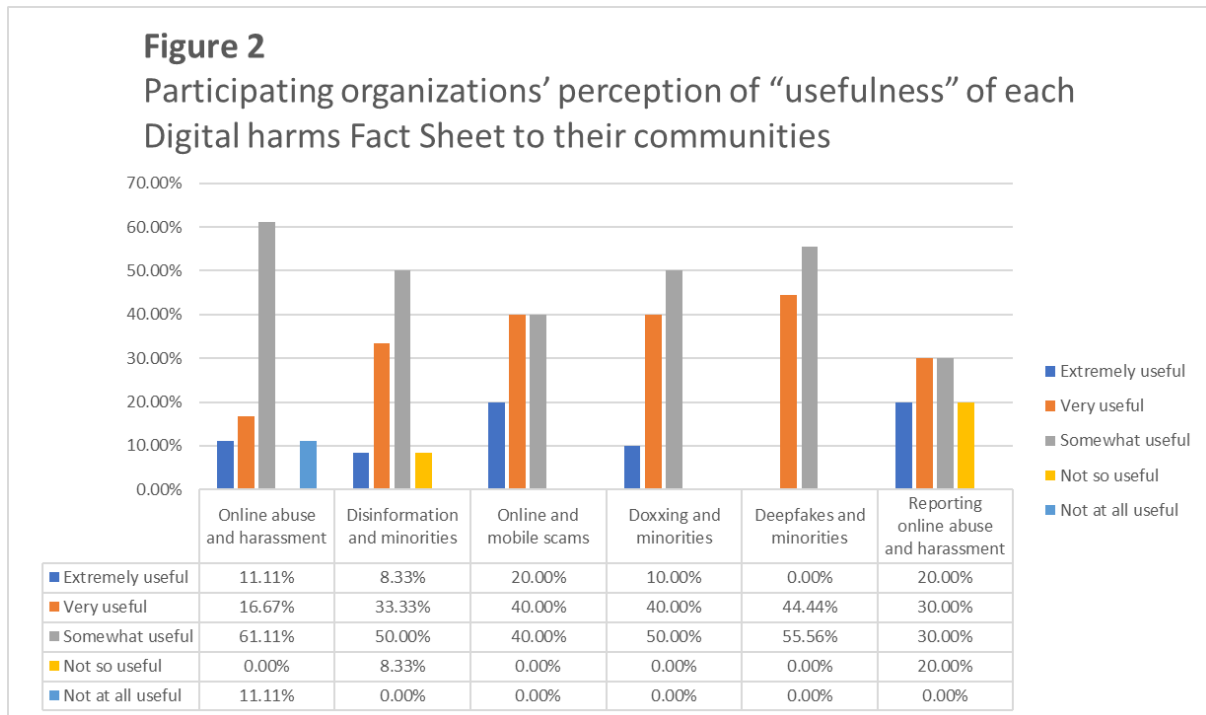
Of the eighteen organizations representing migrants or CALD community members, about one-third provided support or advocacy services or representation to Muslim community members, and *all* responded with "extremely serious" across each of the five digital harms types. This is indicative of the continuing experience of abuse and harassment of Muslim-Australians and disinformation about their community, beliefs or social practices, that has inflected Australian and international culture since the start of the century (Poynting et al., 2004), and the persistence with which harmful content about their community circulates a quarter of a century later, often in relation to global politics.

II: Usefulness of guidance

Despite the very high rating of seriousness given by the participating organizations, the participant group overall was substantially less likely to find the fact sheets in their present form useful to themselves or their community members or clients in their present form. Figure 2 outlines the participants' responses to the question "How useful do you think this fact sheet will be for members of your community (including, if relevant, clients and those for whom your organization may advocate)?" As the data indicates, all fact sheets received responses of "extremely useful" of between 8% and 20%, except deepfakes, and between 30% and 60%

rated as "somewhat useful". Some fact sheets received responses of "Not so useful" (8% for Disinformation, 20% for Reporting), and "Not at all useful" (11% for Online Abuse and Harassment).

Figure 2 Participating organizations' perception of "usefulness" of each Digital harms Fact Sheet to their communities

Figure 2
Participating organizations' perception of "usefulness" of each Digital harms Fact Sheet to their communities

	Online abuse and harassment	Disinformation and minorities	Online and mobile scams	Doxxing and minorities	Deepfakes and minorities	Reporting online abuse and harassment
■ Extremely useful	11.11%	8.33%	20.00%	10.00%	0.00%	20.00%
■ Very useful	16.67%	33.33%	40.00%	40.00%	44.44%	30.00%
■ Somewhat useful	61.11%	50.00%	40.00%	50.00%	55.56%	30.00%
■ Not so useful	0.00%	8.33%	0.00%	0.00%	0.00%	20.00%
■ Not at all useful	11.11%	0.00%	0.00%	0.00%	0.00%	0.00%

None of the six elicitation fact sheets had broad agreement of usefulness, although no greater than 20% of participating organizations rated any one fact sheet as either "Not so useful" or "Not at all useful" in respect to their communities. Only the Scams fact sheet received broad agreement higher than 50%, Doxxing and Reporting at exactly 50%, and the remainder between 28% and 44% broad agreement that they would be useful.

The fact that there was a very high recognition of seriousness of each of the topics but a much lower rate of usefulness of the fact sheets is an important, welcome and significant finding. It indicates the participating organizations' enthusiasm about addressing the issue and its effects not just through *any* guidance, but through guidance which, in their experience and expertise, will 'work' most effectively for their community groups (Ahmed, 2021). It indicates a need for resources and a careful consideration of how resources must be *appropriately tailored* for minorities. That is, the kind of *generic* guidance provided by government agencies and regulators on defining various digital harms and remedying them, and designed for broad population groups, are perceived as less useful by minority actors themselves.

Some of the key reasons given for how to improve the usefulness of each of the fact sheets involved recognizing that experiences of digital harms differ among different minority groups, rather than minority groups' experiences collectively differing from 'mainstream' users. The key finding, then, is that guidelines that 'level out' minority experiences replicate the problems of generic guidelines for all users by addressing 'minorities' as a coherent, homogeneous group. Indeed, many respondents indicated they would like to see the Fact Sheets more tailored to the individual communities they represent. In one case, an important articulation was made about the importance of not collapsing minority communities into a singular formation:
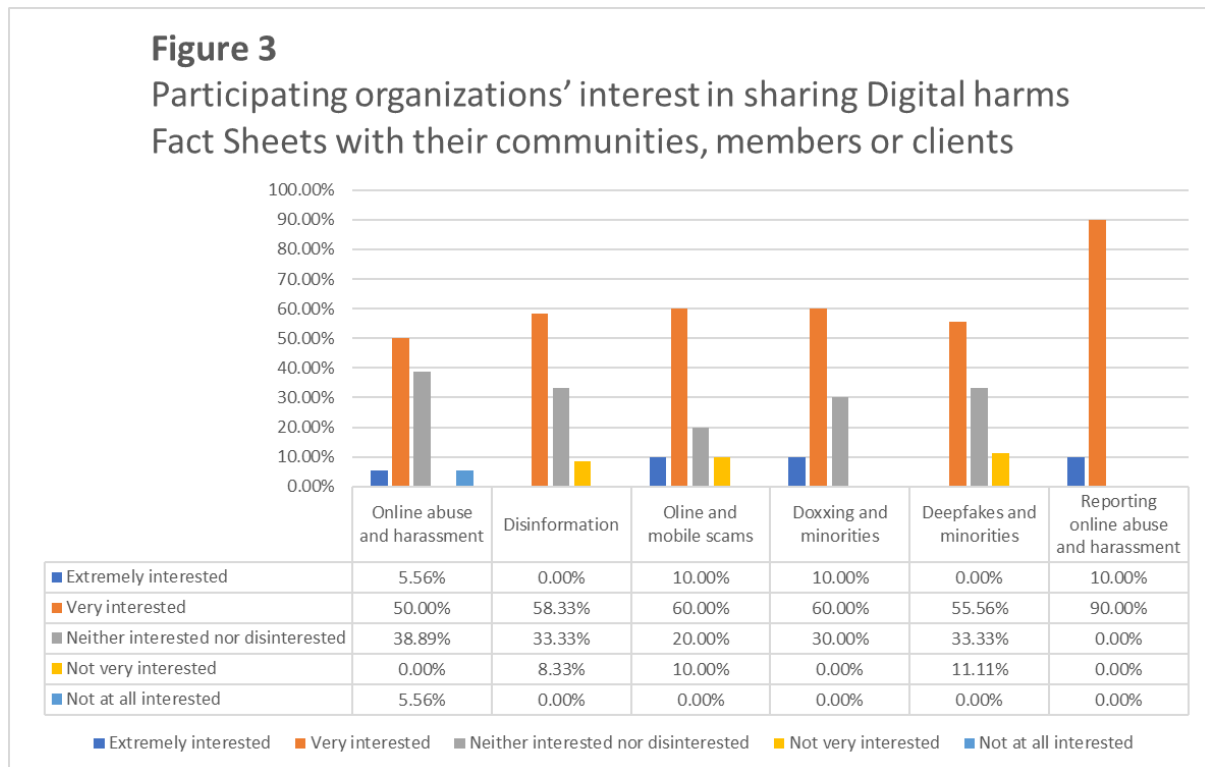
> We encourage more thoughtful framing when using demarcating terms like 'minorities'. We understand the intention, but question the value of this term in some places and encourage the authors to consider more inclusive analysis and/ or being more specific about the communities in question rather collapsing all experiences of marginalization under this term.

To identify or be identified as a member of a minority group is, as we have long known, more than a simple statistical demarcation of a bureaucratized category. Rather, as Dennis Altman (1979) outlined, minority status in a wider society involves an expectation that the greater portion of minority communities accept or tolerate a subordinate position, with only those who have the resources to perceive subordination as an injustice able to disavow that in everyday life. While greater discourse on diversity emerged in the four decades since he theorized this perspective, the late 2010s and early 2020s have seen a very substantial global rise in anti-minority sentiment, policy and withdrawal of anti-discrimination protections, including the cultural formation of pressure on members of majority groups in a population to either see themselves as wounded by the diversity discourse, or to undertake acts of subordination through hurtful language (McRobbie, 2020; The Care Collective, 2020), with much of that subjugation occurring across digital platforms (Terranova 2022).

In this context, it is therefore unsurprising that representations of organizations supporting or advocating for minority groups are at this time finding digital harms to be matters of extremely or very serious concern, but also signifying the importance of guidance documents that are optimized for their usefulness for *discrete* minorities and not a replication of extant guidance for the general population. That is, minorities are experiencing digital harms in ways different from the general population, but the distinctions between different minority groups may be more marked. This may also help explain why better resourced community organizations more willing to advocate for minority rights and wellbeing found the usefulness lower but the seriousness of the issues high. In this respect, usefulness pertains not to the content, advice or framing of the topics and the problems, *but to how the guidelines appropriately address and recognize the different communities targeted in the strategy.*

These points are underscored by our question as to the interest of each organization in sharing a refined version of the fact sheets. Figure 3 indicates broad agreement with an interest in sharing the fact sheets, ranging from 56% for Online abuse and harassment to 100% for Reporting online abuse and harassment. Doxxing and Scams each had 70% in broad agreement, while Deepfakes and Disinformation were the only ones which had no responses of "Extremely interested", perhaps because deepfakes are less well-understood in terms of their implications for harmful activities, although given the public attention to disinformation, the lower result (58% in broad agreement) was surprising. However, in thinking the overall results of seriousness, usefulness and interest together, we are able to discern a high degree of investment in the *need* for fact sheets whereby refinement towards their usefulness and useability for specific communities—as outlined below—will be the key criteria for the uptake of guidance on digital harms.
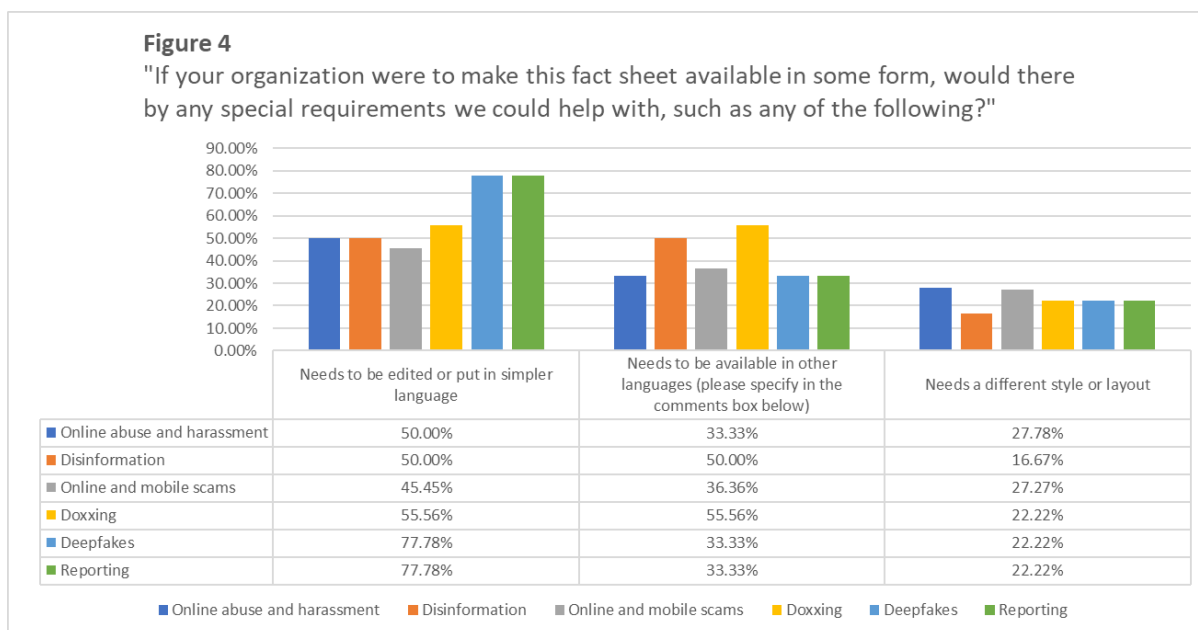
Figure 3 Participating organizations' interest in sharing Digital harms Fact Sheets with their communities, members or clients

Figure 3
Participating organizations' interest in sharing Digital harms Fact Sheets with their communities, members or clients

	Online abuse and harassment	Disinformation	Oline and mobile scams	Doxxing and minorities	Deepfakes and minorities	Reporting online abuse and harassment
Extremely interested	5.56%	0.00%	10.00%	10.00%	0.00%	10.00%
Very interested	50.00%	58.33%	60.00%	60.00%	55.56%	90.00%
Neither interested nor disinterested	38.89%	33.33%	20.00%	30.00%	33.33%	0.00%
Not very interested	0.00%	8.33%	10.00%	0.00%	11.11%	0.00%
Not at all interested	5.56%	0.00%	0.00%	0.00%	0.00%	0.00%

Seriousness and interest speak to the growing evidence that the social burden of minority belonging can, for a greater proportion than the general population, result in added vulnerabilities and poor resilience when confronted with risk, change, unwellness or a sense of lost self-agency (Factor et al., 2011). There is also growing evidence that experiencing digital harms, including those such as insult and offensiveness that fall short of international hate speech definitions, are having a significant negative impact on wellbeing, trust, social participation and liveability (Costello et al., 2019).

We asked participants to provide details of special requirements needed if they were to share the fact sheets with clients or members of their communities. Figure 4 details the key responses. There was a substantially strong response to the need for editing (which we detail in section [A] below), as well as strong indications for a desire for a range of language translations responded to primarily by the migrant and CALD community representative organizations. There was also a small need for a different style or layout, primarily to ensure that the text could be captured by screen readers for users with a visual impairment (which we again deal with in section [A] below). Participants were also invited to provide details as to what will improve the fact sheets in regard to clarity, information, accessibility and framing of the topics. In addition to the points described above in regard to minority positioning and tailoring to communities A number of discrete reasons were given, and we will break three of the key themes below.
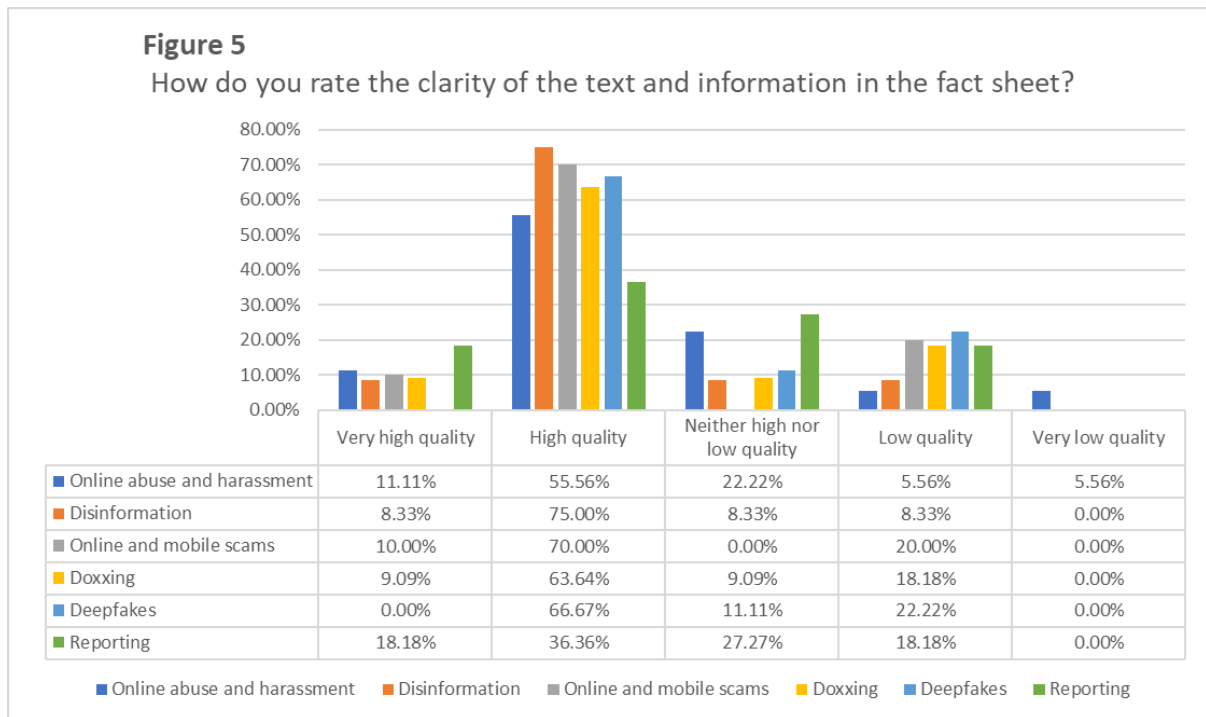
Figure 4 Participating organizations' feedback to enhance usefulness and effectiveness

Figure 4
"If your organization were to make this fact sheet available in some form, would there by any special requirements we could help with, such as any of the following?"

	Needs to be edited or put in simpler language	Needs to be available in other languages (please specify in the comments box below)	Needs a different style or layout
■ Online abuse and harassment	50.00%	33.33%	27.78%
■ Disinformation	50.00%	50.00%	16.67%
■ Online and mobile scams	45.45%	36.36%	27.27%
■ Doxxing	55.56%	55.56%	22.22%
■ Deepfakes	77.78%	33.33%	22.22%
■ Reporting	77.78%	33.33%	22.22%

■ Online abuse and harassment　■ Disinformation　■ Online and mobile scams　■ Doxxing　■ Deepfakes　■ Reporting

Discussion: Improving guidance for minority communities

[A] Simplification, readability and accessibility

A relatively high number of participating organizations indicated a desire for simplification of the language, and for alternative formats, examples, visuals and simplified information. Figure 5 provides a snapshot of the responses of participating organizations to a key question on the clarity of text and information for their communities, and how they rate the quality of the fact sheet in terms of clarity. Targeting a Flesch-Kincaid level of 13 (approximately school-leaver or undergraduate level of readership), the chart indicates success in the goal for clarity in each of the fact sheets, with very high broad agreement that the clarity of the text and information in the fact sheets is of very high or high quality. Nevertheless, a proportion of participating organizations found the clarity to be of less than high quality, with 33% for the Online abuse and harassment fact sheet, and 45% for the Reporting fact sheet.

Figure 5 Participating organizations' rating of clarity of text and information per fact sheet

Figure 5
How do you rate the clarity of the text and information in the fact sheet?

	Very high quality	High quality	Neither high nor low quality	Low quality	Very low quality
Online abuse and harassment	11.11%	55.56%	22.22%	5.56%	5.56%
Disinformation	8.33%	75.00%	8.33%	8.33%	0.00%
Online and mobile scams	10.00%	70.00%	0.00%	20.00%	0.00%
Doxxing	9.09%	63.64%	9.09%	18.18%	0.00%
Deepfakes	0.00%	66.67%	11.11%	22.22%	0.00%
Reporting	18.18%	36.36%	27.27%	18.18%	0.00%

Online abuse and harassment   Disinformation   Online and mobile scams   Doxxing   Deepfakes   Reporting

Organizations which marked lower than high quality for these two fact sheets were uniformly those organizations which provided support or advocated for people living with a disability. This is a finding of considerable significance, not only for the study but for the broader field of support in the context of harmful digital content and behavior. Hatred, discrimination or prejudice directed towards people living with a disability is typically included under hate speech definitions (Chau & Xu 2006; Kilvington 2021), albeit often in ways which is hidden by claims of banter and rumination about people with a disability as a minority group or constructed 'other' (Burch, 2018). People with a disability are often targeted with harmful content or behavior (Burch, 2018), or are exposed to persistent ableist discourse (Cherney, 2019), or may witness the use of anti-disability language applied to others. For example, X platform's owner Elon Musk regularly used the disparaging term "retard" to describe political opponents resulting in a reported 207% increase in posts using the slur subsequent to Musk's first use of it according to one study(Montclair State University 2025). Studies of persons living with an intellectual disability likewise indicate a risk of online abuse and harassment that is well-recognized by victim-survivors with an intellectual disability albeit in ways which do not always utilize the more sophisticated terminology that may circulate in legal and scholarly discourse (eSafety Commissioner, 2022). In this respect, then, there is a high likelihood of conceptual investment in issues of online harms in relation to people with a disability, and a clear need for guidance that is fully inclusive of the broad range of disabilities that may be represented.

Indeed, in pointing out the need for greater simplification, readability and accessibility, several of the organizations provided concrete suggestions, including: (1) A more highly-readable language, tone and style to accommodate people with an intellectual disability or high literacy support needs; (2) providing the guidelines in a video format for those with high literacy support needs; (3) full accessibility including compatibility with the most recent screen readers and image captioning, i.e., not using columns per the ordinary fact sheet layout; (4) An Auslan [sign language] video version for people for whom this is their principal language of communication; (5) greater explanation with examples of some of the newer complex terms

such as doxxing and deepfakes, and simple-to-read explanations for other digital terminology that may be recognized by many users but not necessarily those with an intellectual disability. Several of the organizations included a note to suggest that they were aware of the resource implications of these suggestions, noting too that the onus should not be on a university research project but on both government agencies and digital platforms to provide better guidance in forms that did not exclude this important client area and community.

What these suggestions point to, then, is not the need solely to provide multiple, different versions that enable readability and accessibility by different parts of the community, but that the needs of some of the most vulnerable people in the community ought to be put front-and-centre in the initiation of planning and design of informational, self-care, mutual-care and remedy guidelines in the case of digital harms.

[B] language inclusivity

Several of the migrant advocacy and CALD support organizations pointed to a need we raised deliberately in the questions on improvement: if the fact sheets need to be made available in other languages, and which languages should be prioritized. Naturally, if these were guidance provided by government agencies or transnational platform organizations a diverse range of languages is not only an ethical and inclusive practice, but is often required under statutory law. For independent providers, however, this is not always a requirement but is nevertheless good practice if it can be resourced.

Preferred additional languages raised in the feedback were not indicative, however, of key needs but reflected the organizations who took up the invitation to participate in the study. Tamil, Sanskrit, Vietnamese, Simplified Chinese, Thai and Korean were the languages listed by participant organizations, which nevertheless reflect the broad population and dominant migrant population language groups in Australia.

The language needs, however, were not limited only to the question of translation into a range of other languages, but into how the information was packaged and how much information was contained in each fact sheet. As one organization pertinently outlined:

> We work with recent migrants who are often on Humanitarian Visas.
> Most are not literate in written or spoken English and would require
> someone to translate the information. I feel the amount of information
> on this factsheet would be overwhelming and has some jargon in it that
> may not be familiar even if translated.

Here, an attentiveness to word choices to enable *more accurate* translation, understanding and uptake is important. While the fact sheets were designed to explain jargonistic terms related to digital harms, technological affordances and platform policy and practices, there is a clear need for guidance material to recognize that even definitions and explanations of key terminology may themselves contain jargonistic material that will make sound translation insecure (Gunew, 2000).

An additional element in terms of language inclusivity related not only to translation but to the tone used in the fact sheets. One migrant/CALD advocacy organization raised the positive point of how the fact sheets might be better able to encourage their community members to report or take action when faced with digital harms. It is becoming better recognized in scholarship that a fear "secondary victimization" is a key barrier to reporting or remedying a harm, including harms of injurious language and hate speech (Asquith 2012); this is particularly marked among ethnic and linguistic minorities who fear offense or abuse in response to a failure to follow a reporting procedure adequately or in the preferred terminology. The fear of secondary victimization operates alongside a minority perspective of shame at being victimized, i.e., that a marginalized victim is partly responsible for being victimized in the first

instance (Button et al., 2014). In line with concerns over unwillingness to seek help due to secondary victimization fears and a sense of shame, the organization requested in their feedback the following: "Please address shame as barrier to getting help." We interpret this not only as requiring the guidance documents, particularly the Reporting online abuse and harassment fact sheet, not only outline clearly that shame is both natural and unproductive when it burdens individuals (Probyn 2005), but that the entire language, tone and framing of the issues across all fact sheets needs to be one attentive to the need to avoid any risk that some readers—or, indeed, translators—may read from a shaming perspective.

[C] Seeing themselves explicitly represented

Finally, several of the improvement suggestions offered by participating organizations involved providing examples, focus or risks specific to one of the communities. This was particularly pertinent among those organizations representing older people. Organizations representing, supporting or advocating for older persons in the community noted the usefulness of guidelines, and were particularly interested in their adaptation or refinement to "address the needs of older people" (in the case of Online abuse and harassment), to provide information and details of "aged care scams" (in the case of the Scams fact sheet), and to incorporate information on how deepfakes are being "used to ridicule older people's faces, bodies and minds" (in the case of the Deepfakes fact sheet).

Important here is that while the fact sheets address minorities at their most broad, there was a desire among older groups—arguably perceived as a minority or marginalized community in contemporary western sociality—to see themselves represented. Historically, a number of other marginalized groups and communities have been understood to benefit from representation in communication artefacts and media, particularly LGBTQ+ communities that had traditionally been invisibilized in film and television (Gross, 1998). It has been noted that an ability to see one's own minority community represented is significant for personal and social wellbeing, and that it is potentially harmful to the stability of identity when minority groups are invisibilized except in case when they become the subject of humor, prejudice, stereotyping or exoticized spectatorship (Gross, 2001); the wellbeing implications are beginning to be recognized in relation to the invisibilization of older people (Mordini & De Hert, 2010). Although scholarship on this topic has typically pertained to journalism and entertainment media (Horvat, 2021), there is no particular cultural logic that should disavow the significance of informational guidelines from being a setting in which the dignity of identity is experienced through explicit reference and representation.

One explanation for the particular attentiveness from organizations representing the interests of older persons in the scam fact sheet likely responds to the growing public knowledge that older people may be subject to a higher rate of digital scam targeting, and at higher risk of being defrauded by scams. Those aged sixty-five years and over are regarded as a particularly vulnerable group when it comes to scam susceptibility (Kollmorgen, 2023), and it has been noted in one Australian report that "people over 65 lost more money than other age groups, with $121 million reported lost" (ACCC 2024, p. 16). The report also notes that "older Australians were the only age group that did not experience a decrease in reported losses" from the previous year (16). Vulnerability is also known to increase as people continue to age, with the challenges older adults face in dealing with scam communications can become more pronounced as they enter later life stages (those aged 75-94 and 95+). These challenges are attributable to a range of factors, including reduced levels of digital literacy (James et al., 2014), social vulnerability including loneliness and isolation (Alves & Wilson 2008; Cross, 2016), and changes to cognitive function. In the context of this knowledge, the organizations calling for greater representation and visibility of older people in the scam fact sheet rightly point not only to the dignity of *visibility* in communication for minorities, but to the strategic communication benefit

of ensuring *reference* in order to highlight to the target group the significance of the *risk* (Whitaker et al. 2004)—that is, a more effective and efficient guideline.

Conclusion

In the context of continued growth of incivil, hateful and harmful content and behavior in circulation online, there is a need to continue to utilize the full range of regulatory, pedagogical and justice mechanisms to intervene in, remedy, reduce and prevent digital harms (Kaye 2019). However, given the setbacks in regulatory measures to encourage both platforms and perpetrators to curtail incivility, disinformational, fraudulent or other harmful content and behavior, and the ongoing lack of investment in educational campaigns to encourage all users to recognize that problematic and unethical content may harm others, there is a simultaneous need to provide better guidance to all users on (i) how to recognize and understand digital harms, (ii) how to protect themselves, (iii) how to seek remedy through reporting to statutory bodies and platforms, and (iv) how to engage in self-care and mutual-care practices to survive.

James Porter (2022, p. xx) rightly notes that "Digital aggression, harassment, physical threats, mainly directed at women, at LGBTQ+ persons, at members of racial or ethnic minorities …is, I believe, one of the most serious communication issues currently facing us." That is, while the concerns about general, mainstream and majority population users remain important, our study suggests that these concerns will not be addressed without placing the needs of minority groups—as the groups and identities most targeted by harmful content from incivility to hate speech (Andrews, 2019; Johnson et al., 2019)—at the centre of guidance, information campaigns and other preventative, interventional and care practices. Positioning minority communities as an 'after-thought' in the design and provision of public guidance has been a long-standing problem that is known to increase risk to vulnerable groups rather than capture the entirety of those groups in generic amenities or knowledge frameworks (Wilson & Gutierrez, 1985). When minorities are subject to digital harms, the harm is not necessarily equitable with the kind of abuse, disinformation, scam-targeting, doxxing or misleading imagery that is classified as inter-personal attacks, such as insults, offensiveness or fraud. Rather, it emerges from and carries the force of a longer history of exclusion, prejudice or discrimination, and this is particularly the case for communities of gender and sexuality diversity, CALD and migrant populations and people living with disabilities (Porter, 2021).

Overall, the quantitative and qualitative responses of the participating organizations point to the significance, value and usefulness of guidance material on digital harms for minority groups, but see that usefulness as dependent on refinement that is attentive to the needs of individual communities, noting that there are distinct needs differences highlighted in the feedback between communities of older people, gender- and sexuality-diverse communities, CALD and migrant communities and communities of people living with a disability, as well as intersectionalities and broad diversity within them as raised in some of the responses. It suggests a need for co-design approaches—especially among official providers and government agencies—and a broad team-based approach to a design beyond mere expertise in digital harms or policy but one that can respond to the specific reading and knowledge practices of greatly differentiated minority communities.

What this study has demonstrated is not what needs to be done to refine these particular draft guidelines but, through elicitation, revealed key knowledge frameworks for the provision of guidance in targeted, inclusive ways during a period of increased and damaging digital harms. The study notes that, in addition to a genuine national and international or interjurisdictional harm prevention framework, education, policy and technical intervention practices must be highly attentive to the needs of vulnerable and marginalized communities as a *starting point.* *G*uidance on (i) knowledge, (ii) remedy opportunities and (iii) self-care and

mutual-care practices must (a) be accessible to all users, particularly those with a physical or intellectual disability, (b) use language that is inclusive, and (c) be designed for specific minority communities rather than 'minorities' in general. Co-design of guidance in ways which acknowledge the centrality of community support rather than individual self-responsibility for the management of digital harms is essential (Blaya, 2019). Designing guidance for the very diverse needs of diverse communities does, of course, open a number of key resource issues for independent providers, government agencies and, in some cases, platforms. However, by ensuring inclusivity and representation at the point of guidance design of minority groups— while not necessarily targeting minorities through a broad language of marginalization—is clearly indicated.

Given the study relied on a selective sample of willing participants from Australian minority organizations, further research is necessary to understand the context of minority community guidance needs in the interjurisdictional space in which contemporary digital platforms operate, attentive to the differential regulatory practices in North America, the United Kingdom and Europe. Nevertheless, the study has noted that the underlying need for placing minority needs at the heart of regulatory, prevention and intervention practices in relation to digital harms, and we underscore here the fact that minorities are subject to digital harms in greater proportions exacerbates the urgency of this need.

References

ACCC. (2022, July 6). Scam losses to culturally diverse communities, people with disability and Indigenous Australians almost doubled in 2021. *Australian Competition and Consumer Commission.* https://www.accc.gov.au/media-release/scam-losses-to-culturally-diverse-communities-people-with-disability-and-indigenous-australians-almost-doubled-in-2021#:~:text=%E2%80%9CUnfortunately%2C%20we%20saw%20some%20vulnerable,increase%20in%20losses%20since%20202.

ACCC. (2024, April). *Targeting Scams: Report of the National Anti-Scam Centre on Scams Activity 2023.* https://www.Scamwatch.gov.au/system/files/Targeting-scams-report-2023_0.pdf.

Ahmed, S. (2021). *Complaint!* Duke University Press.

Aldredge, J. (2020, June 9). Is deepfake technology the future of the film industry? *Premium Beat.* https://www.premiumbeat.com/blog/deepfake-technology-future-of-film-industry/.

Altman, D. (1979). *Coming Out in the Seventies*. Wild & Woolley.

Alves, L.M., & Wilson, S.R. (2008). The effects of loneliness on telemarketing fraud vulnerability among older adults. *Journal of Elder Abuse and Neglect, 20*(1), 63-85. DOI: https://doi.org/10.1300/j084v20n01_04

Andrews, K. (2019). Stained Glass Windows and Rainbows: My Journey to Socially Just Perceptions of and Actions toward LGBTQ Students. *American Journal of Qualitative Research, 3*(2), 87-96. https://doi.org/10.29333/ajqr/6436

Asquith, N.L. (2012). Vulnerability and the art of complaint making. In: I. Bartkowiak-Theron & N.L. Asquith (Eds.), *Policing Vulnerability* (pp. 147-161). Federation Press.

Australian Government. (2023, March). *Australian Government Response to the House of Representatives Select Committee on Social Media and Online Safety Report.* https://www.infrastructure.gov.au/sites/default/files/documents/australian-gov-response-to-house-of-reps-select-committee-on-social-media-and-online-safety-report-march2023.pdf.

Blaya, C. (2019). Cyberhate: A review and content analysis of intervention strategies. *Aggression and Violent Behavior, 45*(1), 163-172. DOI: https://doi.org/10.1016/j.avb.2018.05.006

Burch, L. (2018). 'You are a parasite on the productive classes': Online disablist hate speech in austere times. *Disability & Society, 33*(3), 392-415. DOI: https://doi.org/10.1080/09687599.2017.1411250

Butler, J. (2024). *Who's Afraid of Gender?* Allen Lane.

Button, M., Lewis, C., & Tapley, J. (2014). Not a victimless crime: The impact of fraud on individual victims and their families. *Security Journal, 27*(1), 36-54. DOI: https://doi.org/10.1057/sj.2012.11

Chau, M., & Xu, J. (2006). Mining communities and their relationship in blogs: A study of online hate groups. *International Journal of Human-Computer Studies, 65*(1), 57-70. DOI: https://doi.org/10.1016/j.ijhcs.2006.08.009

Cherney, J.L. (2019). *Ableist Rhetoric : How We Know, Value, and See Disability.* Pennsylvania State University Press.

Christchurch Call. (2019). *Christchurch Call to Eliminate Violent and Terrorist Content.* https://www.christchurchcall.com.

Costello, M., Rukus, J., & Hawdon, J. (2019). We don't like your type around here: Regional and residential differences in exposure to online hate material targeting sexuality. *Deviant Behavior, 40*(3), 385-401. DOI: https://doi.org/10.1080/01639625.2018.1426266

Cover R (2022a) Digital hostility: Contemporary crisis, disrupted belonging and self-care practices. *Media International Australia, 184*(1), 79-91. DOI: https://doi.org/10.1177/1329878X221088048.

Cover, R. (2022b). Deepfake culture: The emergence of audio-video deception as an object of social anxiety and regulation. *Continuum: Journal of Media & Cultural Studies, 36*(4), 609-621. DOI: https://doi.org/10.1080/10304312.2022.2084039.

Cover R (2024) Apprehending digital hostility and online abuse: Feminist care ethics in/and digital ecologies. *Thesis 11, 183*(1): 33-48. DOI: https://doi.org/10.1177/07255136241284658.

Cross, C. (2016). 'They're very lonely': Understanding the fraud victimisation of seniors. *International Journal for Crime, Justice and Social Democracy, 5*(4). DOI: https://doi.org/10.5204/ijcjsd.v5i4.268.

Dogan, S.L., & Karaosmanoglu, K. (2025). Maneuvers between Compliance and Transformation: University Students Navigating Minority Perceptions in Turkey. *Journal of Ethnic and Cultural Studies, 12*(4). DOI: https://doi.org/10.29333/ejecs/2182.

Eggensperger, J., & Salvatore, J. (2022). *Strategic Public Relations Writing : Proven Tactics and Techniques*. Routledge.

eSafety Commissioner. (2022). *How bad should it be before I tell someone?' Online abuse experiences of adult Australians with intellectual disability – implications for resource development.* https://www.esafety.gov.au/research/how-adults-intellectual-disability-experience-online-abuse.

eSafety Commissioner. (2025). *Key Topics*. https://www.esafety.gov.au/key-topics.

Factor, R., Kawachi, I., & Williams, D.R. (2011). Understanding high-risk behavior among non-dominant minorities: A social resistance framework. *Social Science & Medicine, 73*(9), 1292-1301. https://doi.org/10.1016/j.socscimed.2011.07.027.

Flew, T. (2021). *Regulating Platforms*. Polity.

Frenkel, S., & Conger, K. (2023, December 2). Hate speech's rise on Twitter is unprecedented, researchers find. *New York Times*. https://www.nytimes.com/2022/12/02/technology/twitter-hate-speech.html.

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. Yale University Press.

GLAAD. (2025, January 7). *GLAAD response: Meta and Mark Zuckerberg remove long-standing anti-LGBTQ hate speech policies after announcing end of fact-checking program.* https://glaad.org/releases/meta-removes-anti-lgbtq-hate-speech-policies-ends-fact-checking/.

Graaff, K. (2021). The Implications of a Narrow Understanding of Gender-Based Violence. *Feminist Encounters: A Journal of Critical Studies in Culture and Politics, 5*(1), Article 12. https://doi.org/10.20897/femenc/9749

Gross, L. (1998). Minorities, majorities and the media. In T. Liebes & J. Curran (Eds.), *Media, Ritual and Identity* (pp. 87-102). Routledge.

Gross, L. (2001). *Up from Invisibility: Lesbians, Gay Men, and the Media in America.* Columbia University Press.

Gunew, S. (2000). Introduction: Multicultural translations of food, bodies, language. *Journal of Intercultural Studies, 21*(3), 227-237. DOI: https://doi.org/10.1080/713678979.

Hancock, A.M. (2011). *Solidarity Politics for Millennials*. Palgrave Macmillan.

Horvat, A. (2021). *Screening Queer Memory: LGBTQ Pasts in Contemporary Film and Television*. Bloomsbury.

Isaac, M., & Schleifer, T. (2025, January 7). Meta to end fact-checking program in shift ahead of Trump term. *New York Times.* https://www.nytimes.com/2025/01/07/technology/meta-fact-checking-facebook.html

James, B.D., Boyle, P.A., & Bennett, D.A. (2014). Correlates of susceptibility to scams in older adults without dementia. *Journal of Elder Abuse and Neglect, 26*(2), 107-122. DOI: https://doi.org/10.1080/08946566.2013.821809.

Jiang, J.A., Scheuerman, M.K., Fiesler, C. and Brubaker, J.R. (2021). Understanding international perceptions of the severity of harmful content online. *Plos One, 16*(8), e2025762. DOI: https://doi.org/10.1371/journal.pone.0256762.

Johnson, N.F., Leahy, R., Johnson Restrepo N., Velasquez, N., Zheng, M., Manrique, P., Devkota, P., & Wuchty, S. (2019). Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature, 573*(1), 261-265. DOI: https://doi.org/10.1038/s41586-019-1494-7.

Kaye, D. (2019). *Speech Police: The Global Struggle to Govern the Internet*. Columbia Global Reports.

Keighley, R. (2022). Hate hurts: Exploring the impact of online hate on LGBTQ+ young people. *Women & Criminal Justice, 32*(1-2), 29-48. DOI: https://doi.org/10.1080/08974454.2021.1988034.

Kilvington, D. (2021). The virtual stages of hate: Using Goffman's work to conceptualise the motivations for online hate. *Media, Culture & Society, 43*(2), 256-272. DOI: https://doi.org/10.1177/0163443720972318.

Kollmorgen, A. (2023, July 25). Older Australians most affected by scams. *Choice.* https://www.choice.com.au/health-and-body/healthy-ageing/ageing-and-retirement/articles/scams-affecting-senior-australians.

Lees, D., Bashford-Rogers, T., & Keppel-Palmer, M. (2021). The digital resurrection of Margaret Thatcher: Creative, technological and legal dilemmas in the use of deepfakes in screen drama. *Convergence, 27*(4), 954–973. https://doi.org/10.1177/13548565211030452.

Lupton, D. (2015). *Digital Sociology*. Routledge.

McRobbie, A. (2020). *Feminism and the Policies of Resilience: Essays on Gender, Media and the End of Welfare*. Polity.

Meikle, G. (2023) *Deepfakes*. Polity.

Montclair State University. (2025, January 10). Use of the slur 'retard' triples on X after Elon Musk shares the word in a post. *News and Announcements*. https://www.montclair.edu/school-of-communication-and-media/2025/01/10/use-of-the-slur-retard-triples-on-x-after-elon-musk-shares-the-word-in-a-post/.

Mordini, E., & De Hert, P. (2010). *Ageing and Invisibility*. IOS Press.

Obermaier, M., Schmuch, D., & Sasleem, M. (2023). I'll be there for you? Effects of Islamophobic online hate speech and counter speech on Muslim in-group bystanders' intention to intervene. *New Media & Society,* 25(9), 2339-2358. DOI: https://doi.org/10.1177/14614448211017527.

O'Connell, S., Ghosh, D., & Reddy, V. (2024). When the Rainbow is Bittersweet: Reflections on Being Queer and Indian in Durban. *Feminist Encounters: A Journal of Critical Studies in Culture and Politics, 8*(1), Article 15. https://doi.org/10.20897/femenc/14226

OfCom. (2023, October 26). *Harmful Online Content: How to Report it and Where to Get Help*. https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/support/.

Poland, B. (2016). *Haters: Harassment, Abuse, and Violence Online*. University of Nebraska Press.

Porter, J.E. (2021). Foreword: Interacting with friends, enemies, and strangers. In: J. Reyman & E.M. Sparby (Eds.), *Digital Ethics: Rhetoric and Responsibility in Online Aggression* (pp. xv-xxii). Routledge.

Poynting, S., Noble, G., Tabar, P., & Collins, J. (2004). *Bin Laden in the Suburbs: Criminalising the Arab Other*. Sydney Institute of Criminology.

Probyn, E. (2005). *Blush: Faces of Shame*. University of Minnesota Press.

Tan, X. (2020). Ethnic Minority Women in Contemporary Chinese Comics: Design, Role and Identity. *Feminist Encounters: A Journal of Critical Studies in Culture and Politics, 4*(1), Article 07. https://doi.org/10.20897/femenc/7911

Taylor, J. (2024, February 9) Publication of Jewish creatives WhatsApp group led to death threats, MP says. *Guardian*. https://www.theguardian.com/australia-news/2024/feb/09/josh-burns-jewish-whatsapp-group-channel-publication-israel-palestine-clementine-ford.

Terranova, T. (2022). *After the Internet: Digital Networks between Capital and the Common*. Semiotext(e)

The Care Collective. (2020). *The Care Manifesto: The Politics of Interdependence*. Verso.

Thomas, C., Mihaela, C., & Rosenberg, N. (2022). *The Hate Speech Crisis: Ways to Start Fixing It: A Toolkit for Civil Society Organizations and Activists*. Minority Rights Group International (MRG). https://minorityrights.org/resources/the-hate-speech-crisis-ways-to-start-fixing-it-a-toolkit-for-civil-society-organizations-and-activists/.

UN Special Rapporteur on Minority Issues. (2022). *Draft 'Effective Guidelines on Hate Speech, Social Media and Minorities.'* https://www.ohchr.org/sites/default/files/2022-06/Draft-Effective-Guidelines-Hate-Speech-SR-Minorities.pdf.

Vincent, N.A. (2017). Victims of cybercrime: Definitions and challenges. In E. Martellozzo & E.A. Jane (Eds.), *Cybercrime and its Victims* (pp 27-42). Routledge.

Wachs, S., & Wright, M.F. (2018) Associations between bystanders and perpetrators of online hate: The moderating role of toxic online disinhibition. *International Journal of Environmental Research and Public Health, 15*(9), 20-30. DOI: https://doi.org/10.3390/ijerph15092030.

Walther, J.B. (2022). Social media and online hate. *Current Opinion in Psychology, 45*(1), 101298. DOI: https://doi.org/10.1016/j.copsyc.2021.12.010.

Waqas, A., Salminen, J., Jung, S.G., Almerekhi, H., & Jansen, B.J. (2019). Mapping online hate: A scientometric analysis on research trends and hotspots in research on online hate. *Plos One, 14*(9), e022194. DOI: https://doi.org/10.1371/journal.pone.0222194.

Whitaker, R.W., Ramsey, J.E., & Smith, R.D. (2004). *MediaWriting: Print, Broadcast and Public Relations*. Lawrence Erlbaum.

Williamson, J.M.L., & Martin, A.G. (2010). Analysis of patient information leaflets provided by a district general hospital by the Flesch and Flesch-Kincaid method. *International Journal of Clinical Practice, 64*(13), 1824-1831. DOI: https://doi.org/10.1111/j.1742-1241.2010.02408.x.

Wilson, C.C., & Gutierrez, F. (1985). *Minorities and Media: Diversity and the End of Mass Communication*. Sage.

Wilson, R.A. (2019). The digital ethnography of law: Studying online hate speech online and offline. *Journal of Legal Anthropology, 3*(1), 1-20. DOI: https://dx.doi.org/10.2139/ssrn.3465225.

Notes on Contributors

Rob Cover is Professor of Digital Communication and Director of the RMIT Digital Ethnography Research Centre. He leads a number of major funded research projects on digital harms, young people and wellbeing, and gender/sexuality diversity in screen contexts. The author of 100+ journal articles and eleven books, he publishes widely on topics related to digital cultures in the context of social identities, young people, suicide prevention and resilience.

Shiyang Zhu is a PhD candidate in the School of Media and Communication at RMIT University. Working at the crossroads of gender studies and political communication, his research explores how Chinese queer communities engage with censorship, visibility, and power through everyday media practices.

ORCID

Rob Cover, https://orcid.org/0000-0002-8815-2126
Shiyang Zhu, https://orcid.org/0009-0000-4165-2599