Responses – Orygen

From Mr Wallace

1. **Do you think that Social Media Companies owe their users, participants, partners, and content creators a duty of care?**

Social media companies should ensure that their products, features, and platforms are designed with safety in mind and are age-appropriate. The Australian Office of the eSafety Commissioner have made suggestions for how to incorporate safety by design principles in the design and implementation of technology so that online harms are adequately anticipated, detected and eliminated. Social media companies should innovate and create in accordance with these principles so that online safety is enhanced for all users.

2. **We talk a lot about effective regulation – we want to get this right, but we also don't want to let perfect be the enemy of the good. The focus then is on measuring the effectiveness of interventions.**

   How do you propose that the government measure the effectiveness of interventions aimed at minimising the risks of social media? Specifically, what metrics or indicators should be used to assess whether the regulation is having the intended impact?

First, current safety tools and features offered by popular platforms are not evidence informed. The first step would be to conduct robust research to identify what safety tools and features are most effective in ensuring online safety, noting that different groups may require different approaches (e.g., younger people versus older people).

Second, a range of research methodologies should be used to assess impact of safety interventions including user self-report, objective measurement of social media usage and interactions, experimental designs assessing new safety features, 'in the moment' research designs that limit recall bias (e.g., ecological momentary assessment studies) and modelling studies that can properly assess the impact of different policy initiatives. Further, longitudinal research that can examine causal explanations for the relationship between social media and youth mental health are critically needed, as are properly resourced trials of safety interventions/approaches.

Finally, improvements to the platforms' current transparency reporting regarding utilisation and impact of safety features, including access to usage data, would improve our knowledge in this area.

3. **Your research indicates that 33% of young people were considered to have problematic social media use. You suggest introducing regulations requiring social media platforms to include features that help young people manage their usage.**

How do you define problematic social media use?

Currently, research tends to associate problematic social media use with time spent online. This is a flawed metric and fails to account for the activities or behaviours that individuals engage in online. There are a number of factors that must be considered when determining the impact of social media: behaviour (e.g., active versus passive), context, subjective experience, and individual differences (e.g., age, gender, development), to name a few. Similar to above, there is a critical need for longitudinal studies that address questions related to the impact of social media use on young people over time and to appropriate identify any causal links between social media and mental health for all users.

What specific design features should be mandated?

As suggested in our joint submission, there are a number of features that would improve the safety and experience of young people on social media. These include co-designing reform and new safety features with young people and other users to ensure that they are likely to be effective and fit-for-purpose; putting an end to 'sticky features' like infinite scroll; and increasing algorithm transparency, giving users control of their algorithms and the content that they see, and minimising exposure to harmful content suggested via recommender systems.

What strategies should the government adopt to identify and support those at risk of developing mental health issues related to their social media use?

As suggested in our joint submission, the youth mental health sector is in strong favour of establishing a system of verification of mental health information on platforms to help users distinguish good advice from misinformation. Improvements to the quality and availability of reliable mental health information for young people on social media would go some way towards supporting those seeking help for their mental health. It would also limit inappropriate or unsafe mental health-related information from circulating.

Further, young people have told us that they want online safety embedded within their school curriculum so that they are taught skills to engage safely. Parents and families have equally called for more support and education. Education and increasing digital literacy will be key to promoting safer online experiences among all social media users.

Do you believe age verification should be a part of these regulations to prevent younger users from accessing harmful content, or from developing an addiction at a young age?

Age verification and the suggestion to ban certain age groups from accessing social media is a one-size fits all approach that lacks evidence and could do more harm than good. Our work continues to show that young people turn to social media for

support with their mental health, to connect with their peers, and to find a sense of belonging. This is reported to a greater extent by those who are unable to do this offline or in face-to-face settings. Developing evidence-based ways to maximise these benefits, whilst minimising risks, will do more to support young people long term.

4. **We have heard from a number of witnesses about the impact of cyberbullying on children – and the way social media platforms have been used to facilitate cyberbullying. You recommended stronger laws to address this issue.**

   What specific legal measures do you believe would be effective and enforceable?

Orygen does not have expertise in effective or enforceable legal measures on this issue.

5. **We know that social media companies are using users' data to target advertising for the advertising customers and clients. Naturally, children under 18 years of age are then in position to be commercially exploited by big tech to boost the bottom line.**

   Do you think that there should be an outright ban on targeted advertising to children?

Orygen does not have expertise in advertising or advertising law, and we do not support the commercial exploitation of social media users. Some of our work has identified opportunities for using targeted advertising to reach vulnerable people with support for their mental health. With the call to verify mental health information / organisations in the future (see response to Question 3c), there is an opportunity to use social media to provide real time, age-appropriate and evidence-based supports to communities.

From Ms Daniel

1. **Have you worked with the social media/tech companies to develop tools to direct people to resources/help/support? Such as pop ups directing to help or support? Similar to Butterfly work with Meta to develop pop ups when people search eating disorder content they will get redirected to support. If so, could you provide examples of this and any data measuring success or not.**

At Orygen, we have created the #chatsafe guidelines. The #chatsafe guidelines and associated resources are intended to support young people to communicate safely on social media about self-harm and suicide. These resources have been created independently through research conducted at the University of Melbourne. The #chatsafe guidelines and resources have been shared by Meta (Facebook) and are available to their users on their Safety Centre. Meta have supported the language translation of some of the #chatsafe resources so that the information is available in countries outside Australia and in languages beyond English. Meta have had no input to the research activities or in the planning, data collection, study conduct or data analysis of our research studies.

Further, we have conducted a range of studies to assess the safety, acceptability and feasibility of using social media to reach young people with information about how to communicate safely on social media about self-harm and suicide (Robinson et al., 2018; Robinson et al., 2023; La Sala et al., 2021; La Sala et al., 2023), and in the involvement of young people in this process (Thorn et al., 2020). Findings from these studies suggest that it can be safe to reach young people with information about their mental health entirely via social media, and that this information can increase a young person's ability to communicate safely about the topics of self-harm and suicide, as well as their willingness to seek help for themselves or their peers. Currently, this approach is being used as a postvention activity across Victoria, where we share targeted social media content that promotes help-seeking to LGA's impacted by suicide, identified by the Victorian Department of Health. We have studies currently underway, and about to begin, that test this approach further and will lead to the development of evidence-based resources that directly inform the safety features currently offered by popular platforms. These findings will be published when they are available.

References

1. Robinson, J., Hill, N. T., Thorn, P., Battersby, R., Teh, Z., Reavley, N. J., ... & Skehan, J. (2018). The# chatsafe project. Developing guidelines to help young people communicate safely about suicide on social media: A Delphi study. PLoS One, 13(11), e0206584.
2. Robinson, J., Thorn, P., McKay, S., Hemming, L., Battersby-Coulter, R., Cooper, C., ... & La Sala, L. (2023). # chatsafe 2.0. updated guidelines to support young people to communicate safely online about self-harm and suicide: A Delphi expert consensus study. PLoS one, 18(8), e0289494.
3. La Sala, L., Teh, Z., Lamblin, M., Rajaram, G., Rice, S., Hill, N. T., ... & Robinson, J. (2021). Can a social media intervention improve online communication about suicide? A feasibility study examining the acceptability and potential impact of the# chatsafe campaign. PLoS One, 16(6), e0253278.
4. La Sala, L., Pirkis, J., Cooper, C., Hill, N. T., Lamblin, M., Rajaram, G., ... & Robinson, J. (2023). Acceptability and potential impact of the# chatsafe suicide postvention response among young people who have been exposed to suicide: pilot study. JMIR human factors, 10, e44535.
5. Thorn, P., Hill, N. T., Lamblin, M., Teh, Z., Battersby-Coulter, R., Rice, S., ... & Robinson, J. (2020). Developing a suicide prevention social media campaign with young people (The# Chatsafe project): co-design approach. JMIR mental health, 7(5), e17520.