Children's exposure to hate messages and violent images online

Daniel Kardefelt Winther[a], Mariya Stoilova[b], Moritz Büchi[a], Rogers Twesigye[a], David Smahel[c], Marie Bedrosová[c], Nikol Kvardová[c], Sonia Livingstone[b]

[a] UNICEF Innocenti – Global Office of Research and Foresight, Italy
[b] The London School of Economics and Political Science, United Kingdom
[c] Masaryk University, Czechia

Background

The digital environment is largely unregulated and might expose children to age-inappropriate or potentially harmful content. There has been particular concern about children's exposure to hate messages and violent images online and how this affects their well-being and rights.[1] Efforts to mitigate exposure to such content are gaining momentum but more evidence is needed to guide policy, regulation and industry practice.

Aims

This brief provides a cross-national comparison of children's exposure to hate messages and violent images online, and analyses it in relation to internet access at the country level. The purpose is to understand what children in different countries are experiencing and if exposure to this content is driven by the level of internet access in the population. The results act as a starting point for exploring which countermeasures may be more effective in terms of policy and legislation, and to encourage industry to develop proactive solutions to protect children in online spaces.

Methods

Data were analysed from 31,790 children aged 12–16 years from 36 countries who took part in either of the surveys Disrupting Harm, EU Kids Online or Global Kids Online between 2016 and 2021. Internet-using children were asked about seeing hate messages or violent images online in the past year. These data were mapped onto public data on the proportion of internet users by country, based on the International Telecommunication Union's ICT Indicators database.

[1] Online hate messages refer to any form of online content that is hateful in nature, for example, targeted at certain groups or individuals based on religion, nationality or sexuality. It does not only refer to hate messages received by children.

Key findings

A considerable proportion of children around the world were exposed to hate messages or violent images online in the previous year. Once approximately half the population has access to the internet, there is a steep rise in children's exposure to hate messages and violent images online, making this an urgent issue for countries with expanding internet access and usage. Differences among countries are substantial and could offer insights into potential good practices to aid countries currently expanding internet access and usage.

Older children are more likely to report exposure to hate messages and violent images online across nearly all countries. This could be because older children use the internet more often and in more varied ways, increasing the likelihood of exposure. Gender differences are small. In line with other research on online risks, exposure to hate messages is associated with exposure to violent images in almost all countries. Children who experience one of these risks are more likely to experience the other, forming low-risk and high-risk country contexts.

Greater internet access is associated with higher risk of encountering hate messages and violent images online. In high-connectivity settings, policies and online safety regulation are a must to ensure children are protected online.

Risks can be substantial even in countries where internet access is low, suggesting the importance of investing in child online protection efforts at an early stage of a country's digital transformation. Countries where more than half of the population have internet access see a steep rise in children's exposure to these risks, reinforcing the need to develop strong child online protection systems, policies and regulation as connectivity increases.

Some countries where internet access is high have managed to maintain low exposure to hate messages and violent images online. Further investigation into the policies and practices of these countries, or the most popular platforms used in these countries, might reveal protective policy or legislative solutions that could be replicated elsewhere. These could include existing laws governing hateful and violent content online, or content moderation practices that often differ by platform or by language.

Children's exposure to hate messages and violent images online

The digital environment is an integral part of children's lives in many parts of the world, and increasingly so in contexts where internet access has traditionally been limited or restricted. During the COVID-19 pandemic, the digital environment was a key source of learning, socializing and entertainment for many children, at a time when they were unable to meet others in person (Mascheroni et al., 2021).

While the internet can stimulate social connection, participation and creativity, it also facilitates the spread of risky content, such as hate messages and images of violence, which can negatively affect children's identity, sense of self-worth and their view of the world. This reinforces the urgency of developing more efficient protection measures in a world where online hate and violence is becoming a global problem (United Nations Children's Fund, 2017).

Research in Europe has shown that hate messages are the most common form of potentially harmful content that children are exposed to online, followed by violent images (Smahel et al., 2020). Similarly, European children have also reported that they find violent content among the most harmful to witness (Livingstone et al., 2013). Children can be exposed to online hate and violent content without necessarily being its target or recipient, but the experience can still negatively influence their well-being (Keipi et al., 2017). For example, with respect to hate messages, such content might create a general culture of hostility and intolerance which will affect children negatively in the longer term, irrespective of whether they are the creators or recipients (Harriman et al., 2020).

There is evidence showing that exposure to different types of harmful content is mutually correlated. In other words, if a child sees one type of risky content, they are more likely to see other types of risky content (Smahel et al., 2020), resulting in accumulating risks for some children.

Most of the evidence on children's exposure to hate messages and violent content online comes from countries in the global North, primarily European countries and the United States of America. The spread and magnitude of the problem beyond the global North remains unclear, preventing a more global analysis of this issue and potentially stalling protective efforts in countries at the early stage of digital transformation.

To assess the extent to which this is a global issue and how it presents in different countries and regions of the world, this research brief provides an overview of children's exposure to hate messages and violent images online by using comparable survey data from children living in 36 countries. These data are also analysed in relation to individuals' internet access at the country level, to determine whether there is a relationship between internet access and children's exposure to hate messages and violent images online.

Methodology

Data

The analysis uses survey data from 31,790 children aged 12–16 years living in 36 countries, generated through three comparable child-centred surveys on children's online experiences and activities. It also draws on secondary data on individuals' internet access from the International Telecommunication Union's ICT Indicators database (International Telecommunication Union, 2022).

Disrupting Harm: 7,486 children aged 12–16 in Cambodia, Ethiopia, Indonesia, Kenya, Namibia, the Philippines, Thailand, Uganda, the United Republic of Tanzania and Viet Nam. Data collected through nationally representative household surveys with internet-using children in 2021.

EU Kids Online: 14,529 children aged 12–16 in Belgium (Flanders), Croatia, Czechia, Estonia, Finland, France, Germany, Italy, Lithuania, Malta, Norway, Poland, Portugal, Romania, the Russian Federation, Serbia, Slovakia, Spain and Switzerland. Data collected through a combination of nationally representative household surveys and school surveys with internet-using children between 2017–2019.[2]

Global Kids Online: 9,775 children aged 12–16 in Bulgaria, Chile, Ghana, South Africa and New Zealand, and children aged 13–16 in Argentina and Uruguay. Data collected through nationally representative household surveys with internet-using children between 2016–2019.[3]

International Telecommunication Union: Public data on individuals aged 5 years and over using the internet (internet users per 100 population). Data collected between 2017–2019.[4]

Measurements

The primary indicators for exposure to hate messages and violent images online is based on the same question used in all three surveys: "In the past year, have you seen online content or online discussions with any of these things: a) Hate messages that attack certain groups or individuals (e.g. people of different religion, nationality or sexuality); b) Gory or violent images".

Variables were recoded to binary yes/no responses if they were originally measured as a frequency (Very often, Often, Sometimes, Rarely recorded as 'yes', Never recoded as 'no').

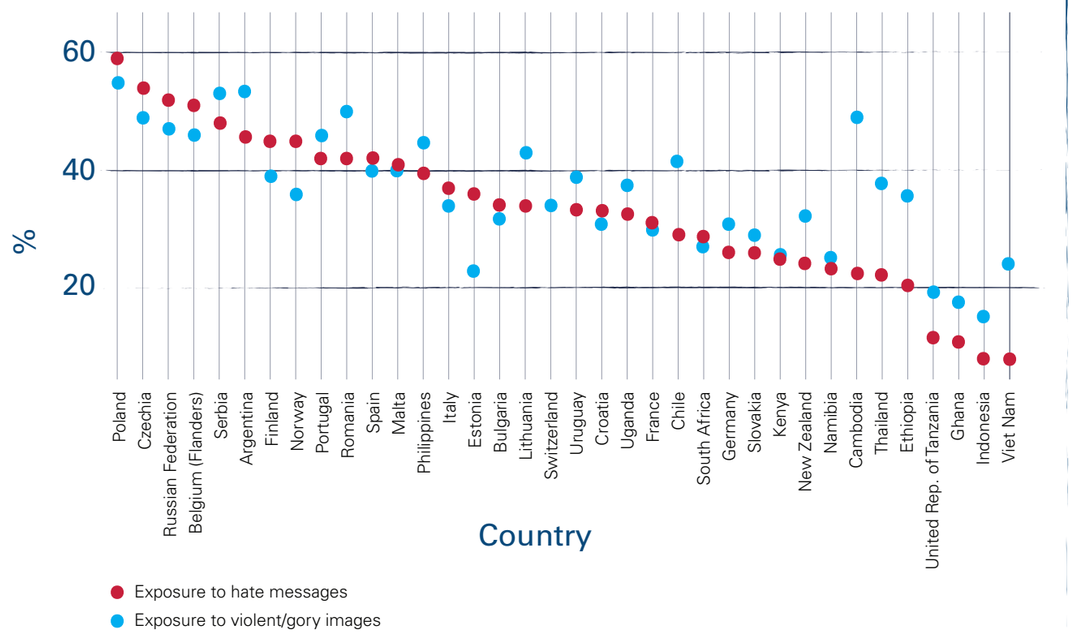
Differences between countries should be interpreted with caution due to slight variations in recruitment methods, sampling, children's ages, missing data (up to 13 per cent in Ethiopia) and different timing for data collection.

[2] For details by country, see Smahel et al., 2020.
[3] Data from South Africa represented children in only three regions.
[4] The International Telecommunication Union notes that there may be countries where data does not exist for internet users as young as 5 years and over, in which case the population estimate represents older age groups only. The estimate provided represents the best possible estimate given available household survey data.

Where are children at higher risk of exposure?

Figure 1 shows the proportion of internet-using children who were exposed to hate messages and violent images online in the previous year, for each of the 36 countries. In most countries, there is an association between children's exposure to hate messages and exposure to violent images, forming low-risk and high-risk country contexts. This has been demonstrated in earlier research in Europe (e.g. Smahel et al., 2020), and these findings show that this is also the case in many non-European countries.

Figure 1: Children's exposure to hate messages and violent/gory images in the previous year, by country



Note: Base: internet-using children aged 12–16 years, 13–16 in Argentina and Uruguay. N=31,790

There is considerable variation among countries in terms of children's exposure. Children's exposure to hate messages ranged from 8 per cent to 58 per cent, being least common in Indonesia and Viet Nam (8 per cent) and most common in Poland (58 per cent). Children's exposure to violent images showed a similar range, between 15 per cent to 55 per cent, again being least common in Indonesia (15 per cent) and most common in Poland (55 per cent).

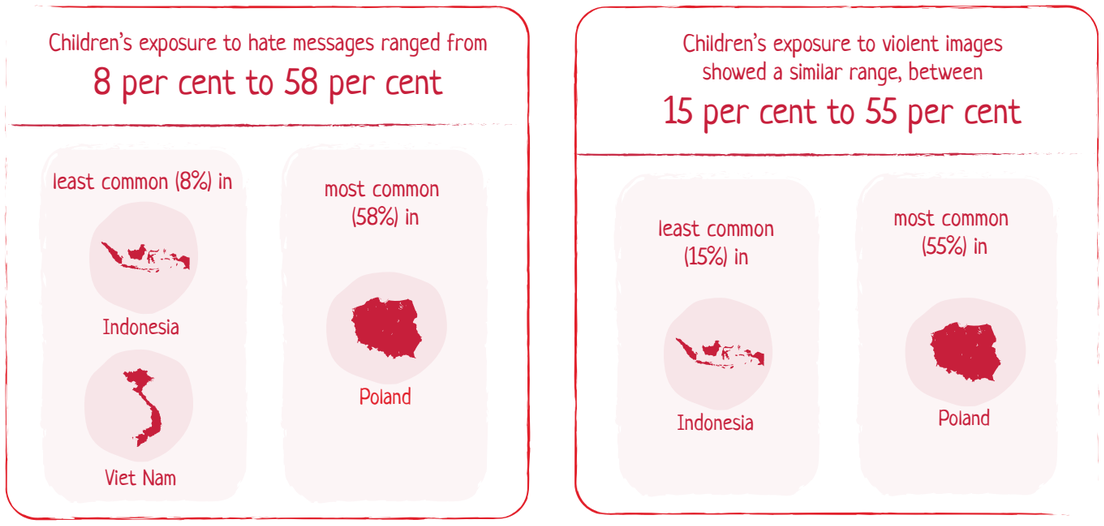
Older children were more likely to report exposure to hate messages and violent images online across nearly all countries, though the differences are not substantial. This is expected as older children engage in a wider range of activities online, which exposes them to both more online opportunities and risks, including exposure to potentially harmful content (Kardefelt Winther, Livingstone and Saeed, 2019).

Gender differences were generally small, and not consistent between countries. This is in line with existing evidence on online risks and harm from Europe and elsewhere which typically display only minor gender differences (e.g. Smahel et al., 2020; Livingstone, Kardefelt Winther and Saeed, 2019).

The analysis also shows that in a few countries, children are much more exposed to one of these two online risks. Countries like Cambodia, Chile, Estonia, Ethiopia, Thailand and Viet Nam have the largest gaps in children's exposure to hate messages and violent images with differences of 10 percentage points or more.

In all these countries except Estonia, it was more common for children to be exposed to violent images than hate messages.

It is important to monitor children's exposure to different types of risks and apply a problem-specific approach when some risks are higher than others. For example, higher exposure to hate messages can indicate negative intergroup attitudes in society which may need to be addressed by educational and awareness-raising efforts that aim to promote tolerance and reduce negative stereotypes and prejudice, among other strategies.



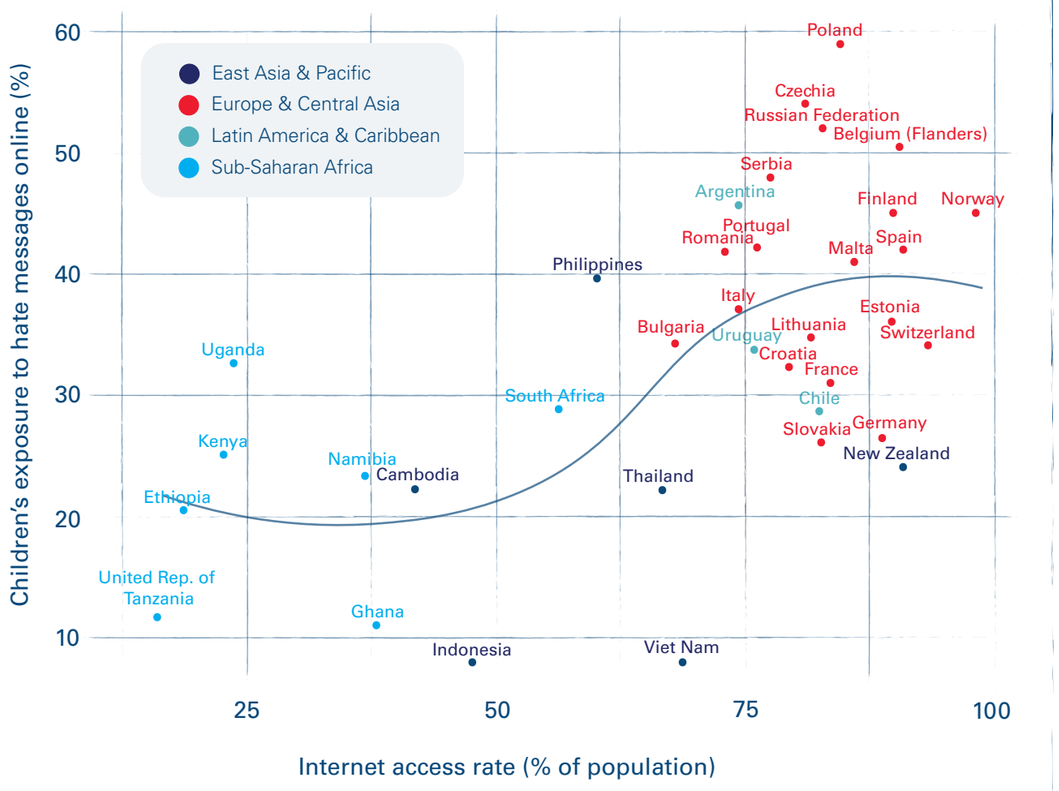
Note: The designations employed in this publication and the presentation of the material do not imply on the part of UNICEF the expression of any opinion whatsoever concerning the legal status of any country or territory, or of its authorities or the delimitations of its frontiers.

Internet access and exposure to hate messages and violent images online

Existing research from Europe suggests that countries where internet use is higher may see children more exposed to some online risks (Lobe et al. 2012). To test this hypothesis in relation to hate messages and violent images, public data from the International Telecommunication Union on individuals aged 5 years and over using the internet (internet users per 100 population) were analysed together with data on children's exposure to hate messages and violent images in each country.

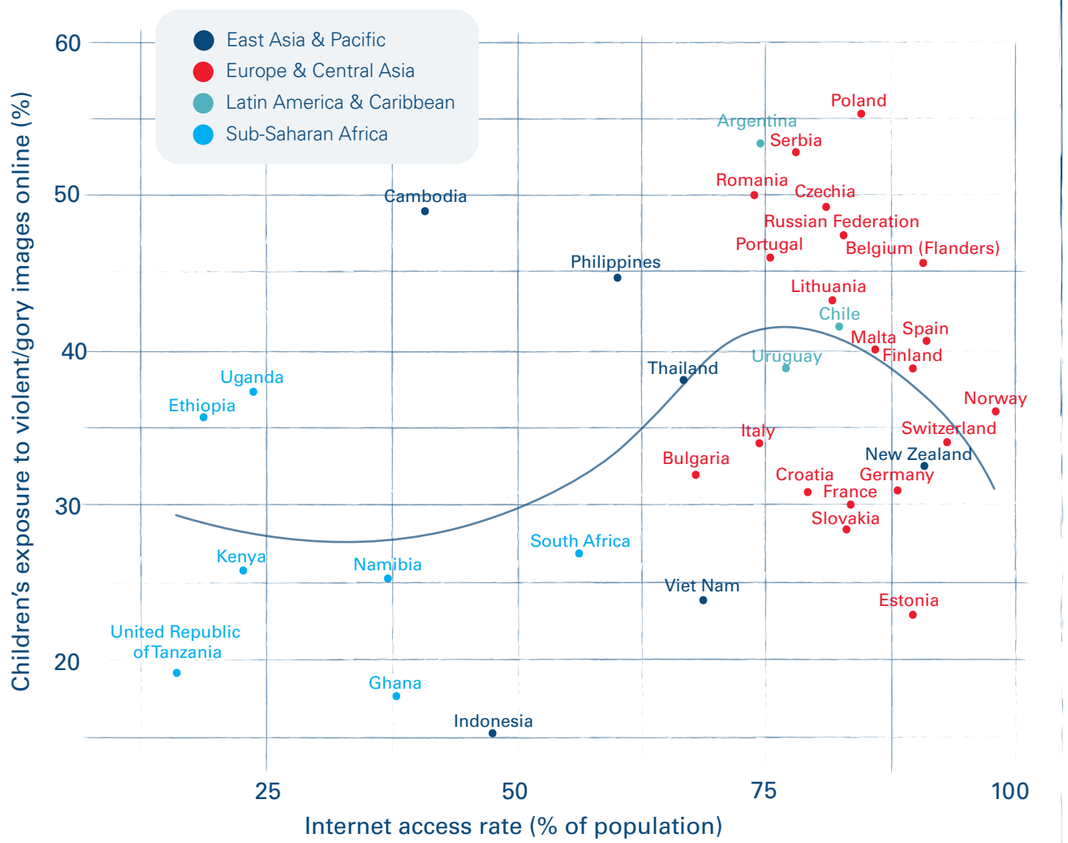
Results show that in general, higher internet access at the country level is associated with higher exposure to hate messages and violent images online for children. However, the effect is not linear. It appears that when approximately half the population has access to the internet, there is a steep rise in children's exposure to hate messages and violent images online (see Figures 2 and 3).

Figure 2: Exposure to hate messages by country's internet diffusion rate (the line shows locally estimated scatterplot smoothing)



Note: Base 1: Internet-using children aged 12–16 years, 13–16 in Argentina and Uruguay. N=31,790
Base 2: Individuals aged 5 years and over using the internet (internet users per 100 population)

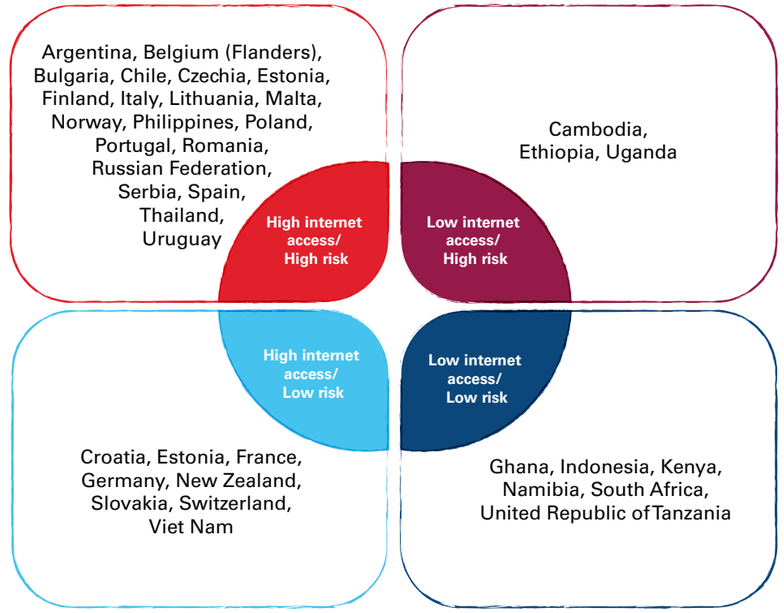
Figure 3: Exposure to violent/gory images by country's internet access rate (the line shows locally estimated scatterplot smoothing)



Note: Base 1: Internet-using children aged 12–16 years, 13–16 in Argentina and Uruguay. N=42,362
Base 2: Individuals aged 5 years and over using the internet (internet users per 100 population)

Observing the two scatterplots, four different clusters of countries emerge (see Figure 4).

Figure 4: Clusters of countries by internet access and children's exposure to harmful content[5]



The least favourable combination among the four clusters is low internet access and high risk, as children in these countries have limited access but are still exposed to relatively higher levels of hate messages and violent images online. As this analysis demonstrates, online risks for children can be substantial even in countries where internet access is low.

It may be that countries with low internet access lack sufficient policies and legislation relating to child online protection, do not yet have mainstreamed digital literacy education and have a population with relatively lower awareness of online risks, which then leads to greater exposure by children to risky content. Parents or caregivers might not use the internet much or at all in these countries, meaning that they lack the skills to support children's internet use and potentially mitigate risk exposure. While relatively few countries are in this category, more research is needed to understand why this happens and how to best prevent it.

In countries where online risks are a relatively new issue and there is no infrastructure to manage them, a good starting point can be to learn from the experience of countries that are slightly ahead in terms of the digital transformation of society, to allow policy, regulation and education to advance alongside further technological developments. Countries at an earlier stage of digital transformation may benefit from introducing policies and education for child online protection, to help mitigate the increased risks of exposure to online harms that may come with greater access.

However, as the high internet access/low risk cluster shows, some countries are managing better than others in mitigating children's exposure to certain online risks, which is an encouraging finding.

[5] Countries were placed in the high internet access cluster if more than 50 per cent of the population had internet access. Countries were considered as high-risk if more than 35 per cent of children had high exposure to at least one of the risks, calculated as being above the median risk percentage (34 per cent) across all countries.

It suggests that internet access in and of itself may not be the main driver of online risk exposure. Research has shown that what children do online and the range of different activities they engage in may be another important driver of online risk exposure (Kardefelt Winther, Livingstone and Saeed, 2019). However, if good policies and practices are not established at the outset, children might be disproportionately exposed to higher levels of risk as internet access and use increases.

Most countries are in a situation of having high internet access and high risk exposure. This demonstrates an online culture where hate messages and violent images are more common, which might be a reflection of existing cultural and social norms, or a history of intergroup conflicts. Efforts need to focus on breaking this pattern and reduce overall risk exposure for children in the digital environment, but also consider the full complexity of children's lives – digital and otherwise – in terms of how these issues are affected by social and cultural dynamics in society, gender equality, religion, exposure to other forms of violence or abuse, hatred, racism and misogyny.

In summary, the analysis shows that the relationship between internet access and online risk exposure is not linear, and that countries can seemingly manage and reduce exposure to hate messages and violent images in different ways. As Figure 4 demonstrates, each region has countries with similar internet access rates that vary in terms of children's exposure to hate messages and violent images. The question is what can be learned from the countries in the high internet access/low risk cluster that manage to maintain a high level of internet access in the population while having relatively lower exposure to harmful content. How can this knowledge be used by those countries in the low internet access/low risk cluster to successfully expand internet access for the general population, without also increasing exposure to potentially harmful content for children?

> Once approximately half the population has access to the internet there is a steep rise in children's exposure to hate messages and violent images online, making this an urgent issue for countries with expanding internet access and usage.

While it is not possible entirely to reduce the risk that children are exposed to potentially harmful content online, countries that have managed to curb exposure can serve as good examples both to countries currently facing greater risk and to those transitioning from low to high connectivity. Further research should establish which factors contribute to these success stories and whether they can be applied in other countries going through a digital transformation.

Roles and responsibilities of industry

Meeting these challenges through social and cultural change will take time, but more rapid changes can be incorporated that directly address the digital platforms where children encounter hate messages and violent images. According to the United Nations Guiding Principles on Business and Human Rights (United Nations, 2011),[6] companies have a responsibility to respect human rights, including children's rights, through policy commitments, due diligence and providing remedy for adverse impacts.

The findings presented here highlight the importance of addressing the issue of children's exposure to hate messages and violent content online within companies' ongoing due diligence processes. There is a need to engage industry stakeholders in different national contexts to report on what data they have of hate messages or violent content proliferating on their platforms and the measures they are taking to prevent it, as recommended in the United Nations Strategy and Plan of Action on Hate Speech (United Nations Secretary-General, 2019). Companies should be requested to conduct child rights impact assessments to determine actual or potential impacts on children and design responses accordingly, as part of an effective and child-friendly corporate strategy (United Nations Committee on the Rights of the Child, 2021).

Industry should use technological expertise to develop procedures and features that identify and mitigate content risks early on and remedy their impact on children. Greater investment in content moderation may be needed for smaller or minority languages. Child-friendly mechanisms for reporting hate messages and violent images online should be developed and supplemented by educational materials to raise children's awareness about these issues and how to report them. There should be clear, transparent and child-friendly terms of service and codes of conduct that discourage the creation and distribution of hateful and violent content, and companies should be encouraged to restrict access to these services for repeat offenders.

Children who have breached company safety policies or the code of conduct should be supported to avoid repeating their mistakes in the future. They should be offered educational resources and guidance on positive and responsible behaviour. In line with the United Nations Convention on the Rights of the Child (United Nations, General Assembly, 1989), tackling hateful and violent content, and supporting children is a critical part of actively promoting children's rights to education, information, participation, civic engagement and citizenship.

[6] Unanimously endorsed by the United Nations Human Rights Council in 2011, the United Nations Guiding Principles are the internationally accepted framework for enhancing standards and practices with regard to business and human rights.

Conclusion

Data presented here show that the spread of online hate and violence is a global issue faced by children across the world, some at a young age. Urgent measures are needed, especially in contexts where the exposure to potentially harmful content is high.

> Countries with growing connectivity would benefit from proactive efforts to mitigate future harm to children.

The substantial variation between the countries in terms of internet access and children's reported exposure to these content risks suggests that there are different possible pathways for countries at an early stage of digital transformation. Countries with growing connectivity would benefit from proactive efforts to mitigate future harm to children. Further research is needed to establish what factors effectively limit children's exposure to harmful content online and how these can be replicated, including effective legislation and content moderation. The evidence presented here calls for early prevention and protective measures while internet access is still low and children's exposure to content risks is still limited.

An immediate step would be for governments to integrate child rights considerations and child rights impact assessments in the process of regulation of online hate and violence, and in technological development. Efforts need to strike a careful balance between all children's rights, including to protection, privacy, access to information, freedom of expression, education and culture, and thus not unduly restrict children's options. This will ensure that children do not lose out on online opportunities because of overly restrictive protection measures. Industry in turn should be supported to understand where to draw this line and how to account for children's rights and best interests in product development and business operations.

> Governments should integrate child rights considerations and child rights impact assessments in the process of regulation of online hate and violence, and in technological development.

References

Harriman, Nigel, et al., 'Youth Exposure to Hate in the Online Space: An exploratory analysis', *International Journal of Environmental Research and Public Health*, vol. 17, no. 22, 17 November 2020, p. 8531, <www.mdpi.com/1660-4601/17/22/8531>.

Helsper, Ellen J., et al., *Country Classification: Opportunities, risks, harm and parental mediation*, LSE Research Online, 27 August 2013, <https://eprints.lse.ac.uk/52023>.

International Telecommunication Union, 'Statistics', <www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>, accessed 13 April 2023.

Kardefelt Winther, Daniel, Sonia Livingstone and Marium Saeed, 'Growing up in a connected world', Innocenti Research Report, UNICEF Office of Research – Innocenti, Florence, 2019, <www.unicef-irc.org/publications/1060-growing-up-in-a-connected-world.html>.

Keipi, Teo, et al., *Online Hate and Harmful Content: Cross-national perspectives*, Routledge, Oxford, United Kingdom and New York, United States, 2017.

Livingstone, Sonia, Daniel Kardefelt Winther and Marium Saeed, 'Global Kids Online: Comparative report', Innocenti Research Report, UNICEF Office of Research – Innocenti, Florence, 2019, <www.unicef-irc.org/publications/pdf/GKO%20LAYOUT%20MAIN%20REPORT.pdf>.

Livingstone, Sonia, et al., *In Their Own Words: What bothers children online? with the EU Kids Online Network*, LSE Research Online, February 2013, <https://eprints.lse.ac.uk/48357>.

Livingstone, Sonia and Mariya Stoilova, *The 4Cs: Classifying online risk to children*, Leibniz-Institut für Medienforschung | Hans-Bredow-Institut (HBI) and CO:RE – Children Online: Research and Evidence, Hamburg, 2021, <www.ssoar.info/ssoar/handle/document/71817>.

Lobe, Bojana, et al., *Cross-National Comparison of Risks and Safety on the Internet: Initial analysis from the EU Kids Online survey of European children*, LSE Research Online, August 2012, <https://eprints.lse.ac.uk/39608/>.

Mascheroni, Giovanna, et al., 'Learning at a Distance: Children's remote learning experiences in Italy during the COVID-19 pandemic', Innocenti Research Report, UNICEF Office of Research – Innocenti, Florence, 2021 <www.unicef-irc.org/publications/pdf/learning-at-a-distance-childrens-remote-learning-experiences-in-italy-during-the-covid-19-pandemic.pdf>.

Resolution adopted by the United Nations General Assembly, 'United Nations Convention on the Rights of the Child', 44/25, 20 November 1989, <www.ohchr.org/sites/default/files/Documents/ProfessionalInterest/crc.pdf>.

Smahel, David, et al., *EU Kids Online 2020: Survey results from 19 countries*, LSE Research Online, 2020, <https://eprints.lse.ac.uk/103294>.

Stoilova, Mariya, Sonia Livingstone and Rana Khazbak, 'Investigating Risks and Opportunities for Children in a Digital World: A rapid review of the evidence on children's internet use and outcomes', Innocenti Discussion Papers, UNICEF Office of Research – Innocenti, Florence, March 2020, <www.unicef-irc.org/publications/1183-investigating-risks-and-opportunities-for-children-in-a-digital-world.html>.

United Nations, United Nations Guiding Principles on Business and Human Rights, 2011, <www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf>.

United Nations Children's Fund, *The State of the World's Children 2017: Children in a digital world*, UNICEF, December 2017, <www.unicef.org/reports/state-worlds-children-2017>.

United Nations Committee on the Rights of the Child, General Comment No. 25 (2021) on Children's Rights in Relation to the Digital Environment, 2 March 2021, <www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>.

United Nations Secretary-General, United Nations Strategy and Plan of Action on Hate Speech, May 2019, <www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf>.

UNICEF works in the world's toughest places to reach the most disadvantaged children and adolescents and to protect the rights of every child, everywhere. Across 190 countries and territories, we do whatever it takes to help children survive, thrive and fulfill their potential, from early childhood through adolescence. And we never give up.

UNICEF Innocenti – Global Office of Research and Foresight tackles the current and emerging questions of greatest importance for children. It drives change through research and foresight on a wide range of child rights issues, sparking global discourse and actively engaging young people in its work.

UNICEF Innocenti equips thought leaders and decision-makers with the evidence they need to build a better, safer world for children. The office undertakes research on unresolved and emerging issues, using primary and secondary data that represent the voices of children and families themselves. It uses foresight to set the agenda for children, including horizon scanning, trends analysis and scenario development. The office produces a diverse and dynamic library of high-level reports, analyses and policy papers, and provides a platform for debate and advocacy on a wide range of child rights issues.

UNICEF Innocenti provides, for every child, answers to their most pressing concerns.

This brief was written by Daniel Kardefelt Winther, Mariya Stoilova, Moritz Büchi, Marium Saeed, Rogers Twesigye, David Smahel, Marie Bedrosová, Nikol Kvardová and Sonia Livingstone.

United Nations Children's Fund
UNICEF Innocenti – Global Office of Research and Foresight
Via degli Alfani, 58
50121, Florence, Italy

Email: innocenti@unicef.org
Website: www.unicef-irc.org

© United Nations Children's Fund (UNICEF), July 2023

Suggested citation: Kardefelt Winther, Daniel, Mariya Stoilova, Moritz Büchi, Rogers Twesigye, David Smahel, Marie Bedrosová, Nikol Kvardová, and Sonia Livingstone, *Children's Exposure to Hate Messages and Violent Images Online*, UNICEF Innocenti – Global Office of Research and Foresight, Florence, July 2023.

This publication was produced with financial support from the Saudi Esports Federation.

for every child, answers

unicef
for every child