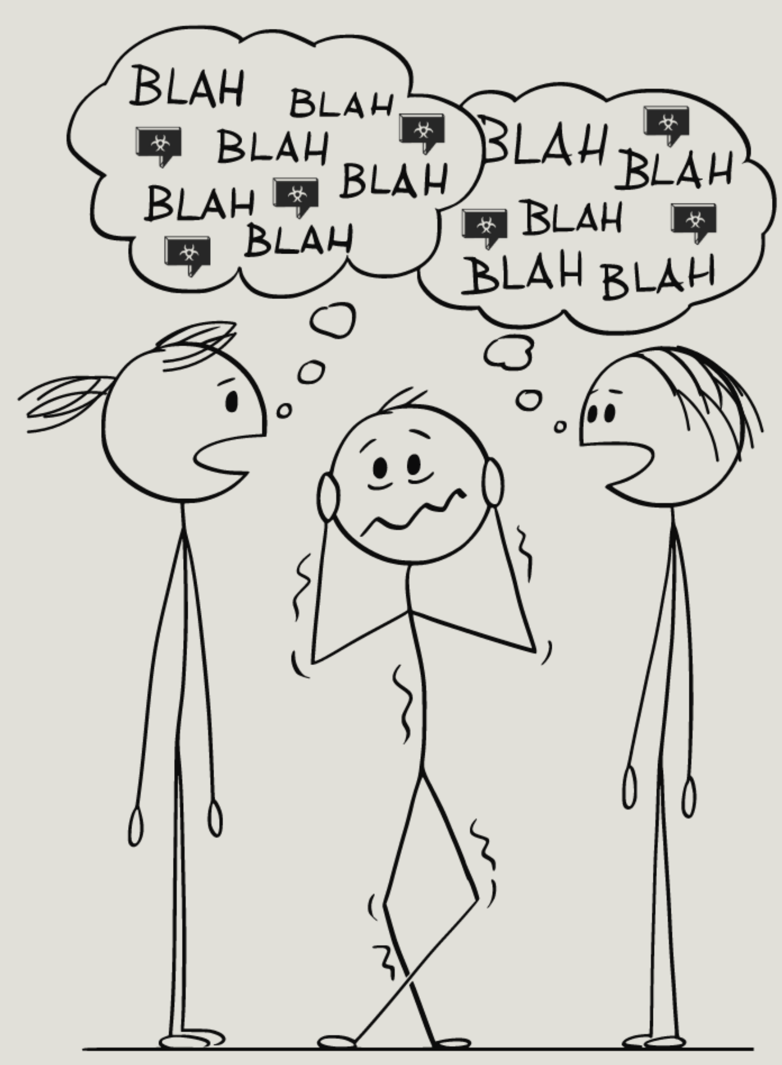
DISCUSSION DOCUMENT

DISINFORMATION

THROUGH A CHILDREN'S RIGHTS LENS



South Africa,
October 2022

Published by the Media Monitoring Africa
https://mma.org.za

South Africa, October 2022

****

This report was prepared by Zahra Abba Omar, Christy Chitengu, and Tina Power with design support from Kwazi Nwana
https://altadvisory.africa/
https://powersingh.africa/

This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License. This means that you are free to share and adapt this work so long as you give appropriate credit, provide a link to the license, and indicate if changes were made. Any such sharing or adaptation must be for non-commercial purposes and must be made available under the same “share alike” terms. Full licence terms can be found at http://creativecommons.org/licenses/by-ncsa/4.0/legalcode

***

Acknowledgements

A special thanks to those who contributed:
This discussion document is in many ways informed by the inputs of a group of phenomenal individuals who generously shared their personal experiences, expertise, insights, and time. A significant thanks go to Professor Sonia Livingstone, Phakamile Khumalo, George Kalu, and Stanley Malematja for setting aside time to engage with us on these critical issues. We are truly thankful to all of you for your invaluable contribution to this work. Your insights, participation, and the information and stories you shared have helped us create resources that we hope will contribute to a growing understanding of disinformation through a children’s rights lens.

CONTENTS

UNPACKING THE INFORMATION ECOSYSTEM 3
- Distinguishing the terminology 3
- Unpacking intention and public harm 4
- Freedom of expression 5
- Intentionally cause public harm 9

UNDERSTANDING WHY DISINFORMATION IS DANGEROUS 11
- The impact on expressive rights 11
- The impact on the right to non-discrimination 12
- The threat to democracy 12
- Causing harm to media sustainability 13
- Potential to be weaponised 14
- The psychological consequences 15

CHILDREN & DISINFORMATION 16
- Children’s rights 16
- Overview of children’s engagement with the online world 16
- Children and disinformation 18
- Impact of disinformation 19

LESSONS FROM AROUND THE WORLD ON RESPONSES TO DISINFORMATION 22
- Regulatory efforts 22
- Policy efforts 24
- Advocacy efforts 26

KEY RECOMMENDATIONS 29
- Children 29
- Parents and caregivers 29
- Civil society 29
- Policymakers 30
- Technology companies 30

CONCLUSION 31

INTRODUCTION

Disinformation is by no means a novel concept; history is littered with examples of the diffusion of deceptive content. Sun Tzu recognised the central role of disinformation in conflict advocating that [a]ll warfare is based on deception."[1] The invention of the printing press saw a rise in the spreading of false content.[2] King Leopold's international disinformation campaign sought to sow discord and distrust amongst his critics,[3] and the KGB of the former Soviet Union had dedicated departments conducting disinformation campaigns.[4]

While the objective of deceiving people is not novel, the nature of the beast is amplified through the use of social media and other online platforms, that are enabling and unleashing new and innovative ways of disseminating content in mass. The speed, reach, and scope of information production and dissemination have expanded exponentially in the digital era. As of late, we have seen online disinformation scourge the fairness and credibility of elections in several countries around the world, threatening democratic political and policy-making processes. We were recently presented with innumerable challenges concerning disinformation amidst the current global health crisis, observed the impact of disinformation and incitement during the July unrest in 2021, and are currently confronted with orchestrated disinformation campaigns targeting foreign nationals.

Disinformation is gaining momentum through the increased number and usage of social media platforms. Social media can spread this content easily and conveniently to varied audiences, meaning that disinformation and misinformation can interfere with the lives of diverse sets of users. This issue is increasingly common in South Africa, where more than 75% of South Africans say they regularly come across political news they think is fake,[5] and 80% believe that disinformation is a problem or a serious problem in the country.[6] The impact presented by disinformation and misinformation on the lives of children presents unique challenges and opportunities for civil society, policymakers, and technology companies.

Through a children's rights lens, the purpose of this discussion document is to unpack evolving understandings of disinformation and its impact on children. This discussion document reflects on how disinformation implicates the right to freedom of expression and access to information and draws on examples from around the world to gain insights into various responses to disinformation. Following engagements with children's rights experts, this document provides recommendations and guidance to various stakeholders addressing key challenges and opportunities for addressing disinformation from a children's rights perspective.

---

[1] Sun Tzu, 'The Art of War' (6th Century B.C.).
[2] Posetti & Matthews, 'A short guide to the history of 'fake news' and disinformation', (2018) International Center for Journalists (accessible here).
[3] D Cley, 'David v Goliath: The Congo Free State Propaganda War, 1890-1909' International History Review.
[4] Cull et al, 'Soviet Subversion, Disinformation and Propaganda: How the West Fought Against it' (2017) London School of Economics and Political Science (accessible here).
[5] H Wasserman and D Madrid-Morales, 'Study sheds light on scourge of "fake" news in Africa', The Conversation (2018) (accessible here).
[6] D Madrid-Morales, 'How do African audiences engage with disinformation and what do they know about fact-checking?, (2021) (accessible here).

At this stage, we do not know the full impact of disinformation on children, but we know that the dangers and harms of such content impacts fundamental rights whether you are an adult or a child.

By adopting a children's rights approach we can grapple with these evolving challenges in a way that prioritises:

- Learning over punishment;
- Access over limitation;
- Education over regulation; and
- Fun over fear.

Core values that underpin children's rights, such as respecting children's evolving capacities, creating safe and nurturing environments for growth and development, enabling children's participation, and finding collective social solutions rather than punitive approaches allow us to consider creative and meaningful solutions to tackle disinformation.

Ultimately, this discussion document seeks to encourage and serve as a starting point for conversations with children, to seek their input and ideas on these themes. Through engaging with children around this document, it is hoped that meaningful and sustainable rights-based solutions are produced which empower and equip children to engage with content online and enable online spaces to be as safe as possible for children.

UNPACKING THE INFORMATION ECOSYSTEM

Distinguishing the terminology

With the flood of information available to citizens of all ages on social media, a collection of various categories of information emerges.[7] Content can be factual, satirical, misleading, or outrightly false. These last two negative forms of communication have often been given the label "fake news", a nebulous term. For greater clarity, these two negative forms of communication can be drawn apart and assigned particular conceptual terms, namely misinformation and disinformation.

'Fake news' – a common but unhelpful term

While a common term, "fake news" is not commonly understood and is often misused to sow division and mistrust.

In recent years the term has been misappropriated by powerful actors who weaponize the term "fake news" to confuse, polarise, mislead, and create distrust of genuine news.[8] Politicians, for example, have been known to call information "fake news" when the information does not align with their views.[9]

Its imprecise nature also means that it encompasses a spectrum of information types, ranging from relatively low-risk forms – such as honest mistakes made by reporters, political discourse, and the use of clickbait headlines – to high-risk forms – malicious fabrications or content that such as content that undermines political processes.[10]

In the South African context, different types of media outputs, with different characteristics and with different aims and objectives, have been labelled as "fake news", speaking to the worldwide definitional vagueness of the term. Some of the varied outputs that are included under the fake news umbrella in South Africa range from mainstream news reports that are considered to be biased and classified as "fake news" in an attempt to dismiss or discredit them, to fake content on social media created with the express aim to mislead and shift political, or social discourse, as has been the case with the 'white monopoly capital' campaign on Twitter.[11]

For these reasons, we avoid using the term "fake news" and rather use misinformation and disinformation.

---

[7] C Wardle, 'Fake news. It's complicated', First Draft (2017) (accessible here).
[8] UNESCO, 'Journalism, 'fake news' and disinformation': Handbook for journalism education and training' (2018) (accessible here) at 9.
[9] M Molina et al, '"Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content' (2019) (accessible here).
[10] European Commission, 'A multi-dimensional approach to disinformation: Report of the independent High Level Group on fake news and online disinformation', (2018) at 10 (accessible here).
[11] H Wasserman, 'Fake news from Africa: Panics, politics and paradigms', Journalism (2020) (accessible here).

While there are no universally accepted definitions of misinformation and disinformation, particularly in the context of the online world, we can draw comparative guidance from evolving understandings of these types of information.

The two concepts relate to types of information that are distinguished on the basis of harm and falseness, as well as on the assumed intention of the person creating or sharing the content. UNESCO distinguishes the two terms as follows:[12]

- Misinformation is generally used to refer to "misleading information created or disseminated without manipulative or malicious intent."

- Disinformation is generally used to refer to "deliberate (often orchestrated) attempts to confuse or manipulate people through delivering dishonest information to them."

The key difference between the terms is the intention, misinformation may be unwittingly shared while disinformation is deliberately created and distributed with the intent to deceive or harm. According to UNESCO, "[b]oth are problems for society, but disinformation is particularly dangerous because it is frequently organised, well resourced, and reinforced by automated technology."[13] Moreover, the consequences of disinformation can be far-reaching, causing public harm. Public harm may, for example, manifest in hampering the ability of the public to make informed decisions or putting the public's health, security and environment at risk.

When unpacking intention and public harm it is necessary to recognise that people engage with content differently and different content may impact people differently. While a post circulating, for instance, on a class WhatsApp group may be regarded as a harmless post when viewed individually, it can have significant consequences on how someone on that WhatsApp group might come to view the world and interact with people, especially those that are different from them. When platforms gather disinformation and scale it, in coordinated, amplified, and directed ways, it can have a significant impact, especially on people who are undecided or vulnerable.

Unpacking intention and public harm

Throughout our engagements with children's rights experts, notions around intention and public harm were a key focus. A distinction was drawn between a disinformer, who has a clear intention to cause public harm through producing content that seeks to knowingly deceive or mislead others, and a misinformer, someone who might not carry the same intent to cause harm that the disinformer carries but becomes concerned by the deceiving or misleading content and shares it out of a desire to create awareness.

While navigating the differences between misinformation and disinformation and unpacking intention and harm it is important to remember that there is a lot of content out there. There is useful, informative and interesting content, there is also amusing and inspiring content. But there is also content that may confuse us, make us sad or anxious, or make us angry.

---

[12] UNESCO (above n 8) at 7.
[13] UNICEF, 'Digital misinformation / disinformation and children 10 things you need to know' (accessible here).

The table below lists some of the types of content we may encounter:

Facts and information	Ideas and opinions	Satire, parody and art	Misleading or confusing	Artistic creativity	Shocking or disturbing
Scary	Hurtful	Controversial	Divisive	Inciting violence	Hateful

When we navigate the various forms of content we encounter there are a few key things to consider:

- Does freedom of expression apply to this content?
- What is the intention of this content?
- Does this content cause public harm?

Freedom of expression

Freedom of expression is a fundamental right recognised by international human rights law[14] and the South African Constitution. Section 16(1) of the Constitution affords every person the right to freedom of expression, which includes:

- freedom of the press and other media;
- freedom to receive or impart information or ideas;
- freedom of artistic creativity; and
- academic freedom and freedom of scientific research.

Our courts explained that "[f]reedom of expression lies at the heart of a democracy" as it enables individuals in our society "to hear, form and express opinions and views freely on a wide range of matters."[15] Given the importance of the right, it applies to various forms of content, "not only to 'information' or 'ideas' that are favourably received or regarded as inoffensive or as a matter of

[14] See Article 19 of the Universal Declaration of Human Rights, article 19 of the International Covenant on Civil and Political Rights; article 9 of the African Charter on Human and People's Rights, and the Declaration of the African Commission on Human and Peoples' Rights on Freedom of Expression and Access to Information.
[15] *South African National Defence Union v Minister of Defence and Another* [1999] ZACC 7 (CC) at para 7 (accessible here).

indifference but also to those that offend, shock or disturb".[16] According to the Constitutional Court, freedom of expression extends "even where those views are controversial".[17]

Section 16(2) of the Constitution lists three types of expression that are not protected under section 16(1):

- propaganda for war;
- incitement of imminent violence; and
- hate speech.

If content falls within these three categories it is not protected by the Constitution and there can be consequences for this content.

Distinguishing between different forms of speech is easier said than done as evidenced by our courts and our lawmakers constantly grappling with questions around speech and when it amounts to freedom of expression and when it is speech that is not tolerated by our constitutional frameworks.

It is important to consider different types of content and their intersection with freedom of expression as it informs how we engage with the content and how we can respond to the content. Attempts to combat disinformation that undermine human rights are short-sighted, counterproductive, and must be avoided.[18] The rights to freedom of expression and access to information are not part of the problem, they are the objective and the means for combating disinformation. It is therefore important to strike a balance and safeguard freedom of expression and access to information.

While some content and forms of disinformation may in enjoy the protection of freedom of expression, when the whole environment of public discourse becomes occupied and dominated by intentional and possibly harmful falsehood, it "frustrates the primary purpose of freedom of expression."[19]

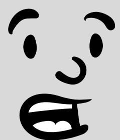
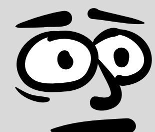
The table below gives high-level guidance on different forms of content and how it intersects with the right to freedom of expression. Given the complexities attached to different forms of speech and evolving legal standards, this table should be viewed as a guide to assist in navigating multifaceted questions relating to freedom of expression and hate speech.

---

[16] *De Reuck v Director of Public Prosecutions (Witwatersrand Local Division) and Others* [2003] ZACC 19 (CC) at para 49, citing *Handyside v United Kingdom* (1976) 1 EHRR 737 (accessible here).
[17] South African National Defence Union v Minister of Defence (above n 15) at para 8.
[18] See further the conclusions and recommendations of Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan, 'Disinformation and freedom of opinion and expression' United Nations Human Rights Council (2021) (accessible here).
[19] European Parliament, 'Disinformation campaigns about LGBTI+ people in the EU and foreign influence' 2021 (accessible here).

Content	Intersection with freedom of expression
	Receiving or imparting information is an important component of the right to freedom of expression. Freedom of expression and access to information "is crucial and indispensable for the free development of the human person, the creation and nurturing of democratic societies and for enabling the exercise of other rights."[20]
	The right to receive or impart information or ideas is an essential element of the right to freedom of expression.[21] Freedom of expression embraces the right to "express and to receive information or ideas freely", which also "encompasses the freedom to form one's own opinion about expression received, and in this way both promotes and protects the moral agency of individuals."[22]
	Satire and parody, which may not be enjoyed by everyone, can be important forms of expression: "A society that takes itself too seriously risks bottling up its tensions and treating every example of irreverence as a threat to its existence. Humour is one of the great solvents of democracy. It permits the ambiguities and contradictions of public life to be articulated in non-violent forms. It promotes diversity. It enables a multitude of discontents to be expressed in a myriad of spontaneous ways. It is an elixir of constitutional health."[23]
	Content can sometimes be misleading, confusing, or false. Sometimes this can be innocuous, but as will be addressed below, where the content is intentionally harmful and threatens or undermines other rights. When this happens, an appropriate balance needs to be struck between competing rights and interests, for example, freedom of expression and dignity and equality.[24] Whether the content is intentionally misleading and created or shared to cause harm such consideration should form part of the balancing exercise.
	Artistic freedom forms part of the right to freedom of expression and includes the right to "imagine, create and distribute diverse cultural expressions".[25] The free artistic expression "nourishes intercultural dialogue and cooperation" and enables curiosity, creativity, innovation, self-realisation and critical thinking.[26]
	Freedom of expression does not protect content that advocates hatred that is based on race, ethnicity, gender or religion, and that constitutes incitement to cause harm. Our Constitution does not protect hate speech. Our Constitutional Court has explained that "[h]ate speech is the antithesis of the values envisioned by the right to free speech", it "marginalises and delegitimises" people and can "can ignite exclusion, hostility, discrimination and violence".27

[20] Declaration of the African Commission on Human and Peoples' Rights on Freedom of Expression and Access to Information (2019) at principle 1.
[21] Id at section 16(1)(b).
[22] *Print Media South Africa and Another v Minister of Home Affairs and Another* [2012] ZACC 22 (CC) at para 53 (accessible here).
[23] *Laugh It Off Promotions CC v South African Breweries International (Finance) BV t/a Sabmark International and Another* [2005] ZACC 7 (CC) at para 109 (accessible here).
[24] *S v Mamabolo (E-TV, Business Day and Freedom of Expression Institute Intervening)* [2001] ZACC 17 (CC) at para 41 (accessible here).
[25] UNESCO, 'Artistic Freedom' (2005) (accessible here).
[26] Council of Europe, 'Manifesto on the Freedom of Expression of Arts and Culture in the Digital Era' (2020) (accessible here).
[27] *Qwelane v South African Human Rights Commission and Another* [2021] ZACC 22 at paras 1 and 78 (accessible here).

	Hurtful and offensive content cannot be limited just by virtue of being hurtful. Our courts have recently grappled with these questions and laws are currently being developed to better navigate the ambit of freedom of expression when it comes to hurtful speech and hate speech. Hate speech is different from hurtful speech and offensive speech – "expressions that are merely hurtful, especially when understood in everyday parlance, are insufficient to constitute hate speech".[28]
	Similarly to hurtful speech, and as noted above, our courts have confirmed that freedom of expression applies "not only to 'information' or 'ideas' that are favourably received or regarded as inoffensive or as a matter of indifference but also to those that offend, shock or disturb".[29]
	Content that sparks fear, anxiety, panic and confusion can be harmful, particularly in times of crisis, such as a pandemic or during civil unrest. Similar to misleading content this form of expression may cause harm and the intention of such content will need to be considered when navigating the bounds of disinformation.
	Sometimes content can maliciously foster division polarise debates and create or deepen tensions in society.[30] Subject to the intention, harm, and possible incitement that this content may trigger, it may or may not be protected speech.
	The right to freedom of expression does not extend to content where there is a clear intention to "incite harm". However, questions around incitement are also presently dealt with by our courts, who consider the literal meanings of content alongside socio-political and historical contexts.[31]

From the above table, it is clear that there are various forms of expression and some grey areas when it comes to a certain form of expression. In the context of mis and disinformation, it is useful to consider the content through the lens of intention and public harm.

The High-Level Group of the European Commission, focuses on the elements of intention and public harm, noting that certain content may not amount to disinformation, noting that disinformation includes:

> "[A]ll forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit. It does not cover issues arising from the creation and dissemination online of illegal content (notably defamation, hate speech, incitement to violence), which are subject to regulatory remedies under … national laws. Nor does it cover other forms of deliberate but not misleading distortions of facts such a satire and parody."[32]

[28] Id at para 103.
[29] *De Reuck* (above n 16).
[30] Joint Declaration on Freedom of Expression and 'Fake News', Disinformation and Propaganda (2017) (accessible here).
[31] *Afriforum v Economic Freedom Fighters and Others* (EQ 04/2020) [2022] ZAGPJHC 599 (accessible here).
[32] European Commission, 'A multi-dimensional approach to disinformation: Report of the independent High Level Group on fake news and online disinformation' (2018) (accessible here).

Intentionally cause public harm

The table below provides some guidance on some of the content that may amount to mis or disinformation when considering intention and public harm:

Content	Description	Intention & Public Harm
	People often have different opinions, beliefs and ideas. Some people think the world is flat, and some people believe that you cannot be friends with someone who is a different race from you. Some beliefs can be harmless like the earth is flat, but some can be very harmful. A person may think it is wrong to be friends with a specific person because of their race. If they say it is illegal to be friends with a specific person because of their race they are portraying their opinion as a fact. This can be harmful as it may lead some people to believe that it is illegal and may cause them to be mean or disrespectful towards people of a particular race.	If sharing an opinion is intended to create a false narrative about a particular group of people and that narrative causes public harm by undermining fundamental rights, the content may be disinformation.
	Sometimes information is designed to confuse us or cause us to mistrust the truth or scientific research. For example, the content may say that if you are sick you should drink detol. This is contrary to medical advice and can be very dangerous. Other content may say that you must not believe your teachers because they are working for an evil organisation that is trying to take over the world. This may cause you to distrust your teacher and affect your ability to learn in class. Sometimes politicians share false content to get people to distrust people who oppose them.	If sharing false content is intended to confuse people, get people to disregard facts, or get people to do something that may harm them, then that content may be disinformation.
	Content can trigger fear, anxiety and panic and be harmful. For example, pictures of burning buildings from a long time ago are shared to make people scared about a current protest. This fear and anxiety may cause people to act in unreasonable ways.	If the content is false and is intended to fuel chaos and cause the public to panic it may be disinformation.
	Sometimes the content may be false and may also hurt someone's feelings. In this instance, disinformation may be used as a tool for cyberbullying, for example, if someone spreads lies or rumours about someone else on social media.	If the content is false and is shared with the intention to mock, ridicule, bully, or defame a person the content may be cyberbullying. Where the intention is to share false content that causes public harm through mocking or ridiculing a group of people it may amount to disinformation.

	Sometimes content can make people change their views, or feel more strongly about pre-existing views. This can lead to people disliking certain things, certain people, or certain ideas, and can sometimes cause harmful division among people with differing views. Content that seeks to trigger strong emotional responses causing people to believe in particular agenda can be harmful. For example, someone may share false content about a certain religion that is intended to make people angry – this anger can manifest into hate and may cause people to be disrespectful, hurtful or even lead to physical violence.	If the content is false and intentionally and maliciously seeks to foster division and breed hatred it is likely to harm the public and may be disinformation.

UNDERSTANDING WHY DISINFORMATION IS DANGEROUS

Far-reaching consequences

Disinformation may have far-reaching consequences, cause public harm, be a threat to democratic political and policy-making processes, and may even put the protection of the public's health, security and environment at risk. It can erode trust in institutions, as well as in the media, and harm democracy by hampering the ability of the public to make informed decisions. It can also polarise debates, create or deepen tensions in society, undermine democratic processes, and impair freedom of opinion and expression.[33]

The impact on expressive rights

While social media has enabled the exercise of expressive rights in an unprecedented way, the dissemination of disinformation can erode the rights to access information and freedom of expression.

Social media is a common space in which disinformation is disseminated, it is crucial to seek to understand how certain activities are enabled by the infrastructure (the code, algorithms, and other aspects which allow a platform to function) on platforms.[34] Platforms and social media sites curate content by using algorithmic personalisation to amplify certain content to particular users based on what platforms believe a given user would find relevant or interesting. By ensuring that a user is kept engaged by their social media feed, platforms can maximise profit through advertising.[35] This poses significant concerns regarding the breadth, quality, and content of the curated information to which users are exposed.

Additionally, research has shown that disinformation spreads faster on social media than on real news. Falsehoods on Twitter, for example, are 70% more likely to be retweeted than facts, and true news stories take up to six times longer to reach people.[36] This has a clear impact on users' ability to access accurate and quality information online.

[33] Joint Declaration (above n 30).
[34] T Gillespie, 'The politics of 'platforms',' New Media and Society (2010) (accessible here).
[35] B Sander, 'Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of Human Rights-based Approach to Content Moderation', Fordham International Law Journal (2020) (accessible here).
[36] S Vosoughi, D Roy, and S Aral, 'The spread of true and false news online', Science (2018) (accessible here).

Disinformation diminishes free expression

"Disinformation causes confusion and has a chilling effect on freedom of expression and information. It directly impacts the level of trust in the public sphere as a space for democratic deliberation. People no longer feel safe to express their ideas for fear of online harassment and of being targeted by disinformation campaigns; others feel paralysed and silenced by the puzzlement and incertitude created by the surrounding information pollution and remove themselves from public debate concerning key issues of public interest."[37]

The impact on the right to non-discrimination

Disinformation may pose risks to people's right to freedom from discrimination. This is because disinformation has been shown to target particular groups in society – such as certain ethnic groups – and is frequently designed to incite violence, discrimination, or hostility. Disinformation involving hate speech has particularly harmful effects on marginalised groups, which may result in physical harm or even loss of life.

Disinformation campaigns about the LGBTQI+ community

In recent years there have been multiple disinformation campaigns targeting the LGBTQI+ community. The campaigns varied in terms of topic and magnitude, ranging from disinformation about the LGBTQI+'s 'threat' to child safety, that members of the LGBTQI+ community are morally corrupted and/or in some way a threat to society, and that the neutral heteronormative family is preferred, with any else being against the natural order.[38] Disinformation aims to polarise society and strengthens narratives that seek to create an image of an external enemy. These campaigns impact the rights to non-discrimination and equality, they can violate human dignity and damage civil and political rights and in some cases can lead to violence.

The threat to democracy

Disinformation has a direct impact on the right to free and fair elections and a discernible influence on democratic processes, including policy-making processes.

For an election to be free and fair, voters need to have accurate information about the parties, candidates, and other factors when they vote. Incorrect information may influence the way that individuals vote, and there are numerous reports which emphasise how the results of elections may be influenced by disinformation.[39]

---

[37] APC, 'Submission to UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression on Disinformation and freedom of expression (2021) (accessible here) at 6.
[38] European Parliament, 'Disinformation campaigns about LGBTI+ people in the EU and foreign influence' (2021) (accessible here).
[39] I Vandewalker, 'Digital Disinformation and Vote Suppression', Brennan Center for Justice (2020) (accessible here).

Tik Tok's influence over elections

In the run-up to Kenya's August 2022 general elections, research revealed that disinformation campaigns and inflammatory rhetoric designed to worsen ethnic divisions were being spread on the country's most downloaded app, TikTok - in violation of the platform's very own policies.[40] Researchers linked this political disinformation to apathy among young voters, which resulted in low turnout for the youth vote, as TikTok's users are typically aged 14-25.

While public discourse is often focused on elections, disinformation and coordinated inauthentic behaviour have the potential to undermine democratic processes more broadly by amplifying anti-democratic narratives or driving polarization.[41]

Causing harm to media sustainability

When social media platforms curate content for users in a way that prioritises disinformation in feeds over credible news, it undermines the diversity of the information landscape. Thus, what is considered newsworthy is determined by an opaque process of algorithmic personalisation, and not what is considered politically or socially important. Should users not have knowledge of the infrastructure of platforms and the content curation at play they might erroneously assume that their feeds are objectively presenting the news.[42]

In this manner, disinformation, frequently geared toward reflecting and reinforcing our existing beliefs and biases, can contribute to a declining public trust in traditional journalism, which might seek to challenge our beliefs by presenting accurate information. Users could be led to believe that accurate information that challenges their beliefs is false, muddying the lines between what is thought to be fact and fiction.

Conversely, without a credible and independent news media sector to provide accurate and reliable information, the public discourse risks becoming further polarised and the challenges of mis- and disinformation online will only continue to be amplified. This further underlines the importance of ensuring media sustainability.

In grappling with questions around freedom of expression, access to information and disinformation, the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression observed as follows:

> "At the core is a human rights challenge, aggravated by an information disorder. There is growing evidence that disinformation tends to thrive where human rights are constrained, where the public information regime is not robust and where media quality, diversity and

[40] O Madung, 'From Dance App to Political Mercenary: How disinformation on TikTok gaslights political tensions in Kenya', (2022) (accessible here).
[41] S Bradshaw, H Bailey and P Howard, 'Industrialized Disinformation: 2020 Global Inventory of Organised Social Media Manipulation', University of Oxford: The Project on Computational Propaganda (2021) (accessible here).
[42] Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, United Nations General Assembly (2018) (accessible here).

independence are weak. Conversely, where freedom of opinion and expression is protected, civil society, journalists and others are able to challenge falsehoods and present alternative viewpoints."[43]

Potential to be weaponised

Gendered disinformation

The Association for Progressive Communication (APC) have categorised the weaponisation of disinformation against women as a form of gender-based violence online.[44] APC cites research that defines gendered and sexualised disinformation as "a subset of online gendered abuse that uses false or misleading gender and sex-based narratives against women, often with some degree of coordination, aimed at deterring women from participating in the public sphere. It combines three defining characteristics of online disinformation: falsity, malign intent, and coordination."[45]

The proliferation of disinformation online poses a number of threats to human rights, but so, too, do laws and policies that are designed to tackle this phenomenon. States have increasingly responded to the spread of disinformation and misinformation online through content-based restrictions and regulation, which in some cases has imposed stricter moderation by platforms. The criminalisation of disinformation may be used as a pretext to clamp down on freedom of opinion and expression and silence dissenting voices.

In a joint declaration on freedom of expression and 'fake news, the United Nations and regional freedom of expression experts stated that: "General prohibitions on the dissemination of information based on vague and ambiguous ideas, including 'false news' or 'non-objective information', are incompatible with international standards for restrictions on freedom of expression."

Criminalising disinformation

In Malaysia, the Anti-Fake News Act of 2018 was passed, which would allow fines of up to $119000 and jail terms of up to six years for disseminating disinformation. Malaysian civil society and international human rights organisations raised concerns that the Act was incompatible with human rights law and would be used to censor speech and punish government critics. In October 2019, the Malaysian Parliament repealed the law.[46]

Such disinformation counter-measures greatly hinder access to information, as people become fearful of sharing information of whatever nature as they fear it may be used against them by the state or flagged as fake or false.

---

[43] Report of the Special Rapporteur 2021 above n 18.
[44] APC above n 37.
[45] Id.
[46] Reuters, 'Malaysia parliament scraps law penalizing fake news', (2019) (accessible here).

Democratic participation can also be stifled due to the disinformation counter-measures that have been deployed by governments. The freedom to express political opinions such as dissatisfaction with the government or the services it renders or government officials may be treated as the wilful spread of false information.

The psychological consequences

Repeated exposure to disinformation can reinforce echo chambers. Once an individual has believed disinformation, it is very difficult to undo this effect, especially when the news is aligned with a person's ideological beliefs.[47] Prior studies have found that disseminating accurate information to counter fake news has been shown to be ineffective[48] and that fact-checking could even have a backfire effect.[49] One possible explanation to demonstrate this phenomenon is the development of echo chambers which suggests that individuals choose to receive news that corroborates their pre-held view.[50]

Overall, and as aptly captured by the UN Special Rapporteur on freedom of expression, disinformation is dangerous and has far-reaching consequences::

> "Interacting with political, social and economic grievances in the real world, disinformation online can have serious consequences for democracy and human rights, as recent elections, the response to the coronavirus disease (COVID-19) pandemic and attacks on minority groups have shown. It is politically polarizing, hinders people from meaningfully exercising their human rights and destroys their trust in Governments and institutions."[51]

[47] S Lewandowsky, U Ecker, and J Cook, 'Beyond misinformation: Understanding and coping with the "post-truth" era,' Journal of Applied Research in Memory and Cognition (2017) (accessible here).
[48] G Appel, LGrewal, R Hadi, and A Stephan, 'The future of social media in marketing', Journal of the Academy of Marketing Science (2020) (accessible here).
[49] E Tandoc, et al, 'Audiences' acts of authentication in the age of fake news: A conceptual framework,' New Media & Society (2018) (accessible here).
[50] S Knobloch-Westerwick et al, 'Selective exposure effects for positive and negative news: Testing the robustness of the informational utility model,' Journalism & Mass Communication Quarterly (2005) (accessible here).
[51] Special Rapporteur report 2021 above n 18.

CHILDREN & DISINFORMATION

Children's rights

Before delving into the ways in which disinformation impacts children it is necessary to briefly set out the children's rights framework. This is important for contextualising children's role in society.

The standard of the best interests of the child is an obvious starting point, internal, regional, and domestic laws give effect to this standard (albeit to slightly varying degrees).[52] It is a "dynamic concept" that must be applied to all and every decision concerning a child.[53] This principle also means that children are bearers of various constitutional rights that must be respected, protected, and promoted.

Giving effect to the best interest principle, our Courts have found:

> "Individually and collectively all children have the right to express themselves as independent social beings, to have their own laughter as well as sorrow, to play, imagine and explore in their own way, to themselves get to understand their bodies, minds and emotions, and above all to learn as they grow how they should conduct themselves and make choices in the wide social and moral world of adulthood. And foundational to the enjoyment of the right to childhood is the promotion of the right as far as possible to live in a secure and nurturing environment free from violence, fear, want and avoidable trauma"[54]

Our Court has further found:

> "Children are precious members of our society and any law that affects them must have due regard to their vulnerability and their need for guidance. We have a duty to ensure that they receive the support and assistance that is necessary for their positive growth and development".[55]

Ultimately, children are developing and need appropriate care and protection to ensure they reach their full potential, which includes enabling them to form opinions, learn and participate in their communities. As such, how children engage with content and how adults elect to respond to disinformation should be informed by progressive standards relating to children's rights.

Overview of children's engagement with the online world

General Comment 25 on children's rights in relation to the digital environment notes that "[m]eaningful access to digital technologies can support children to realize the full range of their civil, political,

---

[52] Convention on the Rights of the Child, 1990 (accessible here); African Charter on the Rights and Welfare of the Child, 1990 (accessible here); Constitution of the Republic of South Africa, 1996 (accessible at here); Children's Act 38 of 2005 (accessible here).

[53] Committee on the Rights of the Child, 'General comment No. 14' (2013) at para 6 (accessible here).

[54] *S v M (Centre for Child Law as Amicus Curiae)* [2007] ZACC 18 (CC) (accessible here).

[55] *Teddy Bear Clinic for Abused Children and Another v Minister of Justice and Constitutional Development and Another* [2013] ZACC 35 (CC) at para 1 (accessible here).

cultural, economic, and social rights."[56] This aligns with our previous discussion document on children's rights in the digital era:[57]

> "the internet is a playground for children: a space where they can learn, develop and participate; where they can experience the world; and where they can dream, imagine and aspire. The internet provides innumerable opportunities for children, and is filled with exciting potential".

Narratives around children's responsibility

During our engagement with Professor Sonia Livingstone,[58] she identified a discourse of children's responsibility which underpins conversations about children's relationship to the online world. A key idea within the United Nations Convention on the Rights of the Child is that children should be able to engage with the world as much as they like but not be burdened with the blame if anything goes wrong. This requires an acceptance of the evolving capacities of children. Children develop at different paces and have had different life experiences which may impact the way in which they make decisions and engage with online content.

In her view, tech companies aim to 'responsibilise' children when it comes to their experiences of the online world. The tech industry emphasises the responsibility of children by championing education as an important tool and by investing in media literacy efforts so that if something goes awry in children's online experiences, blame can be placed on children for using their critical thinking skills and media awareness. Livingstone invites us to read between the lines and recognise that through responsiblising children, platforms are distancing themselves from other solutions, such as government regulation that would infringe upon their business practices.

Additionally, education is a deeply unequal tool – not all children have the same access to it – whereas companies can make their platforms equally safe for everyone so that every individual as unequally prepared are they do not have to battle with issues currently associated with platforms, including mis- and disinformation.

Finally, the tech sector's single-minded focus on education as a tool to tackle mis- and disinformation represents a significant burden placed on the public sector by the private sector.

The language, too frequently repeated, that "children shouldn't be on platforms anyways" fails to recognise that children have a right to be online and that they want to be responsible online actors. Simultaneously, children should not be punished for issues that are far greater than them and that require complex multi-stakeholder solutions.

Our Constitutional Court recently delved into issues relating to children's responsibility, highlighting the importance of finding solutions that promote the best interests of the child, and accept that children may be exposed to hardship but that it is necessary to create conditions that protect the child. In the context of criminal sanctions against children, the Court found that the preferred approach is on that is "child-centred and focused on rehabilitation, rather than punishment" and that social responses can be far more impactful and in line with children's rights than responses aimed at punishment and blame.

[56] Committee on the Rights of the Child, 'General Comment No. 25 on children's rights in relation to the digital environment' (2021) (accessible here).
[57] MMA, 'Children's Rights Online, towards a digital rights charter' (2020) (accessible here).
[58] Professor of Social Psychology at the London School of Economics.

The same logic applies to children's engagement with online content. While indeed they should learn and in some instances take responsibility for their actions the narrative of placing the full weight of responsibility on a child misses the mark of recognising children's evolving capacities and supporting children and creating enabling environments necessary for their positive growth and development.

Since the early 2000s, children have been seen as the "digital generation", with the internet becoming more and more present in young people's lives as it becomes the medium through which they connect, play, and learn about the world. While young people may be more digitally advanced than even their parents, children and adolescents struggle to distinguish between true and false online content.

Children and disinformation

Children are active users of social media, with online platforms and channels frequently the primary information sources for young people.

Falling victim to mis- and disinformation

An important point was raised by Stanley Malematja, from the University of Pretoria's Centre for Child Law, who encourages an understanding of misinformers, especially children who are misinformers, as victims of disinformation.

To help navigate the differences between misinformation and disinformation and to highlight how young people can fall victim to misinformation the Web Rangers have launched a disinformation comic book.[59]

UNICEF has found that children can have various relationships with disinformation – they "can be targets and objects of mis/disinformation, spreaders or creators of it, and opponents of mis/disinformation in actively seeking to counter falsehoods."[60] Young audiences can be affected by and engaged with disinformation in various ways.

- They may be harmed by mis- and disinformation (such as a WhatsApp voice note and fake consent form that warned parents of South African learners to be mindful of the government misleading parents into giving their children a Covid vaccine).[61]

- They may be the targets or objects of mis- and disinformation (for example, false reports spread on WhatsApp in India warning of strangers who were abducting children, which ignited deadly mob attacks).[62]

---

[59] Luthando's MidAventure with Mis- & DisInforamtion (accessible here.)
[60] UNICEF, 'Digital misinformation / disinformation and children' (2021) (accessible here) at 4.
[61] News24, 'FACT CHECK: Is govt trying to trick you into giving your kids a Covid shot? Ignore that voicenote' (2021) (accessible here).
[62] Geeta Anand and Suhasini Raj, 'Rumors on WhatsApp Ignite 2 Mob Attacks in India, Killing 7,' The New York Times (2017) (accessible here).

- They may be responsible for spreading mis- and disinformation among their peers which may undermine their ability to know and distinguish fact from fiction and limit their ability to act as informed, responsible citizens.

The table below highlights these differing engagements with disinformation:[63]

Encountering	Spreading and creating
As with all online users, the likelihood of encountering fake, misleading, and harmful content is fairly common. Various studies suggest that 75% of South Africans say regularly come across political news they think is false, with 80% of South Africans believing that disinformation is a problem or a serious problem in the country.[64] Other studies suggest that in recent years young people between the ages of 14 and 24 reported seeing online disinformation at least once a week.[65]	Children can unwittingly or intentionally, share, amplify or create disinformation. Children may be inclined to reshare content from an influencer they follow, or they may want to be seen to be sharing certain content in order to fit in. In some cases, children may be incentivised by money to reshare or create disinformation, or they may want to share disseminated disinformation in order to bully, harass, and harm others. Children may also unintentionally spread disinformation, as adults do when they are unable to decipher what is false.
Being targeted	**Combatting**
This appears to be an emerging trend where children, particularly those who are activists for certain causes are targeted. For example, climate activist Greta Thunberg has been targeted by smear campaigns and conspiracy theories. In another example, Emma González, a teenage survivor of a school shooting, was photographed for a piece on gun control in which she was tearing up a copy of a shooting range target. A manipulated version, which was widely circulated on social media, showed González ripping up the US Constitution.	On a more positive note, children, as we see through our Web Rangers programme, are becoming central actors in the fight against disinformation. Children are contributing to myth-busting and fact-checking initiatives and are engaging in media and digital literacy training initiatives.

Impact of disinformation

While evidence suggests that children encounter mis or disinformation frequently "there is insufficient data available to researchers and policymakers to get a clear and comprehensive picture of how susceptible children are to mis/ disinformation and how it affects their development, well-being and rights."[66]

Accordingly, whether or not children are more negatively impacted by disinformation than adults is a source of debate. Some research finds young people and older people alike are vulnerable to

[63] Adapted from the UNICEF report above n 60.
[64] H Wasserman and D Madrid-Morales, ' Untangling the web – giving children the right tools to fight fake news' (2022) (accessible here).
[65] UNICEF above n 60.
[66] Id.

disinformation and that people of all ages struggle to replace digested disinformation with corrected facts.[67]

However, it can be argued that children do not always have the cognitive and emotional capacity to identify reliable information and that children may be particularly vulnerable to mis- and disinformation because their maturity and cognitive capacities are still evolving. For example, evidence shows that children typically start to believe in unproven conspiratorial ideas from the age of 14.

Livingstone drew her understanding of children's evolving capacities by comparing the impact of disinformation on junk food advertising. In her research, she found that junk food advertising affects people of all ages equally. However, she argues that adults have had more time to gain awareness and receive or seek guidance and information on the harms of junk food. Children have not had the benefit of this additional time – thus the harm of junk food advertising is not fairly distributed. Children can be less knowledgeable about the ways of the world and how other people might wish to exploit them.

While we do not know the full extent or specific impact of disinformation on children we do know that children can be suspectable to disinformation, and they may not always have the necessary maturity or digital literacy skills to navigate the complex online landscape.[68]

We also know that the impact of disinformation, discussed above, on various rights is equally applicable to children. Hindering access to accurate information, impacting freedom of expression, and sowing division and hatred are all consequences of disinformation that children may experience.

When it comes to children it is important to consider that children's knowledge and awareness of content and disinformation are dependent on various factors such as age, access to digital literacy training, home and educational environments, and children's evolving capacities. For example:

- Primary school children are less likely to be aware of terms such as fake news and disinformation than teenagers. This may make younger children more vulnerable to disinformation.[69]

- Some children lack the confidence to identify disinformation or struggle with knowing what is true and what is not, particularly when information is presented in a way that suggests it is news, or when it is shared by a particular person – for example, a popular influencer, gaming icon, or celebrity.

A unique impact of disinformation in relation to children relates to responses imposed by adults. Similarly to the points raised by Livingstone regarding responsibility, a common response to challenges facing children is simply to remove the perceived challenge. In this instance, impose restrictive measures limiting children's access to the internet and social media.

---

[67] J Guillory and L Geraci, 'The persistence of inferences in memory for younger and older adults: Remembering facts and believing inferences', Psychonomic Bulletin & Review (2010) (accessible here).
[68] UNICEF above n 60.
[69] National Literacy Trust, 'Fake news and critical literacy The final report of the Commission on Fake News and the Teaching of Critical Literacy in Schools' (accessible here).

However, such efforts only succeed in disempowering children and preventing them from being able to take advantage of online opportunities and the ways digital spaces can enable their rights. Importantly, their rights to freedom of expression and to information are infringed upon through such restrictions on their access to the internet – rights that "are fundamental to democracy… [as] children have embraced the internet as a means to learn, share, and participate in civic life".[70]

Children's voices are frequently excluded from traditional forms of media and political participation, including from voting, which emphasises the importance of children's access to the internet and to social media for their freedom of expression. Additionally, restricting children's access to the internet discourages them from taking advantage of online opportunities that can help them acquire the skills they need.[71] Moreover, along with children's rights, they have corresponding responsibilities to be good digital citizens.

To sum up, at this stage, this is insufficient evidence to suggest what the exact impact of disinformation on children is. However, given the unique position that children occupy in society, it is reasonable to assume that they are susceptible to disinformation and its associated harms, and as will be discussed in the following sections, there is an urgent need for various stakeholders to take collective responsibility in ensuring that children can engage in safe and enabling spaces. In doing so, responses must align with progressive understandings of children's rights, and not default to limiting, restricting, or punishing children. As such, we propose the following guidelines be prioritised when considering disinformation through a children's rights lens:

- Learning over punishment;
- Access over limitation;
- Education over regulation; and
- Fun over fear.

[70] C Nyst, A Gorostiaga, and P Geary, 'Industry Toolkit for Children's Online Privacy and Freedom of Expression,' UNICEF (2019) (accessible here).
[71] Id.

LESSONS FROM AROUND THE WORLD ON RESPONSES TO DISINFORMATION

Regulatory efforts

Several governments have introduced regulations to limit the spread of false content online but have differed in their approaches – some countries have focused on curtailing the speech of individuals, whereas other countries have sought to regulate platforms themselves. The implementation of the legislation has often symbolised a battle to balance the curtailing of mis- and disinformation with respect for human rights and freedom of speech. Additionally, taking on tech platforms is a significant legislative challenge in and of itself.

- China: China has some of the strictest laws in the world when it comes to mis and disinformation. In 2016, the government criminalised creating or spreading rumours that "undermine economic and social order."[72] Another law in 2017 requires social media platforms to solely republish and link to news articles from registered news media.[73] Authorities also started requiring microblogging sites to highlight and refute rumours on their platforms.

- European Union: Since 2018, the European Union has had a voluntary code of conduct that asks social media companies to submit reports about their services and products. In May 2021, the European Commission released a complementary Guidance on Strengthening the Code of Practice on Disinformation[74], with an aim to strengthen the application of the Code and expand it beyond large social media platforms, demonetize disinformation, empower users to understand and flag disinformation, expand the coverage of fact-checking, and provide increased access to data for researchers. In April 2022, the EU reached a deal on landmark legislation through a law called the Digital Services Act, which is set to force Meta, Google, and other platforms to combat misinformation, disclose how their services amplify divisive content and stop targeting online ads based on a person's ethnicity, religion or sexual orientation.[75]

- Germany: In 2018 Germany enacted the Network Enforcement Act, which requires large social media companies to remove content that is "manifestly unlawful" within 24 hours.[76] This Act has been used as a prototype for global online regulation, influencing the legislation of at least 13 countries. However, definitions of illegal content vary, and several of these countries, including Venezuela, Vietnam, India, Russia, Malaysia, and Kenya, require intermediaries to remove vague categories of content that include "fake news", "defamation of religions", "anti-government

[72] M Repnikova, 'China's lessons for fighting fake news', Foreign Policy (2018) (accessible here).
[73] Provisions on the Administration of Internet News Information Services (2017) (accessible here).
[74] European Commission, 'Guidance on Strengthening the Code of Practice on Disinformation,' (2021) (accessible here).
[75] A Satariano, 'E.U. Takes Aim at Social Media's Harms with Landmark New Law', The New York Times (2022) (accessible here).
[76] BMJV, 'NetzDG – Gesetz Zur Verbesserung Der Rechtsdurchsetzung in Sozialen Netzwerken', (2017) (accessible here).

propaganda" and/or overly broad definitions of hate speech/incitement, extremism and defamation; in such cases, mainstream news sources might be classified as disinformation so as to dismiss them and prevent the public from accessing this information.[77]

- United States: In the United States, little action has been taken to regulate disinformation spread on tech platforms.[78] While US regulators have filed antitrust cases against Google and Meta, no comprehensive federal laws tackling the power of the tech companies have been passed.[79] Section 230 of the Communications Decency Act grants technology companies immunity for the third-party content they host, and this section of the Act is being contested.

Legislating / criminalising disinformation: old approaches to new problems

A concerning trend has emerged in the Southern African region regarding political commentary around the criminalisation of disinformation. In August 2020, the Heads of State of the Southern African Development Community (SADC) held its 40th Ordinary Summit during which it was resolved that member states are urged to "take pro-active measures to mitigate against external interference, the impact of fake news and the abuse of social media, especially in electoral processes".[80] In South Africa, there have been shifting attitudes of the South African Police Services (SAPS) regarding disinformation. Some members of SAPS have noted that the publication, distribution, disclosure, transmission, circulation or spreading of false information or fake news is an offence,[81] whereas others cautioned the public against the incessant promotion and distribution of malicious untruths that seek to sow panic and pandemonium amongst communities.[82]

These old approaches to new problems do not work - using legislation and regulation to combat the issue does not always address the issue:

- First, efforts to legislate tend to be used to limit more than disinformation but cater, intentionally or unintentionally, to greater incursions on freedom of expression, arguably making the problem worse.

- Secondly, legislation is often slow, and in most instances seeks to punish or deal with offenders who transgress the law. It is not always catered at addressing systemic issues such as access to information and media and information literacy.

- Thirdly, and most importantly, the problem is bigger than any single government, and while states have a critical role to play, they alone are unlikely to meaningfully address the core issues. A multi-stakeholder and multi-disciplinary approach is likely to be more effective. One which includes the state, regional and international bodies, social media platforms, the media, and civil society.

This considerations are equally important when engaging with disinformation through a children's rights lens and align with the principle of learning over regulation.

---

[77] J Mchangama and J Fiss, 'The Digital Berlin Wall: How Germany (accidentally) created a prototype for global online censorship', Justitia (2019) (accessible here).
[78] C Kang, 'As Europe Approves New Tech Laws, the U.S. Falls Further Behind', The New York Times (2022) (accessible here)
[79] ASatariano, 'E.U. Takes Aim at Social Media's Harms with Landmark New Law', The New York Times (2022) (accessible here)
[80] Communique of 40th Ordinary Summit Heads of State of the Southern African Development Community (SADC), (17 August 2020) (accessible here).
[81] South African Police Service, 'Media Statement: 17 September 2020' (accessible here).
[82] South African Police Service, 'Media Statement: 23 September 2020' (accessible here).

Concerns with over-regulation from a child rights perspective:

- Throughout our engagements, stakeholders expressed unease with regulations aimed at curtailing online speech, especially when it comes to children. They believed that overly punitive approaches were not in the best interests of children.

- All stakeholders expressed support for regulating the tech industry, acknowledging the ongoing challenge to regulate this sector.

Policy efforts

Tackling disinformation through a policy lens may include considerations around education and empowerment, there are useful examples from around the world where this approach is proving to be helpful in the fight against disinformation. Moreover, it is necessary to reflect on the role of big tech and their responses to disinformation when it comes to children.

Responses from tech companies

As noted, tech companies have focused on tackling disinformation's impact on children (and other online harms) by supporting efforts to equip children with access to digital literacy resources and training. Additionally, tech companies have sought to create tools within their platforms to address online harm. Meta's Digital Literacy Library contains a collection of lesson plans for educators. These resources aim to help young people develop skills that are essential to navigating the digital world, consuming information critically and responsibly producing and sharing content. These lessons include group discussions, quizzes, games and other engaging activities that have been designed in consultation with teens.[83]

An argument was made during our engagements with stakeholders that tech companies could leverage their power better by designing algorithms that identify and take down disinformation swiftly, by addressing the ways current algorithms curate and amplify misleading or false content to users, and by employing more human content moderators around the world who speak languages other than English and understand cultural contexts outside of the Western paradigm.

The importance of digital literacy

According to UNESCO, digital literacy comprises "a set of competencies that help people to maximize advantages and minimize harms" and "covers competencies that enable people to critically and effectively engage with: communications content; the institutions that facilitate this content; and the use of digital technologies."[84]

In the context of teaching and learning, media literacy has the following benefits:

- It equips teachers with a better knowledge to empower learners as future citizens;

---

[83] Facebook, 'Digital Literacy Library', (accessible here).
[84] UNESCO, Media and information literate citizens: Think critically, click wisely!', (2021) (accessible here).

- It develops knowledge about the function of media and information in democratic societies, an understanding of how those functions are performed effectively and the skills to evaluate media and information providers accordingly; and

- It nurtures an appreciation of free, independent and pluralistic media and information systems.[85]

Insights from Professor Livingstone

Sonia Livingstone reflected critically on the use of digital literacy as a means to address mis- and disinformation, noting that while educators are interested in designing curricula and raising awareness around these issues, little consideration is given to the desired outcomes of these interventions. For instance, is the desired outcome that young people do not spread disinformation, that their decision-making is improved, or that they are empowered to discredit false content? Additionally, she has identified a lack of adequate evaluation of the effectiveness of these interventions.

Examples of programmes and school curricula:

- National curriculum: South Africa lacks a comprehensive national media literacy programme. Often it comes down to individual teachers and schools to make learners more media literate, with skills taught in different subjects, such as life orientation, technology, language, or history. This means media literacy content is fragmented, diffused, and limited. Learners are taught how to use the media, how to stay safe online and how to produce media content, but considerably less attention is devoted to fact-checking and verifying the media.[86]

- Web Rangers: As of 2022, Media Monitoring Africa's Web Rangers programme has trained over 1500 young people across South Africa in critical thinking alongside digital and media literacy skills. The Web Rangers programme is run together with major partners, including Google, Facebook, the Department of Communications and Digital Technologies, the Media Development and Diversity Agency and MTN. In addition to online safety, the programme also equips young people with the skills to critically engage with content, learn how to spot disinformation and develop methods on how to use the internet to build South Africa's democracy. Through the program young digital citizens are empowered to use the internet responsibly and encourage their peers to do the same, and encourages the development of digital literacy skills that "allow young people to gain critical skills and knowledge around online safety that they use to create innovative campaigns that promote safe internet usage and champion their rights in the digital world."[87] These young learners unpack emerging challenges facing young people as well as the exciting opportunities in an increasingly digital world and navigate how best they advance their rights, in line with their evolving capacities and their agency.

[85] C Wilson et al, 'Media and Information Literacy Curriculum for Teachers', UNESCO (2021) (accessible here).
[86] H Wasserman and D Madrid-Morales (above n 5).
[87] For more information see Web rangers (accessible at https://webrangers.co.za/).

Insights from the coordinators of the Web Rangers Programme

Phakamile Khumalo and George Kalu, coordinators of the Web Rangers Programme, shared their takeaways with us.

Khumalo and Kalu explained that some of the forms of teaching they have found to be most effective in reaching children were lessons that focused on children's rights and responsibilities online (in other words, children have a right to access information and equally a responsibility to share only accurate and quality information), encouraging children to interrogate their own intentions when using social media, and empowering children to actively fight against mis- and disinformation.

Web Rangers sessions draw on scenarios from real life and from Real 411, Media Monitoring Africa's portal to report mis- and disinformation (discussed under advocacy efforts), to illustrate the causes, effects, and consequences of false or misleading content.

Additionally, the Web Rangers programme has found success in using art-based participatory methods, most recently through developing a comic book with workshop participants, highlighting a creative method through which these important lessons can be imparted.

Notably, Khumalo and Kalu expressed a desire to scale these teachings to the level of the national school curriculum, but spoke of the near-impossibility of changing school curricula in South Africa and that, when they have attempted to engage with the government on introducing these issues into curricula, there has been a lack of momentum on the part of the government. Instead, the educators suggest that success can be found in presenting Web Rangers' curricula as pre-prepared additional content teachers can work with.

Advocacy efforts

As evidenced by the diverse examples below, activism, advocacy, and fun can play an important role in the fight against disinformation.

Real411

Media Monitoring Africa's Real411 platform is a world-first initiative, bringing together various stakeholders across the media and civil society sectors. The platform was launched in 2019, initially in partnership with South Africa's electoral commission ahead of the country's local government elections to address mis- and disinformation during the election period. Complaints are submitted by members of the public via the online portal, reviewed by experts, and the outcome of the complaint, with an explanation, is published on the Real411 website.[88]

[88]Media Monitoring Africa, 'Real 411: Report Digital Disinformation,' (accessible here)

Luthando's MisAdventure with Mis- & DisInformation

MMA, together with the Impact Amplifier and Web Rangers have developed an informative comic book to assist young people to navigate mis and disinformation.[89] The comic book focuses on misinformation because experience has shown that young people often fall victim to misinformation rather than disinformation. Through the story, it becomes clear that innocent conversations between two people about an unknown issue have the potential to cause great harm.

'Humour over rumour'

Taiwan's Digital Minister Audrey Tang, a self-described civic hacker, implemented a 2-2-2 strategy 'humour over rumour' strategy to deal with mis- and disinformation spread during COVID-19. The strategy involves providing a response to misinformation in 20 minutes, in 200 words or fewer, alongside two fun memes.[90] As Tang explains it, "When we roll out within a couple of hours a funny response, it motivates people to share something enjoyable, rather than something retaliatory or discriminatory, and then people feel much better." Creative teams, including graphic designers and comedy writers, help to challenge this disinformation. Additionally, among the fact-checkers fighting disinformation are many middle-school students connected through the Taipei-based platform for civic participation called g0v (gov-zero), which seeks to "fork" the government, by taking government-issued information and making it more visually and contextually accessible to the average citizens. Tang's efforts represent a creative – and humorous - means of including young people in the fight against disinformation.[91]

Bamse – the world's strongest bear speaks out about disinformation

Bamse – världens starkaste björn (Bamse – the world's strongest bear), a beloved Swedish cartoon bear has been promoting practices aimed at combatting disinformation.[92] In "Bamse and the dark woods", Bamse's friend Little Hop reads on the internet that Bamse can no longer become strong from thunder honey. Shellman the tortoise wisely asks "Have you verified the source?" It then comes to light that the "news" was written by two voles, who are enemies of Bamse. A simple message to be received by over 100 000 children is a creative and effective way of engendering positive narratives around fact-checking.

A fun game about Bad News

Bad News is a fun game developed to assist young people to become more conscious of disinformation distribution practices.[93] In Bad News, a player takes the role of a fake news-monger. whose task is to get as many followers as while slowly building up fake credibility as a news site. "The goal of the game is to expose the tactics and manipulation techniques that are used to mislead people and build up a following," it seeks to build "cognitive resistance against common forms of manipulation that you may encounter online."[94] Bad news was written with an audience in mind of ages 14 and up.

---

[89] S Mandla et al, 'Luthando's MisAdventure with Mis- & DisInformation' (2021) (accessible here).
[90] Rorry Daniels, 'Taiwan's unlikely path to public trust provides lessons for the US,' Brookings Institute (2020) (accessible here)
[91] Jonas Glatthard and Bruno Kaufmann, 'Humour over rumour': lessons from Taiwan in digital democracy,' (2021) (accessible here)
[92] Financial Times, 'Swedish comic strip teaches that internet voles are full of fake news' (accessible here).
[93] Bad New (accessible here).
[94] Bad News, 'About' (accessible here).

Factitious

Another fun learning initiative Factious, a game that challenges players to identify fake and real stories.[95] In the game, players receive a news clipping and are asked to swipe left if they think it is false, and right if they think it is true. The game shares the sources of the clippings and explains when something is true, false or satirical. During the pandemic, the game developers release a new edition of the game focusing on fake articles about COVID-19.[96]

[95] Factitious (accessible here).
[96] Factitious, Pandemic Edition (accessible here).

KEY RECOMMENDATIONS

Numerous approaches have been proposed in the fight against disinformation. One set of approaches focuses on the use of technology to detect disinformation, another on the regulation of online speech, and another yet promotes the use of media literacy through education and training. Reflecting on the theme of media literacy, Sonia Livingstone has warned of the dangers of answering calls for regulation of platforms with counter-calls for more media literacy, noting that is important that the burden of responsibility is not shifted from those creating or amplifying online harms onto their victims, especially when those victims are children.[97]

We have prepared the following recommendations for children, parents and caregivers, civil society, policymakers, and technology companies. We will be using this discussion document and these recommendation to engage further with the Web Rangers and seek their input on these proposals. Following those engagements, we hope to work with the various stakeholders listed to bring the agreed upon recommendations to life.

Children

- Engage in digital and information literacy programmes to the best of your abilities.
- Strive to make your community space for the flow of accurate information.
- Remember that these issues are complex and sometimes scary – they are not yours to solve alone but you can play an important part in fighting mis- and disinformation.

Parents and caregivers

- Engage in improving children's and your own digital literacy and critical thinking abilities.
- Encourage children and all those around you to challenge what they see online.
- Have regular discussions with children about what they see online.

Civil society

- Provide policy guidance to minimise mis- and disinformation for children by working with policymakers and technology companies to rapidly evaluate children's exposure to mis- and disinformation and identify the policy interventions and service and product redesigns needed to minimise its negative impact.
- Advocate for solutions to these issues that do not infringe upon the rights to freedom of expression and access to information.
- Conduct ongoing research on the impact of mis- and disinformation on children and the efficacy of counter-measures, including proper monitoring and evaluation of interventions. The findings should inform continued advocacy for children's rights and evidence-based regulations.

---

[97] S Livingstone, 'Why is media literacy prominent in the UK's draft Online Safety Bill 2021?' (2021) (accessible here).

Policymakers

- Devise child rights-based regulations focused on the classification procedures for online content; enlist the support of technology companies, and require greater transparency, accountability and global responsibility from social media platforms around mis- and disinformation and children.
- Support media and information literacy programmes at the level of national curricula that bring caregivers and educators into the fold, such as those implemented in Estonia[98] and Finland[99].
- Support an independent and diverse media ecosystem committed to fact-based information.
- Utilise and support research conducted by those in industry, civil society, and academia on mis- and disinformation and children and base policy decisions on this body of evidence.
- Establish robust access to public information systems. Public information measures could include frequent public service announcements, increased public-interest reporting and broadcasting, support for fact-checking platforms, and increased spending on the maintenance of libraries.

Technology companies

- Require large technology companies with significant global influence to fund independent, safe oversight of content moderation, digital literacy programmes, and further research.
- Require large technology companies to undergo regular public audits of their algorithms, content policies and mis- and disinformation moderation practices concerning children's use of their services and products.
- Demand that companies pay attention to global trends on mis- and disinformation directed at all countries, not just those with strong regulators, and require more proactive, rights-based responses to minimise the harmful effects on children.

---

[98] A Yee, 'The country inoculating against disinformation,' BBC (2022) (accessible here).
[99] J Henley, 'How Finland starts its fight against fake news in primary schools,' The Guardian (2020) (accessible here).

CONCLUSION

While we do not know the full impact of disinformation on children, the dangers and harms of such content impact fundamental rights whether you are an adult or a child. By adopting a children's rights lens we can grapple with these evolving challenges in a way that promotes learning and engagement rather than punishment and limitation. Core values that underpin children's rights, such as respecting children's evolving capacities, creating safe and nurturing environments for growth and development, enabling children's participation, and finding collective social solutions rather than punitive approaches allow us to consider creative and meaningful solutions to tackle disinformation. It requires collective buy-in and efforts from all stakeholders. A children's rights approach to disinformation prioritises:

- Learning over punishment;
- Access over limitation;
- Education over regulation; and
- Fun over fear.

We are excited to continue working with children and key stakeholders to find meaningful and sustainable rights-based solutions are produced which empower and equip children to engage with content online and enable online spaces to be as safe as possible for children

MEDIAMONITORING
AFRICA