Written evidence submitted by Hadley Newman (OSB0125)

In response to:
> Earlier proposals included content such as misinformation/disinformation that could lead to societal harm in scope of the Bill. These types of content have since been removed. What do you think of this decision?

To address the Scope of the Draft Online Safety Bill

UNDERSTANDING THE DIFFERENCES BETWEEN DISINFORMATION, MISINFORMATION, MALINFORMATION AND INFORMATION – PRESENTING THE DMMI MATRIX

September 2021

Author: Hadley Newman, acting on an individual basis.

Hadley Newman is a senior director with a global public affairs and strategic communication consultancy. He leads national and international communications campaigns across a broad range of policy areas for public and multilateral organisations. He specialises in strategic communication planning and the empirical analysis of target audiences. Hadley is a published commentator on communication governance, and his Ph.D. with Heriot-Watt University explores the influence of targeted communication within information operations.

0

ABSTRACT

This response to the draft Online Safety Bill call for evidence argues that, rather than removing the content such as misinformation/disinformation from the Bill these terms should be clearly defined, and their content made recognisable. In order to understand the distinctions between disinformation, misinformation, malinformation and information (DMMI) clear definitions are offered and the differences between them tabulated. After briefly reviewing the current DMMI environment, where advances in social media have facilitated the rapid dissemination of disinformation, the DMMI Matrix is presented. This matrix helps the understanding and clarity of DMMI and, as a foundation model, supports education efforts. Additionally, it is expected that the DMMI Matrix is relevant to the Ofcom advisory committee on disinformation and misinformation identified in the draft Bill.

Keywords:

Introduction and Background

1.   The published call for evidence noted "earlier proposals included content such as disinformation, misinformation, malinformation and information (DMMI) that could lead to societal harm in scope of the Bill. These types of content have since been removed". The absence, or more accurately the removal, of this content would not achieve the Government's policy objectives as a significant gap in the Bill would exist.

2.   It is questionable, at best, if the differences between DMMI are generally understood, and arguably there is a need provide the means to enable clarity in the fog of confusion. Moreover, malinformation is absent from the government literature, including the draft Online Safety Bill. If DMMI is not properly recognised, the impact of disinformation, misinformation and malinformation cannot be effectively addressed. As this call for evidence notes "British citizens want to feel empowered to keep themselves and their children safe and secure online. Both the government and industry have a responsibility to ensure this is the case" (UK Government, 2020). After all it is stated that the "first duty of the government is to keep citizens safe and the country secure" (UK Home Office, 2021). Indeed, as disinformation is a growing issue for democracy and society, it is beholden on government "to provide more tools for their users to help them identify untrustworthy sources of information" (Collins, 2019). The DMMI Matrix presented below offers one such tool.

Definitions

3.   Though often used interchangeably, disinformation, misinformation and malinformation are different, linguistically, and conceptually. Disinformation refers to the public dissemination of inaccurate, false, and/or misleading information and misinformation refers to the unintentional publication of unreliable content. Malinformation is genuine information that has been shared to cause harm, often by moving information designed to stay private into the public sphere (Wardle and Derakhshan, 2017). In particular disinformation, refers to the deliberate production and/or publication of misleading materials for the sole purpose of deceiving the public. Disinformation, defined by US Congress (1982; p. 8) is "a variety of techniques and activities to purvey false or misleading information, including rumours, insinuation, and altered facts." It differs from overt forms of propaganda as the source of the creator is concealed.

4.   The following definitions have been adopted by this paper:

   • **Disinformation**: Information which is created or disseminated with the deliberate intent to mislead; this could be to cause harm, or for personal, political or financial gain (UK Government, 2019 p. 22).
   • **Misinformation**: The inadvertent sharing of false information (UK Government, 2019 p. 22)

- **Malinformation**: Information that is based on reality, used to inflict harm on a person, organization or country (Wardle and Derakhshan, 2017 p.20)
- **Information:** Is the "knowledge communicated concerning some particular fact, subject, or event; that of which one is appraised or told; intelligence, news" (OED Vol. VII p. 994).

5. As Tatham (2008) notes, information – "in particular its utility, effect and management" – should be considered at the very core of any strategic communication. Meeting the complex challenges of DMMI campaigns will require agility and innovation, as they are fraught with interconnected threats and risks which have the potential to undermine social, national, and international stability. As the Secretary of State for Health and Social Care (2020) noted "a critical part of tackling disinformation is providing accurate, fair and objective positive information."

6. The differences in veracity and intent as well as definition summaries for disinformation, misinformation, malinformation and information are tabulated in the 'DMMI Taxonomy' in Table 1.

	Disinformation	**Misinformation**	**Malinformation**	**Information**
Veracity	False	False	True	True
Intent	Harm Intended	Harm Unintended	Harm Intended	Harm Unintended
Definition	Created or disseminated with the deliberate intent to mislead and cause harm or enable gain	The inadvertent sharing of false information	Information based on reality, used to inflict harm	Facts provided or learned about something or someone

Table 1: DMMI Taxonomy

DMMI Current Environment

7. The dissemination of disinformation has grown rapidly in the age of social media, as the power and potential for disinformation now extends beyond state actors and has fallen into the hands of third-party agents and private citizens. Social media platforms such as Twitter depend on concise communication which enables widespread disinformation. Contemporary concise communication has its foundation in the social media phenomenon, in particular Twitter, where interactions are concise. Twitter has been described as the "greatest relational and communicative phenomenon that has developed on the Internet" (Xifra and Grau, 2010, p. 171). Concise communication has several sub-constructs including information relevance, information equivocality, and system complexity that apply

to the characteristics of social media platforms. When used appropriately, concise communication is an effective source of information and when delivered by a popular platform such as Twitter; it is understood quickly. Disinformation strategists have used mass media to shape the beliefs, perspectives, and, ultimately, the behaviours of the public. Disinformation, as part of wider information operations, has been a powerful weapon to undermine adversaries across social, political, economic, and military lines (Bennett and Livingstone, 2018).

8.   The power to influence people's ideas and behaviours on social media has been well established (Diao et al. 2014). Diao et al. (2014) conceptualised this as a branch of opinion dynamics theory, whereby one's behavioural tendencies and actions are reactions to other members in a multi-agent society. This has since been validated by several studies (Li et al. 2016; Senadheera et al. 2015). For example, prior to responding to a Twitter comment, a user typically interacts with others and may be influenced by the tweets, re-tweets, mentions, and content that emerges as a thread, leading back to the original comment (Vel et al. 2014). In the same way, after interacting and engaging with political content, a user may adopt a behavioural response to a social media campaign (Dwyer, 2012; MacCoun, 2015; Senadheera et al. 2015).

9.   In the era of mass media, the magnitude, potency, and proliferation of disinformation has been amplified – with ever more complex practices emerging as a mechanism of control from within the platforms and constraint from without. The increased adoption of social media has only served to exponentially increase the forms in which the power of disinformation is exerted and the intensity of that power. Disinformation through social media has proven to be a key factor in recent elections throughout the West, including the 2016 US presidential election (Pierri et al., 2020). Also in 2019, a five month-long campaign of social media-based disinformation swept across Italy in advance of the European parliamentary elections (Pierri et al., 2020), to say nothing about the COVID-19 pandemic and the rise of the 'anti-vaxers'.

DMMI Matrix

10.   To clearly demonstrate the differences between disinformation, misinformation, malinformation and information the DMMI Matrix has been developed and is shown in Figure 1.
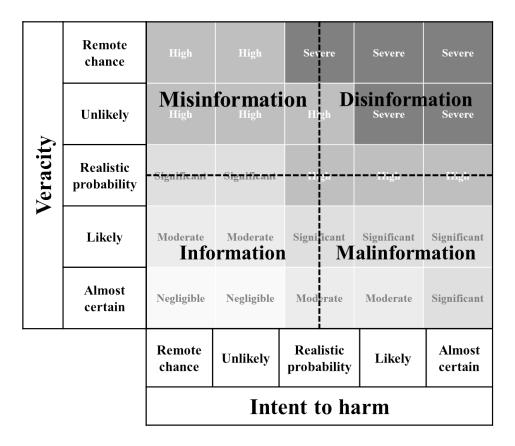
	Remote chance	High	High	Severe	Severe	Severe
Misinformation				**Disinformation**		
Unlikely	High	High	High	Severe	Severe
Realistic probability	Significant	Significant	High	High	High
Likely	Moderate	Moderate	Significant	Significant	Significant
Information				**Malinformation**		
Almost certain	Negligible	Negligible	Moderate	Moderate	Significant

	Remote chance	Unlikely	Realistic probability	Likely	Almost certain

Intent to harm

Figure 1: DMMI Matrix[1]

This style of matrix should be readily recognisable as it has been inspired by well-established risk heat maps such as the Risk Heat Map –Visualization Tool which is designed to "present cyber risk assessment results in an easy to understand, visually attractive and concise format" (Balbix, 2021). The DMMI Matrix uses the PHIA Probability Yardstick – the standard mandated across the UK Intelligence Assessment Community (Irwin and Mandel 2020) – to offer clarity and mitigate subjectivity. The synonyms in the Probability Yardstick establish what the terms within the core of the matrix approximately correspond to regarding the probability (of intent to harm, likelihood of the information being false, level of severity). The more detailed version of the DMMI Matrix includes numerical values (percentage representations) associated to the Probability Yardstick synonyms to ensure that the intended analysis is understood. Addressing 7.31 of the draft Online Safety Bill, the DMMI Matrix is a visualization tool designed to aid understanding and contribute to the protection of users from harm, it also highlights the importance of judging what is true or not, and yet does not discourage freedom of speech online.

Within the DMMI Matrix, veracity and intent are plotted on a two-axis grid. The information rating is where the two points of veracity and intent meet, and the DMMI Matrix affords a view of the information landscape. The DMMI Matrix is divided in to 25 cells, each representing difference levels of severity of the concise communication. The original version of the DMMI Matrix includes a colour scheme similar to a risk heat map where a

---

[1] The original (colour version) of the DMMI Matrix is available upon request from the author. The black and white version was produced solely for this call for evidence.

colour representing 'severity' of the information is appointed to each cell, each of the colours has a distinct meaning, for example cells in the top-right of the DMMI Matrix are red in colour and indicate 'severe', moving diagonally towards the bottom left of the DMMI Matrix are dark orange: high, light orange: significant, yellow: moderate, green: negligible. The DMMI Matrix also presents 'severity' in the different dimensions of information such as the cluster of risks by category of DMMI or by platform, and for example the risks presented by a single piece of concise communication or campaign.

11.   Questions to consider when implementing the DMMI Matrix, a primer:

      i.   Does the information have the intent to harm?

      ii.   Is the information false? This is "to ensure that users can easily flag content that they suspect or know to be false, and which enable users to understand what actions have been taken and why" (UK Government 2020).

      iii.   What, if any, is the level of severity?

      iv.   Does this constitute a material risk?

      v.   What is the range of acceptable variance from (any) established operating metrics?

For example, the Tweet from Chinese Ministry of Foreign Affairs spokesman (Figure 2) is offered as an example of disinformation that can be explained using the DMMI Matrix. The content of this tweet is regarded as 'severe' due to likely-to-almost certain intent to harm with no veracity and therefore would be associated to the top right quadrant of the DMMI Matrix and classified as disinformation.

Figure 2: Tweet from Chinese Ministry of Foreign Affairs spokesman Zhao Lijian

12.   In addressing 7.27. of the draft Online Safety Bill, the DMMI Matrix aids understanding of "the nature and reliability of the information they [users] are receiving, to minimise the spread of misleading and harmful disinformation and to increase the accessibility of trustworthy and varied news content" (UK Government

2020). Furthermore, it addresses 7.28. of the draft Online Safety Bill and will help to "ensure that users can easily flag content that they suspect or know to be false, and which enable users to understand what actions have been taken and why" (UK Government 2020).

13.   Additionally, it is expected that the DMMI Matrix is also relevant to the Ofcom advisory committee on disinformation and misinformation. The Ofcom advisory committee will be established and maintained by Ofcom, in accordance with Chapter 7 of the draft Online Safety Bill and paragraph 14 of the Schedule to the Office of Communications Act 2002, to strengthen the existing work to improve user resilience to disinformation and misinformation through the promotion of media literacy. It is therefore highly recommended that the Ofcom advisory committee considers the inclusion of the DMMI Matrix when writing the report specified in Chapter 7, (5) p.90 of the draft Online Safety Bill.

14.   Finally, as Collins (2019) notes, the focus on DMMI "comes from our belief that there is a genuine danger to democracy and society in the deliberate and malicious targeting of disinformation at citizens, largely using social media to influence what they see and their opinions about politics, society and institution". Building on this, the UK Governments' concern that clarity, education and a better understanding of disinformation/misinformation is required (UK Government, 2020) has been taken into consideration and emphasises the necessity for the DMMI Matrix. The aim of the DMMI Matrix is to support a better clarity and understanding of DMMI and, as a foundation model, support education efforts. The DMMI Matrix should be adopted by the Ofcom advisory committee for inclusion in the report required by the draft Online Safety Bill.

16 September 2021

REFERENCES

Balbix (2021) Risk Heat Map – A Powerful Visualization Tool. https://www.balbix.com/insights/cyber-risk-heat-map/

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. European journal of communication, 33(2), 122-139.

Collins, D. (2019) Sub-Committee on Disinformation. Hansard Volume 657: debated on Thursday 4 April 2019 https://hansard.parliament.uk/Commons/2019-04-04/debates/D33B840A-DC5B-42BB-ADCD-BD4A84337C95/details#contribution-ED9EF775-C72E-450E-B977-31D00D4BFF24

Diao, S. M., Liu, Y., Zeng, Q. A., Luo, G. X., & Xiong, F. (2014). A novel opinion dynamics model based on expanded observation ranges and individuals' social influences in social networks. Physica A: Statistical Mechanics and its Applications, vol. 415, pp. 220-228.

Dwyer, P. (2012). An approach to measuring influence and cognitive similarity in computer-mediated communication. Computers in Human Behavior, 28, 540–551.

Gartner (2021) Gartner Magic Quadrant. Available from https://www.gartner.com/en/marketing/research/magic-quadrants

Irwin, Daniel & Mandel, David. (2020). Variants of Vague Verbiage: Intelligence Community Methods for Communicating Probability.

Li, Y. M., Lai, C. Y., & Lin, L. F. (2016). A Diffusion Planning Mechanism for Social Marketing. Information & Management. Information and Management. In Press. https://doi.org/10.1016/j.im.2016.12.006

MacCoun, R. J. (2015). Balancing evidence and norms in cultural evolution. Organizational Behavior and Human Decision Processes, 129, 93-104.

Pierri F, Artoni A, Ceri S. Investigating Italian disinformation spreading on Twitter in the context of 2019 European elections. PLoS One. 2020 Jan 17;15(1):e0227821.

Secretary of State for Health and Social Care (2020) Anti-vaccination Disinformation Online. Hansard Volume 684: debated on Tuesday 17 November 2020. https://hansard.parliament.uk/Commons/2020-11-17/debates/1E09EF7B-5173-4DA3-AE06-64602176D3CC/details#contribution-8FFE9270-8A84-45C5-801E-D270166CF686

Senadheera, V., Warren, M., Leitch, S., & Pye, G. (2017). Facebook Content Analysis: A Study into Australian Banks' Social Media Community Engagement. In *Social Media Data Extraction and Content Analysis* (pp. 412-432). IGI Global.

Tatham, S. (2008). Strategic Communication: A Primer. https://www.files.ethz.ch/isn/94411/2008_Dec.pdf

UK Government (2019) Online Harms White Paper. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/973939/Online_Harms_White_Paper_V2.pdf

UK Government (2020) Consultation outcome: Online Harms White Paper. Updated 15 December 2020. https://www.gov.uk/government/consultations/online-harms-white-paper/online-harms-white-paper

UK Home Office (2021) About Us. https://www.gov.uk/government/organisations/home-office/about

US Congress (1982) Soviet Active Measures: Hearings Before the Permanent Select Committee on Intelligence, House of Representatives, Ninety-seventh Congress, Second Session, July 13 & 14, 1982. U.S. Government Printing Office, Washington

Vel, Prakash, Salih,A. Brobbey, C. A & Jaheer. H (2014). Emerging trends influencing marketing. 12th International Science Conference, WASET, Belgium, pp. 1-4.

Wardle, C. & Derakhshan, H. (2017). "Information Disorder: Towards an Interdisciplinary Framework for Research and Policy-Making", Council of Europe. ttps://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-forresearc/168076277c.

Xifra, J., & Grau, F. Nanoblogging PR: The discourse on public relations in Twitter. Public Relations Review (2010), doi:10.1016/j.pubrev.2010.02.005