Dealing with misinformation in the digital age

Prevention and Intervention for Aotearoa New Zealand.

March 2021

A Report by:

Stuart McNaughton (Chief Science Advisor, Ministry of Education), John Roche (Chief Science Advisor, Ministry for Primary Industries), Hema Sridhar (Chief Science Advisor, Ministry for Defence), and Jessica Berentson-Shaw (The Workshop).

With input from **The Forum of Chief Science Advisors/ He Rauhinga Tohu Putaiao**

Dealing with misinformation in the digital age: Prevention and Intervention for Aotearoa New Zealand.

Overview

Access to social-media and other digital platforms increases the risks of exposure to and use of misinformation and threats to an informed resilient citizenry, through various distortions of information, untruths, stereotypes, conspiracy theories, harassment, bullying and prejudice. While these platforms pose threats, they are also a means to support an informed inclusive society. Catalytic events such as the Christchurch terror attack and Covid-19 pandemic focus attention on how we can plan for and build resilience in the face of these threats.

In this paper, we review the evidence for how to reduce the risks and increase resilience at individual, community and national levels (capability to adapt positively and resourcefully to changing contexts and disturbances).[1] The evidence contributes to us knowing, generally, how we can be effective, as well as how best to promote the understanding, valuing and practices important to a cohesive and healthy society.

Smart social-media platforms and shared networks are now major sources of information for all ages. Children, adolescents and adults are increasingly exposed to the risks of misinformation and manipulation and misuse of information that can be amplified through these sources. Several implications follow from this. A longer-term preventive strategy is needed that develops the requisite skills and provides for well-being. This strategy needs to recognise that children's learning is enhanced by connections between schools and families / whānau, and, given children comprise the next generations of parents, the preventive approach needs to have a multigenerational perspective.[2] Shorter-term interventions are needed to counteract sudden threats.

We propose a combined intervention and preventive national response to resilience, comprised of: (1) legislation, which limits the spread of misuse of information (e.g., legal frameworks for the use of synthetic media); (2) regulation of online platforms (e.g., increase filters on social media platforms to limit dissemination of conspiracy theories); and (3) socialisation and education of citizenship skills (e.g., critical literacy for all students within the curriculum).

In addition, we outline the role of the scientific community. This has two parts: reducing misinformation, but also modelling reasonable public discourse about science; the latter will increase trust in science and promote understanding of how scientific knowledge can be probabilistic, incomplete and dynamic.

General Features

Misuse of information takes different forms; one analysis distinguishes between misinformation (use of false information that people didn't create, but without the intention to hurt others); disinformation (false information created with the intention of harming a person, group, or

[1] The Treasury Living Standards and well-being frameworks refer to resilience generally as: capability to adapt positively and resourcefully to changing contexts and disturbances (e.g., https://www.treasury.govt.nz/publications/dp/dp-18-05-html)

[2] (see www.who.int/news-room/detail/28-08-2019-who-director-general-statement-on-the-role-of-social-media-platforms-in-health-information)

organization, or even a country), and mal-information (true information used with ill intent).[3] The forms range from information presented as accurate, but which is unreliable, misinformed or wrong, through to computational propaganda, where algorithms, automation, and human curation purposefully distribute misleading information over social-media networks and communities.[4] Self-perpetuating and reinforcing systems of knowledge can be created, whereby misinformation, inaccuracies or untruths are taken as truths through repetition and active dissemination within an online community; these have been called 'filter bubbles', 'homophilous online networks' or 'echo chambers'.

Intentional misinformation has other properties. For example, it has been found to be diffused significantly faster, farther, deeper and more broadly than accurate information.[5] Conspiracy theories are a form that grows rapidly. These are:

> *implausible, unwarranted claims that important social events are caused by malevolent clandestine groups, that usually run in contradiction to the explanations offered by the relevant epistemic authorities and that are embedded in a more general worldview.*[6]

Conspiracy theories range from those that might be relatively benign, such as aliens being hidden in a research facility, through to those that are dangerous to population health (e.g., governments hiding the 'fact' that vaccination causes autism), to those that are extremely malignant, such as the QAnon conspiracy theory.[7]

Understanding risks and resilience

In this paper, we identify ways to support individuals, whānau and family to be resilient. Human characteristics are the basis for our susceptibility and resilience. But these are in part determined by the communities in which we participate, the tools with which we engage, and the policy and system levers that provide an overarching set of conditions.

The role of personal characteristics

We are all susceptible to misuse of information due to common human traits and motives.

General

Susceptibility to misinformation and engaging in conspiracy theorising are associated with general psychological traits, motives and processes. Three broad psychological needs have been proposed that underline susceptibility; the need to understand the world, to feel safe, and to belong

---

[3] Max Soar, Victoria Louise Smith, M.R.X. Dentith, Daniel Barnett, Kate Hannah, Giulio Valentino Dalla Riva, Andrew Sporle (2020). Evaluating the infodemic: assessing the prevalence and nature of COVID-19 unreliable and untrustworthy information in Aotearoa New Zealand's social media, January-August 2020 6 September 2020 (unpublished); Wardle, C. & Derakhshan, H. (2018). https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

[4] Woolley, S. C. & Philip N. Howard, P. N. (2017). Computational Propaganda Worldwide: Executive Summary. In Samuel Woolley and Philip N. Howard, Eds. *Working Paper 2017.11. Oxford, UK: Project on Computational Propaganda*. comprop.oii.ox.ac.uk. 14 pp. http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Casestudies-ExecutiveSummary.pdf

[5] Mair, D., Smillie, L., La Placa, G., Schwendiger, F., Rakkovaska, M., Pasztor, Z., van Bavel, R.. (2019). Understanding our political nature: How to put knowledge and reason at the heart of political decision-making. *EUR 29783 EN, Publications Office of the European Union, Luxembourg*, doi:10.2760/374191,JRC117161; Supovitz, J., Daly, A.J., del Fresno, M., & Kolouch, C. (2017). *#commoncore Project.* Retrieved from http://www.hashtagcommoncore.com; Kahne, J. & Bowyer, B. (2018). 'Educating for Democracy in a Partisan Age: Confronting the Challenges of Motivated Reasoning and Misinformation'. *American Educational Research Journal*, February 2017, Vol. 54, No. 1, pp. 3–34 DOI: 10.3102/0002831216679817.

[6] Stojanov A, Bering JM, Halberstadt J (2020) Does Perceived Lack of Control Lead to Conspiracy Theory Beliefs? Findings from an online MTurk sample. PLoS ONE 15(8): e0237771. https://doi.org/10.1371/journal.pone.023777

[7] QAnon is a wide-ranging conspiracy theory that claims that an elite group of child-trafficking paedophiles have been ruling the world for a number of decades and that President Trump has a secret plan in place to bring this group to justice. The QAnon theory now connects anti-vaccine, anti-5G conspiracies, antisemitic and anti-migrant tropes, Aoife Gallagher, Jacob Davey and Mackenzie Hart (2020). The Genesis of a Conspiracy Theory: Key trends in QAnon activity since 2017. ISD London Washington DC Beirut Toronto Registered charity number: 1141069 © ISD, 2020.

and feel good about oneself and social group.[8] In addition, we are, to varying degrees, all predisposed to: making associations between events and identifying patterns; confirmation bias (seeking out evidence that aligns with our pre-existing views); disconfirmation bias (actively dismissing or finding counterarguments for contradictory information); evaluating arguments like our own as stronger and more accurate than counter arguments; making meaning through narratives; and belonging to an in- group or community.[9]

Other personal characteristics such as levels of stress, self-esteem, anxiety or uncertainty are covariates of conspiracy theory adoption, but more evidence is needed to establish causal links and how these change over time. Increases in these personal characteristics could equally be caused by the experience of believing in a conspiracy theory.[10]

Emotion influences reasoning, both negatively, as in motivated reasoning, and positively. The latter occurs through levels of empathy, which help social functioning and cooperation with others. Levels of empathy are typically higher towards those similar to oneself, and members of the same group.

Personal control

A major hypothesis has been that adopting a conspiracy theory results from threats to personal control, which leads to searching for powerful explanations; but there is little support for this as a general explanation.[11] However, threats to personal control may be important under certain conditions; for example, extreme environmental or health threats. Conspiracy beliefs have also been linked to feelings of powerlessness, anxiety, isolation and alienation, especially under the conditions created by a pandemic.[12]

Traumatic events are often followed by an increase in conspiracy theorising. The initial spike of the uptake of the QAnon conspiracy in late 2017 and early 2018 was followed by a marked spike in 2020 linked to Covid-19. This may reflect people spending more time on social media, a coordinated push to amplify the QAnon theory, or the effects of the threat inducing condition. While initially almost entirely a US phenomenon, growth in other countries has occurred, with Australia now fourth in driving online QAnon discussions (for example tweets increased from 105,545 per six months prior to 2018, to 191,210 in the last six months to June 2020).[13]

Trust

Trust in institutions and authorities, including science, has been identified as a condition for being less susceptible to misinformation and conspiracy theories.[14] Currently, in England there is notable acceptance of conspiracy beliefs about coronavirus in the general population. The conspiracy beliefs are related to widespread mistrust, including mistrust in scientists, doctors and the WHO,

[8] Cichoka, A. (2020). To counter conspiracy theories, boost well-being. *Nature* 587, 177 (2020) doi: https://doi.org/10.1038/d41586-020-03130-6

[9] Kahneman, D. (2011). *Thinking fast and slow*. London: Penguin Books; Mair, D., Smillie, L., La Placa, G., Schwendiger, F., Rakkovaska, M., Pasztor, Z., van Bavel, R.. (2019).op. cit.

[10] Mair, D., Smillie, L., La Placa, G., Schwendiger, F., Rakkovaska, M., Pasztor, Z., van Bavel, R.. (2019).op.cit. ; Stojanov A, Bering JM, Halberstadt J (2020) op.cit.; Stojanov A, Bering JM, Halberstadt J (2020) op.cit.

[11] Stojanov A, Bering JM, Halberstadt J (2020) op.cit.

[12] Cichoka, A. (2020).op. cit.

[13] Aoife Gallagher, Jacob Davey and Mackenzie Hart (2020). Op.cit.

[14] Mair, D., Smillie, L., La Placa, G., Schwendiger, F., Rakkovaska, M., Pasztor, Z., van Bavel, R.. (2019).op.cit

and are associated with less compliance with government guidelines and unwillingness to take up treatment such as vaccination.[15]

Knowledge

Knowledge and access to accurate information are necessary conditions for resilience, being able to identify and respond effectively to misinformation. But are not sufficient for that resilience. For example, knowledge of the science supporting vaccinations is necessary, but is not sufficient to guarantee well-informed decision-making and appropriate action. There is some evidence that more knowledgeable individuals are even more prone to confirmation bias and motivated reasoning, and, under some conditions, confronting misinformation with accurate information can have a 'backfire' effect, leading to greater commitment to the misinformation.[16]

Critical literacy

Critical literacy ('media literacy', 'epistemic vigilance') skills are increasingly recognised as important in resilience. These are not just a predisposition to questioning things or the status quo,[17] but include cognitive skills of 'criticality' (e.g., critical evaluation, reasoned judgements, identifying accuracy and credibility of information). In one study, 15-27 year olds' judgements of the accuracy of online posts about controversial political issues depended on the alignment of the claim with their prior policy position and to a lesser extent on whether the post included an inaccurate statement. However, although level of political knowledge did not improve judgments of accuracy, levels of media literacy education did.[18] A recent test showed that a 'behavioural nudge' to adults to consider the accuracy of information resulted in increased discernment of truthfulness and intention to share misinformation.[19]

Education

Educational levels or employment status also are not sufficient to provide resilience. A study of 13 to 23 year olds' changing beliefs from creationism to evolution showed ongoing commitment to and involvement with a religious group was far more important than educational attainment in predicting changed beliefs.[20] The social network of co-religionists outweighed educational level, and graduating from biology classes with increased knowledge did not outweigh the moderating effect of the community.

Teachers are as susceptible as anyone else to motivated reasoning and confirmation bias. A recent study of US social studies teachers showed that their ratings of the credibility of different news sources mirrored their ideological (political) beliefs. However, those teachers that understood

---

[15] Freeman D et al (2020). Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. *Psychological Medicine* 1–13. https:// doi.org/10.1017/S0033291720001890

[16] Kahne, J. & Bowyer, B. (2018). 'Educating for Democracy in a Partisan Age: Confronting the Challenges of Motivated Reasoning and Misinformation'. *American Educational Research Journal*, February 2017, Vol. 54, No. 1, pp. 3–34 DOI: 10.3102/0002831216679817. Mair, D., Smillie, L., La Placa, G., Schwendinger, F., Raykovska, M., Pasztor Z., van Bavel R. (2019). Understanding our political nature: how to put knowledge and reason at eth heart of political decision-making. EUR 29783 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-08621-5, doi:102760/374191,JRC117161

[17] Kahne, J. & Bowyer, B. (2018). 'Educating for Democracy in a Partisan Age: Confronting the Challenges of Motivated Reasoning and Misinformation'. *American Educational Research Journal*, February 2017, Vol. 54, No. 1, pp. 3–34 DOI: 10.3102/0002831216679817. Mair, D., Smillie, L., La Placa, G., Schwendinger, F., Raykovska, M., Pasztor Z., van Bavel R. (2019). Understanding our political nature: how to put knowledge and reason at eth heart of political decision-making. EUR 29783 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-08621-5, doi:102760/374191,JRC117161

[18] Kahne, J. & Bowyer, B. (2018). Op. cit.

[19] Gordon Pennycook, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. (2020). Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. Psychological Science, 2020, Vol. 31(7) 770–780

[20] Hill, J. P. (2014). Rejecting Evolution: The Role of Religion, Education, and Social Networks *Journal for the Scientific Study of Religion (2014) 53(3):575–594*

'credibility' in terms of higher level journalistic processes such as fact checking and verification (skills related to critical literacy), were less likely to align their judgment to their own political position.[21]

In the final section, we also note that individual scientists are as susceptible as anyone else (e.g., Issue Advocates). The methods of science, in a sense, are designed to mitigate these human traits. But there is plenty of evidence of scientists' failures to resist these risks.

The role of communities

Communities have a major negative role in creating, sharing and disseminating misinformation and conspiracies impacting on the susceptibility of individuals. But, they also have a positive one in developing and sustaining resilience. Social networks are not a new phenomenon; in essence, they reflect the predisposition to belong to a community or networks.[22] What is new, however, is the development of online forms, in the context of social media platforms. This means the personal traits are more immediately, broadly and extensively able to be engaged, and homophilous networks can be created that can rapidly escalate and be relatively impervious to change. The nature and composition of communities of course varies, but there are identifiable features of communities, which contribute to each role.

Community practices

One way of describing the common features is through the idea of 'communities of practice' that exist at various levels of complexity and breadth.[23] In these, values, norms and practices are collectively developed and upheld. In many respects the national response to Covid-19 in Aotearoa New Zealand was like a large positive community of practice, which came to have shared norms and values (e.g., looking after one another, being a team, kindness), shared ways of communicating (e.g., daily briefings spread widely through various media), and practices supported and sanctioned through the community (e.g., social distancing, use of tracing technology).

As already noted, the role of a community of practice in supporting beliefs and practices can be greater than the background characteristics of an individual, such as educational level or level of knowledge.

Online communities

Online communities are particularly powerful. They have changed our communication patterns and this means our traditional forms of responding and addressing these issues needs to change significantly too. In addition to the features already noted for engaging people, they have the potential risk of favouring the rapid sharing of disinformation compared with accurate information. This is because sharing false news is favoured, as a reaction to novelty and the elicitation of strong emotions.[24]

Reasoning within communities

Some communities of practice can be particularly powerful vehicles for the resilience skills. One form occurs when the practices, norms and values entail group reasoning, sometimes called

[21] Clark, C. H., Schmeichel M. & Garrett, H. J. (2020). *Educational Researcher*, Vol. 49 No. 4, pp. 262– 272 DOI: 10.3102/0013189X20909823

[22] Ferguson, N. (2017). *The Square and the Tower: Networks, hierachies and the struggle for global power*. Penguin Random House UK.

[23] A functioning group or network in which there are shared practices, norms, values and beliefs. Different disciplines use different concepts to describe their development and their properties (eg Wenger, E. (1998). *Communities of Practice: Learning, meaning, and identity.* Cambridge: Cambridge University Press; Ferguson, N. (2017). Op. cit.

[24] Mair, D., Smillie, L., La Placa, G., Schwendinger, F., Raykovska, M., Pasztor Z., van Bavel R. (2019). op. cit.

collaborative reasoning or argumentation.[25] Components include cognitive skills, such as producing a claim supported by evidence, critical reflection, evaluation to inform judgements of what to believe or do, and informed rebuttal. Importantly, argumentation also requires perspective taking, bringing into play cognitive and emotional empathy. These are the abilities to understand the views and mental states of others and the capacity to internally simulate and experience the emotions of others.

Experimental demonstrations show face-to-face communities of practice can be built in classrooms, which develop the criticality and reasoning skills, including those that are interpersonal (e.g., sociability, perspective taking), and those that are intrapersonal (e.g., self-regulation), necessary to engage effectively in the practices of the community. There is, as yet, limited evidence for how best to do this in online communities.[26]

Iwi and hapu communities

A strength in Aotearoa New Zealand is the presence of Māori communities at iwi and hapu levels. The strength of these communities, together with their rights and the state's responsibilities under Te Tiriti o Waitangi, mean that partnerships with these communities must underpin a national strategy.

The role of digital tools

Digital tools and platforms have affordances for (increase the likelihood of) certain behaviours. Adults are worse at discerning between true and false content when deciding what to share on social media relative to when asked directly about accuracy, because the social media context focuses attention on factors other than accuracy (e.g., alignment with the community's views); and this distracts from considering accuracy when deciding whether to share news.[27] Mis-, dis-, and mal-information are diffused significantly faster, farther, deeper, and more broadly than accurate information on social media. An online 'disinhibition effect' has been proposed, whereby users of the internet and social media tend to be less inhibited and have reduced capacity to judge the appropriateness of their own behaviour.[28]

However, social media can enhance access to valuable needed information. Supportive networks can provide positive effects for those with health needs and foster social inclusion and community membership for marginalised or excluded groups, such as LGBTI youth.[29] The significance of not being connected is illustrated by the relationship between usage and mental and physical health outcomes for adolescents, which is negative at the extremes of low/no Internet usage and heavy usage (>2 hours/day).[30]

---

[25] Mair, D., Smillie, L., La Placa, G., Schwendinger, F., Raykovska, M., Pasztor Z., van Bavel R. (2019). op. cit.

[26] McNaughton, S, Zhu, T, Rosedale, N., Oldehaver, J, Jesson, R., & Greenleaf, C. (2019). Critical perspective taking: Promoting and assessing online written argumentation for dialogic focus. *Studia Paedagogica* (special issue: Better Learning through Argumentation). 24(4). 119-141.; Rosedale, N., McNaughton, S., Jesson, R, Zhu, T. & Oldehaver, J. (2019) Online written argumentation: Internal dialogic features and classroom instruction. Chapter 15 (pp. 263-278). in Emmanuel Manalo, Yuri Uesaka, Ouhao Chen & Hiroaki Ayabe. (Eds.). (2019). *What It Looks Like: Developing Diagram Use Competencies and Predispositions to Support Problem Solving, Communication, and Thinking*. New York: Routledge

[27] Pennycook, McPhetres, Zhang, Lu and Rand. (2020).op. cit.

[28] Aiken, M. (2016). *The Cyber Effect.* London: John Murray.

[29] Reid Chassiakos Y, Radesky J, Christakis D, et al., AAP COUNCIL ON COMMUNICATIONS AND MEDIA. Children and Adolescents and Digital Media. *Pediatrics.* 2016;138(5): e20162593

[30] Bélanger,R. E., Akre, C., Berchtold, A. & Pierre-André Michaud, P-A. (2011). A U-Shaped Association Between Intensity of Internet Use and Adolescent Health www.pediatrics.org/cgi/doi/10.1542/peds.2010-1235 doi:10.1542/peds.2010-1235.

Three core problems with existing platforms have been identified that threaten democratic processes in Aotearoa New Zealand:[31]

- *Platform monopolies* (a handful of people determine social interactions and access for millions of global citizens);
- *Algorithmic Opacity* (algorithms increasingly influence what is heard and seen without transparency and accountability); and
- *Attention economy* (priority is given on social media to content that grabs attention).

Societal conditions

Societal conditions provide enablers and constraints for individual's resilience and for communities that are supportive of resilience. The degree of trust, of feeling safe, of informed understanding about the world and of feeling good about oneself and one's community are reliant on these conditions. Crisis events like a pandemic exaggerate existing vulnerabilities.

Prevention

We know much more about the nature of mis- dis- and mal information and conspiracy theories and how they are spread than how to counter them; the priority is being able to prevent rather than rely on counter acting.[32]

Jurisdictions are taking various measures to building resilience in the face of the threats of information misuse:

- Educating for and socialisng the skills and practices needed;
- Regulated to manage the threats; and
- Legislation to stop threats.

Local commentators have also identified these three, noting the benefits and risks.[33] The different measures represent a range of longer-term prevention strategies and more immediate interventions.

Building resilience, for example in the face of a pandemic, requires attending to the needs of individuals, including their feelings of grief, uncertainty, powerlessness and marginalization, which can be exaggerated for those who have suffered through negative impacts on health, employment and education.[34]

Longer term prevention

A longer term approach requires developing robust supportive communities of practice, as well as the development of requisite skills and knowledge for individuals.

There is a strong evidence base for how to develop supportive communities for students at the classroom and school levels. There is a developing science of how best to teach the critical literacy skills. Coupled together, these would be part of a longer term preventive approach. However, there are few analyses of how to build communities at a more macro level, apart from descriptions of jurisdictions such as Finland (see below) that can be used as case studies. In essence, the

[31] Marianne Elliot and Jess Berentson-Shaw (2019). Digital threats to democracy. The Workshop. https://www.theworkshop.org.nz/publications/digital-threats-to-democracy-report-2019

[32] Cichoka, A. (2020). op. cit.

[33] Alexander Gillespie (2020). With the election campaign underway, can the law protect voters from fake news and conspiracy theories? *The Conversation*. September 15, 2020 11.48am

[34] Cichoka, A. (2020). op. cit.

response to the Covid-19 pandemic taken in Aotearoa New Zealand illustrates the building of a community at a societal level.

Building and maintaining high levels of trust in public institutions and supporting well-being are central to a preventive approach. The scientific and quality information communities have a critical role to play, through the provision of public advice, informed discussion and information transmission.

Resilience and educational communities

Critical literacy is a focus in many educational systems and is a core part of international assessments.[35] Despite this, students' ability to reason effectively, navigate sources of information confidently, and engage critically with the systems of knowledge in content areas, are not taught well, across schooling and at college level.[36]

Effective approaches, need to guarantee transfer from classroom to everyday use with information sources. Currently, the consensus is that a 'mixed' approach involving explicit instruction in thinking critically embedded in major subjects, coupled with some dedicated course work focused on the critical literacy (media literacy) skills, such as in civics or philosophy courses, likely work best for transfer.[37] In addition to embedded and deliberate forms of instruction, communities of practice at a classroom and school, which feature caring teacher and student relationships, clear norms, expectations and practices, create effective conditions for the acquisition of these skills.[38]

In Aotearoa New Zealand, 15 year olds perform significantly higher than the OECD average in reading literacy generally, but almost one fifth are below the OECD basic proficiency level.[39] In addition, less focus on subject complexity and on higher order skills is evident for Māori and Pasifika students and students from low-decile schools.[40]

Less is known about critical literacy in science, maths and history. The curriculum refers to developing critical and creative thinkers, and, in each learning area, there is reference to students being critical, but there is no elaboration in terms of the critical literacy skills described above.

Changes to the curriculum such as those made by Finland are needed. A mandatory focus across learning areas on critical literacy, which is supported by resources, teacher training and development, similar to the changes being made for Aotearoa New Zealand Histories, is recommended.

---

[35] www.oecd.org/pisa/test

[36] McGrew, S., Ortega, T., Breakstone, J. & Wineburg. S. (2017). The Challenge that's bigger than fake news. Civic reasoning in a social reasoning environment. *American Educator* Fall 2017. 4-9; Fraillon, J., Ainley, J., Shulz, W., Friedman, T. & Duckworth, D. (2018). Preparing for life in a digital world: IEA International Computer and Information Literacy Study 2018 International Report.

[37] Shane Horn & Koen Veermans (2019). Critical thinking efficacy and transfer skills defend against 'fake news' at an international school in Finland. *Journal of Research in International Education*. 2019, Vol. 18(1) 23–41.

[38] Abrami, P.C., Bernard, R.M., Borokhovski, E., Waddington, D. I., Wadwe, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta–analysis. *Review of Educational Research*, *85*(2), 275-314. https://doi.org/10.3102/0034654314551063. Reznitskaya, A., Kuo, L. J., Clark, A. M., Miller, B., Jadallah, M., Anderson, R. C., & Nguyen- Jahiel, K. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education*, *39*(1), 29–48.

[39] educationcounts.govt.nz/publications/series/PISA/pisa-2018

[40] Wilson, W., Madjar, I & McNaughton, S.(2016): Opportunity to learn about disciplinary literacy in senior secondary English classrooms in New Zealand, *The Curriculum Journal*, DOI: 10.1080/09585176.2015.1134339; Jesson, R., McNaughton, S., Rosedale, N., Zhu, T. & Cockle, V. (2018). A mixed-methods study to identify effective practices in the teaching of writing in a digital learning environment in low income schools. *Computers and Education*, 119 (April), 14-30; McNaughton, S, Zhu, T, Rosedale, N., Oldehaver, J, Jesson, R., & Greenleaf, C. (2019). Critical perspective taking: Promoting and assessing online written argumentation for dialogic focus. *Studia Paedagogica* (special issue: Better Learning through Argumentation). 24(4). 119-141.

Resilience with iwi and hapu

Many Māori communities responded to the Covid-19 'lockdowns' with innovation and adaptation.[41] These strengths of community resilience should be built on when considering how they can contribute to a national strategy.

Evidence from one iwi's experience with digital conditions during Covid-19 identified conditions for building their resilience further; these are applicable to a prevention strategy.[42] The most obvious is identifying and building on the capacity of iwi, to promote and enhance digital inclusiveness. Effective partnerships are needed between government agencies and iwi organisations to ensure effective digital inclusion, and mutual design of the strategies to combat the risks of mis- dis- and mal-information. It is also obvious that for rangatahi who are developing their resilience, universal provision of suitable devices and robust connectivity is needed to ensure access to the community online.

Tools

The report on threats to democracy suggests that prevention could include designing and promoting new competitive digital media platforms that have different settings from the existing ones.[43] These platforms would not be driven by monopolies. They would have affordances for more reflective 'dialogue, discussion and debate'.

New designs have been reported to reduce the propensity to limit online interactions in relatively closed communities with similar minded participants and to encourage collective agreement. Collectively owned digital platforms can be designed with affordances for more considered, collaborative dialogue and discussion. The designs can increase civility and promote collective agreement. They have also been shown to create safe environments for marginalised groups. Appropriate designs that afford citizen participation and deliberation can increase trust in public institutions.[44]

Societal systems

In 2014, Finland launched an anti-fake news campaign that has the hallmarks of a community of practice at a societal level aimed at developing the criticality needed by citizens. It is comprehensive and meant to prepare all citizens for threats in the complex digital world, with a particular focus on being able to identify and counter false information designed to undermine the country's politics. The President called on the country to take responsibility for upskilling and, in 2016, a revision to the critical thinking curriculum foregrounded identifying misinformation.

It is as yet not known how successful the Finnish approach has been. As with all prevention and intervention efforts, a research and development approach is needed where strategies are accompanied by evaluation. This is a problem that likely will take many years to fix and it will be critical to know whether strategies are making a difference at a societal level.

---

[41] Ngā Pae o te Māramatanga Comms comms@maramatanga.ac.nz; Rangimarie Hunia, Shazeaa Salim, Stuart McNaughton, Rochelle Menzies, Peter Gluckman and Anne Bardsley (2020). Addressing Rangatahi Education: Challenges after CVovid-19. A partnership report by Ngāti Whātua Ōrākei and Koi Tū: The Centre for Informed Futures.

[42] Rangimarie Hunia, Shazeaa Salim, Stuart McNaughton, Rochelle Menzies, Peter Gluckman and Anne Bardsley (2020).

[43] Marianne Elliot and Jess Berentson-Shaw (2019). Digital threats to democracy. The Workshop. https://www.theworkshop.org.nz/publications/digital-threats-to-democracy-report-2019

[44] Marianne Elliot and Jess Berentson-Shaw (2019). Digital threats to democracy. The Workshop. https://www.theworkshop.org.nz/publications/digital-threats-to-democracy-report-2019

In Aotearoa New Zealand, the Law Commission (2012) proposed the creation of a new criminal offence tailored for digital communication, as well as amendments to various acts to be better suited to digital communications, and new legal requirements for all schools to help combat cyber-bullying.[45] Further proposals for regulating and legislating are summarised below.

Intervention with immediate threats

Building positive well-informed communities and teaching are part of a longer term preventive strategy. But can this approach be activated in a way that deals with immediate threats, such as those posed by the need for universal vaccination against Covid-19?

As noted earlier, communities that curate, spread and amplify misuses of information can form and grow rapidly. The frequency of comments on social media from unreliable sources from before the lockdown period to August 2020 has been relatively constant, but their nature has changed, reflecting the influence of US based disinformation and conspiracy theories, with a noticeable spike in the reference to conspiracies in August.[46] Over a thousand attended a rally in Auckland on September 12, at which speakers promoted this content (e.g., relating to vaccination, QAnon, 5G technology, government tyranny, individual freedom).[47]

Monopoly platforms can contribute as illustrated by Google's response to COVID-19. Their actions to fight disinformation included YouTube developing a specific COVID-19 Medical Misinformation Policy which doesn't allow content that spreads medical misinformation that contradicts the World Health Organization (WHO) or local health authorities' medical information about COVID-19. For example, YouTube does not allow:

- Denial that COVID-19 exists
- Claims that people have not died from COVID-19
- Claims that COVID-19 is caused by radiation from 5G networks
- Videos alleging that the COVID-19 test is a cause of the virus
- Claims that the flu pandemic is a hoax or was manufactured for the purposes of selling vaccines
- Claims that the COVID-19 vaccines will kill people who receive it.

This approach will not solve the pervasive persistent type of disinformation where for example people are just 'asking questions' or posting hundreds of posts a day by bots that go right up to the boundary line of the policy of what's allowed.[48]

Regulating and legislating in Aotearoa New Zealand

Regulating and legislating for mis-, mal- and disinformation are difficult propositions and some jurisdictions have recommended not doing this.[49] Issues to do with free speech are proving difficult to resolve, which has meant developing shared frameworks for action rather than additional regulation and legislation beyond current safeguards. This focus relies on a longer-term response aimed at building societal resilience and the need to increase the transparency of online information sources, promote critical literacy (media and information literacy), and developing tools

[45] Law Commision Te Aka Matua O Te Ture (2012). *HARMFUL DIGITAL COMMUNICATIONS: The adequacy of the current sanctions and remedies*. Wellington, New Zealand |August 2012 Ministerial Briefing Paper.

[46] Max Soar, Victoria Louise Smith, M.R.X. Dentith, Daniel Barnett, Kate Hannah, Giulio Valentino Dalla Riva, Andrew Sporle (2020). Op. cit.

[47] Alexander Gillespie (2020). With the election campaign underway, can the law protect voters from fake news and conspiracy theories? *The Conversation*. September 15, 2020 11.48am

[48] Wardle & Derakhshan, (2018). op. cit.

[49] European Commission (2018). A multi-dimensional approach to disinformation: Report of the independent High Level Expert Group on Fake News and Online Disinformation. https://op.europa.eu/en/publication-detail/-/publication/6ef4df8b-4cea-11e8-be1d-01aa75ed71a1

that increase the agency and resilience of users. It requires filling data deficits and using 'explainers', relying on the quality information providers (media etc) to provide 'quality information' and challenging the incentive structures (e.g., what gets most clicks is what gets published).

The following sections on Regulation and Legislation draw on two recent reports funded in part by the New Zealand Law Foundation - Te Manatū a Ture o Aotearoa - through its Information Law and Policy Project.[50] These are on digital threats to democracy, and on legal and regulatory issues posed by synthetic media technologies (emerging audio visual technologies such as those producing 'deepfakes'). While they have more specific foci, the proposals are potentially generalizable to the broader concerns of this paper.

The former proposes solutions to three core problems in the threats to democracy noted earlier: Platform monopolies; Algorithmic Opacity and Attention economy. The latter proposes a legal framework as a solution to the issues associated with synthetic media. Both identify the fundamental difficulty in legal approaches to balancing the benefits and the risks to society.

The evidence for benefits to citizens is clear. They include enabling those who have been marginalised from full participation in the democratic process to have greater access to democratic processes and making democratic processes more inclusive, transparent and trustworthy. The risks to inclusive democracy are 'the increasing power of private platforms, foreign government interference in democratic processes, surveillance and data protection issues, fake news, misinformation and disinformation, filter bubbles and echo chambers, hate speech and trolling, and distrust/dissatisfaction with democracy and democratic processes.'

Human rights' principles should be applied to policy development in this area, and are particularly useful where there is an absence of specific research evidence. In addition, principles derived from Te Tiriti o Waitangi should be the basis for considering benefits and issues and a primary consideration in regulatory and legislative approaches.

While there may be agreement that regulation is needed, there is little evidence to show how regulation works or, even, whether it does work. Existing government regulations on moderating private intermediaries' practices have not been empirically tested for their efficacy or effectiveness.

An 'adaptive approach' to policy and regulation is proposed which develops and tests 'fit for purpose' regulations. Proposals include new anti-cybersecurity infrastructure and regulating companies' information management practices. Some regulatory measures, like the Singaporean Data Protection Act 2012, would enable formal charges to be laid in situations where there is information mismanagement and abuse. Sites that do not allow anonymisation and require pre-registration have been shown to solicit qualitatively better, but fewer, user comments because of the extra effort required for engaging in discussion and the lack of anonymity. There is evidence that abusive comments are minimised when anonymous commenting is prohibited.

Improvements in and regulation for content moderation are needed, but it is still the case that a combination of automated classification and deletion systems and human input is most effective.

A note on the role of public science discourse

[50] Marianne Elliot and Jess Berentson-Shaw (2019). Digital threats to democracy. The Workshop. https://www.theworkshop.org.nz/publications/digital-threats-to-democracy-report-2019 Curtis Barnes Tom Barraclough (2019) P E R C E P T I O N INCEPTION Preparing for deepfakes and the synthetic media of tomorrow. Brainbox. https://www.brainbox.institute/#report-section

There are lessons to be learned from the Covid-19 experience about appropriate and useful public science debate. Contributing to the public discourse is very important, not least because it can demonstrate features of science as applied to the concerns of the day. It also influences trust.

There can be tensions in public disagreements or changes in the science-based recommendations based as the scientific understanding changes, as it has done during the pandemic. On the one hand, there is low tolerance for alternative views by both the general public and many in the scientific community. A second issue relates to the responsibility of those commenting on and communicating the science accurately as well effectively. There is the tension between conclusions from someone who is an effective communicator of science, but not necessarily an acknowledged expert, and those more qualified, especially if there are differences in their conclusions drawn from the evidence.

Members of the scientific community have been very visible during the pandemic. An analysis by MBIE shows that the responses by scientists demonstrated agility, capability to pivot to new areas, and even work in new disciplines to answer pressing questions.[51] This agility is impressive but has the tensions noted above. Publication rates, mostly through open sources have been high. By using preprints and publishing under open access, researchers were able to share their work prior to formal peer review, and make it generally freely available to the public. We have yet to determine how this may have affected both the trustworthiness of and the trust in the scientific advice. But the very public display of that advice demonstrating the dynamic, contestable and probabilistic nature of the evidence, and its contribution to the generally high levels of public agreement and acting in accordance with that evidence, suggests high levels of trust. But, in some situations, potentially mis-placed trust.

[51] Evidence and Insights branch at the Ministry of Business, Innovation and Employment. New Zealand's COVID-19 research response. MBIE 17 July 2020: E&IScienceandInnovationRequests@mbie.govt.nz

Principles for prevention and intervention[52]:

1. Prevention is better than intervention.

2. Understand varieties of worldviews and values to frame issues through, both weak (eg fear of climate change / fear of becoming infected) and strong (eg collective responsibility and efficacy for reducing green house emissions / infections).

3. Targeting narratives and messaging to frames and concerns of different groups (eg young people , elderly) though values driven stories.

4. Getting in front of the issue (information / evidence socialised early and before 'event' where possible). Creating a positive information environment for: (eg) vaccination uptake; climate change, culling of tar; open borders and immunity. This means frames and communication strategies that deepen thinking and surface existing ideas that people may hold but may be being supressed through various means.

5. Being transparent and truthful about state of evidence and roles -even with the complexity of vaccination for Covid-19 (new, new technologies, multiple vaccines, different susceptibilities).

6. Devolution to communities, which would include the handing over of decision-making and resources. Tino rangatiratanga expressed in the information environment through communities– iwi and hapu; Pasifika communities (eg Church); schools as communities.

7. Modelling through thought, value, opinion leaders.

8. Mokopunatanga / children as 'leaders' through concern for and valuing of koroua and kuia.

9. Inoculation of misinformation by the media and schools: highlighting strategies and motivations and general practices that people will likely encounter when being exposed to false information.

10. Preparedness for 'debunking' through media.

11. Focus communications on middle group of those unsure, or with less firm views, rather than those extremely resistant / opposed or already convinced.

12. Theory driven experimentation. In a rapidly moving information environment, communicators, scientists armed with the principles should experiment in how they communicate without fear of getting a slap on the hand (or worse) from risk adverse communications people who are more about PR than deepening thinking and conversation.

13. Cross agency approaches for greater reach (eg working with education on 'getting in front of the issue' through resources for use in schools).

---

[52] Some additional references for principles:
Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., Lombardi, D., Newman, E. J., Pennycook, G., Porter, E. Rand, D. G., Rapp, D. N., Reifler, J., Roozenbeek, J., Schmid, P., Seifert, C. M., Sinatra, G. M., Swire-Thompson, B., van der Linden, S., Vraga, E. K., Wood, T. J., Zaragoza, M. S. (2020). The Debunking Handbook 2020. Available at https://sks.to/db2020. DOI:10.17910/b7.1182S.
Edith A. MacDonald , Jovana Balanovic , Eric D. Edwards , Wokje Abrahamse , Bob Frame , Alison Greenaway , Robyn Kannemeyer , Nick Kirk , Fabien Medvecky , Taciano L. Milfont , James C. Russell & Daniel M. Tompkins (2020) Public Opinion Towards Gene Drive as a Pest Control Approach for Biodiversity Conservation and the Association of Underlying Worldviews, Environmental Communication, 14:7, 904-918, DOI: 10.1080/17524032.2019.1702568 https://doi.org/10.1080/17524032.2019.1702568