NEWS & MEDIA RESEARCH CENTRE

UNIVERSITY OF CANBERRA

BUILDING RESILIENCE WITH INFORMATION LITERACY AND INFORMATION HEALTH

News and Media Research Centre Submission to Australian Senate Select Committee on Foreign Interference through Social Media

17 February 2023

BACKGROUND

The University of Canberra's News and Media Research Centre (http://www.canberra.edu.au/nmrc) investigates the evolution of news, media, content and communication. The N&MRC is a national leader in the provision of expert commentary and analysis of social media manipulation in Australian politics.[1]

The Australian National University's Virtual Observatory for the Study of Online Networks (http://vosonlab.net/) is a global leader in computational social science and big data analytics. Since 2005 the Lab has advanced the Social Science of the Internet through research, research tool development, teaching and research training.

SUBMISSION AUTHORS

Associate Professor Mathieu O'Neil, News & Media Research Centre, Faculty of Arts and Design, University of Canberra <mathieu.oneil@canberra.edu.au>

Professor Robert Ackland, School of Sociology/VOSON Lab, Research School of Social Sciences, The Australian National University <robert.ackland@anu.edu.au>

Senior Lecturer Rachel Cunneen, Faculty of Education, University of Canberra <rachel.cunneen@canberra.edu.au>

The submission authors are happy to provide further information to the Inquiry if desired.

---

1 Since 2015, the Centre has published the *Digital News Report: Australia*, a national annual online survey of more than 2,000 adult Australians, which monitors changes in news consumption over time. The Australian survey forms part of a global study of 44 news markets by the Reuters Institute for the Study of Journalism at the University of Oxford. *The Digital News Reports: Australia 2015 – 2022* can be downloaded via canberra.edu.au/nmrc

BUILDING RESILIENCE WITH INFORMATION LITERACY AND INFORMATION HEALTH

SUBMISSION TO SELECT COMMITTEE ON FOREIGN INTERFERENCE THROUGH SOCIAL MEDIA

SUMMARY

To counter foreign interference, the information resilience and skills of the Australian public must be increased in a manner that restores trust in public institutions. We define three key resilience principles – non-partisanship, speed, and transparency. We present an information literacy research program conducted in four Canberra schools in 2022, and outline an ongoing research program to develop tools to map the health of online information environments. These research projects implement our three resilience principles and could inform information literacy and information health campaigns and initiatives that will make the Australian public more resilient.

SUBMISSION CONTENTS

1. **Introduction** 4

2. **Background** 5

3. **Information Resilience** 7

4. **Information literacy in schools using lateral reading** 9

5. **Information health in online environments** 12

6. **Recommendations** 16

1. INTRODUCTION

1.1 Previous work by N&MRC

This submission follows an N&MRC contribution to the *Joint Standing Committee on Electoral Matters: Inquiry into and report on all aspects of the conduct of the 2016 federal Election and matters related thereto* (Park et al. 2018) and a joint N&MRC/VOSON contribution to the *Australian Senate Select Committee on Foreign Interference through Social Media* (Ackland et al. 2020) as well as N&MRC reports on disinformation (Jensen & O'Neil 2018, O'Neil & Jensen 2020, Park et al. 2020) and an internal review of media and information literacy programs for the Department of Communication and the Arts in 2019.

These previous submissions and reports mapped to what extent members of the Australian public perceived they had been exposed to misinformation, or how much they trust the news; they also provided empirical evidence of disinformation campaigns by foreign entities aimed at the Australian public sphere, such as Russian Internet Research Agency (IRA) activity on #auspol.

1.2 Countering foreign influence

In broad strokes, there are two main methods available to governments wishing to counter foreign interference in elections via social media. The first is to develop technologies to track hostile influence campaigns. This approach is problematic on several counts. To begin with, it is not always possible to distinguish hostile influence operations which use automated tools to broadcast a large quantity of messages, from the organic sharing of content on social media by legitimate users (e.g., ad agencies or NGOs who employ content schedulers such as Hootsuite). Even if one overcomes this problem by taking into account the content of the messaging using Natural Language Processing, by the time this is detected, the message has already been disseminated.

The same goes for the detection of inauthentic accounts on social media: there is an 'arms race' whereby producers and detectors of fake accounts are constantly seeking ingenious means to outwit one other. These issues could potentially be progressed if social media platforms agreed to make their data more transparent, but the 'Transparency Reports' published by platforms in compliance with the DIGI voluntary code only provide general, non-specific insights about the entities producing disinformation (e.g. Meta's 2021 report states, p, 17: 'In 2019, we took action against a domestic operation in March 2019 that was linked to local political actors related to the New South Wales state election'), let alone information regarding how many people were exposed to the disinformation these entities produced, for how long, etc.

This submission therefore focuses on the second means of countering foreign interference: developing tools to increase the resilience of Australian citizens to disinformation and to consolidate the foundations of trust in democratic institutions. This accords with a report published in 2022 by NATO's Center for Media Literacy:

> Democracy stands or falls on people. The challenge for democracies is to find ways to preserve the freedoms that come with more access to information, while protecting against the threats that come with it. The most democratic way to address this challenge is teaching [members of] society to be wiser information consumers and producers through critical thinking and a pedagogy that empowers them to evaluate, analyse, and choose critically whether to act on information. Media literacy education facilitates this critical thinking and thereby, risk management. (Jolls 2022 p.50)

NATO defines media literacy as key to deep efforts to build resilience in societies, beginning with its own forces, so that NATO soldiers receive training in media literacy (Singer & Johnson 2021). Developing information and media literacy means citizens truly are 'the first line of defence' (Jolls 2022 p.12) as they need to have the knowledge and skills to strategically use media literacy to counteract misinformation.

2. BACKGROUND

2.1 Aims of foreign interference via social media

Recent revelations about the deceptive activities of an Israeli-based disinformation agency (Kirchgaessner et al. 2023), whilst concerning, should be treated with caution as it is unclear what campaigns were run, where, and for what purpose. This submission accordingly addresses documented interference by state entities, whose aims are:

1. Increasing distrust in institutions and societal divisions in liberal-democratic societies in order to weaken them.
2. Increasing positive framings of the hostile foreign entities' interests.

These aims should not be thought about in isolation: they seek to capitalise on, and worsen, already existing social and political fractures in liberal-democratic societies. These fractures include rising distrust in institutions, rising anti-establishment sentiment, conspiratorial thinking, the rejection of democracy, and hate speech. Measures to counter hostile foreign interference in democratic processes can only be effective if they take this context into account. Initiatives seeking to increase the informational resilience and skills of citizens must incorporate the need to restore trust in our public institutions.

2.2 Decline in trust

Indeed, a key strategic objective of foreign interference is worsening distrust in the institutions of liberal democracy (O'Neil & Jensen 2020). This distrust is already high. The Australian Election Study (1987-2022) found that 30% of respondents believe that 'People in government can be trusted' whereas 70% believe 'People in government look after themselves'. Similarly 54% believed the government is run for a 'Few big interests' and only 12% believe it is for 'All the people' (Cameron & McAllister 2022). Trust in news has also been declining globally, as documented by the annual Reuters Digital News Report (Newman et al. 2022).

When asked whether they agreed with the statement: 'I think you can trust most of the news most of the time', positive responses were: 48% of respondents in Brazil (62% in 2015), 44% of respondents in Japan (46% in 2015), 41% of respondents in Australia (43% in 2016), 29% in France (38% in 2015), and (the lowest figure) 26% in the USA (32% in 2015). Why is this decline in trust of traditional news sources occurring? When people think the mainstream media is not holding industries and governments to account, they view it as a mouthpiece for elite interests, and may be more likely to accept information that challenges conventional beliefs.

2.3 Information literacy and media literacy

The establishment of the NATO Strategic Communication Centre of Excellence in Riga, Latvia, in 2014 attests to the importance that strategic communication can play in governmental defence tactics, and media literacy has been identified by NATO and its allies as essential to defence (Jolls 2022). We define *information literacy* as having to do with the correctness of information items, that is to say as seeking to answer a relatively simple question: 'Is this statement true or false?'

In contrast, *media literacy* has to do with the way media framings and representations operate, so the questions it poses are more numerous, and complex: 'Who created this message?' 'What creative techniques are used to attract my attention?' 'What lifestyles, values and points of view are represented in, or omitted from this message?' 'Why is this message being sent?' (Jolls 2022). Whilst these terms are sometimes combined (e.g., 'media and information literacy', MIL) we have chosen to separate them here, and to focus on information literacy, as outlined in our three resilience principles.

2.4 Failure of the current information literacy educational model

Since the advent of the mass Internet in the late 1990s, we have known that the larger the amount of potentially relevant but weakly authoritative information, the more urgent is the need for effective and cognitively viable information processing skills (Taraborelli 2008). Has the Australian public education system met this challenge? Our contention is that it has not, for two main reasons.

The first reason is institutional: there are wide variations between states and territories, and public and private schools, in terms of how information literacy is taught, or not. Australia's states and territories are responsible for implementing the curriculum: implementations vary greatly (Corser et al. 2022). Notley et al. (2020) surveyed 1069 Australian school students aged 8-12 years old (N:545) and 13-16 years old (N:524). They found that one that only one in five of these children and teenagers (20%) said they had received lessons at school in the past year to help them work out if news stories were true and could be trusted.

The second reason is that when information or media literacy is taught, some of the methods used increase cognitive overload: they actually hinder effective information processing. Commonly taught strategies use outmoded concepts about critical literacy that are ineffectual in an online environment, where attention is both precious, and finite. We have developed an alternative framework for information literacy school education which is adapted to the contemporary information environment. This framework is based on three key resilience principles: non-partisanship, speed of execution and transparency.

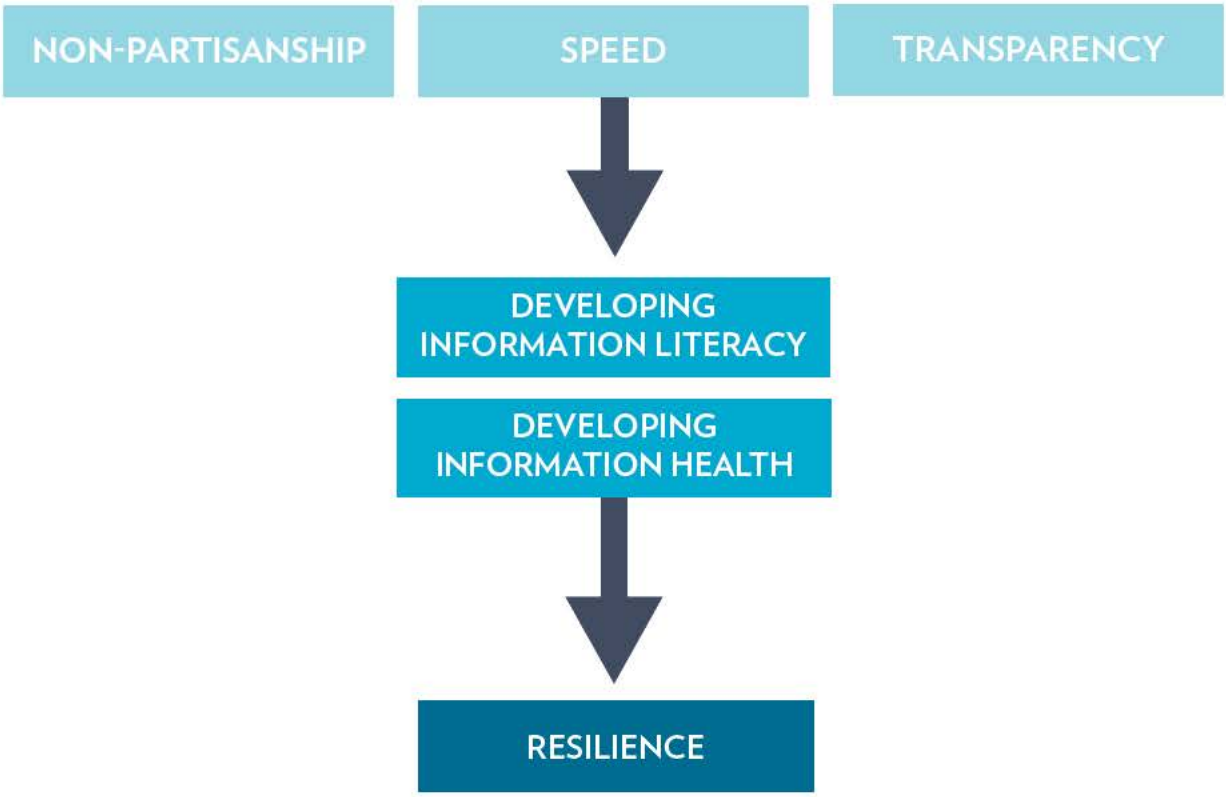

FIGURE 1: Information Resilience

3. RESILIENCE

Resilience can be defined as the capacity to recover quickly from difficulties. At a time of multiple challenges – whether environmental, health, security, or epistemic – it is urgent to develop new practical tools which will help to make people, systems and organisations better able to overcome shocks and crises. How can we create effective information literacy and information health tools that increase resilience?

3.1 First Information Resilience Principle: Non-partisanship

To be accepted by students, teachers, parents, and other stakeholders, information literacy and information health tools should not promote, or appear to promote, a partisan perspective. This first principle contradicts traditional media literacy approaches. For example Douglas Kellner, a pioneering media literacy scholar in the USA, emphasised the role of 'critical media literacy' and media education more broadly, in facilitating social change and democratisation (Kellner & Share 2007). This goes beyond the development of technical skills and competences, or personal responses to texts, to promote engagement with social, cultural, political and economic perspectives, values and ideas, including a critique of how these are created, circulated, used and consumed via media.

What one person defines as 'critical engagement' could very well be defined by another as 'propaganda'. If an educational information literacy strategy is to be developed for young children, and information health tools for the broader population, they must strive to have broad community acceptance and appeal to as wide a variety of people as possible: non-partisanship is key. To be clear, we are not suggesting that perspectives critical of power should be avoided altogether. However, in the first stages of teaching information literacy and fact-checking in schools, the process needs to be uncontroversial. For example, most people would probably agree that social media platforms' primary focus is to maintain their users' engagement, and that deconstructing how and why platforms achieve this aim is a legitimate and useful skill.[1]

3.2 Second Information Resilience Principle: Speed

Disinformation aims to capture our attention. To prevent this, information literacy and information health tools must be fast. In this respect, current media and information literacy instruction is frequently ineffective. One commonly used information-checking methodology uses the memorable acronym of C.R.A.A.P. ('Is it current, relevant, authoritative, accurate? What is its purpose?'). C.R.A.A.P. presents students with a checklist of website design clues, with some questions people might ask themselves when initially arriving at a webpage including: 'Does this webpage look professional? Are there ads? Is it a .com or a .org? Is there scientific language? Does it use footnotes?'

This checklist approach increases cognitive overload, so that students often latch onto the most visible signals, resulting in poor decisions; further, these questions no longer lead to proof of reliability. Anyone can design a professional-looking webpage or use spellcheck; an '.org' URL no longer guarantees the credibility of the content (Caulfield 2020). C.R.A.A.P. is not adequate for our information-rich world, in which a wealth of information creates a poverty of attention (Simon 1971). In the 'attention economy' time is precious: deep engagement with dubious claims is a poor strategy, as it represents time better spent elsewhere. Instead, students must acquire the means to quickly decide which claims are worth their attention.

3.3 Third Information Resilience Principle: Transparency

To be successful, an information resilience strategy must consider first principles. The notion that cabals are secretly manipulating information is a foundational characteristic of the conspiratorial and extremist rejections of 'elite' politics, science and news. Information literacy must therefore incorporate the opposite of conspiracy: transparency.

1 The non-partisanship resilience principle is mentioned in the NATO report: 'There is an urgent need for media literacy programs and approaches to be consistent, replicable, measurable and scalable – and non-partisan – so that media literacy can be effectively and strongly deployed within the NATO Alliance sooner rather than later.' (Jolls 2022 p. 23).

This is not a new idea. Members of the Demos British think tank wrote in 2010: 'Conspiracy theories are a reaction to the lack of transparency and openness in many of our institutions. The more open our institutions, the less likely we are to believe we are living in a conspiring world' (Bartlett & Miller 2010, p. 39).[2]

Transparency has also been identified as a central component for establishing trust in news media. The former Director of the BBC's Global News Division, Richard Sambrook, suggested that transparency has overcome objectivity as the means to deliver trust in the 'new media age'. Sambrook argued that 'news today still has to be accurate and fair, but it is as important for the readers, listeners and viewers to see how the news is produced, where the information comes from, and how it works' (cited in Bunz, 2009).

3.4 Wikipedia: An antidote to distrust?

Fact-checkers need a freely accessible site that can provide reliable information about any topic. Although this may be surprising to some people, we believe that the online encyclopedia Wikipedia is the world's best available resource for this purpose (Cunneen & O'Neil 2021, O'Neil & Cunneen 2022, O'Neil & Jensen 2022). This is because a 'wiki' is a database in which every change is archived. The existence, the author and the date of all modifications to a Wikipedia article appear successively, line after line, in the article's 'History' page. If a reader clicks on a line, the two versions of the article appear side by side. Every article also has a 'Talk' page where 'Wikipedians' collectively resolve disputes about the article's content. They are aided by a host of rules, which the community of editors enforces. Crucially, these include neutrality (no subjective opinion is allowed) and verifiability: all information must be supported by a reliable source, such as an academic article or a book published by a legitimate publisher. In short, the editorial process is auditable.[3]

2 In terms of specific transparency policies to tackle conspiracy theories, these authors suggested 'annual public intelligence reports produced by the new National Security Council, more maximum disclosure policing, increased openness in court proceedings in major terrorism cases, and continued focus on good community relationships in counter-terrorism policing' (Bartlett & Miller, 2010, p. 5).

3 It is necessary to distinguish two types of editorial interventions on Wikipedia. Updating the content of certain categories of scientific articles requires specialised knowledge, so only specialists need apply. It follows that these specialists will strive to maintain article quality. This was the case for articles relating to COVID-19, in which the correctness of information was meticulously curated (Cohen 2020). Numerous studies had previously shown that medical science articles on Wikipedia are as correct as scientific publications (Buchbinder & Bourne 2018, Kräenbring et al 2014, Rajagopalan et al 2011, Thomas et al 2014). For non-scientific articles, and particularly in the case of topics which attract a lot of attention, Wikipedia relies on distributed peer review, on the 'wisdom of the crowd'. In the English version of Wikipedia, there could be hundreds of contributors for a popular article. Some of these contributors will be particularly invested and will include articles of interest on their 'Watch List'; they will then be alerted every time the article is modified, and unverifiable modifications will be eliminated (Morgan 2019). If manipulators persist, Wikipedians with additional administrative privileges will block the article, or the manipulators. Wikipedians can become an 'administrator' by demonstrating good work for the project and by running for election: all editors can vote, but the decision must be consensual, and validated by a 'bureaucrat', which is a higher role in the Wikipedia hierarchy. According to the 'wisdom of the crowd' doctrine, hoaxes are possible, but typically for obscure topics that do not attract popular interest.

4. INFORMATION LITERACY IN SCHOOLS USING LATERAL READING

4.1 Civic online reasoning and lateral reading

We implemented the non-partisanship, speed and transparency resilience principles by co-creating with primary and secondary school teachers in four ACT schools a set of educational resources for children in Years 4, 5, and 6. This project was funded by the ACT Education Directorate - UC Affiliated Schools Research program and the US Embassy in Canberra. We adopted the Civic Online Reasoning framework, developed at the Stanford History Education Group (Wineburg et al. 2016). COR recognises the importance of the Internet as a source of political information and refers to the ability to effectively search for, evaluate, and verify social and political information online. What matters is not what students know, but the steps taken to verify claims: when confronted with a dubious claim, students should 'think like a fact-checker' (Wineburg & McGrew 2018).

In practical terms, this means that students should not engage 'vertically', either by scrolling down the page, or by analysing a claim in depth. Instead, students should learn about a source of information by leaving the webpage, opening another tab on a browser, and searching elsewhere: a concept known as 'lateral reading' (Wineburg & McGrew 2017). If the claim or source is found to be reliable, students can investigate in more depth, but if it is not, they should move on.

A clear advantage of Civic Online Reasoning over other media and information literacy frameworks is that its proponents have engaged in a systematic empirical verification of its effectiveness. In an early project, an assessment of online reasoning was administered to students six weeks prior to the intervention and again five weeks after (Wineburg et al. 2016). The results indicated that students in the treatment group were significantly more likely than students in the control group to have shown gains from pre-test to post-test. Having a gap of several weeks between testing, rather than immediately after, is significant, as it suggests that the students who underwent the training have retained those skills.

4.2 Six Fact-Checking Lessons for Kids

Our educational resources actively engaged students by using vivid language and imagery (O'Neil et al. 2022). The first two resources established the foundations: 'Is the Earth flat?' defined reliable sources of scientific knowledge; 'Is Wikipedia reliable?' explored ways to answer this question. The next four resources presented scenarios intended to trigger a 'fact-checking reflex': 'Street Sandwich' taught students to quickly decide whether a claim should be investigated using lateral reading. 'Why You So Mad' taught students to question information attacking people rather than ideas (e.g., 'ad hominems') and to be careful of sharing information that is emotionally manipulative. 'Red Cars' taught students to be aware of 'The Frequency Illusion': that is, ubiquity does not make information factual. Finally, 'Garage Dragon' taught students to be skeptical of hypotheses that cannot be proven.

Each lesson had a learning intention (what students were expected to learn). 'Street Sandwich', for instance, uses the metaphor of a sandwich found on the street to discuss what kind of new information students should question. Its learning intention was: 'I know when I should check if a claim or person is reliable'. Early results are encouraging: the lessons were popular and fact-checking behaviour has improved (Cunneen & O'Neil 2023).

4.3 Resistance to use of Wikipedia in the teaching community

Long-time Polish Wikimedian and researcher Dariusz Jemielniak (2019) remarked that 'Over time Wikipedia's quality has improved substantially, and yet it is still perceived in a static and dated way, as from the time of its inception'. This was indeed one of the most consistent and persistent findings in our research: negative perceptions of Wikipedia's reliability are widespread in the school teaching community. Many schoolteachers are unaware of the Wikipedia community's strict enforcement of editorial policies: whilst 'anybody can edit' a Wikipedia article, countless trusted volunteers, administrators, and automated type-setting 'bots' ensure that these edits are based on reliable and neutral sources.

This resistance is also present in some sectors of the public, as was apparent in the comments section of our *Conversation* article on the topic (Cunneen & O'Neil 2021), which was shared 7,000 times on social media. The depth of public feeling about using Wikipedia can be attributed to the fact that trusting Wikipedia represents an evolution in our understanding of encyclopaedic knowledge: from placing faith in guarantees offered by authors (e.g., Diderot), to trusting institutions or brands (e.g., Britannica), to probabilities created by transparent, auditable processes – e.g., Wikipedia (Gauntlett 2009). Yet vehement objections about lack of reliability were far from being the dominant view (N:10). The majority of commenters (N:31) were supportive, whilst others (N:11) were neutral, or made comments that were unrelated to our argument.

We observed a similar pattern during a professional workshop organised for schoolteachers at the University of Canberra in April 2022: survey responses collected at the end confirmed that attendees – bar one - were unaware of Wikipedia's internal mechanisms. Further, a majority of participants were persuaded by our presentation, and their opinion about Wikipedia changed. These reactions showed where change needs to occur: teacher education in universities needs to be reformed, and professional development training programs about Wikipedia should be offered to teachers in-service.

4.4 Conclusion: Lateral reading and Wikipedia

Wikipedia is not flawless. It reflects the world, so gender imbalances in terms of number of articles are well documented. The Wiki Education project, financed by the Wikimedia Foundation, seeks to correct these imbalances by encouraging the creation of content about women.[4] At the same time Wikipedia's practical and epistemological benefits are clear. Our project's use of lateral reading and Wikipedia for information literacy and fact-checking rests on their satisfaction of our three resilience principles of non-partisanship, speed, and transparency. The challenge now is to consider whether it is possible to develop media literacy educational programs which are fast, transparent and non-partisan. This is important, as framing and narratives shape perception; they portray lead actors, both 'good' and 'bad'; they have beginnings, middles and ends. Relatedly citizens of all ages should understand how algorithms and bots work, how newsfeeds target people, what business models govern platforms, and how this influences content.[5]

4 See https://wikiedu.org/

5 There are many contemporary initiatives aligned with this thinking and approach. An early example is a Council of Europe report entitled *Information disorder: Toward an interdisciplinary framework for research and policymaking*. The report's appendix includes a full list of initiatives aimed at debunking and fact-checking misinformation. The authors identify the need for a media and information literacy education task-force, for creative thinking about a standardised curriculum and for rigorous testing of new fact-checking strategies. They also suggest the following elements for inclusion in a media and information literacy curriculum: traditional news literacy skills; forensic social media verification skills; information about the power of algorithms to shape what is presented to people, and the possibilities but also the ethical implications offered by artificial intelligence; techniques for developing emotional skepticism to override our brain's tendency to be less critical of content that provokes an emotional response; and statistical numeracy (Wardle & Derakhshan, 2017, p. 70).

Title	Key Concepts	Learning Intentions
1. Is the Earth Flat? 	• How do we know what we know? • Defining why we can trust the scientific process: evidence can be verified. • Defining traditional reliable sources of scientific information.	• I know what a trustworthy source of scientific information is.
2. Is Wikipedia Reliable? 	• Understanding Wikipedia's structure: Article pages, 'Talk" pages, "History" pages. • Understanding some of Wikipedia's key policies (reliable sources, neutrality, no original research) and the community enforcement of these policies. • Recognising warning signs that a Wikipedia article may not be reliable.	• I know when Wikipedia articles are reliable.
3. Street Sandwich 	• Teaches students that new or unfamiliar claims should be fact-checked. • Introduces the lateral reading method: look away, open another tab, verify.	• I know when I should check if a person is credible.
4. Why You So Mad? 	• Learning to question information that attacks people rather than ideas ("ad hominems"). • Understanding the need to reflect before sharing information that is emotionally manipulative.	• I can identify an ad hominem argument. • I know I should pause and think before sharing emotional content.
5. Red Cars 	• Understanding the difference between something being ubiquitous and something being true: cognitive bias may be at work.	• I understand that just because I suddenly see something everywhere online, it does not mean it is true.
6. Garage Dragon 	• Teaching students about testing hypotheses and the possibility of falsification. • Understanding the difference between belief and evidence. • Teaching students to recognise "shifting goalposts" when evidence is being presented.	• I understand the need to be sceptical of hypotheses that cannot be proven.

FIGURE 2: Six fact-checking lessons for kids

5. INFORMATION HEALTH IN ONLINE ENVIRONMENTS

Misinformation is most likely to be accepted in information environments where attitude-challenging content is not tolerated by participants, and methods to empirically map the characteristics of such 'echo chambers' are currently being developed (Fletcher et al. 2021, McKernan et al. 2023). An important aspect of resilience is gaining insights about the quality of one's information environment, so we now present research, funded by the Volkswagen Foundation's Artificial Intelligence and the Society of the Future initiative and led by the VOSON Lab at the Australian National University, which aims to develop tools to map information environment 'health'.

5.1 Conceiving an information health metrics generator

Specifically, we envisage an automated tool that could be used to provide transparent and objective indicators about the 'health' of an online information environment, understood in the following way: to what extent does this environment have a *diverse* range of perspectives; to what extent is it *open* to opposing views; to what extent are debates *informed* by reliable sources; etc?[6] In line with our resilience principles, this tool would provide fast and objective metrics in a transparent fashion, since the methods used would be made public. For example, we could provide an automated measure of the 'health' of particular Twitter discussions. We now outline an example which allows us to both conceptualise indicators of the 'health' of an information environment, and also highlight some of the ethical challenges involved with providing an automated tool. While it is currently possible to collect (via an API) publicly-available Twitter data pertaining to any Twitter user, it would be highly unethical if this tool were used to, for example, identify particular people who were considered to be participating in an echo chamber on Twitter.

Assume that a user of the automated tool is interested to know the quality of the information environment (on Twitter) pertaining to the Indigenous Voice debate. The tool user provides a set of relevant 'field hashtags' (O'Neil & Ackland 2019). These are hashtags that are likely to be used by anyone engaging in public Twitter conversations about the Indigenous Voice, e.g.: #auspol, #voicetoparliament, #voice, #thevoice. The tool user also optionally supplies a set of 'stance' or 'contested' hashtags which are the object of critical evaluations (for or against) and/or may be used as rhetorical tools to make a point against an opposing perspective (O'Neil et al. 2022b). In the case of the Indigenous Voice debate, the stance or contested hashtags might be: #voteyes, #voteno.

The tool would then collect all the publicly-available Twitter data for the hashtag search query, over a specific period of time, and would then apply a privacy-preserving threshold: if there isn't a minimum amount of tweets authored by a minimum number of unique users, then an information health report will not be generated. In doing this, the aim would be to focus attention on Twitter activity that is public, i.e. public personalities (e.g. politicians) using Twitter or private people who are 'signalling publicly'. Twitter activity involving people engaging in private conversations on a public platform are not part of the public information environment. The remainder of this example involves data that were collected using the free (public) Twitter API, using the publicly-available and open source VOSON R software (Gertzel et al. 2023).

5.2 Information environment and stance actors

Let us assume that the user of the tool supplied the following hashtag search: #auspol AND (#voicetoparliament OR #voice OR #thevoice). For the week 5-12 February 2023 we collected 4,871 tweets (of which 82% were retweets) authored by 2,464 unique users. This collection of tweets captures the Twitter activity surrounding the Indigenous Voice debate over that period (we might refer to this as the Indigenous Voice debate's information environment). But it is important to remember that only tweets containing the target hashtags were collected and thus we are missing all the Voice-related discussion where these hashtags were not used. In their study of the first debate of the 2020 US presidential election, Gumbert et al. (2023) found that a collection using hashtags alone (such as what is used in the present report) would have missed 99% of relevant Twitter discussion activity. But as noted above, for the purposes of our tool for measuring information environment health, ethical considerations may necessitate a restriction to hashtag-based collections.

6 We regard Gonzalez-Bailon et al.'s (2010) network-theoretic measures of deliberation in online discussion (breadth and depth of discussion tree networks) as a foundational reference for our own work.

We use the information environment to identify a set of Twitter users who are aligned with a particular side of the debate around the Indigenous Voice. There are several ways that stance actors (typically referred to in the literature as 'partisan actors') can be automatically identified, with the simplest approach being via the use of stance hashtags. A curated list of stance hashtags, assessed to be privacy-preserving since they are used by a large number of Twitter users and clearly interpretable as being related to the field, would then be used to identify stance actors. This curated list of stance hashtags could be supplemented by those supplied by the tool user. In the present example, we identified 150 stance actors (66 'yes', 84 'no'). While it is not necessary for an information environment to contain stance actors, it is our expectation is that this tool would generally be used for assessing discussions around contentious issues, where stance actors are participating.

5.3 The discussion information environment

For this example, the large majority (82%) of the Twitter activity was retweeting - information diffusion, rather than reciprocal communication (including discussion and conflict). Since our tool for assessing the health of an information environment is focused on how people are engaging with information and discussing issues, not how they are spreading information, we remove retweets, quoted tweets and original tweets, and only keep replies (see Gumbert et al. 2023 for more on the importance of distinguishing different types of Twitter activity when researching deliberation and echo chambers). This leaves us with 319 replies authored by 134 unique users. This set of reply tweets encapsulates what we refer to as the *discussion information environment*. We represent this in network form, where nodes are users and directed edges/ties between users indicate replies; this network has 399 nodes and 656 edges (a single reply tweet can generate more than one edge in the reply network, since other users can be mentioned in a reply).

The network representation of the discussion information environment (we refer to this as the discussion network) is in Figure 3 (note that it is possible for someone to reply multiple times to another person, and this is reflected in the width of the edge in the visualisation). It is apparent from the visualisation that the discussion network is quite 'fractured'. There is one large set of connected nodes (the 'giant component') and many smaller components, and a modularity clustering algorithm detects 57 clusters ranging in size from 2 nodes to 54 nodes. This immediately gives us some information about the health of this discussion information environment: if people are only talking to one another in relatively small groups, and these groups tend to only include people sharing the same stance, then this may indicate less opportunity for broad-based consensus building. Network metrics relating to the number and network position of stance actors can also provide insights into the health of the information environment: if actors with a particular stance are greatly numerically outnumbered, or tend to be in more peripheral positions in the network, then this would indicate a potential lack of diversity of voices in the online discussion.

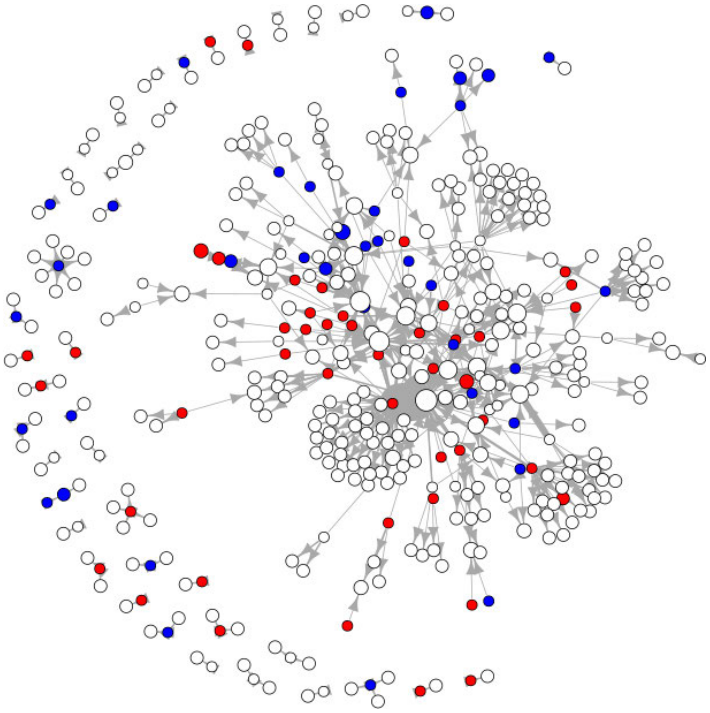

FIGURE 3: Indigenous Voice Twitter discussion network. Red (blue) nodes are users who included anti- (pro-) Voice hashtags in their tweets

By looking more closely at particular clusters within the discussion network (Figure 4), we are able to illustrate additional network indicators for quantitatively assessing the health of online discussion around the Indigenous Voice. Cluster 2 contains 54 users with one 'yes' stance and four 'no' stance users. The majority of the replies are from a single 'no' user, and none of these replies are reciprocated. This user has also directed many replies to an unaligned user (this user is very prominent, receiving replies from all the other users in this network that authored reply tweets). Cluster 2 does not represent a discussion: this looks more like one user 'speaking at' many others, with no two-way interaction.

In contrast, Cluster 6 contains 11 users and there is an even mix of 'yes'/'no' actors, suggesting that balance of actors from opposing sides are in the same vicinity in this part of the discussion network. But what is even more significant, from the perspective of measuring the health of an online information environment, is the fact that in Cluster 6 there is direct interaction between actors with different stances (unlike in Cluster 2) and furthermore, there is reciprocated interaction between two actors on opposing sides. We thus have another quantitative measure of the 'deliberative potential' of this information environment: the extent of reciprocated interaction between actors from different sides of the debate.

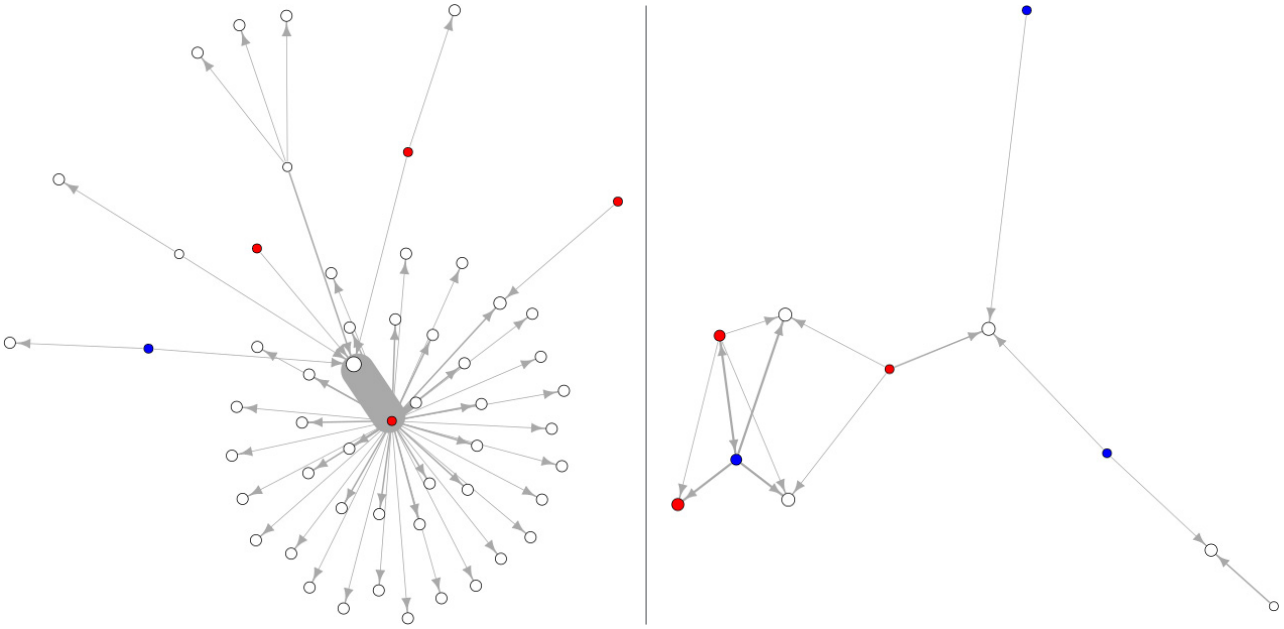

FIGURE 4: Cluster 2 (left) and Cluster 6 (right) in the Indigenous Voice Twitter discussion network

Up until this point we have simply looked at whether actors are interacting with each other, but we have not looked at what they are actually saying to one another. The tweet text content can be used to construct additional metrics of the health of the information environment. While in Cluster 6 there is interaction between two actors with different stances, text analysis could be used to determine whether this interaction is indicative of a fruitful exchange of viewpoints. More generally, our tool could present metrics relating to: (1) presence of hate speech, othering, or mocking of alternative viewpoints; (2) lack of engagement with external or 'authoritative' sources of information, e.g. Wikipedia; (3) usage of conspiracy-related hashtags and keywords.

5.4 Conclusion: Metrics for information health

We have presented some network indicators which could allow people to gain a sense of how 'healthy' or 'balanced' an online information environment is. We are not drawing a direct equivalence between food and information, but we also don't shy away from a useful analogy: eating too much of one particular food may have adverse health effects; and only consuming one type of information may not provide optimal informational benefits. We are instead proposing that empirical measures of the extent of diversity of viewpoints and openness to disagreement in online discussion environments can be of use to both social media users wanting to assess the health of information environments in which they participate, and to policy-makers in Australia focused on the social and political impacts of social media platforms.

6. RECOMMENDATIONS[7]

1. Government must consider online resilience, not just online safety

When it comes to the protection of vulnerable people in the online space, leading Australian government agencies (such as the eSafefy Commissioner) and NGOs (such as the Alannah and Madeline Foundation) highlight the safety of children. It is of course necessary to ensure children are protected from predators and bullying, or do not engage in bullying or predatory behaviour themselves. But it is also important to provide children with effective tools to process the informational claims they encounter every day.

2. Developing the resilience of citizens must be placed at the centre of the Government's defence strategy

It is important to identify principles and values that animate an information resilience program in advance. These purposes should be stated in a strategy plan or project plan. We have suggested that three key principles – non-partisanship, speed and transparency – should inform resilience programs.

3. Creation of a dedicated agency

Establish an independent department in government or organizations or schools to shine a spotlight on media and information literacy and encourage and coordinate others' efforts.

4. Lateral reading must be taught in schools

The school curriculum must include information literacy methods which are suited to the contemporary information ecosystem. Methods relying on deep, 'vertical' engagement with claims must be relinquished. Proven, effective fact-checking methods such as lateral reading must be prioritised.

5. Wikipedia needs to be rehabilitated

The wholesale rejection of Wikipedia by teachers is similarly outdated, and unrealistic, as students will use Wikipedia anyway. A more realistic and effective approach is to recognise the benefits of the online encyclopedia, and to nurture the *Wikipedia literacy* of school children: students should be taught to recognise when a Wikipedia article is reliable, or not.

6. Both Initial Teacher Education (ITE) and Professional Development for teachers in-service need to be reformed to include information literacy

There is a clear lack of teacher training when it comes to media and information literacy.[8] Schools of education at universities bear some responsibility for this challenge. Since Australian teachers must engage with development programs in order to remain registered as teachers, this presents opportunities for increasing the number of media and information literacy teacher development programs.

7. Government should support the development of information literacy and information health tools

Technological methods to track influence operations have benefited from massive investment. It is time to also invest in technological programs that boost the information resilience of citizens, such as information health metrics generators.

---

7 Some of these recommendations are inspired by the Center for Media Literacy *Building Resiliency: Media Literacy as a Strategic Defense Strategy for the Transatlantic* report (Jolls 2022).

8 'Only 18% of the teachers surveyed indicated they had access to professional development opportunities for assistance with teaching about the news. Teachers typically identified professional development as access to specially designed workshops, seminars and information sessions. This low percentage suggests few of these dedicated opportunities exist, and if they do, they are not widely known about, available or accessed by teachers.' (Corser et al. 2022).

8. Information literacy should be provided to professionals and associations

Journalists, media managers and others in the media industry can also benefit from media and information literacy training. These are people who reach huge audiences, and their insight and assistance in helping others gain information literacy skills are invaluable. In parallel information literacy training programs should target peak bodies in the professional (e.g. healthcare workers, library staff) and community sectors (e.g. indigenous groups, disability organisations, old people's homes, youth centres, etc).

9. Support and nurture media and information literacy communities and coalitions

The member organisations of the Australian Media Literacy Alliance,[9] as well as NGOs, are excellent providers of support, but they need the training, resources and financing to be able to offer rich and sustained programs. Moreover, parents, librarians, educators and university researchers are amongst the most important grassroots supporters of media and information literacy, and ways must be found to better support them and provide them with the resources they need. As experience in implementation is gained at the grassroots, policies, regulations or laws can be better informed and based on implementation experience and appropriateness at the local level. Implementation should inform policy, so that policy is better suited to the needs at hand.

10. Measure overall public awareness and progress

Research is needed in two major domains: in understanding the impact of new media technology on people, such as deep fakes or artificial intelligence; and in understanding how to best spread and teach media and information literacy to all citizens. There has traditionally been far more investment in understanding the effects of new technology than in exploring how best to prepare people for resilience through media and information literacy. Evaluation helps establish whether an investment was successful or not, and whether to continue investing in such interventions going forward.

9 https://medialiteracy.org.au/

7. REFERENCES

Ackland, R., O'Neil, M., & Jensen, M. (2020, Mar.) Submission to the Australian Senate Select Committee on Foreign Interference, through Social Media. News and Media Research Centre, University of Canberra / Australian National University.

Ackland, R. & Gwynn, K. (2020). Truth and the dynamics of news diffusion on Twitter. Chapter in R. Greifeneder, M. Jaffé, E. Newman, N. Schwarz (eds) The Psychology of Fake News: Sharing, Accepting and Correcting Misinformation. London: Routledge.

Bartlett J., Miller C. (2010). The power of unreason: Conspiracy theories, extremism and counter-terrorism. Demos. http://www.demos.co.uk/publications/thepowerofunreason

Clarke, I. (2022, Sept. 13) How fake science websites hijack our trust in experts to misinform and confuse, The Conversation.

Cameron, S. & McAllister, I. (2022, Dec.). Trends in Australian political opinion. Results from the Australian Election Study 1987– 2022, Australian National University.

Caulfield, M. (2020, Dec. 1) The truth is in the network [email interview by B. Fister], Project Information Literacy, Smart Talk Interview, no. 31. https://projectinfolit.org/smart-talk-interviews/truth-is-in-the-network/

Cohen, N. (2020, Mar. 15) How Wikipedia Prevents the Spread of Coronavirus Misinformation. Wired. https://www.wired.com/story/how-wikipedia-prevents-spread-coronavirus-misinformation/

Corser, K., Dezuanni, M. & Notley, T. (2022). How news media literacy is taught in Australian classrooms. Australian Education Research 49, 761–777. https://doi.org/10.1007/s13384-021-00457-5

Cunneen, R. & O'Neil, M. (2021, Nov. 4) Students are told not to use Wikipedia for research. But it's a trustworthy source. The Conversation.

Cunneen, R. & O'Neil, M. (2023) Co-developing a new approach to media literacy in the attention economy. Affiliated Schools Report. Faculty of Education, University of Canberra.

Gonzalez-Bailon, S., Kaltenbrunner, A., & R. Banchs (2010). The structure of political discussion networks: a model for the analysis of online deliberation. Journal of Information Technology, 1–14.

Gertzel, B., Borquez, F., and Ackland, R. (2023). VOSON R tools for collecting and analysing online social networks. VOSON Lab, Australian National University.

Gumbert, F., Pütz, O., Muhle, F., & Ackland, R. (2023). Echo Chambers without Conversation? Enriching Research on Polarization and Fragmentation on Twitter with the Analysis of Reciprocal Communication. Forthcoming in the proceedings of the 41st Congress of the German Sociological Association 2022.

Fisher, C., Park, S., Holland, K. & O'Neil, M. (2019) News Media Literacy: What Works and Does it Make a Difference? Department of Communication and the Arts / News & Media Research Centre, University of Canberra. Unreleased.

Fletcher R., Robertson C.T., Nielsen R.K. (2021). How many people live in politically partisan online news echo chambers in different countries? Journal of Quantitative Description: Digital Media. 1. https://doi.org/10.51685/jqd.2021.020

Gauntlett, D. (2009). Case study: Wikipedia. In G. Creeber & R. Martin (Eds.), Digital cultures (pp. 38-45). Open University Press.

Gruzd, Mai, Soares & Saiphoo. (2022). The Reach of Russian Propaganda & Disinformation in Canada. Social Media Lab Toronto Metropolitan University. DOI: https://doi.org/10.6084/m9.figshare.20277855

Jemielniak, D. (2019). Wikipedia: Why is the common knowledge resource still neglected by academics? GigaScience, 8, 1-2.

Jensen, M. & O'Neil, M. (2018) Fake news, real problems: What is disinformation and how do we confront it? Digital News Report: Australia 2018, News & Media Research Centre, University of Canberra.

Jolls, T. (2022) Building Resiliency: Media Literacy as a Strategic Defense Strategy for the Transatlantic. Center for Media Literacy.

Kellner, D., & Share, J. (2007). Critical media literacy: Crucial policy choices for a twenty-first century democracy. Policy Futures in Education, 5(1), 59-69.

Kirchgaessner, K., Ganguly, M., Pegg, D., Cadwalladr, C. & Burke, J. (2023, Feb. 15) Revealed: the hacking and disinformation team meddling in elections. The Guardian.

Kräenbring J, Monzon Penza T, Gutmann J, Muehlich S, Zolk O, Wojnowski L, et al. (2014). Accuracy and completeness of drug Information in Wikipedia: A comparison with standard textbooks of pharmacology. PLoS ONE, 9(9): e106930. https://doi.org/10.1371/journal.pone.0106930;

McKernan, B., Rossini, P., & Stromer-Galley, J. (2023) Echo chambers, cognitive thinking styles, and mistrust? Examining the roles information sources and information processing play in conspiracist ideation. International Journal of Communication, v. 17.

Morgan, J. (2019). Research:Patrolling on Wikipedia/Report - Meta, Sept. 2019. https://meta.wikimedia.org/wiki/Research:Patrolling_on_Wikipedia

Newman, N., Fletcher, R., Robertson, Kirsten, E., Nielsen, R. K. (2022). Digital News Report 2022. Oxford: Reuters Institute for the Study of Journalism.

Notley, T., Dezuanni, M., Zhong, H.F. & Chambers, C. (2020). News and Young Australians in 2020: How young people access, perceive and are affected by news media, Research Report, Sydney, Western Sydney University and Queensland University of Technology.

O'Neil, M. & Ackland, R. (2019) Online field theory. In Allen, M., Hunsinger, J. and Klastrup, L. (Eds.) Second International Handbook of Internet Research, Springer, pp. 445-467.

O'Neil, M. & Cunneen, R. (2022, Apr.) Digital literacy: Using Wikipedia as a fact-checking tool. Connections 121, pp. 8-9.

O'Neil, M., Cunneen, R., Carrigy, B., Cheater, R., Margetts, W., O'Brien, M. & Turner, K. (2022). Six Fact-Checking Lessons for Kids. DCPC/News & Media Research Centre, University of Canberra. https://apo.org.au/node/319173.

O'Neil, M. & Jensen, M. (2020). Australian Perspectives on Misinformation. News & Media Research Centre, University of Canberra.

O'Neil, M. & Jensen, M. (2022, Aug. 12) Three reasons why disinformation is so pervasive and what we can do about it. The Conversation.

O'Neil, M., Khan, I., Holland, K. & Cai. X. (2022) Mapping the connections of health professionals to rumour-based and factual COVID-19 hashtags in the Australian Twittersphere. Information, Communication & Society. doi.org/10.1080/1369118X.2022.2032260

Park, S., Fisher, C., Lee, J.Y., McGuinness, K. (2020). COVID-19: Australian news and misinformation. News & Media Research Centre, University of Canberra.

Rajagopalan, M.S., Khanna, V.K., Leiter, Y., Stott, M., Showalter, T.N., Dicker, A.P., Lawrence, Y.R. (2011). Patient-oriented cancer information on the internet: A comparison of Wikipedia and a professionally maintained database. Journal of Oncology Practice, 7(5): 319-23. https://doi.org/10.1200/JOP.2010.000209;

Simon, H. A. (1971). Designing Organizations for an Information-rich World. Johns Hopkins University Press.

Singer, P., & Johnson, E. B. (2021, February 1). The need to inoculate military service members against information threats: The case for Digital Literacy Training for the force. War on the Rocks. https://warontherocks.com/2021/02/we-need-to-inoculate-military-servicemembers-against-information-threats-the-case-for-digital-literacy-training/

Thomas, G.R., Eng. L., de Wolff, J.F., Grover, S.C. (2013) An evaluation of Wikipedia as a resource for patient education in nephrology. Semin Dial, 26(2): 159-63. https://doi.org/10.1111/sdi.12059

Wardle, C., & Derakhshan, J. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe. https://edoc.coe.int/en/media/7495-information-disorder-toward-aninterdisciplinary-framework-for-research-and-policy-making.html

Wineburg, S., McGrew, S., Breakstone, J., & Ortega, T. (2016). Evaluating Information: The cornerstone of Civic Online Reasoning. Stanford History Education Group. https://purl.stanford.edu/fv751yt5934

Wineburg, S. & McGrew, S. (2017) Lateral reading: Reading less and learning more when evaluating digital information. Stanford History Education Group Working Paper No. 2017-A1.

Wineburg, S. & McGrew, S. (2018, Dec. 18). To avoid getting duped by fake news, think like a fact checker. Huffington Post.