Australian Government

Information resilience

A framework for misinformation interventions

A report prepared for the **National Science and Technology Council**

chiefscientist.gov.au

Copyright

Disclaimer

The purpose of this publication is to provide an evidence base on information resilience for policy makers.

The Commonwealth as represented by the Department of Industry, Science and Resources has exercised due care and skill in the preparation and compilation of the information in this publication.

The Commonwealth does not guarantee the accuracy, reliability or completeness of the information contained in this publication. Interested parties should make their own independent inquires and obtain their own independent professional advice prior to relying on, or making any decisions in relation to, the information provided in this publication.

The Commonwealth accepts no responsibility or liability for any damage, loss or expense incurred as a result of the reliance on information contained in this publication. This publication does not indicate commitment by the Commonwealth to a particular course of action.

Authors

Associate Professor Eryn Newman, Australian National University

Professor Ullrich Ecker, University of Western Australia

Professor Mathieu O'Neil, University of Canberra

Dr Li Qian Tay, Australian National University

Contents

Executive summary

A growing body of multidisciplinary research has examined how to mitigate misinformation spread and impacts, with interventions that empower people to effectively navigate increasingly complex information environments. We provide a **framework highlighting the need for a multipronged approach to information resilience**, which ranges from targeting competencies and knowledge to navigate information environments (pre-emptive interventions) to supporting knowledge and belief updating after encountering misleading and false content (responsive interventions).

Pre-emptive interventions

Teaching skills that help people to quickly assess the quality and credibility of information (especially in online spaces) is a critical education target – improving **information literacy**. A curriculum focus that also targets **media and digital literacies** can enhance information discernment and support civics engagement. A focus in these areas must take into account both inequalities of access to technology and the diversity of Australian communities.

Prebunking supports people to identify deceptive messages, logical fallacies and low-quality content. Delivery modes include gamified applications or short videos. Prompts in the information environment that encourage consumers to consider accuracy and 'think before sharing' are also beneficial. To optimise reach and engagement, platform collaboration is required.

Responsive interventions

When misinformation has reached a broad audience, **debunking** techniques that explain a falsehood or the flawed nature of a misleading argument reduce people's false beliefs. This intervention is a critical tool in correcting misinformation and can be delivered through diverse channels of communication, including trusted community leaders. **Crowdsourcing** can also be used to fact-check claims, drawing on peer knowledge.

Opportunities and innovation

More research is needed to better understand the efficacy of interventions for vulnerable groups, as well as how best to support truth-affirming actors (e.g., Wikipedia, public broadcasters).

Emerging research identifies AI as a potential information resilience tool that may increase reach and efficacy of interventions. We identify AI literacy as a future-facing competency critical for information resilience more broadly.

Introduction

Purpose of report

A growing body of multidisciplinary research has investigated how to mitigate misinformation spread and impacts and empower people to effectively navigate increasingly complex information environments. This report examines the evidence for the efficacy of educational and related interventions in building resilience to mis and disinformation. We limit our report to the evidence from education, communication studies, and psychology; we will refer to the role of technological platforms only as it is relevant in this context.

The report also provides a framework for interventions according to timescale and associated psychological mechanisms. There are a variety of effective entry points for interventions, ranging from long-term educational interventions to build information and media literacy, to short-term interventions that pre-empt misinformation or respond to it to mitigate its impact. The evidence base for each of these interventions is examined by reviewing existing and emerging academic literature providing insights into efficacy and the limitations of existing research.

While there is vast experimental research on the efficacy of interventions, there is a need for more field-based studies. There is also an opportunity for future work to more specifically investigate the efficacy of misinformation interventions in Australia and in the Asia Pacific Region, to consider the role of cultural factors, and to explore how to effectively tailor interventions for vulnerable groups. Implementation challenges are investigated where evidence is available, and we draw on frameworks for intervention evaluation. Key issues for future research into intervention implementation within the Australian context are identified.

Background

Impacts of misinformation

The effects of misinformation can be classified into persuasive and non-persuasive. **Persuasive effects** influence reasoning and decision-making, through the promotion of false beliefs. For example, misleading information about COVID-19 vaccines created misperceptions about their safety and reduced vaccination intentions (Loomba et al. 2023). There is also now convincing evidence that misinformation can influence behaviours (Ecker et al. 2025; see also Lorenz-Spreen et al. 2022). Such persuasive effects can emerge from direct exposure, but also via indirect channels – for example, when online misinformation permeates offline conversations or spreads into mainstream media, influencing public debate and opinion.

Non-persuasive effects include:

- enhanced institutional distrust
- societal division

- citizen disengagement
- the abandonment of epistemic standards (e.g., valuing internal consistency of arguments).

These in turn can boost misinformation vulnerabilities in harmful feedback loops (Lewandowsky et al. 2017; Tay et al. 2024).

Some types of misinformation may not primarily aim to persuade but to sow confusion and distrust, for example when the environment is flooded with low-quality information, or specific groups are targeted with inflammatory content. Even though some misinformation effects may be small, compounding effects at the population level must be considered. An effective toolkit of interventions should therefore address the full range of misinformation types and potential impacts (Kozyreva et al. 2024).

Defining mis and disinformation

We use the term misinformation to broadly refer to all forms of false or misleading content, including but not limited to intentionally created and disseminated disinformation, to align with the psychological literature and the understanding of both experts and laypeople (Altay et al. 2023; Ecker et al. 2022; Hameleers et al. 2022).

Defining information resilience

Information resilience refers to an individual's capacity to ward off or recover from the impact of misinformation when encountered. Across the misinformation field, researchers rarely measure information resilience directly, but rather infer resilience based on specific measures relevant to a given intervention. In the domains of information literacy, media and digital literacy, and science and civics education, critical target measures range from information searching competency, assessing the credibility of content, discernment (the ability to differentiate between factual and misleading content), the ability to assess how people and ideas are represented, to civics engagement. In the domains of cognitive science and political and social psychology, target measures include discernment, effects on cognition (memory, reasoning, decision-making, beliefs, attitudes), and behavioural impacts including intentions and actions. In reviewing evidence across interventions, we will specify target outcome measures of information resilience and identify critical opportunities to extend these measures to better consider ancillary impacts and implementation challenges.

Understanding the information environment

The 'information environment' is a socio-political space in which information is created, stored, and exchanged between individuals, communities, organisations, and governments. The information environment in a functioning democracy needs to privilege accurate information and encourage shared understandings, as these are fundamental to good-faith public debate and normatively optimal policymaking

(Lewandowsky et al. 2023). Australia is a signatory to the Global Declaration on Information Integrity Online, which seeks to promote an 'information ecosystem that produces accurate, trustworthy, and reliable information.'[1] This is threatened by the misinformation effects outlined above.

The information environment constantly evolves. The emergence of online and social media has brought about an almost infinite amount of information, placing greater demands on cognitive processing and information-discernment skills. Online information is also curated by non-transparent algorithms, which favour content that generates emotions, fosters non-analytical thinking and impulsive sharing, and maintains user engagement, with little regard for veracity (Kozyreva et al. 2020). The emergence of generative artificial-intelligence (AI) tools presents new challenges, as they can be used to produce misleading persuasive arguments and 'deepfake' audiovisual content at scale (Clark and Lewandowsky 2026; Nygren et al. 2026). Such content is difficult to detect and easily tailored to vulnerable recipients (Garry et al. 2024; Simchon et al. 2024).

Intervention research

We focus our report on the evidence base regarding the efficacy of interventions to support individuals to:

a) build competencies and knowledge to navigate information environments (mainly **pre-emptive** interventions)

b) update their knowledge, beliefs, and memory to counter the impact of misleading and false content (mainly **responsive** interventions).

We also cover responses to misinformation in the form of in-the-moment interventions that create processing friction, as well as interventions that target misinformation supply (for reviews, see Ecker et al. 2022; Kozyreva et al. 2024)

Figure 1: Conceptual framework for misinformation interventions

Pre-emptive Competencies: long-term	Pre-emptive Competencies: brief	Responsive Contextual	Responsive Updating and evaluation
• Information literacy • Media and digital literacy • Broader benefits of education	• Prebunking • Embedded media literacy tips	• Creating friction	• Debunking • Source discreditation • Crowd-sourcing

Note: Contextual interventions target features of the information environment.

[1] See https://www.dfat.gov.au/news/media-release/global-declaration-information-integrity-online

1 Pre-emptive interventions

These interventions are critical targets, as they seek to endow students and citizens with long-term, robust skills to effectively navigate the information environment. These interventions aim to foster information resilience by building information literacy, critical thinking, as well as digital media and communication competencies. Such skills assist individuals in effectively identifying low-quality information sources, logical fallacies or inconsistencies, and attempts at nefarious persuasion.

1.1 Competencies developed through long-term interventions (education curriculum through to community programs)

In a 2024 study, 2,115 Australian adults were tested on their capacity to verify information on websites and social media posts. Of those adults, 97% were classified as having 'poor' or 'limited' ability to verify information online, and only 3% had an accuracy score of at least 50% across the four key information verification tasks used; yet many respondents overestimated their ability to verify information online (Park et al. 2024).

These data highlight the critical need to build people's competency to evaluate the quality and veracity of content they encounter. While most interventions detailed below target school students, some have been adapted for other populations, such as older adults. Indeed, it is important to conceptualise education interventions as a continuum spanning the school curriculum and ongoing learning opportunities across the lifespan.

At the school level, revising curricula to include evidence-based interventions that improve students' capacity to self-regulate and assess the credibility of information will involve challenging long-held beliefs regarding the nature of 'critical thinking'. Moreover, sustained, active and targeted community engagement will be essential to achieve improved competencies in the population more broadly.

1.2 Information literacy: improving verification

Interventions that teach people to read across sources help them to identify the credibility of information

1.2.1 Introduction to the issue and targeted mechanisms

Australian school children are poorly equipped to deal with a limitless supply of information (Notley et al. 2023). Effective information literacy methods such as 'lateral reading', created in North America, have yet to be embraced by Australian educators; this should change.

Information literacy refers to a person's ability to identify situations in which more information is needed and their knowledge of methods to determine whether a claim is factual, or a source credible. Despite these being crucial skills in the contemporary information environment, a 2023 survey found that only 24% of Australian students aged 8–16 had received lessons in the past year to help them judge the credibility of news stories (Notley et al. 2023). Moreover, information literacy lessons on the school curriculum often focus on 'deep' and critical methods that are likely ineffective in the online environment (McGrew and Breakstone 2023).

For information literacy education to be effective, a greater emphasis on **information-credibility evaluation** is needed: lateral reading thus entails departing from one source or claim and cross-checking information using other sources assessed as reliable. Lateral reading is part of the Civic Online Reasoning (COR) framework, developed by the Stanford History Education Group. COR recognises the importance of the internet as a source of social and political information and refers to the ability to effectively search for and evaluate information online: what matters is not what students know, but the steps taken to verify claims. The term 'civic' in COR is meant to emphasise the key role that finding credible information plays in democratic decision-making (Wineburg and McGrew 2019).

1.2.2 Interventions

School interventions typically involve embedding lateral-reading techniques into subjects (e.g., Geography or Biology). This intervention teaches people to investigate who is behind a piece of information, to evaluate the evidence for a claim, and to compare information across sources. Teachers 'model' or demonstrate the different ways in which information can be verified, both 'vertically' (searching within a source) and 'laterally' (searching across sources), and students then practice techniques themselves. This 'cognitive apprenticeship' is an important part of the lateral reading pedagogy (Axelsson and Nygren 2024; Nygren et al. 2021).

Research studies typically ask students to evaluate the credibility of websites, online claims and photographic evidence both before and after the intervention (Breakstone and McGrew 2022). Programs such as *MediaWise for Seniors* in the U.S. have also applied this approach outside the school context. When participants learned about lateral reading, reverse image searching and trust indicators, their ability to discern true from fake news improved (Moore and Hancock 2022). Video tutorials are an efficient delivery format for such programs for both adolescents and adults (Axelsson et al. 2021).

1.2.3 Strength of evidence and limitations

Building information literacy boosts the capacity to determine the credibility of many types of information, from online and offline news to scientific evidence and social media posts. A growing body of work, initially from Stanford but now also from other research centres, has demonstrated strong evidence of short-term benefits as well as some evidence for longer-term impact of lateral reading interventions (Brodsky et al. 2021; McGrew et al. 2019; McGrew and Breakstone 2023; Wineburg et al. 2022; Weisberg et al. 2022).

Further, intervention studies in education settings have demonstrated impact beyond the laboratory. Jones-Jang et al. (2021) showed that information literacy was positively associated with fake-news identification. However verification studies published in or before 2022 focused only on text verification; other formats, such as images and videos, have seldom been examined (Yu and Shen 2024). One challenge is that verification can be more difficult when consuming entertaining video content (Wineburg and McGrew 2019).

Skills-based information literacy instruction, though highly effective, must be paired with teaching of foundational knowledge about the information and media environment (Weisberg et al. 2022) and teaching of critical thinking skills, topic-specific knowledge, and in particular science education (Osborne and Pimentel 2022). These are now discussed in turn.

1.3 Media and digital literacy: improving understanding

Interventions that increase understanding of the media landscape and help people to gain contemporary digital skills can improve detection of misinformation

1.3.1 Introduction to the issue and targeted mechanisms

Media literacy includes competencies in using media technologies as well as analysing media contexts and representations, and developing a high-level understanding of the media landscape (see Corser et al. 2022).

Digital literacy includes competencies in using digital technology and engaging effectively with digital content. As the digital media landscape evolves, so too will literacy requirements, including the capacity to understand how algorithms impact the information environment, how people's own behaviour informs algorithms, and implications for personal privacy (Hobbs 2020; Kozyreva et al. 2020; Lewandowsky et al. 2024a).

1.3.2 Interventions

The interventions targeting media literacy in schools include a focus on building foundational understanding of journalism and media production. These interventions also examine the role of information-seeking motives and support news-analysis skills (Martens and Hobbs 2015). In Australia the focus lies on the ability to critically examine 'key concepts' such as:

- media language
- institutions creating media
- representations of people

- intended audience.[2]

Key outcome variables of interest are news-analysis and discernment skills, knowledge of the media landscape and sharing behaviours (e.g., see Murrock et al. 2018).

1.3.3 Strength of evidence and limitations

Media literacy instruction in schools can have broad impact, for example increasing engaged citizenship and equity (Burth et al. 2024; Hobbs et al. 2013). Simply promoting news consumption can enhance media trust and support fake-news identification (Altay et al. 2024). Interventions that support broader digital media competencies and engagement can also lead to higher discernment (Lu et al. 2024) and engagement with tools that support detection of misleading content (e.g., reverse image search; Nygren et al. 2021). Locally, there have been a number of initiatives, mainly connected with the Australian Media Literacy Alliance (AMLA[3]).

A 2024 systematic review found that people who had undergone media literacy training were more likely to verify information before sharing on social media (Yu and Shen 2024). A report for the UK's Communication regulator Ofcom (Edwards et al. 2021) found that increasing media knowledge and critical thinking augments people's ability to identify misinformation; however, it also highlighted that the field is dominated by tightly controlled experimental methods with few field studies conducted under 'real-world' conditions or longitudinal studies that assess the longevity of effects.

Further, most studies are carried out with adults, limiting our understanding of the effectiveness of media literacy as a tool to tackle misinformation across different populations – additional research is warranted to assess efficacy with children and diverse ethnic groups.

1.3.4 Broader challenges for media and digital literacy

Possible ancillary effects of media literacy interventions may include people becoming overly sceptical (e.g., of factual news) and overconfident in their literacy skills (Lyons et al. 2024; Nygren et al. 2024). Despite the overall benefits, research with Australian teachers also highlights challenges in implementing information and media literacy education in the classroom, including lack of space in the curriculum, and capacity to prepare content and professional development – aligning with findings of a recent European Commission Report on digital literacy in education (Corser et al. 2022; European Commission 2022; see Webber et al. 2023).

Further, research on teacher training suggests social skills in managing classroom dynamics (in concert with content-specific knowledge) is an important target in supporting teachers to effectively navigate controversial misinformation topics in media

[2] https://medialiteracy.org.au/media-literacy-framework/

[3] Notable examples are the AMLA media-literacy framework, which has been adopted by SquizKids NewsHounds curriculum for primary schools across Years 1–6; the National Film and Sound Archive of Australia's MediaMe Media Literacy program; and the ABC Education's Media Literacy resources. However, whilst all these resources are very useful, none have been formally evaluated.

education (Hansson et al. 2023). Supporting teaching simulations on such topics at scale may help to better position the educational workforce.

Moreover, when considering digital media literacy interventions, it is important to recognise and overcome the societal barriers of digital exclusion (e.g., low initial literacy levels or a lack of access to devices and the internet; Foth et al. 2023). It is also necessary to consider other group-level factors: correlational research shows that older adults (Notley et al. 2024) and those with lower education or academic performance (Martens and Hobbs 2015) have lower levels of media literacy and confidence.

Culturally and linguistically diverse communities also face distinct challenges when accessing and evaluating information, with a higher degree of reliance on family members or social media for important (e.g., health-related) information (Wen et al. 2021). Active equity measures are therefore required, and should fully utilise existing or new infrastructure at the community level. For example, public libraries represent key hubs for access to technology and the dissemination of sound information and media literacy advice, just as teacher-librarians promote information skills in schools (Park et al. 2023).

Broader benefits of education

Critical thinking, knowledge and science education

Critical thinking – engaging in reflective, analytical and evidence-based assessment – is associated with the rejection of unsubstantiated claims (Bensley 2023; Yelbuz, et al. 2022). Domain-specific knowledge can also protect against misleading arguments and work to facilitate critical thinking (Guath and Nygren 2022). In domains such as public health, having knowledge about scientific processes, data evaluation, and standards of evidence can offer protection from pseudoscientific claims (Čavojová, et al. 2024). Such knowledge can also foster a broader appreciation of the value of others' expertise grounded in an understanding of how knowledge is created (Osborne and Allchin 2024).

Civics education

Finding credible information is key to democratic decision-making. This is why in Finland–long recognised as a leader in this field–media and information literacy are part of civics education (OPMCSA 2023). There is also an opportunity in Australia to more closely link civics education to foundational knowledge and literacy skills. Such integration will allow civics education to go beyond learning about Australian democratic values and electoral processes to also highlight the importance of civic engagement and 'political efficacy', the feeling that it is worthwhile to perform one's civic duties (Bowyer and Kahne 2020). From an education perspective, informed democratic debate is enhanced when students learn to engage in reflective and actively open-minded thinking (Curzer and Gottlieb 2019), and when civic-reasoning lessons utilise systematic exercises that require shifts in perspective (Sandahl 2020) and are applicable to real-world problems (Bronkhorst et al. 2020).

Implementation considerations

While such knowledge and skills are important building blocks of an effective response to current challenges, they are insufficient if not supported by positive attitudes towards knowledge, evidence-informed policy, and democracy, which buttress fundamental trust in society and its knowledge institutions (Haider and Sundin 2022; Arechar et al. 2023). One must also acknowledge the limitations of a pure 'knowledge deficit' approach, which assumes that simply providing more and better information is a sufficient solution. It is important to pursue complementary strategies to engage citizens and communities (e.g., scientists engaging with the public; Nadkarni et al. 2019) and support transparency in decision and policy-making as well as communications in order to build community trust (Hyland-Wood et al. 2021). As with literacy, educational interventions must take into account the needs of different groups, in particular those who face structural disadvantages (Dezuanni et al. 2021; Romanova et al. 2024).

1.4 Competencies targeted in brief interventions (information environment interventions)

Brief interventions are deployed through short text or audiovisual formats to address misinformation that is anticipated or to boost key skills that effectively support discernment. The interventions are intended to pre-empt misinformation, such that at exposure, people have the skills and knowledge necessary to identify misinformation and ward off its effects. These interventions can be delivered across a range of contexts–from embedded advertising, pop-ups, and in mainstream content, and can also be embedded in school and community programs.

1.5 Prebunking (fact-based and logic-based)

Interventions that pre-emptively warn about a manipulative or deceptive message or tactic help people detect incoming misinformation.

1.5.1 Introduction to the issue and targeted mechanisms

Prebunking aims to pre-emptively enhance resilience by equipping individuals with tools to resist false or misleading claims (Lewandowsky and van der Linden 2021). Fact-based prebunking involves building protective knowledge about specific misinformation threats (e.g., the U.S. government pre-emptively sharing details about Russia's likely justification for the invasion of Ukraine). In contrast, logic-based prebunking focuses on explaining the logical fallacies and deceptive techniques commonly employed in misleading argumentation (e.g., emotional manipulation, ad hominem attacks, cherry-picking, oversimplification, non-sequitur, etc.), and as such can operate as a brief intervention to boost information literacy and critical thinking skills.

1.5.2 Interventions

One well-studied type of prebunking is psychological inoculation. Drawing on a medical analogy, the idea is to combine a pre-emptive warning of a persuasive 'attack' (e.g., that one is likely to encounter misinformation in a given context) with an explanation of that attack via a weakened example (e.g., explaining the technique of cherry-picking using an easy-to-understand example). Efficacy is typically assessed by comparing a treatment group with a control group in terms of their ability to identify the misleading techniques or their belief in misinformation claims.

1.5.3 Strength of evidence and limitations

A substantive body of work supports the efficacy of prebunking. This includes both lab experiments and field studies on online platforms like YouTube and in educational settings across diverse locations such as Indonesia, the European Union, and Ghana (Cook et al. 2024; Roozenbeek et al. 2022). However, long-term effectiveness can vary

depending on mode of delivery. Specifically, one dose of a brief intervention (e.g., delivered via short texts or videos) can have short-lived effects, as individuals may not retain the learned information over time.

Interventions that include boosters to enhance memory (e.g., delivered across multiple sessions with opportunities for practice and feedback) tend to have greater, more sustained impact (Capewell et al. 2024; Maertens et al. 2021; Maertens et al. 2025). Gamified interventions that teach players about persuasive manipulation by providing the opportunity to play a manipulator have shown promise in online, community, and education contexts (Axelsson et al. 2024; Cook et al. 2023). If used with adolescents, games that encourage and reward mischievous actions are best used in an educational context that provides appropriate guidance.

Potential ancillary impacts of prebunking interventions include boosting individuals' willingness to speak up about key issues and thus the spread of resilience within personal networks (Ivanov et al. 2012); however, there can also be an increase in general scepticism regarding both true and false information (Modirrousta-Galian and Higham 2023; cf. Leder et al. 2024 and a recent review Simchon et al. 2025).

Inoculation: an example

Explaining how the tobacco industry used advertisements with fake experts to promote their products (e.g., '20,679 Physicians say 'Luckies are less irritating'') has been shown to make people less susceptible to the fake-experts strategy in the domain of climate change denial (i.e., the infamous 'Oregon Petition' that allegedly involved thousands of scientists signing a statement denying anthropogenic global warming (Cook et al. 2017).

Prebunking: an example

Through a partnership with Google Jigsaw, researchers have developed prebunking videos on manipulative communication techniques. Across several studies, watching these videos increased content discernment and led to better recognition of persuasion techniques both in the lab and in the field (Roozenbeek et al. 2022).

(Google n.d.)

1.6 Embedded media literacy tips

Short media literacy tips – provided within the information environment (e.g., in the form of pop-ups) – increase detection of misinformation and enhance sharing discernment

1.6.1 Introduction to the issue and targeted mechanisms

Brief media literacy tips embedded in the digital information environment aim to enhance skills in digital-media navigation, critical evaluation and decision-making just prior to information engagement, with the aim of empowering individuals to effectively assess content.

1.6.2 Interventions

Studies assessing literacy tips typically present them as pop-up messages or short videos before participants evaluate the accuracy of a series of headlines. Many studies have sampled from resources developed by Facebook, with tips being as concise as 'Watch for unusual formatting,' 'Investigate the source,' 'Be skeptical of headlines' (Arechar et al. 2023). Impact is typically compared to a control group and measured by examining sharing discernment and accuracy ratings (e.g., Guess et al. 2020).

1.6.3 Strength of evidence and limitations

Results are generally consistent and hold across several countries (including Australia, Arechar et al. 2023): literacy tips enhance discernment in sharing intentions (and perceived accuracy) – largely by reducing the intention to share false content. Further research could better illuminate longer-term impact and the extent to which effects could be amplified by foundational skills built into broader educational curricula. While there is some evidence that the impact of media literacy tips tends to be small, people rate this intervention as helpful (Arechar et al. 2023). The perceived efficacy of interventions is an important area of research, as it may impact public engagement and uptake.

Creating friction in information environments

Friction (including accuracy prompts, social norms, thinking mode/pausing, warnings, source cues)

These interventions focus on providing prompts – often deployed momentarily within the information environment – to shift the way people engage with content in online spaces. Briefly drawing people's attention to the importance of accuracy (Pennycook and Rand 2022; Porter et al. 2024), to social norms (e.g., that most people don't share misinformation and that it is frowned upon, Fazio et al. 2024; Prike at al. 2024), or existing knowledge (Fazio 2020) can enhance sharing discernment.

Explicit tagging of content as contested or potentially misleading can enhance sharing and accuracy discernment (see Fazio et al. 2024). While most of these interventions have been studied in controlled lab settings, there is also emerging evidence from field studies (Lin et al. 2024; Pennycook et al. 2021). The use of such interventions relies on collaboration with platforms, and questions remain regarding their longer-term impacts (Kozyreva et al. 2024).

2 Responsive interventions

These interventions focus on responding to misinformation when it has already been disseminated. They chiefly include 'debunking' messages that expose the falsehood or logical flaws of a misinformation claim by providing structured fact-checks or corrections. This approach can also include crowdsourcing to identify the presence of low-quality content (e.g., on social media) and provide context or counter-evidence.

Discreditation of a misinformation source can also be a responsive intervention. Responsive interventions can mitigate the impact of misinformation, and are a critical tool if misinformation has been broadly disseminated; the obvious downside is that they are only applied once misinformation is already spread or processed.

2.1 Debunking

Responding to misinformation by correcting it with a detailed refutation can reduce people's reliance on it. But efficacy may decay over time and repeated debunking may be necessary – especially if misinformation is widespread

2.1.1 Introduction to the issue and targeted mechanisms

Debunking aims to reduce misconceptions that arise from misinformation consumption. Debunking often targets specific claims, but can also focus on the rhetorical tactics used to mislead. Successful debunking supports belief updating and knowledge revision, and may also provide some temporary protection against future encounters with similar misinformation claims, thus acting as an inoculation (Tay et al. 2024).

2.1.2 Interventions

Debunking can use different formats. The most basic application is simply flagging content as false or misleading. Fact-checking entails an evaluation of evidence in support of a claim, and a classification along a false-to-true scale. Prototypical debunking is often implemented as a refutation that identifies the false claim and explains why it is wrong, while also stating the relevant fact and reinforcing it by summarising the factual evidence.

Logic- or technique-based debunking will focus less on the specific claim content and more on explaining the misleading argumentation strategy or logical flaw applied. Efficacy is usually assessed based on measures of belief or inferences associated with the specific false claim being debunked.

2.1.3 Strength of evidence and limitations

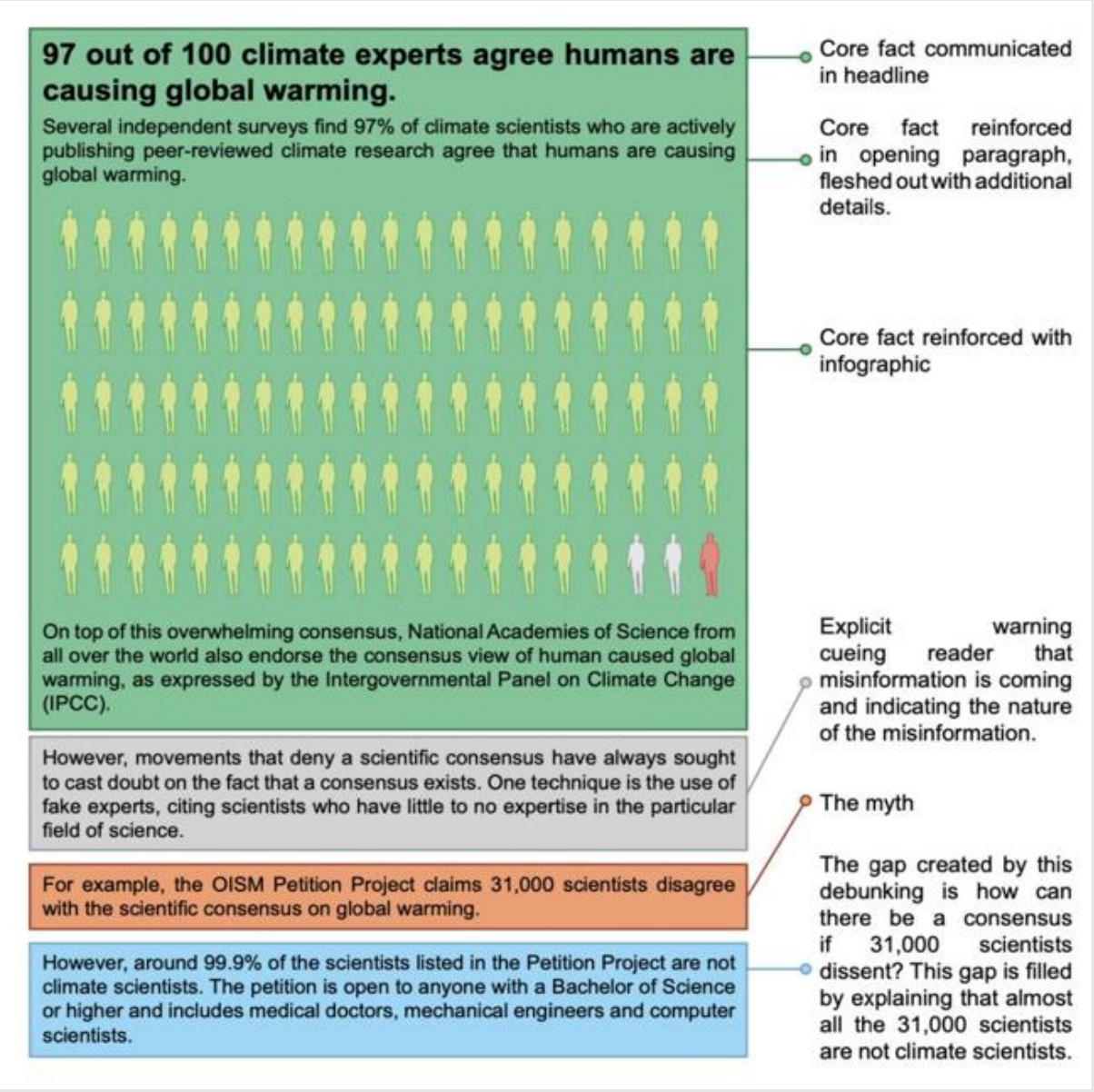
A large literature has established the effectiveness of debunking (Ecker et al. 2022; Kozyreva et al. 2024). The specific format of the intervention is less important than the

level of detail provided (e.g., simply flagging content as false has little impact; detailed refutations are most effective).

Debunking is most effective when originating from a source perceived as trustworthy (Ecker and Antonio 2021). This means that community-level debunking should involve trusted community leaders, but also that efforts to improve the trustworthiness of 'official' sources, especially for marginalised groups, will indirectly boost intervention efficacy. However, debunking is rarely fully effective: it reduces false beliefs without entirely negating the impact of the misinformation. Impact can also be circumscribed (i.e., limited to the specific misconception) and temporary, with beliefs slowly returning back to pre-intervention levels over time (Swire-Thompson et al. 2023). While directly correcting misconceptions tends to be effective (at a group level) even in cases where people are motivated to believe a false claim (e.g., worldview-consistent misinformation, Ecker et al. 2021), debunking is likely to be less useful for controversial issues and can be ineffective for specific individuals who show motivated resistance.

Debunking: an example

Drawing on the 'Oregon Petition' example used above, the first *Debunking Handbook* (2011; see also the updated *Debunking Handbook* 2020) highlighted the structure of debunking, importance of visual communication and warning readers that they were going to be exposed to a 'myth'.



2.2 Source discreditation

2.2.1 Introduction to the issue and targeted mechanisms

Discrediting a source of misinformation can be used as a standalone intervention or in combination with debunking, such that both the message and the messenger are targeted (Ecker et al. 2024). Care should be taken not to engage in ad-hominem attacks, limiting this approach mainly to cases where clear evidence exists of deception (e.g., disinformation).

2.2.2 Interventions

Sources can be discredited by highlighting a clear lack of relevant expertise, significant conflicts of interest, a track record of misleading communication, and evidence of intentional deception (Lewandowsky et al. 2024b).

2.2.3 Strength of evidence and limitations

Source discreditation has primarily been studied in combination with other interventions; only few studies have tested discreditation effects on misinformation belief in isolation, so evidence of impact remains tentative. Possible negative side effects such as antagonisation need to be considered.

2.3 Crowdsourcing

Aggregate trust indicators or added context from consumers can help identify misinformation in the information environment

2.3.1 Introduction to the issue and targeted mechanisms

Crowdsourcing as a responsive intervention refers to the elicitation of peer guidance from content consumers themselves, by enabling them to provide additional relevant information or context.

2.3.2 Interventions

The original crowdsourced knowledge community, Wikipedia, has become a key global resource in the fight against unreliable content. Input from end-users can be useful in identifying misleading information and can thus contribute to fact-checking efforts. Such an approach has been implemented on Twitter/X in the form of 'Community Notes', a community-driven content moderation feature that allows users to attach or vote for fact-checks or provide additional context under potentially misleading posts.

2.3.3 Strength of evidence and limitations

Experimental research has shown the aggregation of individuals' assessments can provide accurate classification of true and misleading headlines across countries including Australia (Allen et al. 2021; Arechar et al. 2023). Implemented systematically

and with appropriate editorial guardrails, crowdsourcing can be used to build substantive bodies of trusted information.

For example, Wikipedia articles must conform to editorial standards such as reliable sources and a neutral point of view, and the editorial process is transparent and auditable. Medical Wikipedia content has been found to be of equal quality to professional publications (Thomas et al. 2013; London et al. 2019) and proved resistant to COVID-19 misinformation (Avieson 2022). However, foreign-influence operations (Miller et al. 2022) and toxicity towards editors (Smirnov et al. 2023) are of concern. Also, evidence for crowdsourcing's effectiveness, when implemented in the field on social media, remains mixed (Allen et al. 2022; Martel et al. 2024; Wirtschafter and Majumder 2023).

Supply-side responses to misinformation

Supply-side interventions to reduce circulation (e.g., through regulation or platform policies) should be considered especially for harmful disinformation. Despite the obvious risks of censorship, there is preliminary evidence from the U.S. showing that the public values freedom from manipulation and is therefore in support of such interventions in appropriate contexts (Kozyreva et al. 2023). Reducing disinformation circulation can indirectly support consumer resilience by creating safer spaces for engagement and counteracting withdrawal from public debate.

3 Assessing interventions

3.1 Intervention deployment

Increasing the information resilience of Australians will require a multi-pronged approach. Current interventions build on foundational skills that can be acquired through education, making curricular and community workshops a critical scaffold. The broad literature on misinformation reviewed here converges on the key psychological and education targets ('what') as displayed in Table 1, articulates the impactful interventions ('how') and has highlighted timing ('when') such measures should be deployed). See Table 2 for a summary of the strengths and challenges of each intervention.

Table 1: Summary of key targets for interventions (what), which interventions align (how) and when to deploy them

What	How	When
Developing foundational knowledge and skills	Education curriculum and community workshops	**Pre-emptive:** Guarding against misinformation
Prompting actionable skills in online spaces	Deploying prebunking and media literacy tips	**Pre-emptive:** Guarding against misinformation
Shifting information consumption in the moment	Creating friction	**Contextual:** Guarding against misinformation
Updating beliefs and knowledge after exposure	Deploying corrections	**Responsive:** Responding to misinformation
Supporting a high-quality information environment	Promoting crowdsourcing Supply-side considerations	**Ongoing:** Maintenance of the quality of information environment

3.2 Limitations of the current literature on misinformation interventions

Many studies, including those reviewed in the current report, focus on average effects between intervention groups relative to a control group. Less research addresses the efficacy of interventions considering critical individual or group-level differences – especially for those who may be most vulnerable (Brashier 2024). For instance, older adults may need more tailored interventions that support long-term retention of intervention messages, and a stronger focus on sharing behaviours. Indeed, research in the US suggests that older adults tend to share more misinformation (see Brashier and Schacter 2020). Further, with rapid technological advances and an evolving information

environment, understanding which findings are likely to hold over time is important (Munger 2023).

Each intervention has specific strengths and limitations. However, across interventions it is challenging to assess relative efficacy, as materials, context and key outcome measures differ substantially between studies. One recent megastudy (Fazio et al. 2024) addressed some of these challenges by sampling pre-emptive and responsive interventions, drawing on the same materials and key outcome measures. This research found that all interventions showed efficacy (for U.S. participants).

A close inspection of assessments of true and false content showed that some interventions reduced belief in false information (e.g., debunking), while others also increased belief in **true** information (e.g., media literacy tips), within the context of the study. However, the paper has not yet undergone peer-review, and findings warrant further study to assess to what extent such findings replicate outside a U.S. context. One further challenge that will require more research is understanding the impact of a given intervention 'in the wild' – it must be assumed that effects obtained in controlled lab experiments will be smaller in real-world application (Roozenbeek et al. 2024). A given intervention can also produce ancillary impacts (e.g. creating general scepticism) and implementation challenges (e.g., scalability and reach). Going forward there is therefore a clear need to draw on (and test) formal evaluation frameworks (Tay et al. 2023).

Table 2a–d: Summarising the current body of evidence: evaluation of interventions

Table 2a: Pre-emptive interventions – long-term competencies

Intervention	Strengths	Challenges
Information literacy / lateral reading	• Fast and effective way for people to improve assessments of information credibility	• Requires an adaptation of traditional teaching focused on critical thinking • Investment in professional development • Questions around systematic embedding across the curriculum
Media and digital literacy	• Aligns with curriculum foci and can boost civic engagement • Through critical analyses of media representations can increase discernment and help people to identify misinformation	• Digital exclusion may hamper access and participation in interventions • Training must evolve with dynamic information environment
Critical thinking	• Helps foster 'informed trust' and an understanding of evidence • Foundational knowledge for all interventions • Critical for information and democratic resilience	• Highly dependent on curriculum structure • Long-term investment • Investment in professional development

Table 2b: Pre-emptive interventions – brief competencies

Intervention	Strengths	Challenges
Prebunking (fact-based and logic-based)	• Can protect people from misinformation that has yet to spread • May boost people's confidence to speak up about target issues and thus build up pre-emptive resilience in networks	• Efficacy over time depends on engagement during intervention; boosters may help • Unclear to what extent prebunking may lead to general scepticism
Embedded media literacy tips	• Highly scalable means of enhancing information discernment • Easily adaptable to different delivery formats such as pop-up messages and short videos	• Simple tips alone may not cover the complexity of modern misinformation tactics • Long-term efficacy can be unclear • To increase reach, there is a need for ongoing collaboration with platforms for deployment

Table 2c: Responsive interventions – contextual

Intervention	Strengths	Challenges
Friction	• Highly scalable means of enhancing information discernment • May have additive impact in combination with other interventions	• Requires collaboration with other platforms • Long term efficacy unclear

Table 2d: Responsive interventions – Updating and Evaluation

Intervention	Strengths	Challenges
Debunking	• Provides an effective targeted response when misinformation has spread widely • Can support belief updating across domains such as science and health	• Can only be applied retroactively • May only have short-term impacts; boosters may be needed • Potential to inadvertently contribute to the amplification of misinformers if using their framing
Source discreditation	• Can easily be used in conjunction with other interventions to enhance their efficacy	• Best applied to cases where clear evidence of intentional deception exists • Risk of antagonisation
Crowdsourcing	• Engages the public in generating knowledge and identifying misinformation • Does not require interveners to act as sole arbiters of truth	• Quality-control guardrails needed • May be prone to bias or manipulation by nefarious actors if not supported

4 Opportunities in the Australian context

4.1 Understanding efficacy in the local context

The large majority of intervention research has been conducted in North America and Europe, meaning the weight of evidence has emerged from the Global North. Findings from other western, liberal-democratic countries are likely applicable to the Australian context, although there is some evidence of variation in sensitivity to interventions in the context of political misinformation (Aird et al. 2018), which may reflect a relatively less polarised Australian political environment. While some interventions have been tested directly in Australia (e.g., Aird et al. 2018; Ecker et al. 2014; Paynter et al. 2019), there remains a need for more locally conducted research to better understand information resilience along with intervention efficacy in the Australian context.

In particular, there is a lack of research focused on understanding interventions in Indigenous communities, and research which is sensitive to diverse communication channels and the needs of multicultural communities in Australia (see Notley et al. 2024). While there is some evidence of variation in intervention efficacy across rural and urban settings, these data are limited (Guess et al. 2020) and little is known regarding the extent to which these factors impact efficacy in Australia or the Asia Pacific Region more broadly. Globally, there remains a significant gap in the current research regarding the cultural and collective factors that may amplify or interact with intervention efficacy (Brashier 2024).

In identifying and addressing misinformation, multilingual approaches should be applied to both monitoring for emerging misinformation themes and the development of interventions. Moreover, sensitivity to the digital divide is critical, as interventions are often technology-based and could amplify existing inequalities in educational access and social inclusion (Lythreatis et al. 2022; see also Thomas et al. 2021). Strategic resource planning and equity are therefore of critical importance.

4.2 Targeting information verification and making space for media and digital literacy in the classroom

Information literacy research suggests that students require a more comprehensive toolkit to effectively navigate the contemporary information environment than is currently provided through the Australian curriculum (e.g., flexible application of fact-checking via lateral reading and/or deep, critical analysis, depending on context and perceived source credibility). However, teaching effective assessments of information credibility is hampered by several factors.

Historically, media literacy education in Australian schools was part of media arts, so the focus was often placed on 'digital media production skills and participation over the development of critical literacies' (Nettlefold and Williams 2018). Further, prevalent conceptual frameworks or 'General Capabilities' and practical resources or 'Curriculum Connections' in the Australian Curriculum are ill-equipped to deal with information over-abundance. A clear illustration of these structural factors was provided when, following the 2019 Digital Platforms Inquiry (ACCC 2019), the Australian Curriculum, Assessment and Reporting Authority (ACARA) was instructed to incorporate in their 2020 review how media literacy was being taught in the curriculum.

The Australian Media Literacy Alliance proposed a Media Literacy Framework to inform the Inquiry. This framework defined media literacy as required to 'receive, verify and share information', but neither its 'Ten Competencies' nor its 'Key Concepts' included verification as operationalised in the evidence-based lateral reading literature (AMLA 2021). Effective verification skills were also lacking in a new version of the Curriculum (released in 2022) informed by AMLA's Framework.

Finnish teachers can access constant upskilling, with their own digital literacy a vital part of professional development (Saari and Säntti 2017). A common, government-supported platform (e.g., modelled on the Estonian 'Sisuloome') for teachers to share lesson plans, or create and adapt lesson plans from templates produced by other teachers, could be very useful in the Australian context.[4] Locally, research has shown that teachers are keen to teach media and information literacy, but they need more time, space, teacher training, and high-quality curriculum-aligned learning resources (Corser et al. 2022). A systematic review of the Australian Curriculum, which is beyond the scope of this report, would help to determine the optimal course of action: either identifying where improved information-credibility assessments would best be articulated with existing General Capabilities and Curriculum Connections, or creating a new General Capability: verification (the latter option would have the benefit of highlighting its vital importance).

While the curriculum is designed at the federal level, its implementation varies between states and territories, and between public and private schools. There is thus a case to be made that, given their central importance for Australian democracy, information, media and digital literacies, as well as civics education ought to be rigorously and consistently applied across regions, schools, year levels, and subject areas. This could be achieved by instituting (for example) a non-partisan 'Civic Information Literacy Certificate' which all Australian school children must obtain (O'Neil et al. 2024). The *Final Report* of the Joint Standing Committee on Electoral Matters' Inquiry into civics education, engagement, and participation in Australia was tabled in Parliament in January 2025 (JSCEM 2025). The report's Foreword by Senator Carole Brown called for a coordinated 'long-term national media and digital literacy strategy' and declared that 'Australia needs a strengthened and standardised approach to formal civics education. The Committee recommends nationally aligned and mandated civics and citizenship content for high-school students, with greater accountability and consistency in how

[4] Ministry of Education and Research (Estonia). https://sisuloome.e-koolikott.ee/

this is implemented, by working more closely with state and territory governments' (Ibid, p.iii–iv).

In brief, educational settings are high-engagement, high-reach contexts that provide important opportunities to challenge and extend the knowledge and skills of all citizens: these settings should be a primary intervention target.

5 Future directions

5.1 Delivery pathways and AI

Existing work on interventions has focussed on a limited set of delivery pathways (e.g., school settings for literacy interventions, text-based content for prebunking and debunking, etc.). However, emerging research highlights innovation and opportunity to better understand the role of embedding interventions in high-engagement contexts such as on YouTube (Roozenbeek et al. 2022).[5]

Recent advances in AI-driven large language models such as ChatGPT have also led to an interest in AI as a resilience tool. For instance, there are now studies testing if conversations with chatbots can durably reduce belief in conspiracy theories (Costello et al. 2024) or whether AI tools can be used to improve the scalability of prebunking interventions (Linegar et al. 2024) or detect and automatically debunk disinformation on social media (Rojas et al. 2024; cf. DeVerna et al. 2024).

Understanding the efficacy of these approaches is an important target for future research. As AI applications become more commonplace, the need to consider AI literacy as a core competency of the future will grow (Efimova and Nygren 2024). Such goals also need to be balanced with an understanding of AI risks in misinformation contexts (e.g., microtargeting of misleading content aimed at vulnerable recipients; Simchon et al. 2024) and strategic and dynamic educational resources (e.g., online resources keeping pace with rapid developments in the field of AI).

5.2 Combining innovative approaches

While much research has focused on specific interventions in isolation, emerging research suggests higher impact may be achieved by combining interventions. For instance, combining pre-emptive foundational competencies with small doses of friction or responsive interventions may lead to greater efficacy (Pennycook et al. 2024).

Likewise, combining debunking with source discreditation could be more effective than the two interventions used in isolation (Ecker et al. 2024). There is also an emerging literature seeking to model the unique and combined effects of different interventions at the level of the information ecosystem, including studies integrating psychological and epidemiological approaches (Bak-Coleman et al. 2022; Pilditch et al. 2022).

[5] A 2024 survey of Australian adults asked what media and information literacy skills respondents wanted, and how they wanted to acquire them: for more than half the adult population the top three topics were ‘how to protect myself from harms’; ‘how to identify misinformation online’ and ‘how to identify high quality news and information sources’. Preferred means include online tutorials, followed by short YouTube/TikTok videos; older Australians favoured libraries (see Notley et al. 2024).

5.3 Supporting truth-affirming actors

Wikipedia's institutional evolution discouraged fringe content believers, who left the project, whilst comforting scientific-content adherents, who remained, so Wikipedia is now referred to as 'the last good place on the internet' (Steinsson 2024). However, distrust of Wikipedia is widespread amongst primary and secondary school teachers. Most are unaware of the community-enforced editorial guidelines, which ensure that Wikipedia articles on popular topics, which attract significant editorial attention, are reliable and neutral (O'Neil et al. 2025).

The value, and use, of Wikipedia can be considered a target in teacher education programs and in school classrooms. This will require consultation with both ACARA and universities delivering pre-service teacher education. In addition, individuals and groups who curate online information environments, such as Wikipedia editors and administrators or Reddit moderators, perform key deliberative roles. These individuals may need to deploy 'counter speech' in response to disinformation and hate speech, but lack adequate identification tools and appropriate response mechanisms. More research is needed to determine how best to support them. Support for truth-affirming actors should also extend to continuing recognition of the role of public broadcasting and targeted support to high-quality independent media.

6 Concluding remarks

Misinformation is an evolving and complex challenge facing democratic societies. With technological advances reshaping communication, interventions must adapt to the dynamic nature of the information environment. Current research provides a significant evidence-base for a range of tools to support individuals to better assess the quality of information they encounter (Kozyreva et al. 2024).

To effectively empower individuals and communities, the full set of interventions presented here ought to be considered, with a range of entry points to address misinformation (see Table 2 for a summary of interventions). This review also highlights opportunities for Australia to better understand efficacy in our cultural context.
To effectively combat misinformation, leveraging technology to enhance engagement with interventions and audience reach should be considered as fundamental to impact.

References

Aird MJ, Ecker UK, Swire B, Berinsky AJ and Lewandowsky S (2018) 'Does truth matter to voters? The effects of correcting political misinformation in an Australian sample', *Royal Society Open Science*, 5(12): 180593, https://doi.org/10.1098/rsos.180593.

Allen J, Arechar AA, Pennycook G and Rand DG (2021) 'Scaling up fact-checking using the wisdom of crowds', *Science Advances*, 7(36): eabf4393, https://doi.org/10.1126/sciadv.abf4393.

Allen J, Martel C and Rand DG (2022) 'Birds of a feather don't fact-check each other: partisanship and the evaluation of news in Twitter's Birdwatch crowdsourced fact-checking program', *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 245: 1–19, https://doi.org/10.1145/3491102.3502040.

Alstyne MWV (2022) 'Free speech, platforms and the fake news problem', https://doi.org/10.2139/ssrn.3997980.

Altay S, Berriche M and Acerbi A (2023) 'Misinformation on misinformation: conceptual and methodological challenges', *Social Media + Society*, 9(1): 20563051221150412, https://doi.org/10.1177/20563051221150412.

Altay S, Hoes E and Wojcieszak M (2024) 'News on social media boosts knowledge, belief accuracy, and trust: a field experiment on Instagram and WhatsApp', *PsyArXiv*, https://doi.org/10.31234/osf.io/hq5ru.

Altay S, Nielsen RK and Fletcher R (2024) 'News can help! The impact of news media and digital platforms on awareness of and belief in misinformation', *The International Journal of Press/Politics*, 29(2): 459–484, https://doi.org/10.1177/19401612221148981.

Arechar AA, Allen J, Berinsky AJ, Cole R, Epstein Z, Garimella K, Gully, A, Lu JG, Ross RM, Stagnaro MN, Zhang Y, Pennycook G and Rand DG (2023) 'Understanding and combatting misinformation across 16 countries on six continents', *Nature Human Behaviour*, 7: 1502–1513, https://doi.org/10.1038/s41562-023-01641-6.

ACCC (Australian Competition and Consumer Commission) (2019) *Digital platforms inquiry: final report*, ACCC.

Avieson B (2022) 'Editors, sources and the "go back" button: Wikipedia's framework for beating misinformation', *First Monday*, 27(11), https://doi.org/10.5210/fm.v27i11.12754.

Axelsson CAW, Guath M and Nygren T (2021) 'Learning how to separate fake from real news: scalable digital tutorials promoting students' civic online reasoning', *Future Internet*, 13(3): 60, https://doi.org/10.3390/fi13030060.

Axelsson CAW and Nygren T (2024) 'The advantage of videos over text to boost adolescents' lateral reading in a digital workshop', *Behaviour and Information Technology*, 1–15, https://doi.org/10.1080/0144929X.2024.2308046.

Axelsson CAW, Nygren T, Roozenbeek J and van der Linden S (2024) 'Bad news in the civics classroom: how serious gameplay fosters teenagers' ability to discern misinformation techniques', *Journal of Research on Technology in Education*, 57(5): 992-1018, https://doi.org/10.1080/15391523.2024.2338451.

Bak-Coleman JB, Kennedy I, Wack M, Beers A, Schafer JS, Spiro ES, Starbird K, West JD (2022) 'Combining interventions to reduce the spread of viral misinformation', *Nature Human Behaviour*, 6: 1372–1380, https://doi.org/10.1038/s41562-022-01388-6.

Bensley DA (2023) 'Critical thinking, intelligence, and unsubstantiated beliefs: an integrative review', *Journal of Intelligence*, 11(11): 207, https://doi.org/10.3390/jintelligence11110207.

Bowyer B and Kahne J (2020) 'The digital dimensions of civic education: assessing the effects of learning opportunities', *Journal of Applied Developmental Psychology*, 69: 101162, https://doi.org/10.1016/j.appdev.2020.101162.

Brashier NM (2024) 'Fighting misinformation among the most vulnerable users', *Current Opinion in Psychology*, 57: 101813, https://doi.org/10.1016/j.copsyc.2024.101813.

Brashier NM and Schacter DL (2020) 'Aging in an era of fake news', *Current Directions in Psychological Science*, 29(3): 316–323, https://doi.org/10.1177/0963721420915872.

Breakstone J and McGrew S (2022) 'Teaching civic online reasoning: across the curriculum', *School Administrator*, 79(5): 22–25.

Brodsky JE, Brooks PJ, Pavlounis D and Johnston JL (2023) 'Instruction increases Canadian students' preference for and use of lateral reading strategies to fact-check online information', *AERA Open*, 9: 23328584231192106, https://doi.org/10.1177/23328584231192106.

Brodsky JE, Brooks PJ, Scimeca D, Galati P, Todorova R and Caulfield M (2021) 'Associations between online instruction in lateral reading strategies and fact-checking COVID-19 news among college students', *AERA Open*, 7: 23328584211038937, https://doi.org/10.1177/23328584211038937.

Bronkhorst H, Roorda G, Suhre C and Goedhart M (2020) 'Logical reasoning in formal and everyday reasoning tasks', *International Journal of Science and Mathematics Education*, 18: 1673–1694, https://doi.org/10.1007/s10763-019-10039-8.

Burth S, Riewestahl E, Ramasubramanian S, Foster B, Johnson PR, Mihailidis P and Tully M (2024) 'Mapping media literacy impact in the US: a review of literature and call for equity', *Annals of the International Communication Association*, 48(4): 336-357, https://doi.org/10.1080/23808985.2024.2354357.

Capewell G, Maertens R, Remshard M, Van Der Linden S, Compton J, Lewandowsky S and Roozenbeek J (2024) 'Misinformation interventions decay rapidly without an immediate posttest', *Journal of Applied Social Psychology*, 54(8): 441–454, https://doi.org/10.1111/jasp.13049.

Čavojová V, Šrol J and Ballová Mikušková E (2024) 'Scientific reasoning is associated with rejection of unfounded health beliefs and adherence to evidence-based regulations during the Covid-19 pandemic', *Current Psychology*, 43: 8288–8302, https://doi.org/10.1007/s12144-023-04284-y.

Clark S and Lewandowsky S (2026) 'The continued influence of AI-generated deepfake videos despite transparency warnings', *Communications Psychology,* 4: 13, https://doi.org/10.1038/s44271-025-00381-9.

Cologna V, Mede N, Berger S, Besley J, Brick C, Joubert M, et al. (2024) 'Trust in scientists and their role in society across 67 countries', *OSF Preprints,* https://doi.org/10.31219/osf.io/6ay7s.

Cook J, Ecker UK, Trecek-King M, Schade G, Jeffers-Tracy K, Fessmann J, Kim SC, Kinkead D, Orr M, Vraga E, Roberts K and McDowell J (2023) 'The cranky uncle game: combining humour and gamification to build student resilience against climate misinformation', *Environmental Education Research*, 29(4): 607–623, https://doi.org/10.1080/13504622.2022.2085671.

Cook J, Lepage C, Hopkins KL, Cook W, Kolog EA, Thomson A, Iddrisu I and Burnette S (2024) 'Co-designing and pilot testing a digital game to improve vaccine attitudes and misinformation resistance in Ghana', *Human Vaccines and Immunotherapeutics*, 20(1): 2407204, https://doi.org/10.1080/21645515.2024.2407204.

Cook J and Lewandowsky S (2011) *The debunking handbook*.

Cook J, Lewandowsky S and Ecker UKH (2017) 'Neutralising misinformation through inoculation: exposing misleading argumentation techniques reduces their influence', *PLoS ONE*, 12(5): e0175799, https://doi.org/10.1371/journal.pone.0175799.

Corser K, Dezuanni M and Notley T (2022) 'How news media literacy is taught in Australian classrooms', *Australian Educational Researcher*, 49: 761–777, https://doi.org/10.1007/s13384-021-00457-5.

Costello TH, Pennycook G and Rand DG (2024) 'Durably reducing conspiracy beliefs through dialogues with AI', *Science*, 385(6714): eadq1814, https://doi.org/10.1126/science.adq1814.

Curzer HJ and Gottlieb J (2019) 'Making the classroom safe for open-mindedness', *Educational Theory*, 69(4): 383–402, https://doi.org/10.1111/edth.12376.

DeVerna MR, Yan HY, Yang KC and Menczer F (2024) 'Fact-checking information from large language models can decrease headline discernment', *Proceedings of the National Academy of Sciences*, 121(50): e2322823121, https://doi.org/10.1073/pnas.2322823121.

Dezuanni M, Notley T and Di Martino L (2021) 'Towards a national strategy for media literacy: national consultation report', Australian Media Literacy Alliance.

Ecker UK and Antonio LM (2021) 'Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect', *Memory and Cognition*, 49: 631–644, https://doi.org/10.3758/s13421-020-01129-y.

Ecker UKH, Lewandowsky S, Chang EP and Pillai R (2014) 'The effects of subtle misinformation in news headlines', *Journal of Experimental Psychology: Applied*, 20(4): 323–335, https://doi.org/10.1037/xap0000028.

Ecker UKH, Lewandowsky S, Cook J, Schmid P, Fazio LK, Brashier N, Kendeou P, Vraga EK, Amazeen MA (2022) 'The psychological drivers of misinformation belief and its resistance to correction', *Nature Reviews Psychology*, 1: 13–29, https://doi.org/10.1038/s44159-021-00006-y.

Ecker UK, Prike T, Paver AB, Scott RJ and Swire-Thompson B (2024) 'Don't believe them! reducing misinformation influence through source discreditation', *Cognitive Research: Principles and Implications*, 9: 52, https://doi.org/10.1186/s41235-024-00581-7.

Ecker UK, Sze BK and Andreotta M (2021) 'Corrections of political misinformation: no evidence for an effect of partisan worldview in a US convenience sample', *Philosophical Transactions of the Royal Society B*, 376(1822): 20200145, https://doi.org/10.1098/rstb.2020.0145.

Ecker UKH, Tay LQ, Roozenbeek J, van der Linden S, Cook J, Oreskes N and Lewandowsky S (2025) 'Why misinformation must not be ignored', *American Psychologist*, 80(6): 867–878, https://doi.org/10.1037/amp0001448.

Edwards L, Stoilova M, Anstead N, Fry A, El-Halaby G and Smith M (2021) *Rapid evidence assessment on online misinformation and media literacy: final report for Ofcom*.

Efimova E and Nygren T (2024) 'AI-optimism, vigilance and biases: support and influence from AI in classroom discussions about social issues', *OSF Preprints,* https://doi.org/10.31219/osf.io/mhpxa.

European Commission, Directorate-General for Education, Youth, Sport and Culture (2022) *Final report of the Commission expert group on tackling disinformation and promoting digital literacy through education and training*, Publications Office of the European Union, https://data.europa.eu/doi/10.2766/283100.

Fasce A and Picó A (2019) 'Science as a vaccine: the relation between scientific literacy and unwarranted beliefs', *Science and Education*, 28: 109–125, https://doi.org/10.1007/s11191-018-00022-0.

Fazio LK (2020) 'Pausing to consider why a headline is true or false can help reduce the sharing of false news', *Harvard Kennedy School Misinformation Review*, 1(2), https://doi.org/10.37016/mr-2020-009.

Fazio L, Rand DG, Lewandowsky S, Susmann M, Berinsky AJ, Guess AM, Kendeou P, Lyons B, Miller J, Newman E, Pennnycook G, Swire-Thompson B (2024) 'Combating misinformation: a megastudy of nine interventions designed to reduce the sharing of and belief in false and misleading headlines', *PsyArXiv* https://doi.org/10.31234/osf.io/uyjha.

Foth M, Osman K and Hourigan A (2023) *The low-income families digital divide: North Queensland community*, Digital Media Research Centre, Queensland University of Technology, https://doi.org/10.25916/faf0-2s66.

Garry M, Chan WM, Foster J and Henkel LA (2024) 'Large language models (LLMs) and the institutionalisation of misinformation', *Trends in Cognitive Sciences*, 28(12): 1078-1088, https://doi.org/10.1016/j.tics.2024.08.007.

Google (n.d.) *Learn to spot online misinformation*, Google website.

Guath M and Nygren T (2022) 'Civic online reasoning among adults: an empirical evaluation of a prescriptive theory and its correlates', *Frontiers in Education*, 7: 721731, http://doi.org/10.3389/feduc.2022.721731.

Guess AM, Lerner M, Lyons B, Montgomery JM, Nyhan B, Reifler J and Sircar N (2020) 'A digital media literacy intervention increases discernment between mainstream and false news in the United States and India', *Proceedings of the National Academy of Sciences*, 117(27): 15536–15545, https://doi.org/10.1073/pnas.1920498117.

Haider J and Sundin O (2022) *Paradoxes of media and information literacy: the crisis of information*, Routledge, https://doi.org/10.4324/9781003163237.

Hameleers M, Brosius A, Marquart F, Goldberg AC, van Elsas E and De Vreese CH (2022) 'Mistake or manipulation? conceptualising perceived mis- and disinformation among news consumers in 10 European countries', *Communication Research*, 49(7): 919–941, https://doi.org/10.1177/0093650221997719.

Hansson PO, Samuelsson M and Höög ML (2023) 'Teaching avatars on controversial issues: lessons learned', *IAFOR Journal of Education*, 11(2), https://doi.org/10.22492/ije.11.2.03.

Hobbs R (2020) 'Propaganda in an age of algorithmic personalization: expanding literacy research and practice', *Reading Research Quarterly*, 55(3): 521–533, https://doi.org/10.1002/rrq.301.

Hobbs R, Donnelly K, Friesem J and Moen M (2013) 'Learning to engage: how positive attitudes about the news, media literacy, and video production contribute to adolescent civic engagement', *Educational Media International*, 50(4): 231–246, https://doi.org/10.1080/09523987.2013.862364.

Hyland-Wood B, Gardner J, Leask J and Ecker UK (2021) 'Toward effective government communication strategies in the era of COVID-19', *Humanities and Social Sciences Communications*, 8: 30, https://doi.org/10.1057/s41599-020-00701-w.

Ivanov B, Miller CH, Compton J, Averbeck JM, Harrison KJ, Sims JD, Parker KA, Parker JL (2012) 'Effects of postinoculation talk on resistance to influence', *Journal of Communication*, 62(4): 701–718, https://doi.org/10.1111/j.1460-2466.2012.01658.x.

Joint Standing Committee on Electoral Matters (2025) *From classroom to community: civics education and political participation in Australia*, Parliament of Australia.

Jones-Jang SM, Mortensen T and Liu J (2021) 'Does media literacy help identification of fake news? information literacy helps, but other literacies don't', *American Behavioral Scientist*, 65(2): 371–388, https://doi.org/10.1177/0002764219869406.

Kozyreva A, Herzog SM, Lewandowsky S, Hertwig R, Lorenz-Spreen P, Leiser M and Reifler J (2023) 'Resolving content moderation dilemmas between free speech and harmful misinformation', *Proceedings of the National Academy of Sciences*, 120(7): e2210666120, https://doi.org/10.1073/pnas.2210666120.

Kozyreva A, Lewandowsky S and Hertwig R (2020) 'Citizens versus the internet: confronting digital challenges with cognitive tools', *Psychological Science in the Public Interest*, 21(3): 103–156, https://doi.org/10.1177/1529100620946707.

Kozyreva A, Lorenz-Spreen P, Herzog SM, Ecker UKH, Lewandowsky S, Hertwig R, et al. (2024) 'Toolbox of individual-level interventions against online misinformation', *Nature Human Behaviour*, 8: 1044–1052, https://doi.org/10.1038/s41562-024-01881-0.

Leder J, Schellinger LV, Maertens R, van der Linden S, Chryst B and Roozenbeek J (2024) 'Feedback exercises boost discernment of misinformation for gamified inoculation interventions', *Journal of Experimental Psychology: General*, 153(8): 2068-2087, https://doi.org/10.1037/xge0001603.

Lewandowsky S, Cook J, Ecker UKH, Albarracín D, Amazeen MA, Kendeou P, Lombardi D, Newman EJ, Pennycook G, Porter E, Rand DG, Rapp DN, Reifler J, Roozenbeek J, Schmid P, Seifert CM, Sinatra GM, Swire-Thompson B, van der Linden S, Vraga EK, Wood TJ and Zaragoza MS (2020) *The debunking handbook 2020*, doi:10.17910/b7.1182.

Lewandowsky S, Ecker UK and Cook J (2017) 'Beyond misinformation: understanding and coping with the "post-truth" era', *Journal of Applied Research in Memory and Cognition*, 6(4): 353–369, https://doi.org/10.1016/j.jarmac.2017.07.008.

Lewandowsky S, Ecker UK, Cook J, van der Linden S, Roozenbeek J and Oreskes N (2023) 'Misinformation and the epistemic integrity of democracy', *Current Opinion in Psychology*, 54: 101711, https://doi.org/10.1016/j.copsyc.2023.101711.

Lewandowsky S, Ecker UK, Cook J, van der Linden S, Roozenbeek J, Oreskes N and McIntyre LC (2024a) 'Liars know they are lying: differentiating disinformation from disagreement', *Humanities and Social Sciences Communications*, 11: 986, https://doi.org/10.1057/s41599-024-03503-6.

Lewandowsky S, Robertson RE and DiResta R (2024b) 'Challenges in understanding human-algorithm entanglement during online information consumption', *Perspectives on Psychological Science*, 19(5): 758–766, https://doi.org/10.1177/17456916231180809.

Lewandowsky S and van der Linden S (2021) 'Countering misinformation and fake news through inoculation and prebunking', *European Review of Social Psychology*, 32(2): 348–384, https://doi.org/10.1080/10463283.2021.1876983.

Lin H, Garro H, Wernerfelt N, Shore JC, Hughes A, Deisenroth D, Barr N, Berinsky A, Eckles D, Pennycook G and Rand DG (2024) 'Reducing misinformation sharing at scale using digital accuracy prompt ads', *PsyArXiv* https://doi.org/10.31234/osf.io/u8anb.

Linegar M, Sinclair B, van der Linden S and Alvarez RM (2024) 'Prebunking elections rumours: artificial intelligence assisted interventions increase confidence in American elections', *arXiv preprint* arXiv:2410.19202, https://arxiv.org/abs/2410.19202v1.

London DA, Andelman SM, Christiano AV, Kim JH, Hausman MR and Kim JM (2019) 'Is Wikipedia a complete and accurate source for musculoskeletal anatomy?', *Surgical and Radiologic Anatomy*, 41: 1187–1192, https://doi.org/10.1007/s00276-019-02280-1.

Loomba S, De Figueiredo A, Piatek SJ, De Graaf K and Larson HJ (2021) 'Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA', *Nature Human Behaviour*, 5: 337–348, https://doi.org/10.1038/s41562-021-01056-1.

Lorenz-Spreen P, Oswald L, Lewandowsky S and Hertwig R (2023) 'A systematic review of worldwide causal and correlational evidence on digital media and democracy', *Nature Human Behaviour*, 7: 74–101, https://doi.org/10.1038/s41562-022-01460-1.

Lu C, Hu B, Bao MM, Wang C, Bi C and Ju XD (2024) 'Can media literacy intervention improve fake news credibility assessment? a meta-analysis', *Cyberpsychology,*

Behavior, and Social Networking, 27(4): 240–252, https://doi.org/10.1089/cyber.2023.0324.

Lyons B, King AJ and Kaphingst KA (2024) 'A health media literacy intervention increases skepticism of both inaccurate and accurate cancer news among US adults', *Annals of Behavioral Medicine*, 58(12): 820–831, https://doi.org/10.1093/abm/kaae054.

Lythreatis S, Singh SK and El-Kassar AN (2022) 'The digital divide: a review and future research agenda', *Technological Forecasting and Social Change*, 175: 121359, https://doi.org/10.1016/j.techfore.2021.121359.

Maertens R, Roozenbeek J, Basol M and van der Linden S (2021) 'Long-term effectiveness of inoculation against misinformation: three longitudinal experiments', *Journal of Experimental Psychology: Applied*, 27(1): 1-16, https://psycnet.apa.org/doi/10.1037/xap0000315.

Maertens R, Roozenbeek J, Simons JS, Lewandowsky S, Maturo V, Goldberg B, Su R, van der Linden S (2025) 'Psychological booster shots targeting memory increase long-term resistance against misinformation', *Nature Communications*, 16: 2062, https://doi.org/10.1038/s41467-025-57205-x.

Martel C, Allen J, Pennycook G and Rand DG (2024) 'Crowds can effectively identify misinformation at scale', *Perspectives on Psychological Science*, 19(2): 477–488, https://doi.org/10.1177/17456916231190388.

Martens H and Hobbs R (2015) 'How media literacy supports civic engagement in a digital age', *Atlantic Journal of Communication*, 23(2): 120–137, https://doi.org/10.1080/15456870.2014.961636.

McCabe SD, Ferrari D, Green J, Lazer DM and Esterling KM (2024) 'Post-January 6th deplatforming reduced the reach of misinformation on Twitter', *Nature*, 630: 132–140, https://doi.org/10.1038/s41586-024-07524-8.

McGrew S and Breakstone J (2023) 'Civic online reasoning across the curriculum: developing and testing the efficacy of digital literacy lessons', *AERA Open*, 9: 23328584231176451, https://doi.org/10.1177/23328584231176451.

McGrew S, Smith M, Breakstone J, Ortega T and Wineburg S (2019) 'Improving university students' web savvy: an intervention study', *British Journal of Educational Psychology*, 89(3): 485–500, https://doi.org/10.1111/bjep.12279.

Miller C, Smith M, Marsh O, Balint K, Inskip C and Visser F (2022) *Information warfare and Wikipedia* [PDF], Institute for Strategic Dialogue.

Modirrousta-Galian A and Higham PA (2023) 'Gamified inoculation interventions do not improve discrimination between true and fake news: reanalysing existing research with receiver operating characteristic analysis', *Journal of Experimental Psychology: General*, 152(9): 2411-2437, https://doi.org/10.1037/xge0001395.

Moore RC and Hancock JT (2022) 'A digital media literacy intervention for older adults improves resilience to fake news', *Scientific Reports*, 12: 6008, https://doi.org/10.1038/s41598-022-08437-0.

Munger K (2023) 'Temporal validity as meta-science', *Research and Politics*, 10(3): 20531680231187271, https://doi.org/10.1177/20531680231187271.

Murrock E, Amulya J, Druckman M, Liubyva T (2018) 'Winning the war on state-sponsored propaganda: results from an impact study of a Ukrainian news media and information literacy program', *Journal of Media Literacy Education*, 10(2): 53–85, https://doi.org/10.23860/JMLE-2018-10-2-4.

Nadkarni NM, Weber CQ, Goldman SV, Schatz DL, Allen S and Menlove R (2019) 'Beyond the deficit model: the ambassador approach to public engagement', *BioScience*, 69(4): 305–313, https://doi.org/10.1093/biosci/biz018.

Nettlefold J and Williams K (2018) *Insight five: a snapshot of media literacy in Australian schools*, University of Tasmania.

Notley T, Chambers S, Park S and Dezuanni M (2024) *Adult media literacy in 2024: Australian attitudes, experiences and needs*, Western Sydney University, Queensland University of Technology and University of Canberra, https://doi.org/10.60836/n1a2-dv63.

Notley T, Chambers S, Zhong HF, Park S, Lee J and Dezuanni M (2023) *News and young Australians in 2023: how children and teens access, perceive and are affected by news media*, Research Report, Western Sydney University.

Nygren T, Al-Afifi M and Axelsson CW (2024) 'Boosting teenagers' skills to fact-check photos and problems of overconfidence: impact and side effects in education against misinformation from Ukraine', *PsyArXiv*, https://doi.org/10.31219/osf.io/dnbz9.

Nygren T, Guath M, Axelsson CAW and Frau-Meigs D (2021) 'Combatting visual fake news with a professional fact-checking tool in education in France, Romania, Spain and Sweden', *Information*, 12(5): 201, https://doi.org/10.3390/info12050201.

Nygren T, Spearing ER, Fay N, Vega D, Hardwick I I, Roozenbeek J and Ecker U K H (2026) 'The seven roles of generative AI: Potential & pitfalls in combatting misinformation' *Behavioral Science & Policy*, 0(0), https://doi.org/10.1177/23794607261417815.

O'Neil M, Cunneen R and Deas M (2025) 'A remedy for epistemic pollution? public and professional reactions to fact-checking with Wikipedia in Australian classrooms', in Song CY, Thussu D and Margolin D (eds), *Checking the factcheckers: a global perspective*, Advances in Internationalizing Media Studies, Routledge.

O'Neil M, Heppner H and Ross A (2024) *Civic information literacy tools*, DCPC/NMRC, University of Canberra, https://doi.org/10.60836/smyk-8669.

Osborne J and Allchin D (2024) 'Science literacy in the twenty-first century: informed trust and the competent outsider', *International Journal of Science Education*, 47(15–16), 2134–2155, https://doi.org/10.1080/09500693.2024.2331980.

Osborne J and Pimentel D (2022) 'Science, misinformation, and the role of education', *Science*, 378(6617): 246–248, https://doi.org/10.1126/science.abq8093.

Park S, Notley T, Thomson TJ, Hourigan A and Dezuanni M (2024) *Online misinformation in Australia: adults' experiences, abilities, and responses*, University of Canberra and Western Sydney University, https://doi.org/10.60836/jpmm-dw04.

Park S, Walsh B and Su J (2023) *Libraries and media literacy education*, News Media Research Centre, University of Canberra, https://doi.org/10.25916/kahm-zr94.

Paynter J, Luskin-Saxby S, Keen D, Fordyce K, Frost G, Imms C, Miller S, Trembath D, Tucker M, Ecker U (2019) 'Evaluation of a template for countering misinformation: real-world autism treatment myth debunking', *PLoS ONE*, 14(1): e0210746, https://doi.org/10.1371/journal.pone.0210746.

Pennycook G, Berinsky AJ, Bhargava P, Lin H, Cole R, Goldberg B, Lewandowsky S, Rand DG (2024) 'Inoculation and accuracy prompting increase accuracy discernment in combination but not alone', *Nature Human Behaviour*, 8: 2330–2341, https://doi.org/10.1038/s41562-024-02023-2.

Pennycook G, Epstein Z, Mosleh M, Arechar AA, Eckles D and Rand DG (2021) 'Shifting attention to accuracy can reduce misinformation online', *Nature*, 592: 590–595, https://doi.org/10.1038/s41586-021-03344-2.

Pennycook G and Rand DG (2022) 'Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation', *Nature Communications*, 13: 2333, https://doi.org/10.1038/s41467-022-30073-5.

Pilditch TD, Roozenbeek J, Madsen JK and van der Linden S (2022) 'Psychological inoculation can reduce susceptibility to misinformation in large rational agent networks', *Royal Society Open Science*, 9(8): 211953, https://doi.org/10.1098/rsos.211953.

Porter E, Wood TJ, Broniatowski DA and Hosseini P (2024) 'Shoves, nudges and combating misinformation: evidence on a new approach', *Behavioural Public Policy*, 1–17, https://doi.org/10.1017/bpp.2024.42.

Prike T, Butler LH and Ecker UKH (2024) 'Source-credibility information and social norms improve truth discernment and reduce engagement with misinformation online', *Scientific Reports*, 14: 6900, https://doi.org/10.1038/s41598-024-57560-7.

Rojas C, Algra-Maschio F, Andrejevic M, Coan T, Cook J and Li YF (2024) 'Augmented CARDS: a machine learning approach to identifying triggers of climate change misinformation on Twitter', *arXiv preprint,* https://doi.org/10.48550/arXiv.2404.15673.

Romanova A, Rubinelli S and Diviani N (2024) 'Improving health and scientific literacy in disadvantaged groups: a scoping review of interventions', *Patient Education and Counseling*, 122: 108168, https://doi.org/10.1016/j.pec.2024.108168.

Roozenbeek J, Remshard M and Kyrychenko Y (2024) 'Beyond the headlines: on the efficacy and effectiveness of misinformation interventions', *Advances in Psychology*, 2: e24569, https://doi.org/10.56296/aip00019.

Roozenbeek J, van der Linden S, Goldberg B, Rathje S and Lewandowsky S (2022) 'Psychological inoculation improves resilience against misinformation on social media', *Science Advances*, 8(34): eabo6254, https://doi.org/10.1126/sciadv.abo6254.

Saari A and Säntti J (2017) 'The rhetoric of the "digital leap" in Finnish educational policy documents', *European Educational Research Journal*, 17(3): 442–457, https://doi.org/10.1177/1474904117721373.

Sandahl J (2020) 'Opening up the echo chamber: perspective taking in social science education', *Acta Didactica Norden*, 14(4): 24-sider, https://doi.org/10.5617/adno.8350.

Simchon A, Edwards M and Lewandowsky S (2024) 'The persuasive effects of political microtargeting in the age of generative artificial intelligence', *PNAS Nexus*, 3(2): pgae035, https://doi.org/10.1093/pnasnexus/pgae035.

Simchon A, Zipori T, Teitelbaum L, Lewandowsky S and van der Linden S (2026) 'A signal detection theory meta-analysis of psychological inoculation against misinformation', *Current Opinion in Psychology*, 67,102194, https://doi.org/10.1016/j.copsyc.2025.102194.

Smirnov I, Oprea C and Strohmaier M (2023) 'Toxic comments are associated with reduced activity of volunteer editors on Wikipedia', *PNAS Nexus*, 2(12): pgad385, https://doi.org/10.1093/pnasnexus/pgad385.

Steinsson S (2024) 'Rule ambiguity, institutional clashes, and population loss: how Wikipedia became the last good place on the internet', *American Political Science Review*, 118(1): 235–251, https://doi.org/10.1017/S0003055423000138.

Swire-Thompson B, Dobbs M, Thomas A and DeGutis J (2023) 'Memory failure predicts belief regression after the correction of misinformation', *Cognition*, 230: 105276, https://doi.org/10.1016/j.cognition.2022.105276.

Swire-Thompson B, Ecker UK, Lewandowsky S and Berinsky AJ (2020) 'They might be a liar but they're my liar: source evaluation and the prevalence of misinformation', *Political Psychology*, 41(1): 21–34, https://doi.org/10.1111/pops.12586.

Tay LQ, Hurlstone MJ, Kurz T and Ecker U (n.d.) 'Do prebunking and debunking protect against novel misinformation?', *PsyArXiv*, https://doi.org/10.31234/osf.io/w7gja.

Tay LQ, Lewandowsky S, Hurlstone MJ, Kurz T and Ecker UK (2023) 'A focus shift in the evaluation of misinformation interventions', *Harvard Kennedy School Misinformation Review*, 4(5), https://doi.org/10.37016/mr-2020-124.

Tay LQ, Lewandowsky S, Hurlstone MJ, Kurz T and Ecker UK (2024) 'Thinking clearly about misinformation', *Communications Psychology*, 2: 4, https://doi.org/10.1038/s44271-023-00054-5.

Thomas GR, Eng L, de Wolff JF and Grover SC (2013) 'An evaluation of Wikipedia as a resource for patient education in nephrology', in *Seminars in Dialysis*, 26(2): 159–163, https://doi.org/10.1111/sdi.12059.

Thomas J, Barraket J, Parkinson S, Wilson C, Holcombe-James I, Kennedy J, Mannell K and Brydon A (2021) *Measuring Australia's digital divide: the Australian digital inclusion index 2021*, RMIT, Swinburne University of Technology, and Telstra, DOI: 10.25916/phgw-b725.

Thorson E (2024) *How news coverage of misinformation shapes perceptions and trust*, Cambridge University Press, https://doi.org/10.1017/9781009488815.

Troy AS, Willroth EC, Shallcross AJ, Giuliani NR, Gross JJ and Mauss IB (2023) 'Psychological resilience: an affect-regulation framework', *Annual Review of Psychology*, 74: 547–576, https://doi.org/10.1146/annurev-psych-020122-041854.

Webber M, McNaughton S, Gerrard J, Meade S, Verdonk C (2023) *He Uru Kahikatea: building young people's resilience through media and information literacy and digital citizenship skills*, Office of the Prime Minister's Science Advisor, New Zealand.

Weisberg L, Kohnen A and Dawson K (2022) 'Impacts of a digital literacy intervention on preservice teachers' civic online reasoning abilities, strategies, and perceptions', *Journal of Technology and Teacher Education*, 30(1): 73–98, Society for Information Technology and Teacher Education, Waynesville, NC USA, https://doi.org/10.70725/926509jtmzpi.

Wen LM, Xu H, Jawad D, Buchanan L, Rissel C, Phongsavan P, Baur LA, Taki S (2021) 'Ethnicity matters in perceived impacts and information sources of COVID-19 among mothers with young children in Australia: a cross-sectional study', *BMJ Open*, 11(11): e050557, https://doi.org/10.1136/bmjopen-2021-050557.

Wineburg S, Breakstone J, McGrew S, Smith MD and Ortega T (2022) 'Lateral reading on the open internet: a district-wide field study in high school government classes', *Journal of Educational Psychology*, 114(5): 893-909, https://psycnet.apa.org/doi/10.1037/edu0000740.

Wineburg S and McGrew S (2019) 'Lateral reading and the nature of expertise: reading less and learning more when evaluating digital information', *Teachers College Record*, 121(11): 1–40, https://doi.org/10.1177/016146811912101102.

Wirtschafter V and Majumder S (2023) 'Future challenges for online, crowdsourced content moderation: evidence from Twitter's Community Notes', *Journal of Online Trust and Safety*, 2(1), https://doi.org/10.54501/jots.v2i1.139.

Yelbuz BE, Madan E and Alper S (2022) 'Reflective thinking predicts lower conspiracy beliefs: a meta-analysis', *Judgment and Decision Making*, 17(4): 720–744, https://doi.org/10.1017/S1930297500008913.

Yu W and Shen F (2024) 'Mapping verification behaviours in the post-truth era: a systematic review', *New Media and Society*, 26(3): 1703–1727, https://doi.org/10.1177/14614448231191138.