The Annual Dennis Moore Oration & 1962 awards

Truth, social. Where are we in the fight against misinformation?

with **Professor Paul Harrigan**

The Australian Computer Society Annual Dennis Moore Oration Dinner. The University Club of Western Australia, 15 October 2025.

Orator: **Professor Paul Harrigan**

Proceedings

Opening and welcome	Dr David Cook FACS CP
Introduction of opening speaker	Dr David Cook FACS CP
Opening address	Government representative
Introduction	1962 Prize and Medal Finalists Dr Bob Cross FACS CP
Presentation of 1962 Awards	Professor Alex Reid FACS CP
Introduction 1962 Educators	Dr Bob Cross FACS CP
Presentation of Educator Awards	Dr Brian von Konsky FACS CP
Sponsor welcome	The Testing Consultancy and Curtin University
Introduction of 2025 Orator	Dr David Cook FACS CP
Oration delivered	Professor Paul Harrigan
Vote of thanks	Dr Brian von Konsky FACS CP
Closing remarks	Dr David Cook FACS CP

Oration Organising Committee

Dr David Cook FACS CP

Dr Bob Cross FACS CP

Dr Brian von Konsky FACS CP

Professor Terry Woodings FACS

1962 Awards Judges

Dr Bob Cross FACS CP

Professor Tanya McGill FACS

Associate Professor Doina Olaru

Professor Terry Woodings FACS

We acknowledge with sadness the passing of Tony Watson FACS, a respected former judge of the Dennis Moore Oration and 1962 Awards. Tony's commitment, integrity, and passion for recognising excellence enriched our community and inspired many. He will be deeply missed and fondly remembered.

Supported by

Jerome Chiew MACS Snr CP (Cyber Security) IP3P MAISA, WA Branch Chair

Darren Roxburgh MACS, WA State Manager

Vicki Leddin MACS, WA Member Services & Events Manager

Annual Dennis Moore Oration Past Orators

Since 2012, to commemorate fifty years of digital computing in Western Australia, the WA Branch of the ACS has invited a distinguished scholar and researcher with a connection to WA to present a lecture on the leading edge of an important and emerging area of information and computer technology.

Previous Orators

Year	Orator
2024	Professor Andreas Wicenec
2023	Professor Tom Gedeon
2022	Associate Professor Vidy Potdar
2021	Associate Professor Doina Olaru
2020	No Oration held due to Covid restrictions
2019	Associate Professor Rachel Cardwell-Oliver
2018	Professor Jinbo Wang
2017	Professor Matthew Bellgard
2016	Dr Adrian Boeing
2015	Professor Svetha Venkatesh
2014	Professor Craig Valli
2013	Professor Ian Reid
2012	Professor Andrew Rohl

1962 Prize

From a suggestion of Dennis Moore (and with his strong support) 2012 also saw the setting up of an annual prize for the best graduating student in ICT from a WA university. Although the primary criteria are based on academic performance, the candidates are also judged on their ability to promote their ideas in computing and contribution so far.

Previous winners of the 1962 Prize are:

Year	Winner
2024	**Tom Sargent**, Curtin University
2023	**Shuang Li**, Murdoch University
2022	**David Adams & Yuval Berman**, University of Western Australia
2021	**Alistair Martin**, Murdoch University
2020	**Samual Heath**, University of Western Australia
2019	**Jarryd Wimbridge**, Edith Cowan University
2018	**Taaqif Peck**, University of Western Australia
2017	**Mark Shelton**, University of Western Australia
2016	**Dalibor Borkovic**, Murdoch University
2015	**Michael Martis**, University of Western Australia
2014	**Anthony Long**, Curtin University
2013	**Laurence Da Luz**, Edith Cowan University
2012	**Kevin Adnan**, Curtin University

The 1962 Prize finalists for 2025 in alphabetical order are:

Kayle Baker – Edith Cowan University

Bae Zane Hardman – Murdoch University

Alastair Kho – Curtin University

Thelisinghe Arachchige Piumi Vanessa Perera – Edith Cowan University

Nunzia Sorrentino – Murdoch University

Julia Szymanski – Edith Cowan University

1962 Medal

In 2019 the 1962 Awards was expanded to include a new award for the most outstanding candidate who completed Doctoral research (eg PhD) in Western Australia in the field of Information Technology and Computer Science.

Previous winners of the 1962 Medals are:

Year	Winner
2024	**Dr Sayma Shammi**, Murdoch University
2023	**Dr Manou Rosenberg**, University of Western Australia
2022	**Dr Uzair Nadeem**, University of Western Australia
2021	**Dr Naeha Sharif**, University of Western Australia
2020	**Dr Anupiya Nugaliyadde**, Murdoch University
2019	**Dr Qiuhong Ke**, University of Western Australia

The 1962 Medal finalists for 2025 in alphabetical order:

Dr Maira Alvi – The University of Western Australia

Dr Kulsoom Saima Bughio – Edith Cowan University

Dr Matthew George Gaber – Edith Cowan University

Dr A S M Mahmudul Hasan – Murdoch University

Dr Khondoker Ziaul Islam – Murdoch University

Dr Duc Minh Chau Nguyen – The University of Western Australia

Dr Sumaiya Pervaiz – Curtin University

Dr Afsah Saleem – Edith Cowan University

Dr Sania Zahan – The University of Western Australia

1962 Educator Recognition

New in 2023, the 1962 awards will recognise teachers and lecturers that have received awards from other organisations. These are the dedicated people who make the other awards possible.

The 1962 Educators for 2025 being recognised are

Petra Trinke - Belridge Secondary College
Jan Kornweibel - CF FACS HLM CP, Pearcey Foundation

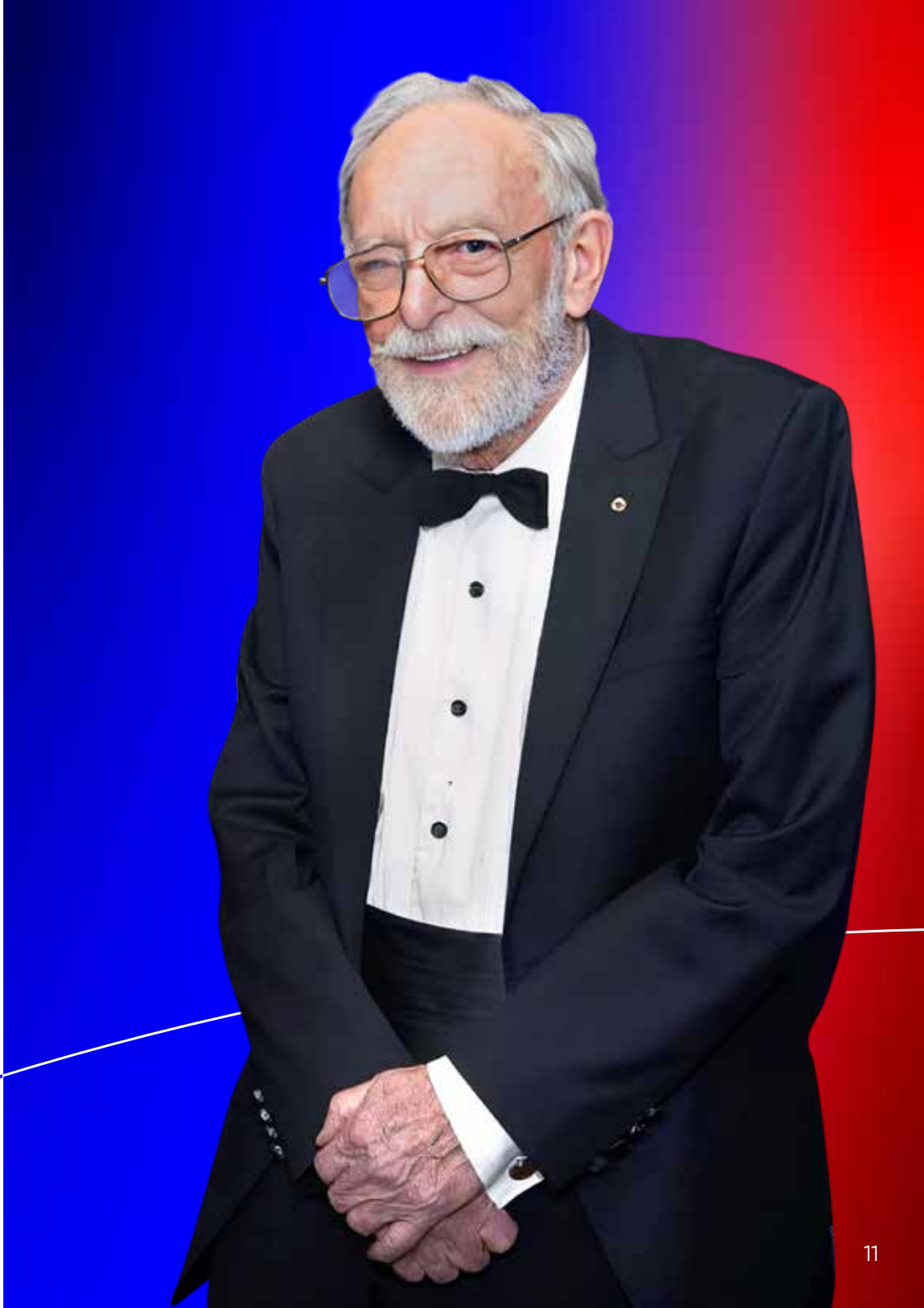
Past Educators Recognised were:

2024

Brett Clarke MACS CP – Catholic Education

2023

Donna Buckley - John Curtis College of the Arts
Michelle Chomiak - St Marks Community School
Dr Michelle Ellis - Edith Cowan University School of Science
Vinicius Madeiros - National Institute of Technology
Dr Chris McDonald - CSSE, University of Western Australia



Dennis Moore AM MA (Cantab) FACS:

Dennis Moore was born in NSW in 1937. He was educated on scholarships at The King's School, Parramatta where he was captain and dux of the school, and at Queens' College Cambridge where he graduated in 1958 in mathematics.

After a period with commerce and industry in computing and operations research in NSW, he pioneered computing in Western Australia, installing the first computer at UWA in 1962. He introduced WA's first computing qualification – the DipNAAC – at UWA.

In 1965, he was responsible for the purchase and installation of the DEC PDP-6. This was the world's first commercial installation of a time-shared computer and Australia's first high precision graphics device.

He was foundation president of the WA Computer Society, which later merged with the Australian Computer Society, becoming the first WA Branch Chairman. He was Director of the Western Australian Regional Computing Centre in the sixties and seventies. This provided computing services to CSIRO and State Government Departments as well as the University.

He was executive director of Government Computing for WA from 1978 to 1984. During this period he promoted the development of inter-departmental systems and was closely associated with the development of the WA Land Information System and the WA Technology Park. This was followed by a two year stint managing a computer company in Malaysia, including a consultancy to the Sarawak Government.

He then undertook research in RAN DATA, an encryption company which he had helped establish, and was appointed foundation Head of School of Computing at Curtin University of Technology in 1987. From 1998 to 2002 he was Director of Academic Planning at Curtin. From 1995 to 1999 he was Chair of the State Government's Information Policy Council.

Dennis Moore was elected a Fellow of the Australian Computer Society in 1970 and was made a Member of the Order of Australia for services to Information Technology in 1997. He retired in 2002 and was made an Honorary Life Member of the ACS in 2014.

Professor Paul Harrigan

Deputy Dean and Professor of Marketing in UWA Business School at The University of Western Australia

Biography

Professor Paul Harrigan is Deputy Dean and Professor of Marketing in UWA Business School at The University of Western Australia (UWA). He is also an Adjunct Professor at IESEG School of Management in France, and Editor-in-Chief of the Australasian Marketing Journal.

Paul has a PhD from Ulster University in the UK (2008), was previously a Lecturer in Marketing at The University of Southampton (2008-12), and has been at UWA since 2012. Paul was also Vice-President of the Australasian and New Zealand Marketing Academy (ANZMAC) 2019-2024.

Paul's research expertise runs across marketing and information systems, specifically digital transformation and misinformation. He has published his research in over 60 international journal articles and books, and presented it at over 60 international conferences.

Current projects investigate the use of debunking and prebunking methods in combatting misinformation on social media, the impact of slow motion video on social media follower engagement, and how to reduce the harm of social media ads on users' body image.

Truth, social. Where are we in the fight against misinformation?

Professor Paul Harrigan

Abstract

The rise of misinformation has become an obvious global concern. We examine why it gets shared on social media. In Study 1, we collected data on 10,201 misinformation narratives and 978 accurate information narratives that were published in more than a quarter million social media posts shared more than 100 million times. The results show that misinformation tends to be more readable and garners higher engagement than accurate information. In Study 2, we complemented our field study with an online experiment in which we found that the perceived cognitive ease (i.e., the perceived ease of processing information) is an underlying mechanism that drives the relationship between readability and sharing. In Study 3, we conducted another online experiment, which shows that cognitive reflection (i.e., people's tendency to engage in effortful thinking) can also influence the effect of perceived cognitive ease and sharing. In Study 4, we show that a social media advertisement that prompts people to think more (that is, engage in effortful thinking) before making a sharing decision reduces misinformation sharing. Our research advances our understanding of how misinformation spreads and lays the groundwork for how to combat it, moving into the use of machine learning.

Keywords: Misinformation, Readability, Sharing, Cognitive Ease, Cognitive Reflection, Social Media, Machine learning.

Introduction

The rise of misinformation on social media has become a global concern. While these platforms have enabled unprecedented connectivity, they have also accelerated the spread of falsehoods, conspiracy theories, and half-truths (Vosoughi et al., 2018; Weismueller et al., 2024). For example, during the COVID-19 pandemic, circulating misinformation had dire consequences, undermining public health efforts and threatening lives (Roozenbeek et al., 2020). Further, misinformation also threatens democratic processes, especially during elections (Brennan Center, 2024). With the emergence of generative AI, the volume and reach of misinformation have grown even further (Ryan-Mosley, 2023). The World Economic Forum now ranks misinformation as the top short-term global threat, ahead of climate change (World Economic Forum, 2024).

A fundamental question remains: why do people share misinformation, particularly on social media? Research points to factors like confirmation bias, lack of critical thinking, and message design (Del Vicario et al., 2016; Pennycook & Rand, 2019; King & Wang, 2021). Readability, referring to how easy content is to process, may also play a role (Pancer et al., 2019) as users tend to prefer simple, "bitesize" content in overloaded digital environments.

However, little is known about whether misinformation tends to be designed in a way that makes it more readable than accurate information, and whether this influences engagement. Misinformation often avoids complex citations and uses emotionally charged, simplified language (Damstra et al., 2021; Mourão &

Robertson, 2019). This makes it more accessible and appealing, especially to users who avoid effortful thinking (Pennycook & Rand, 2019), potentially increasing its spread.

Therefore, we ask:

RQ1. Is misinformation more readable and shared than accurate information?

RQ2. Does perceived cognitive ease (i.e., the extent to which people perceive the information to require little cognitive effort to process) mediate the relationship between readability and sharing?

If users are more likely to share misinformation due to readability and cognitive ease, nudging them toward more thoughtful decision-making may help reduce its spread. Nudges, or subtle environmental cues, have been shown to influence behavior and reduce misinformation sharing (Hwang & Lee, 2024; Moravec et al., 2020). An intervention, such as a social media ad prompting users to think before sharing, could encourage more deliberate and effortful reasoning. This may shift attention away from heuristic cues like readability and promote better evaluation of content accuracy, ultimately reducing misinformation dissemination. Thus, we also ask:

RQ3. Can an intervention reduce the impact of perceived cognitive ease on sharing and overall misinformation sharing?

To answer our research questions, we conducted four studies. In Study 1, we collected over 10,000 URLs linking to news and blog articles rated as either misinformation or accurate by independent fact-checkers working with Facebook. Using CrowdTangle, we gathered over 400,000 Facebook posts linking to these URLs, which had received more than 100 million shares. We analyzed the readability and engagement (i.e., number of shares) of each post.

In Studies 2 and 3, we conducted experiments exposing participants to Facebook posts containing misinformation with varying levels of readability. We measured perceived cognitive ease, cognitive reflection (the ability to engage in effortful thinking), and sharing intentions. In Study 4, we tested the effectiveness of a Facebook ad (intervention) designed to reduce reliance on cognitive ease and lower misinformation sharing.

Background

Given the harmful impact of misinformation, research has increasingly focused on understanding why people share it (see Table 1; Chen et al., 2023; King & Wang, 2021; Miller et al., 2024). Studies show that personal factors, such as trust in online information, perceived information overload, and lack of analytical thinking, can drive unverified sharing (Laato et al., 2020; Pennycook & Rand, 2019). However, fewer studies have examined how content characteristics, such as readability, influence misinformation sharing.

To address this gap, we propose that misinformation is often more readable than accurate information, and that this readability drives sharing. In fast-paced digital environments, users are exposed to a constant stream of headlines and snippets. With limited time and cognitive resources, they tend to prefer content that is easy to process. Readability reduces cognitive effort, making content more appealing and more likely to be shared.

This preference for cognitive ease, which is especially relevant in social media contexts, means that users may share misinformation because they find it easier to understand. Our model highlights readability as a key factor in misinformation spread, complementing existing research on user traits and offering new insights into how content design influences online behavior.

Table 1. Selected overview of information systems (IS) literature on misinformation sharing

Authors (Year)	Method	Focus	Key Findings
Chen et al. (2023)	Systematic Review	Drivers of misinformation sharing	• A systematic and structured overview of the factors that influence the spread of misinformation by analyzing the four vital elements of information communication, namely, source, message, context, and receiver • A summary of the current state of research on strategies against the spread of misinformation on social media from various perspectives and discuss their advantages, disadvantages, and effectiveness.
Laato et al. (2020)	Online Survey	Drivers of misinformation sharing	• A person's trust in online information and perceived information overload are strong predictors of unverified information sharing.
Weismueller et al. (2024)	Field Study and Online Experiment	Drivers of misinformation sharing	• Misinformation and extreme partisan information are associated with higher levels of negative emotions and greater engagement than accurate information and non-extreme partisan information. • Exposure to misinformation and extreme partisan information elicits stronger negative emotions than exposure to accurate information and non-extreme partisan information. These negative emotions, in turn, contribute to attitude polarisation.
King and Wang (2021)	Field Study	Drivers of misinformation sharing	• People are more inclined to retweet misinformation that is negative as compared to misinformation that is positive or neutral. • People are more likely to retweet misinformation as compared with authentic news.
Miller et al. (2024)	Online Experiment	Drivers of misinformation sharing	• Users with higher literacy in new media were less likely to engage with disinformation featuring low veracity language. • Users with higher literacy in traditional media were more likely to engage with disinformation, regardless of the article's linguistic veracity.
Kim and Dennis (2019)	Online Experiment	Combatting misinformation sharing	• The presentation format of highlighting the source in (fake) news posts had a main effect; it made users more skeptical of all articles, regardless of the source's credibility. For unknown sources, low source ratings had a direct effect on believability. • Believability, in turn, influenced the extent to which users would engage with the article (e.g., read, like, comment, and share).
Moravec et al. (2020)	Online Experiment	Combatting misinformation sharing	• Fake news flags were effective in reducing believability and engagement with fake news. • Awareness training on the meaning of the flags increased the effectiveness of the System 2 intervention but not the System 1 intervention.
Hwang and Lee (2024)	Field Study	Combatting misinformation sharing	• After Twitter's nudging policy introduction, a news article that contains misinformation is less likely to start a diffusion process on Twitter. • The observed reduction is driven by the decrease both in original tweet and in those resharing the misinformation, although the reduction is more significant in resharing posts.

Theoretical Development

The role of falsehood and readability on sharing

Research shows that misinformation often garners more engagement than accurate information (King & Wang, 2021; Vosoughi et al., 2018). One reason is that misinformation tends to be more negative and sensationalistic, which increases its emotional appeal and shareability (Tucker et al., 2018). While accurate content can also be sensational, these traits are more common in misinformation.

Another key factor is readability. On social media, users scroll quickly and prefer content that is easy to digest (Pancer et al., 2019). Readability, measured by tools like the Automated Readability Index (ARI), reflects how simple a message is to understand. Misinformation often simplifies complex topics, avoids technical language, and

omits citations, making it more accessible and engaging (Carrasco-Farré, 2022; Mourão & Robertson, 2019). This increased readability makes misinformation more accessible and likely to be shared, as users prefer information that is more readable (Pancer et al., 2019). Thus, we argue that:

H1. Misinformation tends to be more readable than accurate information, which, in turn, increases sharing.

To explain why readability increases sharing we draw from cognitive load theory (Sweller, 2010), to argue that more readable content will be perceived as easier to process. According to cognitive load theory (Sweller, 2010), human working memory has limited capacity, and how information is presented can either ease or strain that capacity (Sepp et al., 2019). Content features, such as readability, play a key role in shaping how cognitively demanding content feels (Klepsch et al., 2017; Zumbach & Mohraz, 2008). Readable content creates processing fluency and reduces extraneous cognitive load, not only making content easier to process but also making it feel more trustworthy (Alter & Oppenheimer, 2009; Reber et al., 2004). As a result, users are more likely to engage with and spread misinformation, not necessarily because they believe it, but because it is easier to process.

This aligns with processing fluency theory, which suggests that people prefer information that feels easy to process (Alter & Oppenheimer, 2009). On social media, where users are bombarded with information, they tend to favor content that is quickly understood and requires minimal cognitive effort (Davis et al., 2019; Yu et al., 2024). Consequently, more readable posts are more likely to be shared.

Particularly in fast-paced, hedonic environments like social media, people's preference for cognitive ease means that readability becomes a powerful heuristic. When content is clear, structured, and simple, it feels more fluent and trustworthy, even if it is inaccurate (Reber et al., 2004). Readability not only enhances comprehension but also increases the likelihood of sharing, contributing to the spread of misinformation. Thus:

H2. More readable information tends to be perceived as easier to process than less readable information, which, in turn, increases sharing.

The moderating role of cognitive reflection

Cognitive reflection, defined as the ability to override intuitive, automatic responses in favor of deliberate reasoning, moderates the link between cognitive ease and sharing behavior (Evans & Stanovich, 2013; Kahneman, 2011). It reflects one's capacity to resist System 1 thinking and engage System 2 thinking (Pennycook & Rand, 2018, 2019). Most social media users operate in a hedonic, System 1-driven mindset (Moravec et al., 2020), preferring easily processed content. In contrast, individuals with higher cognitive reflection are more comfortable engaging with effortful information, as they can suppress the tendency to rely on mental shortcuts. This makes them less susceptible to sharing misinformation based on readability alone. Thus, we argue that:

H3. The positive effect of high perceived cognitive ease on sharing is stronger when people have low cognitive reflection ability as opposed to high cognitive reflection ability.

The moderating role of an intervention

If cognitive reflection reduces reliance on heuristics like readability when deciding to share content, an important question arises: can this tendency be influenced? Since people often share posts based on ease of processing rather than factual accuracy, encouraging more deliberate thinking could help curb misinformation. One promising approach is the use of interventions. Interventions can act as subtle cues that influence behavior, otherwise known as nudges (Hwang & Lee, 2024; Moravec et al., 2020). Examples include warning labels or training users to detect false claims (Pennycook & Rand, 2022; Roozenbeek et al., 2022). These nudges can prompt users to engage in System 2 thinking, promoting more effortful reasoning and reducing reliance on readability as a sharing cue. By shifting attention from how easy content is to process toward its truthfulness and consequences, such interventions can help reduce the spread of misinformation and foster more responsible information sharing on social media. Thus, we argue:

H4. The positive effect of high perceived cognitive ease on sharing is weaker when people are nudged to engage in cognitive reflection.

The full research model is presented in Figure 1.

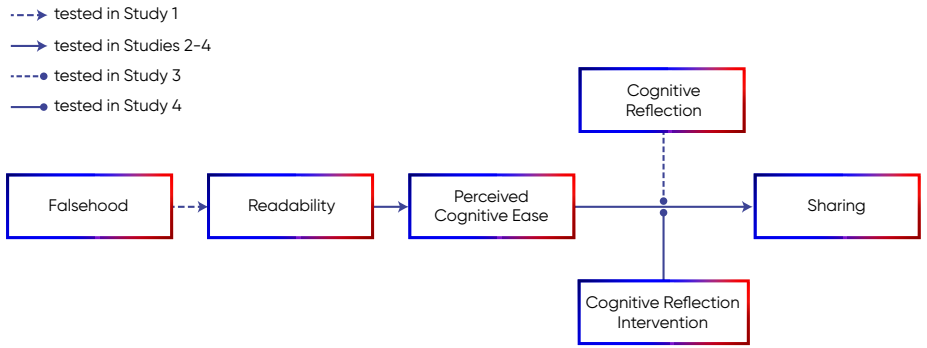


Figure 1. Research model

Empirical Studies

To empirically examine parts of the proposed model, we collected data from Facebook. Facebook is among the most used social media platforms in the world, and there is evidence that, like many other social media platforms, it is subject to the spread of misinformation (Guess et al., 2019). Our dataset contained 10,201 web pages (URLs) shared on Facebook starting January 1, 2017, up to and including December 31, 2021. The dataset contains information on whether the webpage (URL) shared on Facebook during that timeframe has been fact-checked by at least one of the independent fact-checking websites that Facebook works with (Meta, 2021). URLs rated as 'false' from fact-checking websites or reported as false information were identified as misinformation. URLs rated as 'true' from fact-checking websites were identified as accurate information. We used this dataset for an initial field study (Study 1) and then followed up with online experiments (Studies 2 – 4).

In Study 1, we used the list of URLs that were identified as misinformation and accurate information to collect 378,488 (misinformation) and 32,536 (accurate information) Facebook posts that linked to those URLs (see Figure 2 and 3 for examples). We then analyzed their readability and engagement data, as well as empirically examined the proposed relationships. Once we confirmed that misinformation was indeed more readable than accurate information and that readability served as an underlying mechanism that drives higher engagement with misinformation when compared to accurate information, we conducted an online experiment in Study 2. We complemented our findings from the field and explored how people perceive misinformation of varying readability, which could not be done in the field study (see Figures 4 and 5 for stimuli). In Study 3, we replicated Study 2 in another context while also exploring participants' natural tendency to engage in effortful thinking (see Figures 6 and 7 for stimuli; Mosleh et al., 2021; Pennycook & Rand, 2019; Roozenbeek, Maertens, et al., 2022). Lastly, in Study 4, we replicated Study 3 but randomly exposed half of the participants to an intervention in the form of a Facebook ad (see Figure 8). The intervention encourages people to think (that is, engage in effortful thinking) before making a sharing decision. In so doing, we explored the effectiveness of an intervention in mitigating the spread of misinformation. Table 1 summarizes the results in relation to each hypothesis.

Figure 2. Example of Facebook post including URL

Figure 3. Example of misinformation story from URL

Figure 4. High readability stimulus

Figure 5. Low readability stimulus

Figure 6. High readability stimulus

Figure 7. Low readability stimulus

Figure 9. Intervention

Table 1. Hypotheses testing

	Hypotheses	Supported/not supported	Significance (p)
1	Misinformation tends to be more readable than accurate information, which, in turn, increases sharing.	Supported	< .001
2	More readable information tends to be perceived as easier to process than less readable information, which, in turn, increases sharing.	Supported	p < .05, < .001
3	The positive effect of high perceived cognitive ease on sharing is stronger when people have low cognitive reflection ability as opposed to high cognitive reflection ability.	Supported	< .05
4	The positive effect of high perceived cognitive ease on sharing is weaker when people are nudged to engage in cognitive reflection.	Supported	< .01

General Discussion

Our findings reveal several key insights:

1. Misinformation posts were more readable and more widely shared than accurate posts.
2. Readability increased perceived cognitive ease, which in turn increased sharing intentions.
3. Cognitive reflection moderated this relationship, where individuals with higher reflection ability were less influenced by cognitive ease.
4. An intervention prompting users to think before sharing reduced misinformation sharing and reliance on cognitive ease.

These findings contribute to the literature on misinformation and social media engagement by highlighting readability as a key content characteristic influencing sharing behavior. While prior research has focused on user traits and emotional content, our study emphasizes how linguistic features such as readability can drive engagement with misinformation.

We also show that cognitive reflection acts as a boundary condition: individuals who avoid effortful thinking are more likely to share readable misinformation. Finally, we extend IS research on interventions by demonstrating that nudges encouraging critical thinking can reduce misinformation spread.

Practical Implications

Misinformation is a major societal threat in the digital age, contributing to polarization, violence, and even death (van der Linden et al., 2024). While platforms have introduced tools like detection algorithms and fact-checking, addressing misinformation requires broader collaboration. As some platforms shift strategies, for example, Meta replacing fact-check labels with community notes (Watt et al., 2025), the role of policymakers and non-profits becomes increasingly important.

Platforms could be encouraged to monitor readability as a potential signal of misleading content. Algorithms might flag highly readable but suspicious posts for review. Educational initiatives, integrated within school and university courses, could raise awareness about how readability and cognitive biases influence sharing behavior.

Digital literacy campaigns could teach users to critically assess content that feels easy to process, highlighting the risks of relying on heuristic cues like simplicity or emotional appeal. Platforms might also create information hubs to explain how cognitive ease can lead to misinformation spread and societal harm.

Finally, interventions such as nudges or ads promoting critical thinking can reduce impulsive sharing. Policymakers could mandate such features, while governments might fund public campaigns to encourage thoughtful engagement. Ultimately, individuals must take responsibility for verifying content and resisting the urge to share based on readability alone.

Future Research

Much work remains to be done if we are to truly understand how and why misinformation spreads on social media and, more important still, what to do about it. Our latest research lies in the use of advanced machine learning algorithms as an additional tool for detecting misinformation (e.g., Zhang et al., 2019; Zhou & Zafarani, 2020). We again leverage the real-life Meta misinformation dataset to set up a large-scale benchmarking experiment using encoder-only (BERT, DistilBERT, ModernBERT), encoder-decoder (T5, BART) and decoder-only (GPT-4o-mini) models, in their general-purpose and domain-adapted state. Through this research, we investigate the comparative effectiveness of these tools in detecting misinformation. We also propose a framework to explain to decision makers whether and why a post contains misinformation. This research contributes to explainable artificial intelligence by introducing a novel framework that integrates prediction with interpretable reasoning, thus addressing the black-box nature of powerful language models.

This is an abridged version of papers that have been published or are under review, in collaboration with Dr Jason Weismueller, Lecturer in Marketing, The University of Western Australia.

References

1. Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. Personality and Social Psychology Review, 13(3), 219-235.

2. Brennan Center. (2024). Election Rumors in 2024. https://www.brennancenter.org/series/election-rumors-2024

3. Carrasco-Farré, C. (2022). The fingerprints of misinformation: how deceptive content differs from reliable sources in terms of cognitive effort and appeal to emotions. Palgrave Macmillan, 9(1), 1-18.

4. Chen, S., Xiao, L., & Kumar, A. (2023). Spread of misinformation on social media: What contributes to it and how to combat it. Computers in Human Behavior, 141, 107643. https://doi.org/10.1016/j.chb.2022.107643

5. Damstra, A., Boomgaarden, H. G., Broda, E., Lindgren, E., Strömbäck, J., Tsfati, Y., & Vliegenthart, R. (2021). What does fake look like? A review of the literature on intentional deception in the news and on social media. Journalism Studies, 22(14), 1947-1963.

6. Davis, S. W., Horváth, C., Gretry, A., & Belei, N. (2019). Say what? How the interplay of tweet readability and brand hedonism affects consumer engagement. Journal of Business Research, 100, 150-164. https://doi.org/10.1016/j.jbusres.2019.01.071

7. Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. Proceedings of the National Academy of Sciences, 113(3), 554-559. doi:10.1073/pnas.1517441113

8. Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition:Advancing the Debate. Perspectives on Psychological Science, 8(3), 223-241. 10.1177/1745691612460685

9. Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances, 5(1), eaau4586.

10. Hwang, E. H., & Lee, S. (2024). A Nudge to Credible Information as a Countermeasure to Misinformation: Evidence from Twitter. Information Systems Research, 0(0), null. 10.1287/isre.2021.0491

11. Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

12. Kim, A., & Dennis, A. R. (2019). Says who? The effects of presentation format and source rating on fake news in social media. MIS Quarterly, 43(3), 1025-1039.

13. King, K. K., & Wang, B. (2021). Diffusion of real versus misinformation during a crisis event: a big data-driven approach. International Journal of Information Management, 102390.

14. Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load [Original Research]. Frontiers in Psychology, 8. 10.3389/fpsyg.2017.01997

15. Laato, S., Islam, A. K. M. N., Islam, M. N., & Whelan, E. (2020). What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? European Journal of Information Systems, 29(3), 288-305. 10.1080/0960085X.2020.1770632

16. Meta. (2021). How Meta's third-party fact-checking program works. https://www.facebook.com/formedia/blog/third-party-fact-checking-how-it-works

17. Miller, S., Menard, P., & Bourrie, D. (2024). I'm not fluent: How linguistic fluency, new media literacy, and personality traits influence fake news engagement behavior on social media. Information & Management, 61(2), 103912. https://doi.org/10.1016/j.im.2023.103912

18. Moravec, P. L., Kim, A., & Dennis, A. R. (2020). Appealing to Sense and Sensibility: System 1 and System 2 Interventions for Fake News on Social Media. Information Systems Research, 31(3), 987-1006. 10.1287/isre.2020.0927

19. Mosleh, M., Pennycook, G., Arechar, A. A., & Rand, D. G. (2021). Cognitive reflection correlates with behavior on Twitter. Nature Communications, 12(1), 921.

20. Mourão, R. R., & Robertson, C. T. (2019). Fake News as Discursive Integration: An Analysis of Sites That Publish False, Misleading, Hyperpartisan and Sensational Information. Journalism Studies, 20(14), 2077-2095. 10.1080/1461670X.2019.1566871

21. Pancer, E., Chandler, V., Poole, M., & Noseworthy, T. J. (2019). How readability shapes social media engagement. Journal of Consumer Psychology, 29(2), 262-270.

22. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. Psychological Science, 31(7), 770-780. 10.1177/0956797620939054

23. Pennycook, G., & Rand, D. G. (2018). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. Journal of personality.

24. Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. Cognition, 188, 39-50. https://doi.org/10.1016/j.cognition.2018.06.011

25. Pennycook, G., & Rand, D. G. (2022). Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. Nature Communications, 13(1), 2333. 10.1038/s41467-022-30073-5

26. Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? Personality and Social Psychology Review, 8(4), 364-382.

27. Roozenbeek, J., Maertens, R., Herzog, S. M., Geers, M., Kurvers, R., Sultan, M., & van der Linden, S. (2022). Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking. Judgment and Decision Making, 17(3), 547-573.

28. Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L., Recchia, G., Van Der Bles, A. M., & Van Der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. Royal Society open science, 7(10), 201199.

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

30. Ryan-Mosley, T. (2023). How generative AI is boosting the spread of disinformation and propaganda. https://www.technologyreview.com/2023/10/04/1080801/generative-ai-boosting-disinformation-and-propaganda-freedom-house/

31. Sepp, S., Howard, S. J., Tindall-Ford, S., Agostinho, S., & Paas, F. (2019). Cognitive Load Theory and Human Movement: Towards an Integrated Model of Working Memory. Educational psychology review, 31(2), 293-317. 10.1007/s10648-019-09461-9

32. Sweller, J. (2010). Cognitive load theory: Recent theoretical advances.

33. van der Linden, S., Ecker, U. K. H., & Lewandowsky, S. (2024). Misinformation is a threat to society – let's not pretend otherwise. https://blogs.lse.ac.uk/impactofsocialsciences/2024/10/08/misinformation-is-a-threat-to-society-lets-not-pretend-otherwise/

34. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146-1151.

35. Watt, N., Riedlinger, M., & Montaña-Niño, S. (2025). Meta is abandoning fact checking – this doesn't bode well for the fight against misinformation. https://theconversation.com/meta-is-abandoning-fact-checking-this-doesnt-bode-well-for-the-fight-against-misinformation-246878

36. Weismueller, J., Gruner, R. L., Harrigan, P., Coussement, K., & Wang, S. (2024). Information sharing and political polarisation on social media: The role of falsehood and partisanship. Information Systems Journal, 34(3), 854-893.

37. World Economic Forum. (2024). Global Risks 2024: Disinformation Tops Global Risks 2024 as Environmental Threats Intensify. https://www.weforum.org/press/2024/01/global-risks-report-2024-press-release/

38. Yu, J., Hong, W. C. H., & Egger, R. (2024). The Art of Post Captions: Readability and User Engagement on Social Media. Journal of Travel Research, 0(0), 00472875241228822. 10.1177/00472875241228822

39. Zhou, X., & Zafarani, R. (2021). A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. ACM Computing Surveys, 53(5), 1–40. https://doi.org/10.1145/3395046

40. Zumbach, J., & Mohraz, M. (2008). Cognitive load in hypermedia reading comprehension: Influence of text type and linearity. Computers in Human Behavior, 24(3), 875-887. https://doi.org/10.1016/j.chb.2007.02.015

Thank you to our sponsors

Innovation sponsor

ttc

ttcglobal.com

Cocktail sponsor

Curtin University

Powering Australia's technology brilliance

www.acs.org.au