1

2

3

4

5

6

7   **Thinking Clearly About Misinformation**

8

9   Li Qian Tay[1,2], Stephan Lewandowsky[3,1,4], Mark J. Hurlstone[5], Tim Kurz[1],

10   and Ullrich K. H. Ecker[1,6]

11

12   [1]School of Psychological Science, University of Western Australia, Perth, Australia

13   [2]School of Medicine and Psychology, Australian National University, Canberra, Australia

14   [3]School of Psychological Science, University of Bristol, Bristol, UK

15   [4]Department of Psychology, University of Potsdam, Potsdam, Germany

16   [5]Department of Psychology, Lancaster University, Lancaster, UK

17   [6]Public Policy Institute, University of Western Australia, Perth, Australia

18

19

20

21

22

23

24

25   Word Count: 1797 (excl. preface and display items)

26 **There is concern that many social problems in Western societies have been caused by**

27 **misinformation. However, some researchers argue that misinformation is merely a**

28 **symptom of such problems. We argue that (1) this is a false dichotomy, (2)**

29 **misinformation has had clear impacts, and (3) researchers should consider the different**

30 **dimensions of misinformation when making such evaluations.**

31   In Western societies, misinformation concern is at an all-time high. Recently,

32 however, debate has ensued regarding the level of concern that is warranted. Some

33 researchers note the potential for misinformation to incur significant costs on individuals and

34 societies, and call for interventions to reduce misinformation susceptibility and impacts[1,2].

35 Others warn against "alarmist" narratives, arguing that misinformation exerts only limited

36 influence over beliefs and behaviours. This view proposes that problematic behaviours, such

37 as vaccine hesitancy, are caused by systemic socio-economic and psycho-social issues, and

38 thus calls for interventions to target those societal issues rather than misinformation creation

39 and consumption, which represent only "symptoms" of these deeper issues[3]. Similarly,

40 assuming low prevalence of misinformation, researchers have argued that interventions

41 should focus on increasing trust in factual information[4].

42   A principled way to resolve these contradicting analyses is needed, to better inform

43 policies and minimise the risk of enshrining a problematic status quo or investing resources

44 to address a perhaps negligible problem. Here, we engage with two key questions: (1) Is

45 misinformation a "symptom" rather than a "cause"? (2) Is misinformation consumption low

46 and therefore not a reason for concern? In answering these questions, we argue that

47 misinformation has had clear impacts; that depending on individual and contextual factors, it

48 can be both a symptom and a cause; and that its multidimensionality (i.e., topic, type, and

49 depth of dissemination) ought to be more fully considered when making such evaluations.

50   **A Call for Causal Clarity**

51   Societal issues can shape individuals' beliefs and produce problematic behaviours.

52   Behaviours such as vaccine hesitancy and climate-change denial have been facilitated by

53   factors such as populism, inequality, disenfranchisement, political polarisation, and the

54   concentration of media ownership[5]. These factors are amplified by low institutional trust,

55   which is a wicked problem because even if many institutions are generally trustworthy, some

56   politicians, scientists, media outlets, and corporations have engaged in unethical behaviours

57   that do warrant scepticism.

58   Yet, even if the misinformation problem is symptomatic of such deeper issues, this

59   does not negate the fact that "symptoms" can cause outcomes of their own. To illustrate: A

60   factor such as inequality might increase the "symptom" of misinformation susceptibility,

61   while misinformation itself might cause belief changes or behaviours (unrelated to inequality)

62   in a causal chain; alternatively, a factor such as polarization or institutional distrust might

63   causally affect misinformation susceptibility, which in turn might further entrench

64   polarization or distrust in a vicious cycle.

65   A counterfactual perspective can provide further clarification: causation is essentially

66   the difference between a world in which a putative cause is present and a counterfactual

67   world in which all is equal except for the absence of the cause. Thus, if misinformation were

68   merely a "symptom," then nothing in the world would change if all misinformation were to

69   disappear. This is clearly implausible. Observational and experimental studies have

70   demonstrated that misinformation can causally alter beliefs and behaviours[1,6,7,8,9], even

71   though measurement of misinformation impacts is often impeded by ethical considerations

72   (e.g., exposing individuals to potentially harmful misinformation) or lack of access to

73   relevant data (e.g., historical or transnational data; data from social-media platforms or closed

74   channels such as offline communications and encrypted chat applications). Indeed, in a

75 counterfactual world without any misinformation, false beliefs could only emerge via

76 spontaneous generation. Such spontaneous generation is not uncommon (e.g., stereotypes and

77 superstitions can result from social processes or illusory correlations). However, it would be

78 inadequate as an all-encompassing explanation for the spread of false beliefs that go beyond

79 individuals' immediate experiences or observations. For example, the widespread false belief

80 that the mumps-measles-rubella (MMR) vaccine causes autism would be unlikely to gain

81 traction had fraudulent MMR-vaccine research not received high-profile media coverage.

82   Critically, the counterfactual perspective can account for multicausality. Consider a

83 situation in which an individual is influenced by a claim that a vaccine is harmful. Both the

84 misinformation and the existing susceptibility of the individual (e.g., low trust in science) are

85 causal factors, if, without either, the individual would not have been misinformed to the same

86 extent (e.g., formed a weaker misconception). Whether the misinformation or the existing

87 susceptibility is a better explanation then depends on their relative prevalence and the

88 probability of sufficiency. For example, in case of a fire breaking out after an individual

89 lights a match, match-lighting may be a better explanation for the fire than the presence of

90 oxygen, because oxygen is more prevalent than match-lighting and the individual lighting the

91 match should have anticipated the presence of oxygen (such analyses are used in legal

92 reasoning to determine damages)[10]. Thus, even if institutional mistrust can partially explain

93 some individuals' tendency to be affected by vaccine misinformation (alongside other

94 individual-specific factors such as perceived plausibility, worldview congruence, utility for

95 behaviour justification, etc.), it does not absolve the causal responsibility of misinformants,

96 nor negate the potential effects of vaccine misinformation on public health.

97   One way to capture the complexity of such causal networks is through directed

98 acyclic graphs, as shown in Figure 1. This approach can also illustrate how existing research

99 has focused on specific direct effects within limited timeframes, often neglecting more

100   indirect causal factors and potentially important context variables. For example, the existing

101   misinformation literature is biased towards a liberal-democratic, Western framework and has

102   largely overlooked the potential influence of environmental context factors such as state

103   capacity and the presence of ethnic conflicts or historical grievances, which may co-

104   determine misinformation impacts.

105   **Figure 1**

106   *Directed Acyclic Graphs Illustrating Causal Networks of Misinformation Effects*

107   [Figure 1]

108   *Note.* (a) Directed acyclic graphs are graphical causal models characterized by nodes
109   representing variables and edges representing direct causal effects. In the example, both low
110   institutional trust and misinformation can cause outcomes such as vaccine hesitancy.
111   Additionally, low trust and misinformation can have cross-lagged effects (e.g., low trust at
112   Time 1 causes more misinformation at Time 2), and there are likely other relevant factors (U_1
113   and U_2; e.g., technological and economic conditions, state capacity, or specific events); (b)
114   Research leveraging randomization, on average, controls for spurious factors and allows
115   causal identification for a subset of misinformation (*Misinformation$_R$*). However, many
116   studies tend to focus on a limited timescale, estimating only specific direct effects and not
117   "total" effects (e.g., nodes and arrows within the red box, where the effects of prior
118   misinformation and other context factors (U_2) might not be captured).

119   In sum, it is important to avoid a false dichotomy. The key question is not whether

120   misinformation is better framed as a symptom or a cause of social issues, but rather under

121   what conditions one framing is more appropriate than the other. We urge researchers to

122   precisely define the boundary conditions of their claims to avoid hasty generalizations, to

123   explicate any causal assumptions, and to systematically examine both individual and

124   contextual factors as conditions for causal influence. Although an extended discussion of

125   such factors is beyond scope, we discuss in the next section a selection of misinformation

126   dimensions that should be considered to appropriately recognize misinformation

127   heterogeneity.

Recognizing Heterogeneity

128

129   Objectively and easily identifiable misinformation, typically referred to as "fake

130   news," represents only a small portion of the average person's information diet in Western

131   societies[11]. However, in our view, (1) the misinformation problem should not be considered

132   negligible because a subset of obvious misinformation has low prevalence, and (2) it is

133   unreasonable to expect all types of misinformation to always have strong effects on all

134   outcomes. Some studies will find misinformation has minimal effects, others may suggest the

135   opposite[8,9]—as a generalization, both characterizations will be inaccurate unless qualified

136   with explicit recognition of heterogeneity.

137   To this end, we direct attention to three key dimensions of misinformation—topic,

138   type, and depth—that will influence its real-world reach and impact. The first dimension,

139   topic, refers simply to the subject matter of the information. For instance, individuals in

140   Country A will be impacted more by misinformation about a specific situation (e.g., an

141   election) in Country A than similar misinformation regarding Country B.

142   Second, with regards to type, we follow McCright and Dunlap[12] in distinguishing

143   between truthiness (misleading information that simply "feels" true), systemic lies (carefully

144   crafted misinformation advancing ideological interests), bullshit (persuasive misinformation

145   used opportunistically with total disregard for evidence), and shock-and-chaos (large volumes

146   of content that aim to confuse or fatigue). Note that not all information captured in this

147   framework will need to be literally false; for example, some information that is "truthy" or

148   part of a shock-and-chaos approach might not be objectively false or even falsifiable (e.g., in

149   a conflict situation, the narrative that the adversary is scared). Similarly, the selective,

150   slanted, or miscontextualized presentation of true information can be used to mislead, an

151   approach sometimes referred to as "paltering." Table S1 applies this categorization to a

152   selection of real-world misinformation. Considering this diversity, it becomes clear that much

153   misinformation is advanced—intentionally or unintentionally—by sources that would

154   typically not be categorized as dubious by researchers estimating misinformation prevalence.

155   For example, Grinberg and colleagues[11] focussed exclusively on websites known to publish

156   fabricated stories. This leaves subtler types of misinformation outside of researchers' tallies;

157   if these neglected types are considered, misinformation will be found to occupy a greater

158   portion of the information landscape.

159   The third key dimension, depth, relates to both distribution and repetition. The

160   distributional aspect refers to whether the misinformation is dispersed haphazardly (e.g.,

161   individual social-media posts or headlines) or if content is systematically bundled and/or

162   targeted (e.g., an organized disinformation campaign; a revisionist history curriculum). The

163   repetitional aspect relates to the well-known finding that repeated and thus familiar

164   information is more likely judged to be true regardless of veracity[1]. Misinformation depth is

165   important to consider because pieces of misinformation can have compound impacts[13]. Much

166   like a river can be fed from multiple *tributaries*, multiple information sources can contribute

167   to the same false narrative. This "narrative gist" can then be shared by downstream

168   *distributaries*, which can include individuals never exposed to any initial misinformation, or

169   news organizations that would never refer to the original low-quality sources. In this manner,

170   misleading narratives can infiltrate mainstream news coverage and influence public discourse

171   (e.g., conspiratorial claims influencing public debate during "Pizzagate"). Thus, assessing

172   prevalence without accounting for narrative gist will systematically underestimate the scale

173   of the misinformation problem.

174   Critically, potential outcomes can differ across misinformation types and depths, and

175   can be undesirable even if the misinformation is identified. For example, a Republican

176   correctly identifying *bullshit* from a Democrat might have lowered opinions of Democrats (or

177   vice versa), which can fuel polarization even without any direct impact on beliefs. Even the

178   discourse surrounding misinformation itself can have negative effects (e.g., erode satisfaction

179   with electoral democracy[14]). Figure 2 presents an idealized illustration of some potential

180   misinformation impacts across types, depths, and outcomes.

181   **Figure 2**

182   *Potential Misinformation Effects Across Types, Depths, and Outcomes*

183   [Figure 2]

184   *Note*. Graphical illustration of some potential misinformation effects. Plotted data are
185   hypothetical and on an arbitrary scale, to illustrate that different misinformation types at
186   various depths can have different impacts across outcomes. For example, bullshit, even at
187   high depth, may have minimal effects on beliefs, but may drive polarization and mistrust
188   (even if identified as misleading). By contrast, paltering may affect beliefs without affecting
189   trust (because fewer individuals will identify the misinformation).

190   A final point is that active forces can drive misinformation consumption. For instance,

191   a vaccine-hesitant individual seeking vaccine information will encounter more vaccine

192   misinformation than someone who is incidentally exposed. Moreover, vulnerable individuals

193   may be targeted with misinformation tailored to their psychological vulnerabilities. If this has

194   the potential to cause harm (to the individual or the public good), then it should be of

195   concern, even if overall consumption is low. Caution is therefore needed when making

196   general claims of prevalence and (lack of) impacts based on limited data.

197

Box 1. *Recommendations for Future Research*

First, as we have argued in more detail elsewhere[15], a focus shift in misinformation-intervention evaluation is recommended. To illustrate: One of the most popular paradigms presents participants with large sets of true and false claims, with the difference in truth or belief ratings between the two taken as a measure of discernment. This paradigm limits studies to short-format misinformation (e.g., headlines, tweets), as tasking participants to engage with lengthier misinformation (e.g., articles, videos) in large sets can be impractical. This favours light-touch interventions that might not address persuasive misinformation at higher depth, even though such misinformation could be more impactful.

Second, future research should make more use of observational causal-inference strategies. Regardless of how realistic or incentivized laboratory-based measures can be, it remains true that many factors are not manipulable due to ethical or feasibility considerations. For example, researchers have used the positioning of cable-TV channels (which varies randomly across localities in the US) in instrumental-variable analyses showing that exposure to unreliable news sources reduced social-distancing behaviours during the early stages of the COVID-19 pandemic[9]. However, further studies in other domains are needed.

Finally, as an integrative account of false beliefs is lacking, another promising direction is to borrow from the broader cognitive-science literature. For instance, cognitive research has shown that individuals preferentially rely on "gist" representations of quintessential meanings[13]. Future research attempting to delineate the evolution of narrative gist at a societal level might therefore benefit from first examining gist processing at the individual level. Cognitive models of decision making could also be used to explore misinformation impact beyond observable outcomes. For example, evidence-accumulation models could be used to decompose choice and response-time data into cognitively interpretable parameters (e.g., response boundaries represent the varying levels of evidence individuals require to make decisions and could be interpreted as caution).

Conclusions and Recommendations

Taken together, a clear implication of our discussion is that the standard paradigms used for evaluating the impacts of misinformation and misinformation interventions are likely insufficient. Some recommendations for changes to current research practice in the field are provided in Box 1. We have argued that the evaluation of misinformation impacts is an important, but complex, research question, particularly in the current era of rising geopolitical tensions and rapid technological change. We hope that the current Comment will contribute

233   to increasingly nuanced debates about the impact of misinformation and potential

234   interventions.

References

1. Ecker, U.K.H. et al. *Nat. Rev. Psychol.* **1**, 13–29. https://doi.org/10.1038/s44159-021-00006-y (2022).

2. Lazer, D.M.J. et al. *Science.* **359**, 1094–1096. https://doi.org/10.1126/science.aao2998 (2018).

3. Jungherr, A., & Schroeder, R. *Soc. Media + Soc.* **7**. https://doi.org/10.1177/2056305121988928 (2021).

4. Acerbi, A., Altay, S., & Mercier, H. *Harv. Kennedy Sch. Misinf. Rev.* https://doi.org/10.37016/mr-2020-87 (2022).

5. Lewandowsky, S., Ecker, U.K.H., & Cook, J. *J. Appl. Res. Mem. Cogn.* **6**, 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008 (2017).

6. Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. *Nat. Hum. Behav.* **7**, 74–101. https://doi.org/10.1038/s41562-022-01460-1 (2023).

7. Motta, M., & Stecula, D. *PLOS ONE.* **16**, e0256395–e0256395. https://doi.org/10.1371/journal.pone.0256395 (2021).

8. Loomba, S., de Figueiredo, A., Piatek, S.J., de Graaf, K., & Larson, H.J. *Nat. Hum. Behav.* **5**, 337–348. https://doi.org/10.1038/s41562-021-01056-1 (2021).

9. Simonov, A., Sacher, S., Dube, J.-P., & Biswas, S. *Mark. Sci.* **41**, 230–242. https://doi.org/10.1287/mksc.2021.1328 (2022).

10. Pearl, J. *J. Causal Inference.* **7**, 567–573. https://doi.org/10.1515/jci-2019-0026 (2019).

11. Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. *Science.* **363**, 374–378. https://doi.org/10.1126/science.aau2706 (2019).

12. McCright, A.M., & Dunlap, R.E. *J. Appl. Res. Mem. Cogn.* **6**, 389–396. https://doi.org/10.1016/j.jarmac.2017.09.005 (2017).

260   13. Reyna, V.F. *Proc. Natl. Acad. Sci.* **118**, e1912441117.

261   https://doi.org/10.1073/pnas.1912441117 (2021).

262   14. Nisbet, E.C., Mortenson, C., & Li, Q. *Harv. Kennedy Sch. Misinf. Rev.*

263   https://doi.org/10.37016/mr-2020-59 (2021).

264   15. Tay, L.Q., Lewandowsky, S., Hurlstone, M., Kurz, T., & Ecker, U.K.H. *Harv.*

265   *Kennedy Sch. Misinf. Rev*. https://doi.org/10.37016/mr-2020-124 (2023).

266   **Competing Interests**

267   The authors declare no competing interests.

(a)

U_1

$Misinformation_{T1}$ → $Misinformation_{T2}$

$Low\ Institutional\ Trust$

$Outcomes$

U_2

(b)

U_1

$Random$

$Misinformation_{T1}$

$Misinformation_{R}$

$Low\ Institutional\ Trust$

$Outcomes$

U_2