The Influence of Cues to Consensus Quantity and Quality on Belief in Health Claims

Benjamin Simmonds (benjamin.simmonds@adelaide.edu.au)
The University of Adelaide, Faculty of Health and Medical Sciences, School of Psychology

Rachel G. Stephens (rachel.stephens@adelaide.edu.au)
The University of Adelaide, Faculty of Health and Medical Sciences, School of Psychology

Rachel Searston (rachel.searston@adelaide.edu.au)
The University of Adelaide, Faculty of Health and Medical Sciences, School of Psychology

Nusrat Asad (nusrat.asad@sydney.edu.au)
The University of Sydney, Faculty of Science, School of Psychology

Keith J. Ransom (keith.ransom@adelaide.edu.au)
The University of Adelaide, Faculty of Health and Medical Sciences, School of Psychology

Abstract

Many people turn to social media for public health information, but such platforms often contain conflicting and inaccurate medical advice. To assess complex health claims online, people may consider the prevailing consensus; however, previous work suggests that people may not be very sensitive to important cues to consensus "quality". To explore further, across two experiments we tested people's sensitivity to the consensus-quality cues of source diversity and source expertise. Via a mock Twitter platform, participants rated their belief in a series of health claims both before and after reading various kinds of tweets about the claims. Experiment 1 showed that experts (both individual medical experts and health organisations) were more persuasive than non-experts. Additionally, stances that were supported by a diverse set of sources were more persuasive. Experiment 2 showed that participants continue to favour experts even when outnumbered in tweet quantity by non-experts. When experts were not present, however, participants favoured high tweet quantity. Both experiments suggest that cues to consensus quality (namely, expertise and source diversity) and consensus quantity (tweet quantity) are salient cues in belief revision. These findings are important in understanding how socially acquired health information (and misinformation) shifts opinion, and the role that experts can play.

Keywords: reasoning; consensus quality; consensus quantity; induction; expertise; source diversity.

Introduction

In times of crises, such as the COVID-19 pandemic, effective communication of credible health advice plays a critical role in public health – in reducing the spread of infection, for example. The widespread consumption of online social media offers the potential to positively transform the way that public health information is communicated. However, because social media represents a different information environment than traditional media, it is prudent to consider how the different properties of the environment may impact the way public health information is evaluated. For example, there is some evidence to suggest that the proportion of non-expert versus expert opinion has increased to the point that misleading or false information contradicting expert opinion is already a serious issue; Suarez-Lledo and Alvarez-Galvez (2021) found that approximately 40% of the most shared links about common health topics contained false information. Indeed, the prevalence of false health information online, especially concerning COVID-19, has risen to the extent that the WHO (2022) has since declared an "infodemic". A further salient property of the online information environment is the ease with which people can share, forward or otherwise republish information. This changes the likelihood of encountering content: some content is republished many times - and therefore frequently encountered - while other content goes largely ignored. Unfortunately, republishing-frequency is not necessarily an indicator of accuracy (e.g., Vosoughi, Roy & Aral, 2018) or even whether the person sharing the information has considered it carefully, or read it at all (Gabielkov, Ramachandran, Chaintreau & Legout, 2016). These properties of the information environment mean that the need to carefully consider the credibility of information is all the more important, while at the same time determining what advice is credible is increasingly problematic.

In an information environment where information quantity is high but quality is variable, assessing claims such as "Poor sleep is linked to Alzheimer's" can be challenging, particularly for the average reasoner who lacks the knowledge and expertise to directly evaluate the evidence for themselves. To circumvent such challenges, people may come to rely on salient features of the decision-making environment, such as the consensus – or lack of – among their trusted online

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

sources. A majority opinion is often more persuasive than a minority one, especially if the difference in size is significant (Lewandowsky, Cook, Fay & Gignac, 2019). Similar “consensus” effects have been found in consumer behaviour research (i.e., Alba, Broniarczyk, Shimp, & Urbany, 1994) and in classic social conformity studies (i.e., Asch, 1951). However, consensus information can be misleading - it is often unclear whether each “data point” comprises evidence of equivalent merit. Additionally, people may confound consensus with high message quantity by failing to consider the number of unique sources providing those messages. For example, Harkins and Petty (1981) suggested that under high levels of distraction, a series of unique arguments provided by three unique sources can be equally persuasive as the same series of arguments presented by a single source. Therefore, considering the highly distracting environments of social media (Pennycook, McPhetres, Zhang, Lu & Rand, 2020), a key question is the extent to which people are sensitive to the quality of an apparent consensus, as opposed to simply weighing the relative quantity of information on either side of a claim. Troublingly, recent research has shown people are often insensitive to the independence of original evidence in terms of the primary sources that are cited (Alister, Perfors & Ransom, 2022; Desai, Xie & Hayes, 2022; Yousif, Aboody & Keil, 2019). Similarly, Ransom, Perfors and Stephens (2021) showed people’s beliefs across a variety of claims were influenced more by the relative quantity of supporting and opposing “evidence” (posts) presented, than by potential indicators of evidence quality (the diversity of the people posting, or the diversity of arguments offered). The persuasiveness of quantity and repetition effects have been well studied in the literature, with explanations ranging from increased processing fluency (Begg, Anas & Farinacci, 1992; McGlone & Tofighbakhsh, 2000; Reber & Schwarz, 1999), to elaboration likelihood (Petty & Cacioppo, 1986).

Reliance on quantity is likely to be particularly problematic when people reason about important health information online, where non-experts can be prolific and there is uncritical sharing of large volumes of information. Therefore, the current study aimed to investigate the impact of different cues to consensus within the domain of health-related claims. First, we sought to extend the work of Ransom et al. (2021) to determine whether people would be sensitive to source diversity (same vs. different authors) beyond consensus quantity (the relative number of messages for vs. against a claim) in this domain. Second, we investigated the extent to which people are sensitive to source expertise. Experts’ authority status typically affords them a significant level of persuasiveness within their domain (Maddux & Rogers, 1980). Additionally, adult decision-makers have shown to be well-calibrated when asked to identify the expert most applicable to a subject-matter (Bromme & Thomm, 2016), making them a clear point of reference when making a decision. However, given the power of consensus quantity (e.g., Yousif et al., 2019), it is unclear whether people will be more persuaded by the consensus views of experts over non-experts. With widespread social media use, non-experts are increasingly being given powerful voices (Sunstein, 2007). Simultaneously, medical mistrust has risen, becoming especially pronounced during worldwide responses to emerging diseases, such as COVID-19 and Ebola (Knobel, Zhao & White, 2022; Richards, Mokuwa, Welmers, Maat & Beisel, 2019). High medical mistrust can lead to a favouring of non-expert opinion and personal anecdotes over the scientific information provided by experts (Burrows, Nettleton, Pleace, Loader & Muncer, 2000). We therefore compared the persuasiveness of experts versus non-experts, when expert individuals form either the majority (Experiment 1) or minority (Experiment 2) view.

As a final aim, we investigated the confluence of source diversity and expertise. In addition to testing the impact of individual medical experts on people’s beliefs, we also examined the impact of health organisations (Experiment 1). A message from a health organisation may carry more weight than a message from either an individual expert or a non-expert, as organisations usually represent the view of multiple experts – that is, each organisation represents a consensus view within itself. Indeed, non-experts have been found to self-report preferentially accessing official health organisation sources (e.g., WHO) when seeking vaccine information (Volkmer, 2021). Lin, Spence and Lachlan (2016) examined the perceived credibility of an expert health group against unknown non-experts and peers and found that expert groups were evaluated most favourably. However, health organisations, like individual medical experts, have also been associated with varying degrees of favour by the public. An analysis of Facebook user discourse, for example, found that many users associated a prominent public health body (the U.S. Centre for Disease Control and Prevention) with corruption and fearmongering (Laurent-Simpson & Lo, 2019).

Our aims were addressed across two online experiments, using a mock Twitter platform. We examined people’s belief change for various health claims, in response to messages with different levels of source diversity (same vs. different author) and source expertise (non-experts, individual experts or expert groups). Experiment 1 varied these two consensus quality cues across the majority stance for a given claim, testing whether people are sensitive to the quality cues or to consensus quantity only. Experiment 2 focused on sensitivity to the expertise of individuals and considered an alternative situation where experts had the minority stance; thus, we directly pitted consensus quantity against quality.

Experiment 1

Method

Experiment 1 was administered via a mock Twitter interface where participants rated their belief in various health claims (e.g., “Medicinal cannabis usage reduces symptoms of depression") before and after viewing tweets in favour or against the claim. For each trial, five Target Tweets (T) all endorsed or refuted the claim and were authored by either non-experts, individual experts (e.g., doctors), or health expert groups. The Target Tweets were also written by diverse or non-diverse sources. In opposition to the Target Tweets was a single Non-Target Tweet (T') written by a non-expert.

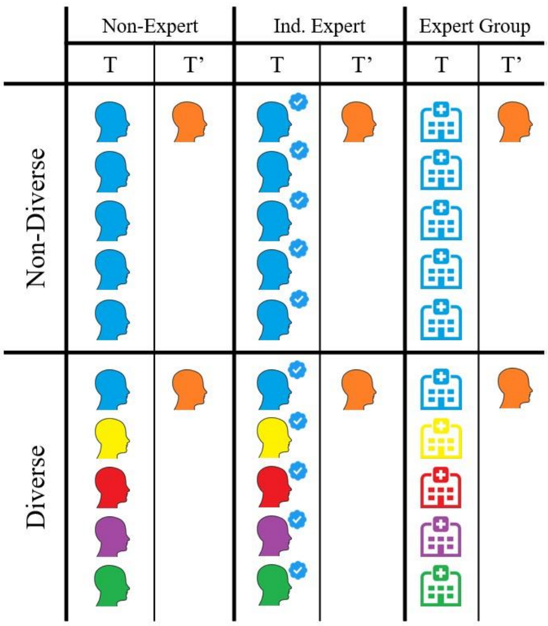

Figure 1**:** Experiment 1 design. Target (T) and Non-Target (T') Tweets. Each colour is a unique source.

Design Experiment 1 employed a 3 (source expertise: non-expert vs. individual expert vs. expert group) × 2 (source diversity: diverse vs. non-diverse) within-participants experimental design (see Figure 1). Each trial presented a health claim followed by related tweets presenting a 5:1 majority consensus. Five Target Tweets formed the majority and were either for (Pro) or against (Con) the claim (randomly assigned). The remaining Non-Target tweet opposed the Target Tweets. Source expertise and source diversity varied in the Target Tweets only. The Non-Target tweet was always presented by a non-expert. The diverse conditions presented Target Tweets from five unique sources, whereas the non-diverse conditions presented Target Tweets from a single repeated source. All Target Tweets presented the same general argument or reason to support their given stance (but worded differently). Participants engaged in two trials per cell of the design present in Figure 1, for a total of twelve trials.

Note the design is such that if participants attend purely to the quantity of tweets for/against a claim, beliefs should shift towards the Target Tweet stance but see no effect of source diversity or expertise across conditions. The 5:1 majority tweet structure was of particular interest because it allowed us to explore how participants would perceive the non-diverse conditions. In these conditions, one source with a single tweet (the Non-Target Tweet) is pitted against a source that is repeated five times (the Target Tweets). Hence, while the number of unique sources and arguments was equal, they differed purely in number of tweets. The dependent variable of Experiment 1 was the difference between people's initial and updated belief ratings (delta), calculated on a trial-by-trial basis, and sign-adjusted so a positive value reflected belief revision in the direction consistent with the stance of the Target Tweets.

Participants A total of 103 participants were recruited online via Amazon's Mechanical Turk platform in August 2022. Each participant received $5 AUD as compensation. The inclusion criteria for the study were that all participants must be 18 or above and able to read English. The age range was between 22 and 71 (M = 38). 61.17% of the sample identified as male, 33.98% identified as female, and the remaining 4.85% identified as neither. 85.44% had attained a high-school-level education or above. Ethics approval to conduct this research was obtained from the Human Research Ethics Subcommittee at The University of Adelaide (Ref. 22/75).

Materials – Claims and Related Tweets Participants were presented with a set of claims[1] covering various health-related topics. Claims were based on articles published in the Lancet and news articles published in the New York Times International's health section to maximise ecological validity. Before conducting the main experiment, a larger set of potential claims were piloted via Amazon Mechanical Turk to assess initial belief ratings and select the 12 that people were most uncertain about, maximising the scope for belief revision during the main experiments. The related tweets were written in a manner such that they could be plausibly produced by non-experts and experts. Six related tweets were written for both the Pro and Con stances for each claim[1], with similar arguments presented within their given stance (e.g., Pro: "Vitamin D helps to activate immunity cells, which helps to prevent respiratory infections" vs. Con: "Vitamin D has not been shown to reduce respiratory infections any more than placebo tablets do").

Materials – Related Tweet Sources The sources presented in the mock Twitter interface consisted of non-experts, individual experts and expert group users[1]. Individual experts were signalled by a verified tick[2] and randomly assigned credentials in their titles, such as Dr. as prefix and/or MD as suffix. Avatars and names for the non-experts and individual

[1] See OSF link: https://osf.io/wzef8/?view_only=faafb7e837b8448a9ba56de23d0f78ad

[2] Note that both experiments were conducted prior to changes in the verification system that occurred in October 2022.

experts were fictional and randomly selected per trial. Avatar photos were generated by artificial intelligence or obtained online, and source names were based on random name generators. For the expert groups, official health organisations were used and allocated randomly (e.g., British Medical Association; Health Research Authority). The expert groups were purposely chosen to avoid including mainstream healthcare organisations (e.g., WHO), thus preventing familiarity effects (which was important given that the other types of individual authors were unfamiliar).

Procedure Before beginning the experiment, participants completed demographic questions and were shown task instructions. Participants were informed arguments would be shown from users of various backgrounds. Participants were also required to correctly answer three verification questions to ensure they understood the instructions before proceeding. Each participant was then presented with all twelve trials in a random order. Each trial began with the health claim shown on its own (e.g., "Vitamin D reduces the risk of respiratory infection.") along with a neutral framing post and image[1], taking the form of a single tweet linking a fictional article about the associated health topic. Participants gave an initial belief rating, between 0 (do not agree at all) and 100 (complete agreement) on a numeric slider scale. Next, the set of related tweets were presented below the framing post in random order[1]. Participants had to click on each tweet to progress, as a means of encouraging them to read them. They then reported their updated belief rating (their initial rating could be seen).

Results and Discussion

Figure 2 presents the mean change in belief (delta) across each of the six conditions. As the figure shows, belief ratings generally shifted towards the majority stance. However, belief change was greater in trials that presented diverse tweets compared to non-diverse tweets. Overall, expert groups and individual experts were more persuasive than non-experts. Interestingly, similar belief ratings were found for diverse non-experts and non-diverse individual experts.

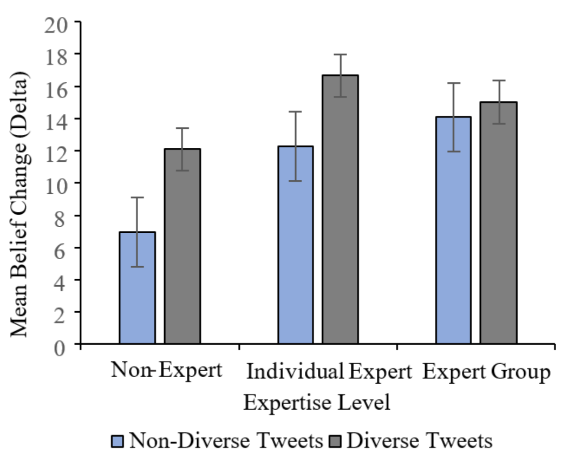

Figure 2: Mean Belief Change for each Source Expertise (x-axis) and Source Diversity (colour) condition.

To analyse the magnitude and significance of these findings, a factorial 2 (source diversity: non-diverse, diverse) × 3 (expertise: non-expert, individual expert, expert group) ANOVA and post hoc pairwise t-tests were conducted. The main effect of source diversity was significant, $F(1, 102) = 7.84$, $p = 0.006$, $\eta^2 = 0.006$. This indicates the average belief change was significantly higher in the diverse tweets condition ($M = 14.59$, $SD = 22.64$) than in the non-diverse tweets condition ($M = 11.10$, $SD = 21.84$), but with a small effect size. The main effect of expertise was also significant but with a small effect size, $F(2, 204) = 8.24$, $p < 0.001$, $\eta^2 = 0.011$. Post hoc analyses using pairwise t-tests with a Holm adjustment indicated the mean belief change was significantly higher ($p = 0.001$ and $p = 0.001$, respectively) in both the individual expert ($M = 14.45$, $SD = 22.80$) and expert group conditions ($M = 14.55$, $SD = 23.98$) compared to the non-expert condition ($M = 9.53$, $SD = 19.58$). There was no statistically significant difference in the mean belief change between the individual expert condition and the expert group condition ($p = 0.952$). The interaction between expertise and source diversity was not significant and had a small effect size, $F(2, 204) = 1.68$, $p = 0.19$, $\eta^2 = 0.002$. A follow-up linear mixed effect model analysis was conducted to account for individual differences, which confirmed the findings of the ANOVA and post-hoc t-tests.

Experiment 1 found that people did not treat the six different consensus conditions as equivalent. Instead, Experiment 1 suggests that people are – at least to a small extent – more persuaded by diverse than non-diverse sources (cf. Ransom et al., 2021), and by both expert groups and individual experts than non-experts when evaluating medical information (cf. Richards et al., 2019; Sunstein, 2007). These effects extend beyond the consideration of consensus quantity (i.e., the relative number of messages for vs. against). However, the findings suggest people do not count health organisations themselves as representing a diverse number of expert voices, weighting them similarly to individual experts.

Experiment 2

Method

In Experiment 2, we turned to exploring the weight given to expert messages when they are directly pitted against a non-expert majority, thereby allowing a direct comparison of consensus quality cues and consensus quantity cues. The minority status held by experts in this design mirrors the more realistic minority status often held by experts in actual social media environments, enhancing ecological validity. Given that individual experts and health organisations had similar effects in Experiment 1, Experiment 2 focussed on individual experts only. Experiment 2's method was identical to Experiment 1, with any differences noted below.

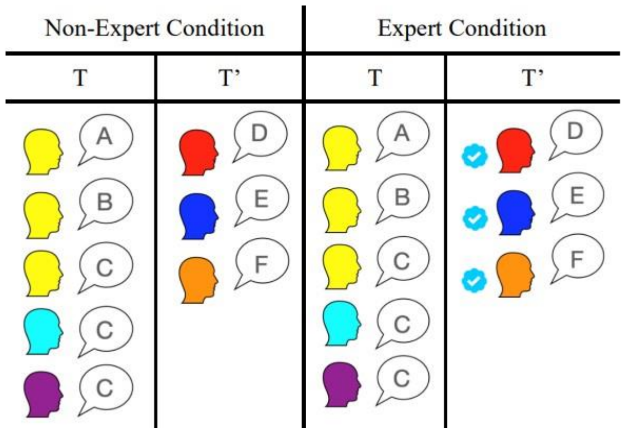

Figure 3: Experiment 2 design. Target (T) and Non-Target (T') Tweets. Each colour represents a unique source, and each letter represents a unique argument.

Design Expertise was manipulated within subjects and between trials (non-expert vs. expert)[1]. Within each trial, source and argument diversity were controlled across Target and Non-Target Tweets (see Figure 3), such that each stance was conveyed by a total of three unique sources (differently coloured heads in Figure 3), with three unique arguments/reasons given for their stance (different letters in the speech bubbles). Of the eight tweets shown per trial, five were Target Tweets (T). Three Non-Target Tweets (T') took the opposing stance. The stance of the Target Tweets (Pro vs. Con) was counterbalanced across trials. In the expert condition, there were always fewer tweets from individual experts (three) than non-experts (five). In the non-expert condition, all tweets were written by non-experts. Participants made initial and updated belief ratings as per Experiment 1 and chose a message to retweet (the retweet data is not analysed here). There were 8 trials, shown in random order (4 per expertise condition). The dependent variable in Experiment 2 was the updated belief rating (standardised such that a higher number indicated greater belief in the Target Tweets' stance); we were interested in whether ratings were above 50 (siding with the Target Tweets) or below 50 (siding with the Non-Target Tweets).

Participants A total of 101 people participated in the study. The age range was between 23 and 70, (M = 37.91), with 67.33% males, 31.68% females and 0.99% other. All participants had at least a high-school equivalent education.

Results and Discussion

Updated belief ratings were examined to determine the extent to which they differed across expertise conditions. Figure 4 shows that the updated belief ratings were, on average, lower in the expert than the non-expert condition, indicating that participants tended to report beliefs that favoured the expertise present amongst the Non-Target Tweets. To further investigate this finding, a series of four nested regression models were compared.

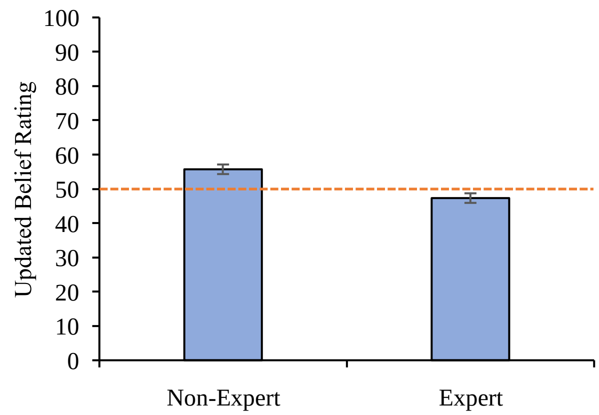

Figure 4: Mean Updated Belief Ratings by Expertise Condition. Values above the red line prefer the Target Tweets while values below the red line prefer the Non-Target Tweets.

These models sequentially added various predictors in order to explain the variance in updated belief ratings. From a comparison of the Akaike Information Criterion (AIC), the preferred model was one that comprised the following significant predictors: expertise (B = -12.54, t = -2.99, p = 0.003), Target Tweet stance (B = -8.32, t = -2.12, p = 0.034), initial belief rating (B = 0.78, t = 15.71, p < 0.001), and the three-way interaction between initial belief rating, Target Tweet stance and expertise (B = -0.23, t = -2.32, p = 0. 021) This was such that updated belief ratings favoured the Target Tweets during non-expert trials (i.e., ratings were above 50) and favoured the Non-Target Tweets during expert trials (i.e., ratings were below 50), as displayed in Figure 4. Additionally, participants more strongly favoured the Target Tweets when their stance was Con, perhaps indicating that Con arguments were stronger than Pro ones. Overall, this model suggests participants tended to report belief ratings that favoured experts when they were present, despite their minority stance. In contrast, when experts were not present, the larger tweet quantity provided by the Target Tweets were preferred.

Experiment 2 expanded upon Experiment 1 by investigating whether the persuasiveness of expertise would still hold when experts held a minority number of tweets in support of their stance. Encouragingly, Experiment 2 showed that experts are able to maintain their persuasiveness, suggesting that expertise can outweigh consensus quantity (cf. Richards et al., 2019; Sunstein, 2007). However, when experts were not present, participants preferred the higher tweet quantity, which is consistent with longstanding demonstrations of the persuasiveness of consensus quantity (e.g., Asch, 1951; Lewandowsky et al., 2019).

General Discussion

Across two experiments, we examined the persuasiveness of consensus quantity and quality in reasoning about health claims online. Our finding that people's belief in health claims somewhat shift towards the majority view extends previous demonstrations of the persuasive power of consensus quantity in other domains (e.g., Lewandowsky et al., 2019; Ransom et al., 2021; Yousif et al., 2019) and broader literature (e.g., Alba et al., 1994; Asch, 1951). Partially in contrast with previous findings which suggest that people are often insensitive to levels of independence among primary sources of evidence (Alister et al., 2022; Desai et al., 2022; Yousif et al., 2019), we find some sensitivity to source diversity, albeit a modest effect. It is important to note that the current study differs from these aforementioned ones in terms of social mediation. For example, Yousif et al. (2019) explored people's sensitivity to the independence of primary sources upon which various secondary sources were basing their opinion. The current study differs from this design in that we explored people's sensitivity to the diversity of authors (more equivalent to secondary sources in the Yousif et al., 2019, design), which could explain why our results differed. Our results align with other research suggesting that people associate high source diversity with informational independence, diversity of reasoning and a level of consensus broader than the messages seen (i.e., Kim & Spelke, 2020; Mercier & Miton, 2019; Ransom et al., 2021). One important direction for future research will be to examine people's sensitivity to other types of source diversity by including, for example, unique sources (people posting) who vary in the similarity of their demographics, training, social connections, and so on.

Our results demonstrate that people are also somewhat sensitive to source expertise when assessing health claims encountered online; people appear (at least marginally) more persuaded by both individual medical professionals and health organisations than non-experts. This result is consistent with prior studies that have highlighted the influence of medical expertise on public decision-making—emphasising its association with credibility and epistemic authority (see Lin et al., 2016; Maddux & Rogers, 1980). There has been some evidence of a decline in public trust of medical professionals and health organisations amongst the literature (see Knobel et al., 2022; Laurent-Simpson & Lo, 2019), and while the current study is not able to comment on whether the persuasiveness of expertise has declined, it shows that participants were favouring experts, both when they held a majority stance (i.e. Experiment 1) and a minority stance (i.e. Experiment 2). This observation is further evidence that people do not completely ignore consensus quality; when consensus quantity and quality cues are in opposition, people sometimes prioritise quality cues in assessing their belief in a claim. However, it is noteworthy that in Experiment 1, diverse non-experts were as persuasive as an individual expert producing the same number of comparable messages. This particular result highlights a challenge for expert health communication and could be especially problematic in cases where experts may disagree with non-experts.

An important finding (Experiment 1) was that health organisations were no more persuasive than individual medical professionals. This is somewhat surprising, as health organisations represent the confluence of source diversity and expertise; a message from such a source would usually represent the view of multiple experts. One possible explanation to test in future research is that people may not perceive a message from an organisation as conveying the view of diverse, independent sources. Another direction for future research will be to consider the effect of familiarity; our stimuli included only unfamiliar individuals, experts and organisations, but familiarity may be important if public trust has declined or been bolstered for particular sources (e.g., cf. Laurent-Simpson & Lo, 2019; Volkmer, 2021).

Other directions for future research include varying the target populations and procedures. The current samples were recruited primarily from the US, but other populations may show different effects on belief revision. Additionally, other procedures for measuring belief revision can be explored. Changes in belief rating may have been influenced by the way in which the belief scale was presented twice to participants. After reading the tweet responses to the health claim, participants were immediately prompted with an opportunity to update their initial belief rating. Prompting them in this way may have encouraged the participants to change their belief rating even under instances where there was no actual change in belief. Future research could use other procedures to confirm the current findings, such as collecting belief ratings only after reading the response tweets (not before), or having a larger gap between initial and updated ratings.

The current findings have several applied implications for online health reasoning. In particular, they suggest (all else being equal) claims supported by a large number of messages tend to be more influential than claims supported by a smaller number of messages. This finding provides a potential explanation for the increasing dominance of misinformation; social media users are affected by the relative quantity of consensus information about a given claim. The ease with which health advice can appear more believable by simply posting more tweets is concerning, especially given the rise in bot activity on platforms such as Twitter to sway the popularity of certain content (Gilani, Farahbakhsh & Crowcroft, 2017). However, the current findings also suggest several means of combating misinformation. Experts and expert groups can be persuasive, perhaps even when producing a smaller quantity of tweets than non-experts. Therefore, increasing the prevalence or visibility of a variety of qualified subject-matter experts within social media spaces would likely benefit the increasing number of people who rely on online health information.

Acknowledgements

This research was supported by: 1) a 2022 Digi+ FAME grant from The University of Adelaide; 2) a 2022 Collaborative Research Fund grant from Defence Innovation Partnership.

References

Alba, J. W., Broniarczyk, S. M., Shimp, T. A., & Urbany, J. E. (1994). The influence of prior beliefs, frequency cues, and magnitude cues on consumers' perceptions of comparative price data. *Journal of Consumer Research, 21*(2), 219–235.

Alister, M., Perfors, A, & Ransom, K. J. (2022). Source independence affects argument persuasiveness when the relevance is clear. In *Proceedings of the 44th Annual Meeting of the Cognitive Science Society.*

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men; research in human relations* (pp. 177–190). Carnegie Press.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, *121*(4), 446-458.

Bromme, R. & Thomm, E. (2016). Knowing Who Knows: Laypersons' Capabilities to Judge Experts' Pertinence for Science Topics. *Cognitive Science, 40*(1), 241–252.

Burrows, R., Nettleton, S., Pleace, N., Loader, B., & Muncer, S. (2000). Virtual community care? Social policy and the emergence of computer mediated social support. *Information, Communication & Society*, *3*(1), 95–121.

Desai, S. C., Xie, B., & Hayes, B. K. (2022). Getting to the source of the illusion of consensus. *Cognition, 223*, 105023.

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social Clicks: What and Who Gets Read on Twitter? *Performance Evaluation Review, 44*(1), 179–192.

Gilani, Z., Farahbakhsh, R., & Crowcroft, J. (2017, April). Do bots impact Twitter activity?. In *Proceedings of the 26th International Conference on World Wide Web Companion* (pp. 781-782).

Harkins, S. G. & Petty, R. E. (1981). The Multiple Source Effect in Persuasion: The Effects of Distraction. *Personality & Social Psychology Bulletin, 7*(4), 627–635.

Kim, S., & Spelke, E. S. (2020). Learning from multiple informants: Children's response to epistemic bases for consensus judgments. *Journal of experimental child psychology, 192*, 104759.

Knobel, P., Zhao, X., & White, K. M. (2022). Do conspiracy theory and mistrust undermine people's intention to receive the COVID-19 vaccine in Austria? *Journal of Community Psychology, 50*(3), 1269–1281.

Laurent-Simpson, A. & Lo, C. C. (2019). Risk society online: Zika virus, social media and distrust in the Centers for Disease Control and Prevention. *Sociology of Health & Illness, 41*(7), 1270–1288.

Lewandowsky, S., Cook, J., Fay, N., & Gignac, G. E. (2019). Science by social media: Attitudes towards climate change are mediated by perceived social consensus. *Memory & Cognition*, *47*(8), 1445–1456.

Lin, X., Spence, P. R., & Lachlan, K. A. (2016). Social media and credibility indicators: The effect of influence cues. *Computers in human behavior, 63*, 264-271.

Maddux, J., & Rogers, R. (1980). Effects of source expertness, physical attractiveness, and supporting arguments on persuasion: A case of brains over beauty. *Journal of Personality and Social Psychology, 39*(2), 235–244.

McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a Feather Flock Conjointly (?): Rhyme as Reason in Aphorisms. *Psychological Science, 11*(5), 424–428.

Mercier, H., & Miton, H. (2019). Utilizing simple cues to informational dependency. *Evolution and Human Behavior, 40*(3), 301-314.

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.

Petty, R., & Cacioppo, J. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 123–205). San Diego, CA: Academic Press.

Ransom, K. J., Perfors, A., Stephens, R. (2021). Social metainference and the evidentiary value of consensus. *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 833-839).

Reber, R., & Schwarz, N. (1999). Effects of Perceptual Fluency on Judgments of Truth. *Consciousness and Cognition, 8*(3), 338–342.

Richards, P., Mokuwa, E., Welmers, P., Maat, H., & Beisel, U. (2019). Trust, and distrust, of Ebola Treatment Centers: A case-study from Sierra Leone. *PLoS One, 14*(12), e0224511.

Suarez-Lledo, V., & Alvarez-Galvez, J. (2021). Prevalence of health misinformation on social media: systematic review. *Journal of medical Internet research, 23*(1), e17187.

Sunstein. (2007). On the Divergent American Reactions to Terrorism and Climate Change. *Columbia Law Review, 107*(2), 503–557.

Volkmer, I. (2021). *Social media and COVID-19: A global study of digital crisis interaction among Gen Z and millennials.* World Health Organisation.

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

World Health Organization. (2022). *Infodemic.* World Health Organization.

Yousif, S. R., Aboody, R., & Keil, F. C. (2019). The Illusion of Consensus: A Failure to Distinguish Between True and False Consensus. *Psychological Science, 30*(8), 1195–1204.