Why Do People Share Disinformation On Social Media?

POLICY BRIEF
SEPTEMBER 2020

Professor Tom Buchanan
Why Do People Share Disinformation On Social Media?

POLICY BRIEF

Professor Tom Buchanan, University of Westminster

This is the Full Report from the *Why Do People Spread Disinformation On Social Media?* project, commissioned by CREST. The project looks at the human factors influencing the spread of disinformation, in order to better understand how to counteract it.

[www.crestresearch.ac.uk/projects/disinformation-on-social-media](http://www.crestresearch.ac.uk/projects/disinformation-on-social-media)

This research was funded by the Centre for Research and Evidence on Security Threats - an independent Centre commissioned by the Economic and Social Research Council (ESRC Award: ES/N009614/1) and which is funded in part by the UK Security and intelligence agencies and Home Office.

[www.crestresearch.ac.uk](http://www.crestresearch.ac.uk)
TABLE OF CONTENTS

INTRODUCTION .......................................................................................................................... 4
THE RESEARCH .......................................................................................................................... 6
IMPLICATIONS ............................................................................................................................ 10
COMMENTARY ............................................................................................................................ 12
READ MORE ............................................................................................................................... 13
INTRODUCTION

This brief summarises a research project that explored some of the processes underlying the spread of false material on social media. It positions the research and findings within the broader literature relevant to online political disinformation and considers what we currently do and don’t know about the problem.

It conceptualises disinformation as false material created and disseminated with the intent of deceiving others and causing harm and focuses specifically on the distribution of that material via social media platforms. This is distinct from other elements of online information operations, such as selective presentation of true material or other forms of polarising or ‘hyperpartisan’ communication.

CONTEXT

21st July 2020 was notable for the publication of two reports from UK Parliamentary committees. The first was from the House of Commons Digital, Culture, Media and Sport Committee (DCMS Committee, 2020) on ‘Misinformation in the COVID-19 Infodemic’, which expanded on ‘Disinformation and Fake News’ previously from the Committee (2019); pointed to significant real-world harms arising from the spread of false information; and discussed the role of hostile actors in generating it. The second was the long-awaited ‘Russia Report’ from the Intelligence and Security Committee (2020). While considerably broader in scope, it again noted the problem of political disinformation. It pointed to the use of social media as a tool of disinformation and influence campaigns in support of Russian foreign policy objectives.

It is not just parliamentary committees or academic researchers who think this matters. A recent survey indicated that around the world, the majority of people were “concerned about what is real and fake on the internet when it comes to news.” Ironically, much of the spread of false information can be attributed to those very people.

ORGANIC REACH OF DISINFORMATION

Once false material is published online, it can then be spread to very wide audiences through the phenomenon of ‘organic reach’. While bot networks and ‘coordinated inauthentic activity’ are important, everyday human behaviour may actually account for a great deal of the spread of false material. People who encounter the material may share it with their own networks of contacts, friends and family. Even if they don’t deliberately share it, interacting with the material in other ways, such as ‘liking’ it, may cause a social network’s algorithms to increase its visibility. The people who subsequently see it may then share it to their own social networks, potentially leading to an exponential rise in its visibility.

Research has suggested that relatively few people (maybe less than 10% of social media users) actively share false material they encounter online. Even so, this small proportion of users can significantly increase the reach of the false information. Therefore, it is important to understand who extends the organic reach of disinformation, and why. The answers to those questions will help inform the development of effective interventions.

1 Newman et al., 2020.
3 Guess, Nagler and Tucker, 2019.
MESSAGE CHARACTERISTICS

A growing body of research has explored the psychological processes associated with encountering and sharing disinformation online. Much of this revolves around mechanisms associated with information processing, cognitive ability and digital literacy, demographic factors, and political partisanship.  

One influential line of research draws on dual-process models of social information processing. In some situations, we carefully consider the information available to us. In others, we make rapid decisions, based on peripheral cues and cognitive rules-of-thumb. Interactions with social media are very likely to fall into this second category. 

When we decide to ‘like’ or ‘share’ a story we come across, we may do so very rapidly, without a great deal of deliberation. In such situations, use of heuristics is common: mental shortcuts that enable us to act or make decisions rapidly, without conscious deliberation. Three potentially important heuristics in this context are consistency, consensus, and authority.  

CONSISTENCY

Consistency is the extent to which sharing a message would be consistent with past behaviours or beliefs of the individual. For example, in the USA people with a history of voting Republican would be more likely to endorse and disseminate right-wing messaging because it is consistent with their past political behaviour.  

CONSSENSUS

Consensus is the extent to which people think their behaviour would be consistent with that of most other people. For example, seeing that a message has already been shared widely might make people more likely to forward it on themselves. This is sometimes described as the ‘social proof’ effect. The bot activity documented by Shao et al. (2018) could be an attempt to exploit this effect.  

AUTHORITY

Authority is the extent to which the communication appears to come from a credible, trustworthy source. Again, there may have been attempts to exploit this effect. In 2018, Twitter identified fraudulent accounts that simulated those of US local newspapers, which may be trusted more than national media. These may have been sleeper accounts established specially for the purpose of building trust prior to later active use. 

Thus, manipulating these characteristics may be a way to increase the organic reach of disinformation messages. There is real-world evidence of activity consistent with attempts to exploit them. If these effects do exist, they could also be exploited by initiatives to counter disinformation.  

RECIPIENT CHARACTERISTICS

Of course, not all individuals react to encountering disinformation in the same way. In fact, leveraging the ‘consistency heuristic’ outlined above would rely on an interaction between the characteristics of the message and the person receiving it (e.g. consistent or inconsistent with their political orientation). Findings suggesting that only a minority of people share political disinformation also show that there must be some role for individual factors. 

Guess et al. (2019) found that older adults were much more likely to share false material. An explanation advanced for this is that they may have lower levels of
digital media literacy and are thus less able to detect a message as being untruthful, particularly when it appears to come from an authoritative or trusted source. The ‘digital literacy’ account has a strong hold, not least in political circles as a basis for interventions.\footnote{DCMS Committee recommendations.} The logic is that educational interventions will enable people to be able to detect disinformation (and there is evidence that such interventions do have positive effects).

It is also possible that personality characteristics may influence disinformation-related behaviour. This would be unsurprising, given that social media behaviour is influenced by personality in a number of ways.\footnote{Gil de Zúñiga et al., 2017.} However, evidence is somewhat patchy on this front. One study has found that lower Agreeableness was associated with increased likelihood of spreading a potential disinformation message (Buchanan and Benson, 2019), while other recent work found the same not for Agreeableness but for Conscientiousness (Buchanan, 2020). Other studies have looked for links with personality but found none.

### THE RESEARCH

The project ‘Why do people spread disinformation on social media?’\footnote{https://crestresearch.ac.uk/projects/disinformation-on-social-media/} was commissioned by CREST in 2019. Its overall aim was to explore some of the potential factors influencing organic reach: Why do people share false information online? In four scenario-based self-report studies with over 2,600 participants, the project explored the effect of message attributes (authoritativeness of source, consensus indicators, consistency with recipient beliefs) and viewer characteristics (digital media literacy, personality, and demographic variables) on the likelihood of spreading disinformation. The same procedure was used in all four studies, which respectively examined UK samples with Facebook (Study 1), Twitter (Study 2) and Instagram (Study 3) materials, and a US sample with Facebook materials (Study 4).

In each of the four studies, paid participants were drawn from commercial research panels. They completed demographic items; a five-factor personality test (measuring Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to experience); a measure of political orientation (conservatism); and a measure of digital media literacy. They were then shown three genuine examples of false information found on social media and asked how likely they would be to share each to their own social networks. The stories people saw were manipulated so that they appeared to come from sources more or less authoritative.

They were also manipulated to appear as if they had high or low levels of endorsement from other people (consensus / social proof). They all had a right-wing orientation, while participants came from across the political spectrum, so people varied in the level of consistency there was between the stories and their own political beliefs.

As well as their likelihood of sharing, people were asked about the extent to which they thought the stories were likely to be true, and how likely they thought it was that they had seen them before. Participants also reported whether they had in the past shared false political information online, either unwittingly or on purpose.
Perhaps the most encouraging finding was that most people rated their likelihood of sharing the disinformation examples as very low. This is consistent with real-world findings\textsuperscript{14} indicating that only a minority of people share false material.

However, people did vary in their reported likelihood of sharing, so it is important to try to understand why. The findings of all four studies are summarised in Table 1, which shows the variables that were positively or negatively related to self-reported likelihood of information sharing.
Table 1. Predictors of Self-reported Likelihood of Sharing, Retweeting or Liking Stimuli Across all Studies.

<table>
<thead>
<tr>
<th></th>
<th>Study 1 (Facebook sharing)</th>
<th>Study 2 (Twitter retweeting)</th>
<th>Study 3 (Instagram liking)</th>
<th>Study 4 (Facebook sharing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Media Literacy</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Conservatism</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Openness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Authoritativeness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consensus</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belief stories true</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Age</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gender (M=1, F=2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Education</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Platform use</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Likelihood seen before</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+ denotes a statistically significant positive association, - a negative association.

Table 2. Predictors of Unknowingly Sharing False Political Stories Across all Studies.

<table>
<thead>
<tr>
<th></th>
<th>Study 1 (Facebook sharing)</th>
<th>Study 2 (Twitter retweeting)</th>
<th>Study 3 (Instagram liking)</th>
<th>Study 4 (Facebook sharing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Media Literacy</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Conservatism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Openness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gender (M=1, F=2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Education</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Platform use</td>
<td>+</td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+ denotes a statistically significant positive association, - a negative association.
Table 3. Predictors of Knowingly Sharing False Political Stories Across all Studies.

<table>
<thead>
<tr>
<th>Study 1 (Facebook sharing)</th>
<th>Study 2 (Twitter retweeting)</th>
<th>Study 3 (Instagram liking)</th>
<th>Study 4 (Facebook sharing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Media Literacy</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Conservatism</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Openness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Gender (M=1, F=2)</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platform use</td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

+ denotes a statistically significant positive association, - a negative association.

Table 2 shows the variables that were associated with participants’ reports of whether they had previously shared political stories online that they later found out were false (13.1% to 29.0% reported doing this across the four studies).

Table 3 shows variables that were associated with participants’ reports of whether they had shared political stories that they knew at the time were false (6.2% to 20.7% of respondents).

In terms of the message characteristics, self-reported likelihood of sharing was not influenced by authoritativeness of the source of the material, nor indicators of how many other people had previously engaged with it (consensus). What was important, however, was the extent to which the stories were consistent with people’s political orientation, and the extent to which they thought the stories were likely to be true. There was also a relationship between likelihood of sharing and previous familiarity with the stories. People who thought it was likely they had seen the stories before were more likely to share them.

In terms of individual characteristics, effects were weak and inconsistent across populations, platforms, and behaviours. The people more likely to share disinformation were typically male, younger, and less educated. In terms of personality, they were lower in Agreeableness and Conscientiousness, and higher in Extraversion and Neuroticism. These effects were found in multiple studies, but only had very small effects.
Overall, the project found that spreading disinformation is a relatively rare behaviour. Only a minority of individuals forward false material they encounter. Prevalence estimates vary in this and other sources, but the proportion who are influenced into sharing false material because they think it is true may be around 10%. In practical terms, this suggests that any interventions at the ‘information operation’ level should be targeted at specific segments who actually share the material rather than the entire population. Based on findings discussed further below, this should be done along the lines of ‘interest group’ membership or political affiliation rather than based on personality traits or demographics. While the latter are feasible, the strength of the effects found is so low that it is unlikely to be worthwhile.

**CONSISTENCY**

Consistency with pre-existing attitudes, and particularly the belief that the material is true, increases the likelihood of sharing it. Therefore, to maximise the chance of messages being propagated, it is desirable to target messaging at individuals likely to be sympathetic to the cause. This is consistent with the idea of ‘activating the base’ rather than trying to change the minds of opponents or those with different or sceptical views.

**FAMILIARITY WITH MATERIAL**

Familiarity with material – operationalised here as higher self-reported likelihood of having seen it before – increases willingness to share it. Taken together with other recent work, current findings suggest that repeated exposure to disinformation materials may increase our likelihood of sharing it (even if we don’t believe it). A practical recommendation arising from this is that social media platforms should try to reduce the visibility of such information, rather than just flagging it as disputed. The DCMS Committee coronavirus misinformation report (2020) suggests that warning labels can be a useful and proportionate alternative to simply removing material. Findings here would suggest it is important not to ‘repeat the lie’ even when countering it. It would therefore be preferable to completely remove the false material, to prevent repeated exposure to it. Of course, for interventions based on counter-messaging, the same principles would dictate that one should try to maximise exposure to the true material one wants to spread.

**AUTHORITATIVENESS**

The authoritativeness of the original source of the material, as operationalised here, appears not to matter. This implies that there is no point in developing or spoofing ‘trustworthy’ personae if they are no more likely to have their posts shared than a random individual.

**CONSENSUS INDICATORS**

Consensus indicators (‘social proof’), operationalised as visible numerical indicators of engagement with a post, appeared not to matter here. It is possible to use bot networks or purchase likes from ‘manipulation service providers’ to artificially boost the numbers next to a post. However, there may be little value in doing so. What seems to matter is actual exposure – repeated ‘eyes on’ the material may be what makes people more likely to share even if they know it is not true. This ‘actual exposure’ principle also has implications for why high-profile figures such as celebrities may be influential amplifiers or conduits for disinformation: not because they are trusted, but because lots of people see the things they share. This finding may seem to run counter to other research: Innes, Dobreva and Innes (2019) talk about the role of social proof, operationalised using the feedback mechanisms of
social media sites, in stimulating the spread of false material online following terrorist incidents. According to current findings, the ‘social proof’ effect described by Innes et al. may reflect the fact that people had repeatedly seen messages rather than were influenced by numerical indicators of shares, likes and so on. However, a recent study based on a simulation game\textsuperscript{18} did find effects of engagement indicators, so it may be that further work is required to understand this.

In the current project, some people reported deliberately forwarding material they knew to be untrue. The proportion of people within the population who do this is not known but based on the numbers in these studies a percentage in the low single figures is not unrealistic. Of course, even a very low percentage of social media users is still a very large number of people, so the role of these individuals in spreading disinformation may be important. Attempts to influence or convince such individuals not to behave in this way are unlikely to be effective. This is likely to be motivated, intentional, partisan behaviour. One potential action social media companies could take is to block such accounts.

DIGITAL MEDIA LITERACY,\textsuperscript{19}

This finding also casts doubt on prevailing narratives about the importance of digital media literacy, at least with respect to the onward propagation of disinformation. In these studies, lower digital media literacy was not associated with increased self-reported likelihood of sharing disinformation stories. In fact, there were some indications that higher digital media literacy might actually be associated with a higher likelihood of sharing false material. If this finding is correct, it could be due to the individuals who choose to share false material on purpose: higher digital media literacy levels could be an enabling factor in that behaviour. Other recent work has produced similar insights, finding that older people did not share more disinformation due to lower (political) digital literacy or critical thinking: lower political digital literacy scorers actually shared less.\textsuperscript{19} Overall, the implication is that lower digital media literacy is not what causes people to share disinformation. Furthermore, some people are sharing known disinformation on purpose, not because they are unable to identify false information online. This implies that attempts to raise digital media literacy at a ‘whole population’ level, as recommended by DCMS Committee reports for example, are unlikely to stop the onward spreading of disinformation. This is not to say that such an intervention wouldn’t have positive effects on audience belief or reception of the material – it most likely would. However, it will not solve the problem of some people sharing false material and thus exposing others to it.

PERSONALITY

There was evidence from each of the studies that personality variables – particularly Low Agreeableness and Conscientiousness – are relevant to sharing disinformation. However, the effect sizes found were low, and effects were inconsistent across platforms and populations. This implies that targeting social media users for interventions on the basis of personality is unlikely to be of value. There most likely is a ‘personality effect’, which is of scientific interest, but knowing about it may have little applied value for professionals seeking to counter disinformation.

DEMOGRAPHIC VARIABLES

Similarly, demographic variables – particularly being young and male – are relevant to sharing disinformation. Again however, effect sizes are low and inconsistent. The current pattern of findings contradicts other research that has suggested older adults are more likely to share false material (but those claims assume digital media literacy is lower in such individuals, whereas it was controlled here). Again, the practical implication is that targeting social media users for interventions on the basis of general demographic characteristics is unlikely to be of value.

\textsuperscript{18} Avram et al, 2020.
\textsuperscript{19} Osmundsen et al, 2020.
WHY DO PEOPLE SHARE DISINFORMATION ON SOCIAL MEDIA?

When interpreting the current findings, a number of issues need to be borne in mind. The first is that these were self-report, scenario-based studies. While experiments such as these are able to test hypotheses about important variables in a controlled manner, they fall short of real-world behavioural observations in a number of ways. Thus, it would be desirable to evaluate whether the current findings on individual user characteristics (digital literacy, personality, political preference) hold true in a real-world context, measuring actual behaviour rather than self-reports or simulations. This would also enable critical questions about the likelihood of people sharing false information in real-world contexts to be addressed.

Exploring the motivation of individuals who share false material (both deliberately and unwittingly) is likely to be critical. Is it done with good intentions, or maliciously? Little is known about motivations; what work there is seems mostly speculative. The DCMS Committee coronavirus misinformation report (2020) again points to the importance of understanding this. It is very likely that different individuals will have different motivations for spreading different kinds of information. Treating all sharers – and indeed recipients – of false information as if they are the same would be a mistake. The same false material can be viewed as either disinformation (when it is spread deliberately for malicious reasons) or misinformation (when it is spread by people who may believe it is true and may be well-intentioned). The key distinction is the intent of the spreader. Understanding this will inform the kinds of interventions most likely to be effective with particular types of sharer.

A comprehensive anti-disinformation strategy should probably consider deliberate onward sharers, accidental onward sharers, and ultimate recipients (the wider audience), as well as the initial originators of the material. All are likely to require different intervention tactics.

Some of the comments above may seem to diminish the importance of digital media literacy. The current findings should not be taken to imply that it should be dismissed entirely. Observations here on digital literacy refer only to its effect on the onward spreading of false material: whether people lower in digital literacy are more likely to believe online falsehoods, and act on them in any way, requires further exploration.

Finally, it should be stressed that social media disinformation cannot be considered as an isolated phenomenon. It is part of a much wider information ecosystem, that in fact extends beyond hostile information operations or computational propaganda. It forms part of the same virtual sphere as online conspiracy theories, medical misinformation, deliberately polarising ‘hyperpartisan’ communication, humorous political memes, selective or decontextualized presentation of true information, and other phenomena. The bidirectional links between all these online phenomena, and also traditional broadcast and print media, need to be considered.

The lines between different behaviours are increasingly blurred. For example, the coronavirus ‘infodemic’ saw influential political actors sharing medical misinformation online and in traditional media, while online rumours found expression in physical-world actions. As amply demonstrated by the Intelligence and Security Committee’s (2020) ‘Russia Report’, there are big questions about how, and by whom, these types of phenomena should be addressed. But given the potential harms to society, there is a nettle here that needs to be grasped.

ABOUT THE AUTHOR

Tom Buchanan, Professor of Psychology, University of Westminster, London


WHY DO PEOPLE SHARE DISINFORMATION ON SOCIAL MEDIA?


For more information on CREST and other CREST resources, visit

www.crestresearch.ac.uk