

Trust, personality, and belief as determinants of the organic reach of political disinformation on social media

Tom Buchanan

To cite this article: Tom Buchanan (2021): Trust, personality, and belief as determinants of the organic reach of political disinformation on social media, The Social Science Journal, DOI: [10.1080/03623319.2021.1975085](https://doi.org/10.1080/03623319.2021.1975085)

To link to this article: <https://doi.org/10.1080/03623319.2021.1975085>



© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 15 Sep 2021.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Trust, personality, and belief as determinants of the organic reach of political disinformation on social media

Tom Buchanan 

School of Social Sciences, University of Westminster, London, UK

ABSTRACT

False political information spreads far and fast across social media, with negative consequences for society. Individual users play a key role in sharing such material, extending its range through the phenomenon of organic reach. An online experiment tested the hypotheses that higher trust in the source of false information, and lower agreeableness of the person encountering it, would predict their likelihood of extending its reach. One hundred and seventy-two participants saw real examples of disinformation stories that had been posted to social media and rated their likelihood of sharing and interacting with it in other ways. Neither trust in the source nor agreeableness influenced organic reach. However, people lower in conscientiousness rated themselves as more likely to extend its reach, as did people who believed the stories more likely to be true.

ARTICLE HISTORY

Received 13 April 2021
Revised 17 August 2021
Accepted 18 August 2021

KEYWORDS

Disinformation;
misinformation; fake news;
personality; trust; social
media

Introduction

Social media disinformation is a critical problem for society (Buchanan & Benson, 2019). Online spaces are awash with false political stories, created with the express intent of deceiving others and causing harm. Typical definitions of ‘disinformation’ revolve around false material which is shared with harmful intent, e.g., Freelon and Wells (2020). A recent worldwide survey found that the majority of people were ‘concerned about what is real and fake on the internet when it comes to news’, rating it as the number one internet-related risk (Newman et al., 2020, p. 18).

Once false material is initially posted online – to social network platforms or dedicated websites – it can then be spread to huge audiences (‘going viral’). Much of the spread of disinformation can be attributed to human action (Vosoughi et al., 2018). People encounter the material on social media, then share it to their own networks of contacts, friends and family, increasing its reach. They may also engage with it in different ways – ‘liking’ it, commenting on it, or performing other actions that cause a social network site’s algorithms to increase the visibility of the material to others. This is described as the ‘organic reach’ of a post (Buchanan & Benson, 2019).

The number of people who actively share false material online is relatively low – perhaps lower than 10% (Guess et al., 2019). However, the vast scale of social networks means that if even a small minority of users behave in this way, they can significantly increase the reach of the false information. This in turn can lead to troubling outcomes for society: for example, influencing electoral processes or damaging public health. Therefore, it is important to consider who shares false information they encounter online: what sets them apart from those who do not? Knowing this has both scientific and practical values. Understanding the processes and people involved will inform development of effective interventions.

CONTACT Tom Buchanan  T.Buchanan@westminster.ac.uk  School of Social Sciences, University of Westminster, 115 New Cavendish St, London W1W 6UW, UK

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

A growing body of work is examining the psychological processes associated with sharing false material online. Much of this revolves around mechanisms associated with use of heuristics, cognitive ability and digital literacy, demographic factors, and political partisanship (Anthony & Moulding, 2019; Celliers & Hattingh, 2020; Pennycook et al., 2018; Pennycook & Rand, 2019; Talwar et al., 2019). One potential determinant of whether people spread material online is the extent to which they trust the source of that information (Deng et al., 2017). For instance, are we more likely to share material that we see has come from a reputable news source than an unknown user?

Trust

There is evidence that when we see information that comes from a trusted sharer, we are more likely to recommend the source of that news to others (The Media Insight Project, 2017), and that higher levels of trust in online information in general are associated with self-reports of sharing ‘fake news’ online (Talwar et al., 2019). It has been found that individuals rated themselves as more likely to extend the organic reach of a potentially untrue story when they had been asked to share it by a trusted contact (a close friend) as opposed to a less trusted acquaintance (Buchanan & Benson, 2019). However, while that study found that participants did rate the friend as more trustworthy in a manipulation check, the effect found could actually have been due to the closeness of the social relationship with the contact. It is known that people are more likely to comply with requests from friends as opposed to strangers (e.g., Boster et al., 1995). The manipulation used by Buchanan and Benson (2019) confounded trustworthiness with relational intimacy. Therefore, while it seems likely that people will be more ready to share disinformation that comes from a source they trust, further experimental examination is required to verify this.

Agreeableness

Relatively little research has examined personality traits as determinants of sharing false information. Personality is relevant here, because it is known that a number of personality traits are associated with aspects of social media behavior (Hall et al., 2013) to the extent that personality can be reliably detected from the ‘data footprints’ left in our social media profiles (Azucar et al., 2018). Past work has found people lower in trait agreeableness rated themselves as more likely to extend the organic reach of a potentially untrue story (Buchanan, 2020; Buchanan & Benson, 2019). A possible explanation is that political disinformation items are normally overtly negative about their target, often being outright attacks on specific individuals, groups, nations, or issues. Disagreeable people are themselves more critical and hostile in their outlook. They are also less concerned about offending others. A more agreeable person might hesitate to forward a story to their family and friends if they knew it would be received negatively, even though they privately believed it.

A ‘social’ explanation such as this might be more likely than an explanation revolving around inability of less agreeable people to identify false information. Findings of previous research on links between agreeableness and discernment of false from true news information have been inconsistent. Wolverton and Stevens (2019) suggest that individuals lower in agreeableness are better able to identify false news headlines.¹ Conversely, however, in a larger study with more reliable measures that controlled for other variables, Calvillo et al. (2021) reported that higher agreeableness was related to better ability to discern true from false headlines. Meanwhile, in a larger study still, Sindermann et al. (2021) found no effect of agreeableness on ability to distinguish true and false headlines.

¹The findings of Wolverton and Stevens (2019) should be interpreted with caution, given criticisms by Calvillo et al. (2021) and Sindermann et al. (2021) on a number of grounds.

There are also indications that agreeableness may interact with trust (more agreeable people are typically more trusting). Effects of agreeableness on general online information sharing may be attributable to higher levels of generalized trust (Deng et al., 2017). Therefore, there is a strong rationale for considering agreeableness alongside trust.

Aims and hypotheses

This study aims to extend our understanding of how trust in the source of false material, and the personality of the individual encountering it, influence likelihood of them extending its organic reach. To increase generalizability and ecological validity, it will use multiple examples of genuine disinformation stories as stimuli. It will address some of the limitations of previous research (Buchanan & Benson, 2019) in that interpersonal social distance between the participant and the person asking them to share material will be held constant, but the trustworthiness of the original information source will be manipulated. Two specific hypotheses are advanced:

Hypothesis 1: Organic reach will be higher for information portrayed as originating from a more trusted source; and

Hypothesis 2: Participants lower in agreeableness will be more likely to increase the organic reach of the disinformation stories.

A number of other variables will be measured (demographic data and personality traits, and the extent to which people believed the stories were true), for use as controls and in exploratory analyses. No specific hypotheses are advanced concerning these.

Materials and methods

The study adopted an experimental design. Participants were randomly allocated to one of the three conditions. In each, they saw three ‘news’ items that had previously been posted on social media. The items were presented as if they had been originated by sources of high, medium, or low trustworthiness according to the experimental condition.

The study was conducted online using the Qualtrics online research platform, with paid participants recruited through the Prolific research panel. Participants were paid UK £1 for their participation. All were UK citizens, were currently resident in the UK, and used Facebook at least once per month. The Prolific participant pool is heavily biased towards younger individuals. Therefore, to ensure that older adults were represented in the sample, recruitment was run in two phases: one restricted to people aged 60+, and one open to anyone aged 18+ . Ethical approval for the study came from the Psychology Department Research Ethics Committee of the host university.

Materials

All materials were presented online, on the Qualtrics platform. Materials, data, and analysis syntax can be seen at <https://osf.io/dt6vj>.

Personality

Personality was measured using a 41-item International Personality Item Pool-derived measure of Extraversion, Neuroticism, Openness to Experience, Agreeableness and Conscientiousness. This has been validated for use on the internet (Buchanan et al., 2005).

Disinformation items

Participants each saw the same three ‘news’ items that had previously been posted on social media. One was titled ‘REVEALED: UN PLAN TO FLOOD AMERICA WITH 600 MILLION MIGRANTS’; one was titled ‘US sends 3,600 tanks against Russia – Massive NATO deployment underway’; and one dealt with the 2018 ‘Novichok’ poisonings in Salisbury, UK. All of these had been identified as examples of disinformation (Nimmo, 2017; Nimmo et al., 2019; Valverde, 2018). Rather than full articles, excerpts (screenshots) were used that had the size and general appearance of what respondents might expect to see on social media sites. Three disinformation items were used, and responses compiled across all three, to ensure that any findings should not be unique to any specific item. The stimuli can be seen at <https://osf.io/dt6vj>.

Sources

The disinformation items were presented to participants as having been originated by sources high, medium, or low in trustworthiness (three of each). In each case, the source was presented as a combination of an avatar and a username, as they would normally be encountered in social media.

The high trustworthiness sources were the social media usernames of reputable UK news organizations. These were ‘BBC News (World)’, ‘The Sunday Times’, and ‘ITV News’.

The medium trustworthiness sources were the usernames of accounts that had been identified by Twitter as having been fakes set up by the Russian ‘Internet Research Agency’ (US House of Representatives Permanent Select Committee on Intelligence, n.d.). These resembled US local news organizations: ‘Chicago Daily News’, ‘Los Angeles Daily’ and ‘El Paso Top News’.

The low trustworthiness sources were fictitious usernames based on accounts that had shared material identified as disinformation. Modified usernames were used to reduce the likelihood of individuals being identified in future searches. They were ‘applesofeden_73 #IamSOLDIERX’, ‘wtf!!!’ and ‘Information Wars’. The sources used can be seen at <https://osf.io/dt6vj>.

To establish that the sources really did differ in the extent to which they would be perceived as trustworthy by participants, they were compared in a pilot study using participants drawn from the same population as the main study. A sample of 20 individuals rated all sources on the item ‘How trustworthy is this user as a source of information?’ using a scale from 0 to 10 anchored at ‘Not at all trustworthy’ and ‘Very trustworthy’. Mean scores for each of the nine sources used are shown in Table 1. A repeated measures ANOVA indicated that there were statistically significant differences between the three groups of sources (Wilks’ Lambda =.052, $F_{(2,18)}=162.86$, $p<.001$), with pairwise comparisons indicating that each of the three groups differed from both the others in the expected manner.

Organic reach

The key-dependent measure was the participants’ rating of their likelihood to extend the organic reach of the set of stimuli. For each stimulus, participants were asked how likely they would be to share it to their own public timeline, to ‘like’ it, to comment on it (either positively or negatively), and to react to

Table 1. Pilot ratings of trustworthiness for sources.

	N	Min	Max	Mean	SD
ITV News	20	6	10	8.50	1.47
BBC News (World)	20	6	10	8.50	1.57
The Sunday Times	20	6	10	8.40	1.76
Chicago Daily News	20	2	10	6.30	2.18
Los Angeles Daily	20	2	9	6.15	2.06
El Paso Top News	20	0	7	4.50	2.26
Information Wars	20	0	5	1.65	1.31
wtf!!!	20	0	4	0.80	1.11
applesofeden_73 #IamSOLDIERX	20	0	3	0.75	0.97

Ratings came from 20 individuals drawn from same population as eventual sample, rating all sources on ‘How trustworthy is this user as a source of information?’ using a 10-point scale.

it by posting an emoji. For each, they rated their likelihood on a 5-point scale anchored at ‘Not at all likely’ and ‘Very likely’. These ratings were summed to create a composite measure of organic reach: the overall likelihood of the participant behaving in such a way as to extend the organic reach of all three source/story pairings they saw.

Procedure

Participants gave informed consent, then completed demographic items (gender, country of residence, education, age, occupational status, Facebook use), and the personality inventory. Participants were then randomized to one of the three experimental conditions, where each saw the same three stories but attributed to the high, medium, or low trustworthiness sources depending on the condition. Within each condition, six different source/story combinations were used with participants being randomized to different combinations. For example, in one variant, participants saw the ‘Novichok’ story attributed to ITV News, in another they saw it attributed to The Sunday Times. This was done to avoid any effects arising from specific source/story combinations.

For each of the three source/story combinations each participant saw, they were told ‘A friend of yours recently shared this on Facebook, commenting that they thought it was important and asking all their friends to share it.’ They were then asked how likely they would be to share it to their own public timeline, to ‘like’ it, to comment on it (either positively or negatively), and to react to it by posting an emoji. When they had responded to all three stories (each on a separate page), they were shown all three again on the same page and asked to indicate on a 5-point scale how likely they thought it was that the message was truthful. For analysis, ratings were summed across all three stimuli to create compound measures applying to the whole set of materials rather than single stories. Finally, participants reconfirmed consent, then were debriefed.

Data screening and processing

The data were examined for evidence of inauthentic or ineligible responding. Multiple completions by the same individual were precluded by Prolific’s screening processes and Qualtrics’ proprietary technologies. One participant gave their age as 12 years, one gave their country as Eire. One 18-year old claimed to have a master’s degree. Four people indicated they did not ever use Facebook. These seven were removed from the sample.

Participants

Target sample size was initially set at 179 participants. This was sufficient to achieve 80% power to detect an R^2 of .081 (as reported in a comparable analysis by Buchanan & Benson, 2019) in a multiple regression with eight predictors (source trustworthiness, the five personality dimensions, age, and gender). Following data screening, 172 participants were included in the final sample, reducing power slightly. Demographic characteristics are shown in Table 2. All participants were UK citizens and residents. While there was considerable heterogeneity, the majority were employed women with post-secondary education. Ages ranged from 18 to 82, with 23.8% being aged 60 years or older.

Results

Descriptive statistics for key predictor and outcome variables are shown in Table 3. In common with other work (Buchanan & Benson, 2019) the four items indexing organic reach (share, like, comment, react) were summed to create an overall measure of organic reach with acceptable reliability. Overall, participants tended to rate themselves as unlikely to engage with the material in these ways: the score

Table 2. Demographic data.

<i>N</i>		172
Sex	Men	43 (25.0%)
	Women	128 (74.4%)
	Other	1 (0.6%)
Highest level of education	Less than high school	3 (1.7%)
	High school/secondary school or equivalent	31 (18.0%)
	Some college/University	45 (26.2%)
	College/University Graduate	71 (41.3%)
	Master's degree	17 (9.9%)
	Doctoral degree	3 (1.7%)
	Professional degree (JD, MD)	2 (1.2%)
Occupation	Employed for Wages	91 (52.9%)
	Self-employed	26 (15.1%)
	Unemployed	8 (4.7%)
	Home-maker	6 (3.5%)
	Student	13 (7.6%)
	Retired	23 (13.4%)
	Unable to work	5 (2.9%)

Percentages may not sum exactly to 100% due to rounding errors.

distribution for the compound measure was positively skewed. This is consistent with other research finding that the proportion of people likely to share false material online is relatively low (Guess et al., 2019). One reason for this might be that of course not everyone is interested in politics. The combined indices of organic reach and belief that the materials were likely to be true are broken down by condition in Table 4.

The hypotheses that personality variables and trustworthiness of the message originator would influence organic reach were simultaneously tested using multiple regression. The analysis, summarized in Table 5, indicated that this set of variables explained little variance in the self-reported likelihood of extending the reach of the three stimuli, with Adjusted $R^2 = 0.04$. The only variable which statistically significantly predicted reach was conscientiousness ($\beta = -.24$, $t_{(171)} = -2.50$, $p = .013$). People lower on conscientiousness reported themselves as more likely to propagate the

Table 3. Descriptive statistics for personality scales, organic reach, likely truthfulness of materials, and age.

Variable	<i>N</i>	<i>M</i>	<i>SD</i>	Cronbach's α	Range of scores		Skew	Kurtosis
					Potential	Actual		
Openness	172	25.36	5.04	.72	10–35	7–35	−0.48	−0.10
Neuroticism	172	22.77	7.01	.87	8–39	8–40	0.04	−0.38
Extraversion	172	26.92	7.23	.89	9–43	9–45	−0.26	−0.50
Conscientiousness	170	36.36	6.78	.87	21–50	10–50	−0.02	−0.75
Agreeableness	172	27.94	4.11	.78	11–35	7–35	−1.19	1.90
Organic Reach	172	17.62	8.29	.86	12–60	12–49	1.93	3.52
Likely Truthfulness	172	6.15	2.80		3–15	3–15	1.02	0.85
Age	172	41.94	16.54			18–82	.50	−.89

Table 4. Organic reach and likely truthfulness scores by trust condition.

Trust Condition	<i>n</i>	Organic	Reach	Likely	Truthfulness
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Lowest	57	17.53	7.36	5.72	2.54
Middle	59	16.27	6.85	6.12	2.53
Highest	56	19.14	10.22	6.61	3.26

disinformation messages. No other variables significantly predicted likelihood of expanding organic reach. Collinearity statistics (Tolerance and VIF) did not indicate any issues with multicollinearity, and the Durbin–Watson statistic (2.13) did not indicate a problem with independence of residuals. Given the skewed distribution of the dependent variable, the relationship between conscientiousness and reach was checked using a non-parametric Spearman correlation. This was again statistically significant ($r_s = -.28, p < .001, N = 169$).

Additional exploratory analyses examined the role of belief in the disinformation stories. Ratings of the likelihood that the stories were likely to be truthful were positively correlated with the self-reported likelihood of extending the reach of the three stimuli ($r = .22, n = 172, p = .004$). Repeating the multiple regression analysis summarized in Table 5, but using belief in the stories as the outcome variable, indicated that only age significantly (negatively) predicted belief in the stories (Table 6): older people were less likely to believe them ($\beta = -.30, t_{(170)} = -3.84, p < .001$). Notably, trustworthiness condition did not influence belief ($\beta = .11, t_{(170)} = 1.47, p = .14$): there was no evidence that trustworthiness of source influenced ratings of how likely the stories were to be true. Collinearity statistics (Tolerance and VIF) did not indicate any issues with multicollinearity, and the Durbin–Watson statistic (2.15) did not indicate a problem with independence of residuals. The relationship between age and belief was checked using a non-parametric Spearman correlation. This was again statistically significant ($r_s = -.36, p < .001, N = 171$).

Table 5. Effect of trustworthiness condition, personality, age, and gender on organic reach.

	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	Tolerance	VIF
Constant	26.05	8.25		3.16**	.002		
Trust Condition	0.41	0.77	0.04	0.54	.59	.95	1.05
Openness	−0.13	0.12	−0.08	−1.04	.30	.95	1.05
Neuroticism	0.06	0.12	0.05	0.47	.64	.50	2.02
Extraversion	0.09	0.10	0.09	0.95	.34	.72	1.34
Conscientiousness	−0.28	0.11	−0.24*	−2.50	.013	.62	1.61
Agreeableness	0.19	0.17	0.10	1.13	.26	.78	1.28
Age	0.00	0.04	0.00	−0.06	.96	.85	1.18
Gender	−2.79	1.48	−0.15	−1.88	.06	.89	1.13
R^2		0.08					
Adjusted R^2		0.04					
<i>F</i>		1.83			.08		

Trust condition coded as 1 (lowest) to 3 (highest). Gender coded as 1 (m) or 2 (f). One individual describing gender as ‘other’ was excluded from this analysis.

* $p < .05$, ** $p < .05$.

Table 6. Effect of trustworthiness condition, personality, age, and gender on rated truthfulness of stories.

	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	Tolerance	VIF
Constant	10.05	2.76		3.64***	<.001		
Trust Condition	0.38	0.26	0.11	1.47	.14	.95	1.05
Openness	−0.07	0.04	−0.13	−1.72	.09	.95	1.05
Neuroticism	0.05	0.04	0.12	1.21	.23	.50	2.02
Extraversion	−0.02	0.03	−0.06	−0.72	.47	.72	1.34
Conscientiousness	0.01	0.04	0.03	0.34	.73	.62	1.61
Agreeableness	−0.04	0.06	−0.06	−0.67	.50	.78	1.28
Age	−0.05	0.01	−0.30	−3.84***	<.001	.85	1.18
Gender	−0.38	0.50	−0.06	−0.77	.44	.89	1.13
R^2		0.18					
Adjusted R^2		0.13					
<i>F</i>		4.24***			<.001		

Trust condition coded as 1 (lowest) to 3 (highest). Gender coded as 1 (m) or 2 (f). One individual describing gender as ‘other’ was excluded from this analysis.

*** $p < .001$.

Finally, the regression summarized in Table 5 was repeated with perceived truthfulness as an additional predictor. Summarized in Table 7, this showed that conscientiousness remained a significant predictor of reach ($\beta = -.25$, $t_{(170)} = -2.64$, $p = .01$), indicating that its influence was independent of belief in the stimuli. Belief in the stimuli also significantly predicted reach ($\beta = .24$, $t_{(170)} = 2.96$, $p = .004$). People who rated the stories as more likely to be true also rated themselves as more likely to engage with them in ways that would extend organic reach. Collinearity statistics (Tolerance and VIF) did not indicate any issues with multicollinearity, and the Durbin–Watson statistic (2.01) did not indicate a problem with independence of residuals. The relationship between belief and reach was checked using a non-parametric Spearman correlation. This was again statistically significant ($r_s = .22$, $p = .003$, $N = 171$).

Discussion

Hypothesis 1 predicts that participants seeing disinformation stories portrayed as coming from more trustworthy sources would rate themselves as more likely to extend its organic reach in various ways. However, the data indicated that a story attributed to a trustworthy news source such as the BBC News service was no more likely to be engaged with than one attributed to a username such as ‘Information Wars’. Trustworthiness of the source also did not affect participants’ ratings of how likely the stories were to be true. This runs counter to previous findings (Buchanan & Benson, 2019) where a more trusted contact (close friend) was more likely to elicit organic reach than a less trusted one (virtual stranger). It suggests that those results could actually be attributed to interpersonal distance rather than trustworthiness, with which it was confounded. This finding is consistent with suggestions that how much an individual trusts content is determined less by who originally created the news story and more by the person who actually shares it with you (The Media Insight Project, 2017). In the current study, the putative person sharing it was described as ‘a friend of yours’ in each case. Thus, as argued by The Media Insight Project (2017), one might hypothesise that the proximal (person who brought it to our attention; in this case, the ‘friend’ described as sharing it with participants) rather than distal (in this case the original publisher) sharers of disinformation stories are the important factor in how we react to it. If we trust the proximal sharer, or are interpersonally close to them, we may not pay attention to the original publisher. However, the framing of the materials here as shared by ‘a friend’ could be interpreted in multiple ways (e.g., Facebook contact vs. ‘real’ friend). A way to test this in future research could be a two-way experimental design, where original publisher (with different levels of trustworthiness, as in the current design) was crossed with immediate proximal sharer (close friend vs. virtual stranger).

Table 7. Effect of trustworthiness condition, personality, age, gender, and likely truthfulness on organic reach.

	<i>B</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	Tolerance	VIF
Constant	19.19	8.38		2.29*	.02		
Trust Condition	0.16	0.75	0.02	0.21	.84	.94	1.07
Openness	−0.08	0.12	−0.05	−0.66	.51	.93	1.07
Neuroticism	0.02	0.12	0.02	0.20	.84	.49	2.04
Extraversion	0.11	0.10	0.10	1.14	.26	.72	1.39
Conscientiousness	−0.29	0.11	−0.25	−2.64**	.009	.62	1.61
Agreeableness	0.21	0.16	0.11	1.31	.19	.78	1.28
Age	0.03	0.04	0.07	0.81	.42	.78	1.29
Gender	−2.53	1.45	−0.14	−1.75	.08	.88	1.13
Likely Truthfulness	0.68	0.23	0.24	2.96**	.004	.83	1.21
<i>R</i> ²		0.13					
Adjusted <i>R</i> ²		0.08					
<i>F</i>		2.68**			.01		

Trust condition coded as 1 (lowest) to 3 (highest). Gender coded as 1 (m) or 2 (f). One individual describing gender as ‘other’ was excluded from this analysis.

* $p < .05$, ** $p < .01$.

Hypothesis 2 predicts that people lower in agreeableness would rate themselves as more likely to extend the organic reach of disinformation material. However, there was no effect of agreeableness. Again, this is inconsistent with the previous work (Buchanan, 2020; Buchanan & Benson, 2019).

In exploratory analyses, perceived truthfulness of the disinformation stories proved to be an important factor. Participants who rated the messages as more likely to be truthful also rated themselves as more likely to extend their organic reach. Intuitively, this result is unsurprising: we are more likely to share, or otherwise engage with, stories that we believe in. This principle is long established. For example, Knapp (1944) suggested that ‘No rumor will travel far unless there is already a predisposition among those who hear it to lend it credence’ (Knapp, 1944, p. 27). It is also consistent with findings from other disinformation research (e.g., Buchanan, 2020) that people ‘share fake news stories that support their pre-existing beliefs’ (Marwick, 2018, p. 507). However, some research also suggests there is at least a partial disconnect between believing and sharing false material. There is evidence (e.g., Buchanan, 2020; Buchanan & Kempley, 2021; Chadwick et al., *in press*) that some people report sharing material they knew at the time to be untrue. This makes motivations for sharing important to consider in future work considering why some people might deliberately amplify false material (e.g., Chadwick et al., *in press*). Related to this, it has been found (Pennycook et al., 2020) that priming people to think about accuracy increases their likelihood of sharing true relative to false headlines. However, this would not create a methodological problem for the current study, given that people were not asked to rate the stories’ likelihood of being true until after they had provided their ratings about likely engagement.

Among the personality variables measured, only conscientiousness was associated with participants’ ratings of their likelihood to engage with the material. More conscientious people were less likely to extend its organic reach. This is plausible, as more conscientious individuals might be expected to ‘fact check’ what they share more carefully. As more careful individuals, they might be less likely to share potentially dubious material. This is in line with evidence from Buchanan (2020) that less conscientious people were more likely to report having shared false material accidentally in the past. Calvillo et al. (2021) also report that people high in conscientiousness are better able to discern false from true news – this is consistent with the current finding that those higher in conscientiousness would be less likely to amplify falsehoods.

Finally, it is worth noting one effect that was not observed. Influential past work (Guess et al., 2019) has flagged the role of older individuals in spreading disinformation, with people over 65 showing much higher rates of sharing stories from ‘fake news’ domains. Similarly, Brashier and Schacter (2020) argue that older adults are more susceptible to disinformation. However, in this study age had no effect on the composite measure of organic reach. Other work has argued that older adults may actually be less susceptible: for example, Buchanan (2020) reports some evidence of a negative relationship between age and self-reports of sharing false material. Calvillo et al. (2020) report that in the context of COVID-19 disinformation, older adults were better able to discern true from false headlines. That is consistent with the finding in this study that older people rated the stimuli as less likely to be true. Overall, the relationship between age and engagement with disinformation appears to be complex. It is unlikely to be a linear relationship that can easily be attributed to factors such as digital literacy (e.g., Brashier & Schacter, 2020). There may be a mismatch between intentions and actual behaviours, given that Brashier and Schacter (2020) report that age is negatively associated with self-rated willingness to share false material. Furthermore, age-related variables such as political orientation and attitudes or concern about specific health issues are likely to play a role.

Limitations

This study clearly has limitations. It is based on self-reports of the likelihood of engaging in various behaviours, rather than actual real-world observations of those behaviours. Work examining whether the relationships found here are replicated for actual sharing of material online would be desirable. However, there is evidence that self-reports of social media sharing do parallel actual behaviour (Mosleh et al., 2020).

The framing of the stimulus material as having been posted by ‘a friend of yours’ could also be seen as problematic. As noted above, this could be perceived in multiple ways, and participants’ responses could be influenced by whether they considered the person a real friend or just a casual acquaintance. This might be especially pertinent when considering polarizing topics such as politics, where people might well be concerned about the impressions they make on others. Again, observational behavioural research on how people react to such materials in naturalistic settings would be a good way of eliminating this potential demand characteristic.

Another issue is that sharing disinformation – or engaging with it in other ways – is a relatively rare behaviour, leading to a skewed distribution in the measure of organic reach used here. A number of factors might contribute to this. It may be that participants clearly recognised the stories as false, and most did not want to say they would engage with them for that reason. It may be that many participants were simply not interested in political issues. Or it may simply be the fact that in reality, relatively few people share false material online (Guess et al., 2019). All these factors could contribute to skewed distributions, restricted range, and floor effects on the variables of interest, with the effect of attenuating any associations that might exist between variables. It may be that different research strategies (e.g., purposive sampling of individuals known to have shared false material) would be of value to counter this issue.

Given that this study was an unsupervised online experiment, it is possible that participants ‘cheated’ – for example, by searching for the news stories online, before deciding whether they would share them or thought they were true. However, the limited evidence available militates against, rather than supports, this suggestion: the typical mean time people spent examining and responding to each of the three stories was around 30 seconds. This does not really allow time for ‘fact checking’ during the study.

Most participants had at least some post-secondary education. Allcott and Gentzkow (2017) have shown that education was positively correlated with ability to judge the veracity of news stories. While education levels within this sample did not correlate significantly with either reach ($r_s = .03$, $p = .72$) or belief ($r_s = .03$, $p = .74$) in the stories, it is possible that scores on the dependent variables were lower, and their relationships with personality variables attenuated, relative to a more representative population sample.

Implications and conclusions

In this study, the trustworthiness of their source did not appear to matter as a determinant of whether people would engage with false stories. This suggests that interventions intended to teach social media users to identify trustworthy and untrustworthy sources would have little impact. This chimes with work by Helmus et al. (2020) who found that revealing the source of Russian propaganda materials had no effect on participants’ ratings as to whether they would share them.

Rather, the finding that conscientiousness was negatively associated with likelihood to increase organic reach lends support to interventions associated with ‘inattention’ accounts of why people share false material (e.g., Pennycook et al., 2020). People high in conscientiousness are more likely to pay attention to details (e.g., Witt, 2002). The findings for conscientiousness also suggest a hypothesis about people who may be more likely to engage with fake news. Among others, Wolverton and Stevens (2019) and Buchanan (2020) have argued that it is important to identify personality traits that may confer vulnerability to being deceived by false information, in order to inform the development of targeted interventions as well as cast light on underlying processes. Conscientiousness may thus be a worthwhile focus of future behavioural research.

In conclusion, it appears that the trustworthiness of the original publisher of a disinformation story does not affect willingness to extend its organic reach by sharing or otherwise interacting with it. Rather, it may be the relationship between the person seeing the story and the person who forwards it to them that affects their likelihood of them engaging with it. On the basis of the current data, people low in agreeableness are not more likely to engage with disinformation. However, people low in conscientiousness are. While this pattern of results was inconsistent with predictions, it does reinforce the notion that individual differences in personality are relevant to our interactions with disinformation. Finally, it was found that people are more likely to extend the organic reach of disinformation if they believe it is true. Given that the stories used as stimuli in this study were patently *untrue*, this suggests that factors influencing belief in false information are likely to be very relevant to the likelihood of spreading it. Future research should examine the consistency of material with individuals' pre-existing beliefs as a predictor of them sharing false material.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Tom Buchanan  <http://orcid.org/0000-0002-8994-2939>

References

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Anthony, A., & Moulding, R. (2019). Breaking the news: Belief in fake news and conspiracist beliefs. *Australian Journal of Psychology*, 71(2), 154–162. <https://doi.org/10.1111/ajpy.12233>
- Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*, 124, 150–159. <https://doi.org/10.1016/j.paid.2017.12.018>
- Boster, F. J., Rodriguez, J. I., Cruz, M. G., & Marshall, L. (1995). The relative effectiveness of a direct request message and a pregiving message on friends and strangers. *Communication Research*, 22(4), 475–484. <https://doi.org/10.1177/009365095022004005>
- Brashier, N. M., & Schacter, D. L. (2020). Aging in an era of fake news. *Current Directions in Psychological Science*, 29(3), 316–323. <https://doi.org/10.1177/0963721420915872>
- Buchanan, T. (2020). Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *PLoS One*, 15(10), e0239666. <https://doi.org/10.1371/journal.pone.0239666>
- Buchanan, T., & Benson, V. (2019). Spreading disinformation on Facebook: Do trust in message source, risk propensity, or personality affect the organic reach of “Fake news”. *Social Media + Society*, 5(4), 1–9. <https://doi.org/10.1177/2056305119888654>
- Buchanan, T., Johnson, J. A., & Goldberg, L. R. (2005). Implementing a five-factor personality inventory for use on the Internet. *European Journal of Psychological Assessment*, 21(2), 115–127. <https://doi.org/10.1027/1015-5759.21.2.115>
- Buchanan, T., & Kempley, J. (2021). Individual differences in sharing false political information on social media: Direct and indirect effects of cognitive-perceptual schizotypy and psychopathy. *Personality and Individual Differences*, 182, 111071. <https://doi.org/10.1016/j.paid.2021.111071>
- Calvillo, D. P., Garcia, R. J. B., Bertrand, K., & Mayers, T. A. (2021). Personality factors and self-reported political news consumption predict susceptibility to political fake news. *Personality and Individual Differences*, 174, 110666. <https://doi.org/10.1016/j.paid.2021.110666>
- Calvillo, D. P., Ross, B. J., Garcia, R. J. B., Smelter, T. J., & Rutchick, A. M. (2020). Political ideology predicts perceptions of the threat of COVID-19 (and susceptibility to fake news about it). *Social Psychological and Personality Science*, 11(8), 1119–1128. <https://doi.org/10.1177/1948550620940539>

- Celliers, M., & Hattingh, M. (2020). A systematic review on fake news themes reported in literature. In M. Hattingh, M. Matthee, H. Smuts, I. Pappas, Y. K. Dwivedi, & M. Mäntymäki. (Eds.), *Lecture notes in computer science: Responsible design, implementation and use of information and communication technology* (pp. 223–234). Springer International Publishing. https://doi.org/10.1007/978-3-030-45002-1_19
- Chadwick, A., Vaccari, C., & Kaiser, J. (in press). The amplification of exaggerated and false news on social media: The roles of platform use, motivations, affect, and ideology. *American Behavioral Scientist*.
- Deng, S., Lin, Y., Liu, Y., Chen, X., & Li, H. (2017). How do personality traits shape information-sharing behaviour in social media? Exploring the mediating effect of generalized trust. *Information Research: An International Electronic Journal*, 22(3). <https://eric.ed.gov/?id=EJ1156382>
- Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, 37(2), 145–156. <https://doi.org/10.1080/10584609.2020.1723755>
- 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. <https://doi.org/10.1126/sciadv.aau4586>
- Hall, J. A., Pennington, N., & Lueders, A. (2013). Impression management and formation on Facebook: A lens model approach. *New Media & Society*, 16(6), 958–982. <https://doi.org/10.1177/1461444813495166>
- Helmus, T. C., James, V. M., Marek, N. P., & Danielle, S. (2020). Russian propaganda hits its mark: Experimentally testing the impact of Russian propaganda and counter-interventions. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA704-3.html
- Knapp, R. H. (1944). A psychology of rumor. *Public Opinion Quarterly*, 8(1), 22–37. <https://doi.org/10.1086/265665>
- Marwick, A. E. (2018). Why do people share fake news? A sociotechnical model of media effects. *Georgetown Law Technology Review*, 2(2), 474–512. <https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Marwick-pp-474-512.pdf>
- The Media Insight Project. (2017). 'Who shared it?': How Americans decide what news to trust on social media. The Media Insight Project. Retrieved June 26, 2020, from <https://apnorc.org/projects/who-shared-it-how-americans-decide-what-news-to-trust-on-social-media/>
- Mosleh, M., Pennycook, G., & Rand, D. G. (2020). Self-reported willingness to share political news articles in online surveys correlates with actual sharing on Twitter. *PloS One*, 15(2), e0228882. <https://doi.org/10.1371/journal.pone.0228882>
- Newman, N., Fletcher, R., Schulz, A., Andi, S., & Nielsen, R. K. (2020, June 16). *Reuters institute digital news report 2020*. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf
- Nimmo, B. (2017). *Three thousand fake tanks*. Atlantic Council's Digital Forensic Research Lab. Retrieved October 24, 2019, from <https://medium.com/@DFRLab/three-thousand-fake-tanks-575410c4f64d>
- Nimmo, B., Buziashvili, E., Sheldon, M., Karan, K., Aleksejeva, N., Bandeira, L., Andriukaitis, L., & Hibrawi, R. (2019, October 24). *Top takes: Suspected Russian intelligence operation*. Atlantic Council's Digital Forensic Research Lab. <https://medium.com/dfrlab/top-takes-suspected-russian-intelligence-operation-39212367d2f0>
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880. <https://doi.org/10.1037/xge0000465>
- 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*. PsyArXiv. <https://doi.org/10.31234/osf.io/uhbk9>
- 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>
- Sindermann, C., Schmitt, H. S., Rozgonjuk, D., Elhai, J. D., & Montag, C. (2021). The evaluation of fake and true news: On the role of intelligence, personality, interpersonal trust, ideological attitudes, and news consumption. *Heliyon*, 7(3), e06503. <https://doi.org/10.1016/j.heliyon.2021.e06503>
- Talwar, S., Dhir, A., Kaur, P., Zafar, N., & Alrasheedy, M. (2019). Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. *Journal of Retailing and Consumer Services*, 51, 72–82. <https://doi.org/10.1016/j.jretconser.2019.05.026>
- US House of Representatives Permanent Select Committee on Intelligence (publisher). (n.d.). *Exposing Russia's effort to sow discord online: The Internet Research Agency and advertisements*. Retrieved October 24, 2019, from <https://intelligence.house.gov/social-media-content/default.aspx>
- Valverde, M. (2018). *Misleading Infowars story says UN plans to 'flood America with 600 million migrants'*. The Poynter Institute. Retrieved October 24, 2019, from <https://www.politifact.com/punditfact/statements/2018/aug/01/infowars/misleading-infowars-story-says-un-plans-flood-amer/>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Witt, L. A. (2002). The interactive effects of extraversion and conscientiousness on performance. *Journal of Management*, 28(6), 835–851. <https://doi.org/10.1177/014920630202800607>
- Wolverton, C., & Stevens, D. (2019). The impact of personality in recognizing disinformation. *Online Information Review*, 44(1), 181–191. <https://doi.org/10.1108/oir-04-2019-0115>