



Open Research Online

The Open University's repository of research publications and other research outputs

Pathway to a Human-Values Based Approach to Tackle Misinformation Online

Conference or Workshop Item

How to cite:

Piccolo, Lara; Puska, Alisson; Pereira, Roberto and Farrell, Tracie (2020). Pathway to a Human-Values Based Approach to Tackle Misinformation Online. In: Human-Computer Interaction. Human Values and Quality of Life. HCII 2020., Lecture Notes in Computer Science, Vol 12183, Springer, pp. 510–522.

For guidance on citations see [FAQs](#).

© [not recorded]



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Accepted Manuscript

Link(s) to article on publisher's website:

http://dx.doi.org/doi:10.1007/978-3-030-49065-2_36

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Pathway to a Human-Values Based Approach to Tackle Misinformation Online

Lara S G Piccolo¹[0000-0001-6503-8016], Alisson Puska²[0000-0001-9965-3770], Roberto Pereira²[0000-0003-3052-3016], and Tracie Farrell¹[0000-0002-2386-4333]

¹ Knowledge Media Institute, The Open University, Milton Keynes MK7 8FE, UK
{lara.piccolo, tracie.farrell}@open.ac.uk

² Department of Computer Science, Federal University of Paraná, Curitiba, Brazil
alisson.puska@gmail.com, rpereira@inf.ufpr.br

Abstract. Echoing what matters to us, our values pervade the criteria we apply in the judgment of the information we receive on social media when assigning to it a degree of relevance. In this era of “fake-news”, understanding how the values of a social group influence perception and intentions for sharing pieces of (mis)information can reveal critical aspects for socio-technical solutions to mitigate misinformation spreading. This particular study contrasts the reasoning of a group in the United Kingdom and another in Brazil when judging and valuating the same set of headlines. The results confirm the influence of dominant values in the group in the interpretation of the headlines and potential motivations for sharing them, pointing out directions to advance with the human values-based approach to fight misinformation.

Keywords: Human values · Misinformation · Disinformation · Fake News.

1 Introduction

With more and more people around the world relying on social media as a source of information and actively promoting what they judge as relevant, distinguishing whether the information received is reliable, accurate and shared in good faith is becoming increasingly harder. In this global scenario, false or manipulated information has commonly played a role in challenging or distorting values that shape the public opinion in different sectors of the society, such as health, science, politics, etc.

The information intentionally created to trigger, mislead, generate decision errors, manipulate beliefs or deceive is characterised as *disinformation* [34]. Illustrating that, persuasive strategies similar to those used in cyber attacks exploring cognitive hacking [6], which persuades people to fall into spear phishing and malware installation [4, 20] have been applied in the creation and dissemination of false and manipulative news and hoaxes [1].

Misinformation, in turn, can be defined as misrepresented information that causes confusion and is not always intentionally created [34]. Despite the difference in intention, misinformation can be equally harmful and challenging both

for human judgement and computational detection. As a matter of simplification, in this paper we refer to *misinformation* as any type of false information, which includes disinformation.

Events worldwide notably influenced by the power of social media, as the results Brexit referendum in the UK and 2016 presidential elections in the US [3,7], have alarmed the world about the need to review practices and regulation on information spreading online, and better prepare the society to deal with eventual manipulation. Beyond the political arena, the spread of misinformation against vaccines in the form of myths and conspiracy theories reinforced by individuals' beliefs have undermined public health programs to immunise citizens. As a consequence, some diseases such as yellow fever, measles and poliomyelitis, which have been under control for decades, are infecting and causing deaths again [11]. Similarly, cancer treatments which have been proven ineffective keep being promoted online [11], usually with good will.

We argue that the problem of misinformation spreading can not be understood from a single perspective. In different countries, regions or even social groups, communication is established in different ways, including meta-factors in the communication process, how they perceive and appropriate received messages [14]. Social media mediated communication is not apart from the socio-cultural influence [3]. People's judgement on whether a piece of information is relevant and how it should be communicated is influenced by several conditions, such as social and economic factors, education, cultural traits [31] and human values [9].

Expanding the view presented by Piccolo et al. [22], we understand the problem of misinformation spreading in three different layers that coexist and interact with each other as a social information system. Figure 1 illustrates this view by applying the metaphor of an 'onion' [30]. The *technical* layer is where computational solutions and platforms are situated; the *formal* layer with policies and regulation in place, which also influences ethics, an aspect with formal elements. Motivations, beliefs and human values are mainly in the *informal* layer of the society.

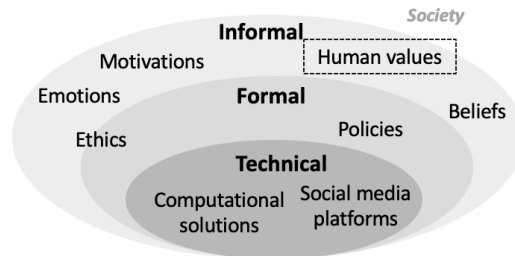


Fig. 1: Three layers of misinformation in a social system

Although computational advances are essential to deal with misinformation spreading online, focusing on technical aspects alone captures a minimal perspective of this information system. A lack of a comprehensive understanding on how people perceive and deal with misinformation online can potentially lead to limited or inefficient solutions, and the risk to produce non-desired impacts to the society; therefore, it is not a responsible way to tackle a social issue [21].

In this paper, we present the first steps towards informing with human values technical solutions for tackling misinformation spreading on social media. We argue that understanding values embedded in some pieces of information and the possible impact of this information upon some cultural groups (i.e. potential to trigger fear, violence, altruism) can inform automated detection of misinformation online in early stages of dissemination, providing an opportunity to social media platforms to prioritise the credibility check or to nudge users before they actually share [9].

We build this study on the results of a survey in the UK that investigated how people interpreted a set of news headlines in terms of veracity and potential to be shared [9]. To this end, we analyse and contrast results of similar tasks with a group of participants in Brazil, evidencing that dominant values according to the Schwartz's theory [25] influence judgment, perception and possible intentions for sharing. Therefore, considering social values embedded in pieces of information should not be a one-size-fits-all approach.

In the next section, the literature review is organised as approaches focused on technical, formal and informal levels and then human values are discussed from a theoretical perspective. In the sequence, the value-based study is described followed by the discussion of the findings using socio-technical lenses. After that, the paper is concluded pointing to future works.

2 Background

At the *technical level*, most popular approaches to fight misinformation support credibility assessment [13] or target the detection of misinforming content [10]. This content include claims, statements containing information/facts that can potentially be assessed; clickbaits, persuasive content leading people to click on particular links [18]; bots, agents that communicate on social media promoting specific information [19], naming a few. Other studies seek to understand the dynamics of misinformation online as echo chambers and filter bubbles [7]. Detection of human values has been addressed by Chen et al. [5], where Reddit posts were analysed to identify a number of word categories that are associated with values. Drawing on Schwartz's basic values theory [26], the research confirmed word-use in social media as a potential predictor of people's values.

At the *formal level*, policies and regulation substitute meaning and intention. According to [16], there are no clear policies regulating misinformation spreading on social media and no guidance towards legal frameworks or ethical issues to be considered. Criticism and public accusations of contributing to the spread of misinformation led big social media players to invest in solutions to tackle it. Like

Facebook, many platforms created internal policies for communities and users’ content management to deal with content considered not adequate. However, for [15], these policies are the result of centralised policy-making by small groups of experts, platform managers, and developers, failing to include and empower platform users. Another formal and crucial approach to fight misinformation is fact-checking, journalistic practices determined by a shared code of principles to verify information [17]. However, as a laborious work performed by groups of experts, existing fact-checking initiatives struggle to cope with the speed, broadness and volume of information spread on social media.

At the *informal level*, Thorson [32] investigated people’s beliefs and evidenced that only informing that the information they believed was incorrect is not enough. Exposing explanations or demographic similarity of the opposing group are some possible strategies to potentially influence changes in opinion. Understanding sharing behaviour is another typical approach. Researches with this objective typically know very little about the characteristics of users, especially regarding subjective aspects. Few studies have considered the social and demographic contexts and their influence on misinformation consumption and spreading. Bedard et al. [2] found that age, education, sex, and political affiliation predict distinguishing “fake news” and satire, and Goyanes [12] applied demographic and situational factors like perception of responsibility, for example, to predict the probability of sharing misinformation. Trilling et al. [33] and Vousoughi et al. [36] investigate characteristics of pieces of news that can suggest its potential for dissemination. Subjective aspects like an emotional language, as well as positivity or negativity, were found to have an influence. Current approaches to detect emotions are mainly based on patterns referring to a small set of basic human emotions [36], limited when referring to emotions as a social construct, culturally built [23].

Human values were addressed by Verma et al (2019) [35] with demographic information to understand individuals’ trust behaviour online. Their findings shed light on the potential role values play in shaping people’s interaction with hyperlinks (and potentially other features) on social media posts.

2.1 Human Values

Values reflect what is important to people in life; as a concept, they have been used in the social sciences and psychology literature to characterise social groups, individuals, to trace changes over time, and to explain the motivational bases of attitudes and behavior [29]. For Schwartz [26], human values are “the criteria people use to select and justify actions, to evaluate people (including the self) and events”. His theory is based on ten basic personal values shared universally [29]:

- *Self-direction*: Independent thought and action;
- *Stimulation*: Excitement, novelty and challenge in life;
- *Hedonism*: Pleasure or sensuous gratification for oneself;
- *Achievement*: Personal success, demonstrating competence;
- *Power*: Social status, control or dominance over people and resources;

- *Security*: Safety, national security, family security, social order;
- *Conformity*: Obedience, self-discipline, politeness;
- *Tradition*: respect, commitment, and acceptance of customs and ideas;
- *Benevolence*: preserving and enhancing the welfare of the ‘in-group’;
- *Universalism*: Understanding, appreciation and protection for the welfare of all people and for nature.

This set of values are represented in a circular structure that illustrates conflicts and compatibility among them. Self-direction, Stimulation and Hedonism are values related to *Openness to change*, which contrasts the *Conservation* values of Conformity, Tradition and Security. Similarly, *Self-transcendence* values of Universalism and Benevolence refer to the importance of enhancing others, beyond selfish interests, opposing to *Self-enhancement* of Hedonism, Achievement and Power.

Judging the veracity of a piece of information, according to this theory, is related to personal beliefs. Values come into play by adding feelings to this judgment, setting it a level of importance and, possibly influencing people’s decision to share it with others or not. Typically, people consider possible consequences for their most important values as a criteria to decide what is good or bad, justified or illegitimate. This decision, though, is often not conscious [29].

As the basic values are understood as universal with varying degrees of influence over human understanding and behaviour [24], previous knowledge acquired on priority values for specific social groups could point directions on how specific pieces of information will be judged and potentially shared. As an example, the European Social Survey¹ systematically assesses the priority values in several European Countries along the years. This approach can help to understand or predict, in a situated context, underlying motivations for sharing and potential social impact due to triggered emotions.

3 Values-Based Study

As fully described in [9], an online survey in the UK recruited 97 library professionals to analyse how they judge a piece of information they see online. This particular group of participants was targeted for sharing a similar background on information literacy.

As part of the survey, without consulting any other source of information, the participants were asked to: (i) judge whether ten headlines were true or false; (ii) select three of the headlines to be shared and justify why; (iii) if they could verify the information before sharing, select three of them and justify the choices. No additional data beyond the headlines was presented, simulating a typical social media behaviour where people quickly browse only the headlines. The ten headlines presented to the participants did not address political issues but touched on themes related to national security, natural world, etc. [9]. They were selected from a commercial card game², as described below:

¹ www.europeansocialsurvey.org

² Fake News Card Game by The Takeover Game.

1. Bearded London hipster mistaken for a member of ISIS and assaulted by nationalists. FALSE. *The Sunday*
2. Man high on drugs rescues dog from imaginary house on fire. TRUE. *The Telegraph*
3. Nigerian restaurant serves human flesh. FALSE. *BBC News*
4. Fish survives six months without half its body. TRUE. *The Independent*
5. Neighbour from hell eats girl's guinea pig. FALSE. *USA Today*
6. Man allowed to board plane after bomb found in his baggage. TRUE. *CTV News*
7. NHS purchases Gluten-free bread for £32.27 per loaf. FALSE. *The Express, the Sun and the Telegraph*
8. In Switzerland, it is illegal to own only one guinea pig. TRUE. *The Mirror*
9. The Bluegill fish is one of the most dangerous fish in North America. When the bluegill are feeding in a school, they can completely dismantle a human body in less than 15 minutes. FALSE. *Facebook*
10. Britain has the highest rate of cocaine use among young adults in Europe, their consumption being almost double that of other nations on the continent. TRUE. *The Times*

The justifications were analysed qualitatively by a group of independent researchers and associated with Schwartz's basic values. Declared interest in the topic and avoidance of risk of consequences were the dominant reasons for choosing what to be shared. Therefore, the headlines (8) 'Guinea pig in Switzerland', (4) 'Fish survives' and (2) 'Rescue dog', which are mainly associated to the values of Stimulation and Universalism, were top-ranked for sharing. While those headlines related to close values of Security and Power (1) 'Hipster', (7) 'NHS bread', and (6) 'Bomb' had the strongest demand for fact-checking, but a low intention for sharing. The results of the study also suggest that, in some cases, the values associated to the headlines are even more critical in the decision of sharing than the judgment of it is true or false.

The analysis of priority human values according to Schwartz's theory [27] on the data collected by the European Social Survey in 2018³ [8] supports participants' expressed rationale. Power was found the predominant value in the UK, followed by Achievement and Stimulation.

³ European Social Survey Round 9 Data (2018). Data file edition 1.2. NSD - Norwegian Centre for Research Data, Norway - Data Archive and distributor of ESS data for ESS ERIC

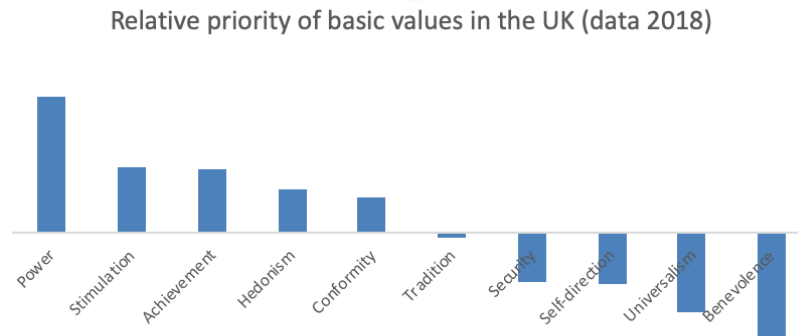


Fig. 2: Relative priority values in the UK in 2018

Avoiding headlines that challenge Power can be seen as a strategy to avoid risks on triggering negative social impact, and also, for the sender, to be perceived by the social network as someone that challenges dominant societal values. Favourite headlines to be shared support Stimulation or Self-enhancement, also dominant aspects of the social group.

3.1 Survey in Brazil

A group of 15 media literacy postgraduate students in Brazil were invited to: (i) complete the short version of Schwartz' Portrait Values Questionnaire (PVQ) [29], the same one used by the European Social Survey (ESS) to capture and compare priority values in different countries [8, 27]; (ii) point out which values they believe are priority among their social network online; (iii) judge the same ten headlines used in the UK-based survey whether they are true or false; (iv) associate headlines with predominant value(s); (v) select three headlines they would share if they had to; (vi) select which ones they would fact-check before sharing; (vii) select those they would never share. They were asked to justify the choices in the last three tasks.

Participants had a brief introduction to Schwartz's theory of Human Values before completing the online survey. Eventual questions about values or the headlines were answered as the survey was applied in English, a second language for the participants. Headline (7) 'NHS' was briefly explained to make it clear that it refers to a national-health entity in the UK.

Results The group of students in Brazil had Universalism as the primary value, followed by Benevolence and Self-direction. Power and Tradition were the least significant values, as illustrated in Figure 3. This resulting ranking is very similar to the 'dominant values they perceived in their social network', with the only

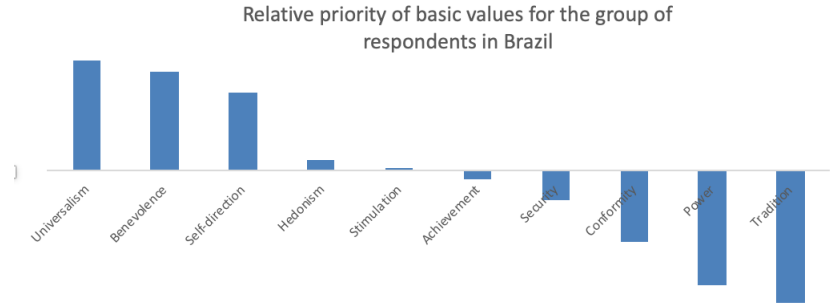


Fig. 3: Relative priority of values as calculated from PVQ survey

distinction that Power is perceived as slightly more important than assessed via questionnaire.

Results obtained for the headlines classifications are described in Table 1 below. For each headline, the percentage of participants that correctly classified the headline as true or false and the values most frequently associated with the headlines are presented.

Table 1: Results of data collected in the survey

| Headline | T or F | Values Associated |
|-----------------|--------|------------------------------|
| 1 Hipster | 7% | Security; Power |
| 2 Rescue dog | 80% | Stimulation; Benevolence |
| 3 Nigerian | 87% | Security; Tradition |
| 4 Fish survives | 40% | Universalism |
| 5 Neighbour | 67% | Security |
| 6 Bomb | 27% | Security; Power/Universalism |
| 7 NHS | 53% | Power; Benevolence |
| 8 Guinea pig | 13% | Universalism; Tradition |
| 9 Bluegill | 60% | Universalism; Security |
| 10 Cocaine | 67% | Security; Power/Hedonism |

Six out of the ten headlines had a very similar percentage of people that judged it correctly. Other four had the true or false judgment notoriously distinct, as described below. These and a few other headlines also differ in the way they have been connected to values:

- (8) ‘Guinea pig’ was correctly judged as true by 13% of the Brazilian participants and by 61% in the UK, where it has been associated mostly to Power instead of Universalism in Brazil;

- (1) ‘Hipster’ was judged correctly as false only by 7% and associated to Security and Power, while in the UK 48.5% got it right and connected mostly to Universalism;
- (2) ‘Rescue dog’ correctly judged as true by 80% of the Brazilian participants which understood it as Stimulation, and the 50.5% in the UK that connected it to Conformity.
- (4) ‘Fish survives’ and (9) ‘Bluegill’ were both associated to Stimulation due to the learning element in the UK. They were understood as Universalism due to the connection with the natural environment. In Brazil, 40% understood the ‘Fish survives’ headlines as true, while 72% in the UK.
- (6) ‘Bomb’ was related to Security and Power in Brazil, while in the UK was connected to Universalism as a social concern.

Table 2 presents the five most cited headlines participants would choose if they must share three of them, those they could share if fact-checked before, or those they would never share. The percentage refers to the proportion of choice considering all selected headlines.

Table 2: Top-five ranking of selected headlines

| MUST share | CAN share | NEVER share |
|--------------------|------------------|--------------------|
| 1 Hipster (18%) | 1 Hipster (20%) | 5 Neighbour (25%) |
| 10 Cocaine (18%) | 10 Cocaine (20%) | 3 Nigerian (18%) |
| 3 Rescue dog (18%) | 3 Nigerian (13%) | 8 Guinea pig (16%) |
| 9 Bluegill (11%) | 7 NHS (11%) | 7 NHS (11%) |
| 7 NHS (11%) | 9 Bluegill (11%) | 6 Bomb (9%) |

The majority of justifications for choosing headlines to be shared refers first to their perceived plausibility, relevance to society, or humour ‘because it sounds funny’. Interestingly, headlines (1) ‘Hipster’, (10) ‘Cocaine’, and (3) ‘Rescue dog’ had multiple justifications based on Universalism ‘a way to prevent drugs consumption’, ‘an alert about intolerance’, ‘expression of solidarity’ or ‘to be used as a comparison, evidencing that developed countries have similar problems to ours’, or as example of protecting the nature or welfare.

Social relevance, personal interest and the suspicious it is a false news due to the text style led the decisions on the headlines to be fact-checked before sharing. Headlines associated with Security and Power were the most consensual ones regarding the need for fact-checking as a way to prevent a negative societal impact. The lack of relevance for the society dominated decisions to never share some specific news. The ‘bad news’ tone of some headlines, for example (3) ‘Nigerian flesh’, was also considered a reason for not sharing.

Similar to the rationale applied by UK-based participants, the decision for sharing seems to be more influenced by the perception of the social network in terms of interest and impact. For this particular public with a degree of information literacy, plausibility is a significant criterion in both scenarios. The

concern to avoid societal harm by spreading challenging headlines associated with Security is also present, especially as a principle to never share negative news. Part of Brazilian respondents introduced a new perspective in the way to judge and value some headlines, different from the results found in [9]. The stories that were somehow challenging important values were seen as a way to alert, raise awareness or compare realities, and not necessarily as a risk or threat.

As Schwartz explains [29], Self-transcendence values (Universalism and Benevolence), which are dominant in this group, are more focused on the social than personal aspects, with motivation to preserve and enhance the welfare of others [27, 28], a characteristic that can explain these results.

4 Discussion

Contrasting results obtained in the UK and Brazil illustrates the role of values in adding feelings or setting the level of importance to the headlines, as per Schwartz’s theory [29]. Therefore, the perceived ‘shareworthiness’ seems to be tuned to the values (and interest) of the participants’ social network.

Headlines considered ‘neutral’, with no potential to assert any societal harm, as (4) ‘Fish survives’ and (9) ‘Bluegill’ were associated with Stimulation in the UK and with Universalism in Brazil. Although it sounds similar, Schwartz [28] details the drives behind these values: while Stimulation is related to learning (know, understand, by exercising curiosity...), Universalism refers to ‘connecting plus learning’, where connecting means to build social relationships and develop mutual commitment, suggesting that people may be still prone to share things that are not necessarily connected to their individual values, but to communicate with their ‘audience’.

For Schwartz [29], Self-transcendence values (Universalism and Benevolence) lead to motivation for preserving and enhancing the welfare of others. They are opposed to Self-Enhancement values (Achievement and Power) which justify self-serving behaviour instead. For the participants with dominant Self-transcendence values, even headlines challenging societal values as Security and Power were many times seen through the lenses of Universalism, as an expression of solidarity, or as an example to (not) be followed.

Both in the UK and in Brazil, the intention to avoid triggering a negative societal impact by sharing potential risky headlines was evident. In fact, *risk of harm* has been a credibility signal considered in assessing credibility of online information in the Web⁴.

Building on the idea of how informal, formal and technical layers in Figure 1 can be connected for providing solutions to tackle misinformation on line, we argue that the notion of human values could inform the characterisation of potential societal harm as a credibility signal, and applied as a way to prioritise assessment by fact-checkers or the social platform or to nudge users about the potential to assert risks.

⁴ *Credibility Signals*: a live document by the W3C Credible Web Community. <https://credweb.org/signals-beta/claim-risk-of-harm>

4.1 Limitations

As limitations for this study, it is acknowledged in the literature that researches that ask people how likely it is that they would share some stimulus hardly mimics a real-life situation [33]. For the survey, participants received instruction that they had to share three headlines and should select which ones. Yet, the headlines not necessarily match their personal interests. Then, the most important point for the analysis is how they perceived the headlines, their reasoning, and not the real intention of sharing them or not.

For the Brazilian survey, the group size (n=15) limited the range of statistical analysis that could be performed like correlations or trends. This sample size is not enough either to compare relative priority values across countries. For this reason, the PVQ results presented here must NOT be understood as representative of the Brazilian population, referring only to the group of participants in the study. Cross-national studies within Europe could rely on the ESS data for comparisons, which was not the case for Brazil.

For ensuring compatibility in terms of information literacy, this study restricted the participation to groups of people with a known background. On the one hand, a broader audience could lead to different results, for example, on the importance of plausibility. But on the other hand, it would add extra variables related to the understanding and objectives of the survey.

As a preliminary study, results can be applied to inform the setup for broader comparative analysis, suggesting what needs to be taken into account to understand better and predict the role of human values in perceived ‘shareworthiness’, and to consider these factors when designing technical solutions to mitigate misinformation spreading.

5 Conclusion

This research explored the role of human values in perceiving and judging a set of headlines, simulating pieces of news accessed through social media. Groups with different cultural background perceived and valued headlines in different ways, evidencing the influence of dominant values in this subjective process. This study points direction to broader research across nations where dominant values have been periodically assessed. Future work comprises analysing real social media data, especially on how similar topics have been perceived and spread across countries to build a more systematic knowledge on how human-values can inform existing pipelines for fact-checking or contextual strategies to nudge social media users.

Acknowledgment. This paper has been supported by the EC within the Horizon 2020 programme under grant agreement 770302 - Co-Inform.

References

1. Bazan, S.: A new way to win the war. *IEEE Internet Computing* **21**(4), 92–97 (2017). <https://doi.org/10.1109/MIC.2017.2911419>

2. Bedard, M., Schoenthaler, C.: Satire or fake news: Social media consumers' socio-demographics decide. In: Companion Proceedings of the The Web Conference 2018. pp. 613–619 (2018)
3. Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D., Marlow, C., Settle, J.E., Fowler, J.H.: A 61-million-person experiment in social influence and political mobilization. *Nature* **489**(7415), 295–298 (2012)
4. Caputo, D.D., Pflieger, S., Freeman, J.D., Johnson, M.: Going spear phishing: Exploring embedded training and awareness. *IEEE Security Privacy* **12**(01), 28–38 (jan 2014). <https://doi.org/10.1109/MSP.2013.106>
5. Chen, J., Hsieh, G., Mahmud, J.U., Nichols, J.: Understanding individuals' personal values from social media word use. In: Proceedings of the 17th ACM CSCW & Social Computing. pp. 405–414 (2014)
6. Cybenko, G., Giani, A., Thompson, P.: Cognitive hacking: a battle for the mind. *Computer* **35**(8), 50–56 (Aug 2002). <https://doi.org/10.1109/MC.2002.1023788>
7. DiFranzo, D., Gloria-Garcia, K.: Filter bubbles and fake news. *XRDS* **23**(3), 32–35 (Apr 2017). <https://doi.org/10.1145/3055153>
8. European Social Survey, y.: ESS-9 2018 Documentation Report. Edition 1.2. Bergen, European Social Survey Data Archive, NSD - Norwegian Centre for Research Data for ESS ERIC
9. Farrell, T., Piccolo, L., Perfumi, S.C., Alani, H.: Understanding the role of human values in the spread of misinformation. In: Proc of Truth and Trust Online (2019)
10. Garrett, R.K., Weeks, B.E.: The promise and peril of real-time corrections to political misperceptions. In: Proc of CSCW 2013. pp. 1047–1058. ACM (2013)
11. Ghenai, A., Mejova, Y.: Fake cures: user-centric modeling of health misinformation in social media. *Proceedings of the ACM on Human-Computer Interaction* **2**(CSCW), 1–20 (2018)
12. Goyanes, M.: The sociology of fake news: Factors affecting the probability of sharing political fake news online. LSE Working Paper Series (06 2018)
13. Gupta, A., Kumaraguru, P., Castillo, C., Meier, P.: Tweetcred: Real-time credibility assessment of content on twitter. In: International Conference on Social Informatics. pp. 228–243. Springer (2014)
14. Hall, E.T.: The silent language. greenwich, conn. Fawcett Publications. Inc **1**, 959 (1959)
15. Han, O., Baris, I., de Nigris, S., Staab, S.: Democratic policy-making for misinformation detection platforms by git-based principles
16. Hosseini, A.: Content Management Policies for Combating Misinformation (2020), <https://coinform.eu/content-management-policies-for-combating-misinformation/>
17. Ireton, E.C., Posetti, J.: Journalism, 'Fake News' & Disinformation Handbook for Journalism Education and Training. United Nations Educational, Scientific and Cultural Organization - UNESCO (2018)
18. Karadzhov, G., Gencheva, P., Nakov, P., Koychev, I.: We built a fake news & click-bait filter: what happened next will blow your mind! arXiv preprint arXiv:1803.03786 (2018)
19. Kudugunta, S., Ferrara, E.: Deep neural networks for bot detection. *Information Sciences* **467**, 312–322 (2018)
20. Nguyen-Vu, L., Park, J., Chau, N.T., Jung, S.: Signing key leak detection in google play store. In: Proceedings of the 2016 International Conference on Information Networking (ICOIN). p. 13–16. IEEE Computer Society, USA (2016)

21. Pereira, R., Baranauskas, M.C.C.: A value-oriented and culturally informed approach to the design of interactive systems. *International Journal of Human-Computer Studies* **80**, 66–82 (2015)
22. Piccolo, L.S.G., Joshi, S., Karapanos, E., Farrell, T.: Challenging misinformation: Exploring limits and approaches. In: *Human-Computer Interaction - INTERACT 2019*. LNCS, vol. 11749, pp. 713–718. Springer (2019), <http://oro.open.ac.uk/68822/>
23. Richerson, P.J., Boyd, R.: *Not by genes alone: How culture transformed human evolution*. University of Chicago press (2008)
24. Rokeach, M.: *The nature of human values*. Free press (1973)
25. Schwartz, M.: *Guidelines for Bias-Free Writing*. ERIC, Bloomington, IN, USA (1995)
26. Schwartz, S.H.: Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, vol. 25, pp. 1 – 65. Academic Press (1992)
27. Schwartz, S.H.: A proposal for measuring value orientations across nations. *Questionnaire package of the european social survey* **259**(290), 261 (2003)
28. Schwartz, S.H.: Les valeurs de base de la personne: théorie, mesures et applications. *Revue française de sociologie* **47**(4), 929–968 (2006)
29. Schwartz, S.H.: An overview of the schwartz theory of basic values. *Online readings in Psychology and Culture* **2**(1), 2307–0919 (2012)
30. Stamper, R., Liu, K., Hafkamp, M., Ades, Y.: Understanding the roles of signs and norms in organizations—a semiotic approach to information systems design. *Behaviour & information technology* **19**(1), 15–27 (2000)
31. Theng, Y.L., Goh, L.Y.Q., Lwin, M.O., Shou-Boon, S.F.: Dispelling myths and misinformation using social media: A three-countries comparison using the case of tuberculosis. In: *2013 IEEE International Conference on Healthcare Informatics*. pp. 147–152. IEEE (2013)
32. Thorson, E.: Belief echoes: The persistent effects of corrected misinformation. *Political Communication* **33**(3), 460–480 (2016)
33. Trilling, D., Tolochko, P., Burscher, B.: From newsworthiness to shareworthiness: How to predict news sharing based on article characteristics. *Journalism & Mass Communication Quarterly* **94**(1), 38–60 (2017)
34. Tudjman, M., Mikelic, N.: Information science: Science about information, misinformation and disinformation. *Proceedings of Informing Science+ Information Technology Education* **3**, 1513–1527 (2003)
35. Verma, N., Fleischmann, K.R., Koltai, K.S.: Understanding online trust and information behavior using demographics and human values. In: *International Conference on Information - iConference*. pp. 654–665. Springer (2019)
36. Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. *Science* **359**(6380), 1146–1151 (2018). <https://doi.org/10.1126/science.aap9559>