



A Cross-Domain Study of the Use of Persuasion Techniques in Online Disinformation

João A. Leite
University of Sheffield
Sheffield, UK
j.leite@sheffield.ac.uk

Carolina Scarton
University of Sheffield
Sheffield, UK
c.scarton@sheffield.ac.uk

Olesya Razuvayevskaya
University of Sheffield
Sheffield, UK
o.razuvayevskaya@sheffield.ac.uk

Kalina Bontcheva
University of Sheffield
Sheffield, UK
k.bontcheva@sheffield.ac.uk

Abstract

Disinformation, irrespective of domain or language, aims to deceive or manipulate public opinion, typically employing advanced persuasion techniques. Qualitative and quantitative research on the weaponisation of persuasion techniques in disinformation narratives, however, has been mostly limited to specific topics (e.g., COVID-19). To address this gap, our study conducts a large-scale, multi-domain analysis of the role of 16 persuasion techniques in disinformation narratives, by leveraging a state-of-the-art persuasion technique classifier. We demonstrate how different persuasion techniques are employed disproportionately in different disinformation domains. We also include an in-depth case study on climate change disinformation, which demonstrates how linguistic, psychological, and cultural factors shape the adaptation of persuasion strategies to fit unique thematic contexts.

CCS Concepts

• **Human-centered computing** → **Empirical studies in collaborative and social computing**; • **Computing methodologies** → **Discourse, dialogue and pragmatics**; *Information extraction*.

Keywords

Disinformation, Persuasion Techniques, Domain Adaptation

ACM Reference Format:

João A. Leite, Olesya Razuvayevskaya, Carolina Scarton, and Kalina Bontcheva. 2025. A Cross-Domain Study of the Use of Persuasion Techniques in Online Disinformation. In *Companion Proceedings of the ACM Web Conference 2025 (WWW Companion '25)*, April 28-May 2, 2025, Sydney, NSW, Australia. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3701716.3715535>

1 Introduction

Disinformation campaigns, amplified through social media platforms, are often weaponising sophisticated **persuasion techniques** to manipulate public opinions [11]. Existing research on disinformation discourse has predominantly examined the narratives and tactics employed to influence public opinion, often touching on

elements related to persuasion techniques, such as propaganda, which has a broader scope [11]. For instance, previous studies have examined pro-Kremlin propaganda in the Russo-Ukrainian war, uncovering tactics such as oversimplifications, distractions, and associations, which are aimed at maintaining domestic support and justifying aggression [2]. Similarly, research into climate disinformation has highlighted that logical fallacies and conspiracy theories are used to discredit scientific consensus [4], while analyses of COVID-19 disinformation have identified narratives such as fear-mongering and conspiracy theories [12]. In natural language processing, research has focused on creating datasets and models for automatic detection of persuasion techniques [5, 18]. As highlighted in Srba et al., one of the major limitations of the existing studies is a narrow focus on individual domains and the lack of a unified taxonomy. This leads to a limited understanding of the different ways in which persuasion techniques are employed in the different domains targeted by disinformation.

To address this gap, our study focuses on a large-scale cross-domain analysis into the types of persuasion techniques typically employed in disinformation narratives. In particular, we analyse existing disinformation datasets from various domains (COVID-19, climate change, anti-Islam, and the Russo-Ukrainian war) by applying a state-of-the-art classifier trained to identify sixteen different kinds of persuasion techniques. This allows us to answer the following research questions: **(RQ1)** Do disinformation narratives from different domains employ persuasion techniques differently? **(RQ2)** Are persuasion techniques adapted linguistically, psychologically, and culturally to fit the context of domain-specific disinformation?

By comparing statistical and qualitative patterns in the use of persuasion techniques across different domains, our study provides a deeper understanding of the ways in which disinformation seeks to manipulate and offers actionable insights for countering its effects. Our code and supplementary material are made openly available to facilitate reproducibility (<https://github.com/joaoaleite/pmd>).

2 Methodology

2.1 Identifying Persuasion Techniques

In this study, we leverage the state-of-the-art persuasion classifier introduced in Razuvayevskaya et al. [17], which uses a RoBERTa-Large pretrained model fine-tuned jointly on several different languages translated into English. The persuasion classifier was trained using the largest dataset of human annotated persuasion techniques,



This work is licensed under a Creative Commons Attribution-NonDerivs International 4.0 License.

WWW Companion '25, April 28-May 2, 2025, Sydney, NSW, Australia
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1331-6/2025/04
<https://doi.org/10.1145/3701716.3715535>

which was introduced in SemEval-2023 [15]. The data consists of news articles in 9 languages, collected from a variety of mainstream and alternative sources. It is annotated with 23 persuasion techniques at the sentence level, with the task framed as a multi-label classification problem, allowing multiple techniques to be tagged simultaneously within each sentence. Further details about the classifier can be found in Razuvayevskaya et al. [17]. The classifier is currently ranked first on the SemEval task's post-competition leaderboard, and is available through an API¹.

The study is focused on 16 persuasion techniques which are: *Loaded Language* (22% of instances in the SemEval training data), *Name Calling-Labeling* (16%), *Doubt* (13%), *Questioning the Reputation* (7%), *Exaggeration-Minimisation* (6%), *Appeal to Fear-Prejudice* (5%), *Repetition* (3%), *Appeal to Authority* (3%), *Slogans* (3%), *Conversation Killer* (3%), *Appeal to Hypocrisy* (3%), *Guilt by Association* (2%), *Appeal to Values* (2%), *False Dilemma-No Choice* (2%), *Flag Waving* (2%), *Causal Oversimplification* (2%).

To improve the reliability of our analysis, we discard the following seven persuasion techniques produced by the SemEval-2023 classifier, due to their low frequency in the training set (below 2%): *Causal Oversimplification* (1%), *Appeal to Time* (1%), *Straw Man* (1%), *Appeal to Popularity* (1%), *Obfuscation-Vagueness-Confusion* (1%), *Red Herring* (1%), *Whataboutism* (1%).

2.2 Disinformation Domains

We study four disinformation datasets from diverse domains (see Table 1): CIDII (Islamic issues) [9], COVID-19 [14], Climate Fever (climate change) [6], and EUvsDisinfo (Russo-Ukrainian war) [13]. All texts are in English, except for EUvsDisinfo, which spans 41 additional languages where we translated the non-English texts to English using GPT4o. EUvsDisinfo also includes topics beyond the Russo-Ukrainian war, which we filtered out. The articles in all these datasets are human labelled. In this study specifically, we analyse articles human labelled as disinformation, and exclude all trustworthy articles across all datasets. The texts are split into sentences using NLTK, and the persuasion classifier is used to identify persuasion techniques within each sentence.

Table 1: Disinformation domains studied.

Domain	Islamic Issues	COVID-19	Climate Change	Russo-Ukrainian War
# Sentences	1, 687	7, 369	254	10, 240
Avg. Sent. Length	102.3	95.9	109.0	150.7
Avg. PTs per Sent.	1.1	1.1	1.2	1.4

3 Analysis

Figure 1 shows the proportion of persuasion techniques in each dataset, offering a visual comparison. To complement this and enable a more detailed quantitative analysis, we calculate the odds ratios (ORs) for each technique between datasets in Table 2. We calculate ORs by comparing the odds of a technique appearing in one domain to its average proportion in the other three domains.

¹<https://cloud.gate.ac.uk/shopfront/displayItem/persuasion-classifier>

Throughout our analysis, we only discuss statistically significant ORs, based on Fisher's exact test with $p < 0.05$.

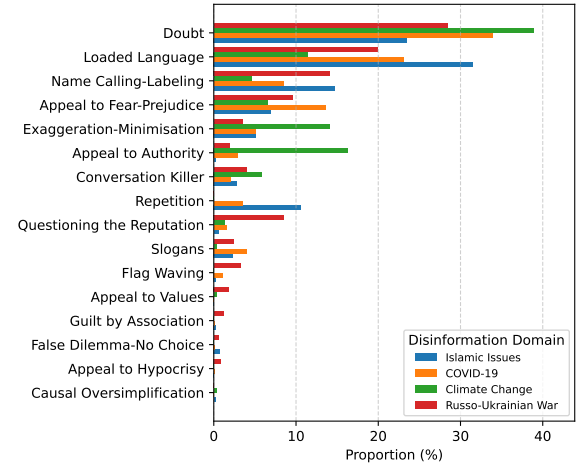


Figure 1: Proportion of persuasion techniques in the different disinformation narratives.

Table 2: Odd ratios of occurrence of persuasion techniques in one domain vs. the average proportion in the others. Statistically significant ratios are underlined (with $p < 0.05$ – Fisher exact test). A→Islamic issues, B→COVID-19, C→Climate change, D→Russo-Ukrainian war.

Persuasion Technique	A	B	C	D
Doubt	<u>0.60</u>	1.18	<u>1.59</u>	<u>0.84</u>
Loaded Language	2.07	1.14	<u>0.39</u>	0.88
Name Calling-Labeling	<u>1.72</u>	<u>0.74</u>	<u>0.34</u>	<u>1.61</u>
Appeal to Fear-Prejudice	<u>0.67</u>	<u>1.90</u>	<u>0.63</u>	1.06
Exaggeration-Minimisation	0.66	0.66	<u>3.39</u>	<u>0.41</u>
Appeal to Authority	<u>0.03</u>	<u>0.46</u>	<u>11.37</u>	<u>0.28</u>
Conversation Killer	0.70	<u>0.49</u>	<u>2.06</u>	<u>1.10</u>
Repetition	<u>10.15</u>	<u>0.98</u>	<u>0.01</u>	<u>0.01</u>
Questioning the Reputation	<u>0.15</u>	<u>0.43</u>	<u>0.36</u>	<u>8.06</u>
Slogans	1.04	<u>2.47</u>	<u>0.11</u>	<u>1.06</u>
Flag Waving	<u>0.19</u>	<u>0.98</u>	<u>0.01</u>	6.99
Appeal to Values	<u>0.06</u>	<u>0.06</u>	0.51	<u>13.26</u>
False Dilemma-No Choice	<u>3.33</u>	<u>0.37</u>	0.01	<u>1.72</u>
Guilt by Association	<u>0.64</u>	<u>0.22</u>	0.01	<u>9.45</u>
Appeal to Hypocrisy	<u>0.14</u>	<u>0.30</u>	0.01	<u>19.28</u>
Causal Oversimplification	<u>1.97</u>	0.13	3.88	<u>0.01</u>

We observe that *Loaded Language* and *Doubt* are used ubiquitously across all disinformation domains, comprising more than 20% of the techniques in each, except *Loaded Language* in climate change (11%). Specifically *Doubt* attempts to undermine trust in credible sources or established facts, creating confusion and making audiences more susceptible to accepting misleading or false claims [16]. *Loaded Language* aims to evoke strong emotional responses, such as fear or anger, which can override rational analysis and lead individuals to accept false information without critical scrutiny [8].

Next, we discuss the persuasion techniques with ORs greater than 1, with statistical significance, indicating that their usage in one domain is more prevalent than in the others.

3.1 Domain-Specific Use of Persuasion Techniques (RQ1)

In disinformation on Islamic issues, *Name Calling-Labeling*, *Causal Oversimplification*, *Loaded Language*, *False-Dilemma-No Choice*, and *Repetition* are the most prevalent techniques, with ORs of 1.72, 1.97, 2.07, 3.33, and 10.15 respectively. *Repetition* capitalises on the illusory truth effect, a psychological phenomenon where repeated statements are more likely to be perceived as true [7, 10]. In this domain, **repetition is often used to associate events and/or groups**: “All Muslim terrorists kill for the same reason the Saudi terrorists did on 9/11, the same reason ISIS killed [...], the same reason Boko Haram is killing [...], the same reason Muhammad waged wars [...]”. **False Dilemma forces binary choices to portray Islam as inherently extremist**, as in “There’s no middle ground, nothing like moderate Islam.”

In the context of COVID-19 disinformation, *Loaded Language*, *Appeal to Fear-Prejudice*, and *Slogans* appear significantly more frequently (ORs of 1.14, 1.90, 2.47, respectively). *Slogans* use short and memorable phrases to convey key ideas, as in “We need to open up, our lives depend on it”. Both *Loaded Language* and *Appeal to Fear-Prejudice* are designed to evoke strong emotional responses, which can bypass rational analysis and promote the acceptance of false information [8]. **Fear is linked to both the virus—“#coronavirus this is extremely scary and terrifying”—and vaccines**, as in “According to Bill Gates the COVID-19 RNA vaccine will permanently alter our DNA”. Notably, fear of COVID-19 has been linked to mental health issues during the pandemic period [1].

Climate change disinformation uses more frequently *Doubt* (1.59), *Conversation Killer* (2.06), *Exaggeration-Minimisation* (3.39), and *Appeal to Authority* (11.37) than the rest. **Appeal to Authority exploits trust in credible entities like the IPCC**: “Latest IPCC Reports show global temperature forecasts exceeded actual readings.” References to NASA and the UN are also common. **Exaggeration-Minimisation downplays the urgency of climate change**: “In the past, warming has never been a threat to life on Earth”. **Conversation Killer dismisses the debate with absolute certainty, leaving no room for debate**: “Sea level rise is not going to happen.”

Disinformation about the Russo-Ukrainian war prominently features *Flag Waving* (6.99), *Questioning the Reputation* (8.06), *Guilt by Association* (9.45), and *Appeal to Hypocrisy* (19.28). These techniques target Western and Ukrainian credibility. **Appeal to Hypocrisy highlights perceived inconsistencies**, e.g., “They said one thing and did another.”—referring to NATO’s territorial advancement. **Guilt by Association links entities to controversial groups, with 28% of tagged sentences referencing the word ‘Nazi’**, e.g., “Nazi symbolism is actively utilised in their daily life”—referring to Ukrainian soldiers. *Questioning the Reputation* undermines trust in political entities: “As for Europe, it lost its political independence after World War II”. **Flag Waving invokes patriotism to justify Russia’s actions**, e.g., “It is against this Evil that our soldiers bravely fight side by side”.

3.2 Contextual Domain Adaptation (RQ2)

To investigate the contextual adaptation of persuasion techniques to specific linguistic, cultural, and psychological patterns across disinformation domains, we compute the correlation between LIWC features [3] and persuasion techniques. Due to space constraints, we limit our analysis to the climate change dataset only, focusing on the four persuasion techniques that occur disproportionately in this context: *Appeal to Authority*, *Conversation Killer*, *Doubt*, and *Exaggeration-Minimisation*. Figure 2 presents the ten highest correlations between these techniques and the LIWC features. We compare these correlations across the other three domains to highlight patterns unique to climate change disinformation, focusing on features highly correlated within climate change but not within other domains. We provide the results for the other domains in the supplementary material^{??}.

Appeal to Authority in climate change disinformation uses distinct linguistic and psychological features to emphasise credibility and logic. These texts have a higher overall and per-sentence word count, creating longer, more complex sentences that convey authority. A high **Analytic** score reflects formal, logical language, reinforcing a well-reasoned tone. Frequent use of **prepositions** and **numbers** adds detail to enhance the legitimacy of claims. **Temporal markers** like ‘when’ and ‘now’ situate arguments in time, adding urgency or inevitability. Psychologically, this technique appeals to **curiosity** and **reward**, engaging readers intellectually and offering positive outcomes. Terms like ‘enough’ and ‘full’ (i.e., **fulfillment**) suggest solutions framed as complete and authoritative. Notably, these texts avoid **causation** language (e.g., ‘because’, ‘why’), favouring definitive statements over explanations. These patterns align with the rhetorical context of climate change narratives, where the discussion of scientific topics requires a tone of credibility and rigour to persuade audiences.

Exaggeration-Minimisation employs specific linguistic and psychological features to amplify or downplay issues strategically. A high correlation with **certitude** words (e.g., ‘really’, ‘of course’) reinforces claims with a tone of absolute confidence, making arguments appear definitive and irrefutable. Terms related to **moral behaviour** (e.g., ‘honour’, ‘deserve’) inject ethical undertones, framing the issue as one of right versus wrong. Cultural references, particularly related to **politics** (e.g., ‘govern’, ‘congress’) and broader **culture** (e.g., ‘american’, ‘chinese’), ground the discussion in societal and geopolitical contexts, often linking climate change to governance or national responsibility. Linguistically, the technique relies heavily on **determiners** (e.g., ‘the’, ‘that’) and **function words** (e.g., ‘to’, ‘I’), creating a conversational and relatable tone. The frequent use of **adverbs** (e.g., ‘so’, ‘just’) adds nuance or emphasis to descriptions, subtly influencing how issues are perceived. Emotional cues, especially those tied to **anxiety** (e.g., ‘worry’, ‘fear’), heighten the sense of urgency or concern, tapping into the audience’s fears about climate change. Together, these features amplify certainty, invoke morality, embed discussions within cultural and political frameworks, and simplify complex issues while evoking emotional responses.

Both *Conversation Killer* and *Doubt* exhibit a lesser degree of contextual adaptation, as most of the LIWC features that display high correlation in climate change are also highly correlated in other

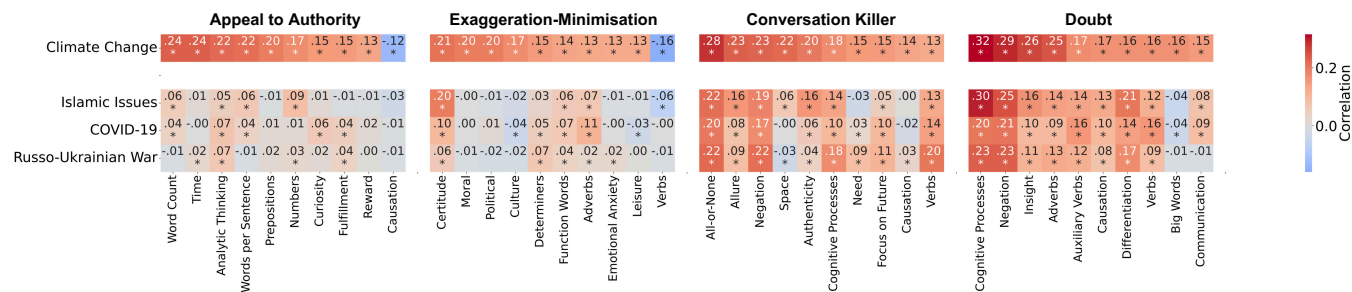


Figure 2: Top 10 most correlated LIWC features for each of the four domain-specific PTs in climate change compared to other domains. Statistically significant ($p < 0.05$) coefficients are indicated with an asterisk (*).

domains (e.g., Cognitive Processes, Negation). Nevertheless, some domain-specific features are present. *Conversation Killer* leverages **spatial context (Space)** (e.g., ‘in’, ‘there’) to ground arguments in specific locations, **need-related states** (e.g., ‘need’, ‘have to’) to emphasise urgency and necessity, and **causation** (e.g., ‘because’, ‘why’) provides justifications that reinforce the authority of dismissive rhetoric. Similarly, *Doubt* amplifies skepticism with the prominent use of **long words** to create an impression of sophistication and **communication** features (e.g., ‘say’, ‘tell’) that reference ambiguous sources, subtly eroding trust in credible information.

4 Discussion and Conclusions

This study conducted a large-scale, cross-domain analysis of sixteen persuasion techniques in disinformation narratives on COVID-19, climate change, Islamism, and war. Our findings reveal that while some techniques—such as *Doubt* and *Loaded Language*—are common across all domains, others are more context-dependent. For instance, narratives on Islamic issues often rely on *Repetition* and *False Dilemma-No Choice*, whereas those on the Russo-Ukrainian war frequently use *Appeal to Hypocrisy* and *Guilt by Association*. We also observe stylistic adaptations; in climate change narratives, *Appeal to Authority* tends to be more formal and analytic, while *Exaggeration-Minimisation* frequently incorporates moral and cultural references. These observations illustrate how disinformation narratives are tailored to specific audiences and how their rhetorical and psychological features shift with thematic and cultural factors.

Our analysis can inform media literacy training, helping citizens, fact-checkers, and journalists recognise the presence and evolving forms of these persuasion techniques. Future research studying persuasion techniques in disinformation narratives can explore how detection models can incorporate domain- and technique-sensitive features to improve their ability to identify subtle, context-dependent cues. Additionally, future studies may examine how persuasion techniques evolve over time within specific domains.

Ethical Use of Data

We adhere to ethical principles in data use by employing well-established and publicly available datasets that do not include personal or identifiable information. Due to the focus on disinformation narratives, the data may inherently contain sensitive content.

Acknowledgments

This work has been co-funded by an Innovate UK grant 10039055 (approved under the Horizon Europe Programme as vera.ai) and the European Media and Information Fund (grant 291191 ExU). João Leite is supported by a University of Sheffield EPSRC Doctoral Training Partnership Scholarship.

References

- [1] Zainab Alimoradi et al. 2022. Fear of COVID-19 and its association with mental health-related factors. (March 2022). <https://doi.org/10.1192/bj0.2022.26>
- [2] Samy Amanatullah. 2023. Tell Us How You Really Feel: Analyzing Pro-Kremlin Propaganda Devices & Narratives to Identify Sentiment Implications. (2023). <https://www.illiberalism.org/tell-us-how-you-really-feel-analyzing-pro-kremlin-propaganda-devices-narratives-to-identify-sentiment-implications/>
- [3] Ryan L Boyd et al. 2022. The development and psychometric properties of LIWC-22. (2022). <https://doi.org/10.13140/RG.2.2.23890.43205>
- [4] John Cook. 2019. Understanding and Countering Misinformation About Climate Change. <https://doi.org/10.4018/978-1-5225-8535-0.ch016>
- [5] Giovanni Da San Martino et al. 2020. A Survey on Computational Propaganda Detection. <https://doi.org/10.24963/ijcai.2020/672>
- [6] Thomas Diggelmann et al. 2021. CLIMATE-FEVER: A Dataset for Verification of Real-World Climate Claims. <https://doi.org/10.48550/arXiv.2012.00614>
- [7] Lisa K. Fazio et al. 2015. Knowledge does not protect against illusory truth. (Oct. 2015). <https://doi.org/10.1037/xge0000098>
- [8] Gloria Gennaro and Elliott Ash. 2022. Emotion and Reason in Political Language. 132, 643 (April 2022). <https://doi.org/10.1093/ej/ueab104>
- [9] Suhaib Kh Hamed, Mohd Juzaidin Ab Aziz, and Mohd Ridwan Yaakub. 2023. Disinformation Detection About Islamic Issues on Social Media Using Deep Learning Techniques. (July 2023). <https://doi.org/10.22452/mjcs.vol36no3.3>
- [10] Lynn Hasher, David Goldstein, and Thomas Toppino. 1977. Frequency and the conference of referential validity. (Feb. 1977). [https://doi.org/10.1016/S0022-5371\(77\)80012-1](https://doi.org/10.1016/S0022-5371(77)80012-1)
- [11] Garth S. Jowett and Victoria O'Donnell. 2018. *Propaganda & Persuasion*. <https://books.google.com/books?id=vIC92PdJ0l0C>
- [12] Bonka Kotseva et al. 2023. Trend analysis of COVID-19 mis/disinformation narratives—A 3-year study. (2023). <https://doi.org/10.1371/journal.pone.0291423>
- [13] João A. Leite et al. 2024. EUvsDisinfo: A Dataset for Multilingual Detection of Pro-Kremlin Disinformation in News Articles. <https://doi.org/10.1145/3627673.3679167>
- [14] Parth Patwa et al. 2021. Fighting an Infodemic: COVID-19 Fake News Dataset. https://doi.org/10.1007/978-3-030-73696-5_3
- [15] Jakub Piskorski et al. 2023. SemEval-2023 Task 3: Detecting the Category, the Framing, and the Persuasion Techniques in Online News in a Multi-lingual Setup. <https://doi.org/10.18653/v1/2023.semeval-1.317>
- [16] Robert N Proctor and Schiebinger Londa. 2008. *Agnostology: The making and unmaking of ignorance*. <https://philarchive.org/archive/PROATM>
- [17] Olesya Razuvayevskaya et al. 2024. Comparison between Parameter-Efficient Techniques and Full Fine-Tuning. (2024). <https://doi.org/10.1371/journal.pone.0301738>
- [18] Ivan Srba et al. 2024. A Survey on Automatic Credibility Assessment of Textual Credibility Signals in the Era of Large Language Models. <https://doi.org/10.48550/arXiv.2410.21360>