Using Focus Groups to Improve Counterarguments Against Misinformation

Janice Butler
EUI Department of Political and Social Sciences

Introduction to Qualitative Methods
April, 2022

Professor: Dr. Jeffrey Checkel

Whilst misinformation refers to the unintentional spread of false or inaccurate information, disinformation refers to its intentional distribution, the information being created or deliberately shared in order to deceive or cause harm (Freelon & Wells, 2020). The problem of misinformation is in itself serious, but has been exacerbated by the advent of social media.

There is much being done to detect the most egregious cases, but little is being undertaken to intervene and there is indeed a lack of clarity on the best way to intervene (Kan et al, 2021). E, Sakura and Li (2021) highlight the need for work on online misinformation correction, identifying it as a poorly researched area. Most research into bias and persuasion, coming typically from the disciplines of communications and psycholinguistics, points to the necessity of early interventions, the correct persuasiveness of intervention, the need for repetitive interventions and – due to the nature of social networks with the attendant simplicity of creation and mass propagation of misinformation – the need to do this at scale. I intend to contribute to the research by improving on the incompleteness and ineffectiveness[1] of current counter measures, by researching how the gap at the intervention level can be narrowed.

Previous Interventions. In response to the proliferation of misinformation over recent years, a media ecosystem has been created, focussing on countering misinformation. The Royal Society's report highlights different approaches, beginning with automated detection systems as one of the chief instruments used in combatting the spread of misinformation (The Royal Society, 2022). Apart from detecting misinformation, media literacy campaigns[2], legal action[3], platform agreements[4], arrests[5] and internet shutdowns[6] are amongst the measures aimed at countering misinformation (Funke & Flamini, 2021).

Most techniques designed to stop the spread of misinformation begin with fact-checking services. The mere existence of fact checking, does not in itself remedy the misinformation problem (Lewandowsky et al., 2012). It is an essential aid to those who are motivated to combat misinformation and does provide some of the remedial means (i.e. relevant truths), but the next step of actively addressing the problem requires a broader approach.

A New Contribution. I intend to improve on the incompleteness of current counter measures to misinformation by addressing the gap at the intervention level (i.e. the need for rapid interventions, persuasive language tailored to the political topic, bolstered by repetition with differing arguments

[1] Continued Influence Effect (Kan et al., 2021; Ecker et al.,2011, Fein et al., 1997, Lewandowsky et al., 2012, Wilkes & Leatherbarrow, 1988)
[2] in the Netherlands and Singapore
[3] In China, France and Germany etc.
[4] in Brazil
[5] e.g. in Bahrain and Egypt
[6] e.g. in India

and wording). As identified by Roozenbeek, Freeman and van der Linden (2021) evidence-based interventions are in low supply, with more research needed. It is therefore necessary to understand, and subsequently create, persuasively optimal strategies for debunking political misinformation of different types. The creation of automated counterarguments using text generation for timely/immediate interventions is proposed. Particularly necessary for use in social media, addressing the typical, rapid flood of misinformation that can arise, is a high degree of automation. With this in mind, the presumed advantage of debunking the same message multiple times is then also achievable.

The thesis addresses the two following principal research questions:

RQ1: What factors[7] are most pertinent to successful misinformation debunking?
RQ2: How do recipients of political misinformation on social media react to debunking?

The following two hypotheses are made in light of the illusory truth effect (ITE) addressing the timeliness argument of H2, the continued influence effect (CIE) and the dual systems theory concern the persuasiveness of messages (thus resulting in H1). H2 in particular is derived from the theories of emotion and build thematically on the work of Sangalang, Ophir & Cappella (2019) and Shi et al. (2020). Debunking is typically factual in nature and does not contain any emotion whilst this proposal posits making factually correct arguments more persuasive. H2 thus combines ideas from both ITE, CIE, the theories of emotional persuasion and dual system theory (Okuhara et al.2020; Stanovich, 2005).

H1: Debunking of misinformation is most effective when counterarguments are using techniques found to be politically persuasive.

H2: Debunking is most effective when combining timeliness, matching counterarguments and persuasive language.

Research Design

To improve upon the current misinformation counter measures, debunking needs to be almost immediate (Brashier et al., 2021), factually well targeted and in appropriately persuasive language (Scanell et al., 2021). Additionally, such processes must work at scale; this is essential since disinformation is being produced in ever more copious amounts, spreads faster than the truth (Vosoughi, Roy & Aral, 2018) and fact checking – by itself a time-consuming process (Shaar et al., 2020) – cannot keep step when manually administered. The practical consequence of this necessitates a very high degree of automation, which is the central ambition of the project/thesis and to my best knowledge has until now not been demonstrated either in research or in commercial applications.

[7] Potential factors: timeliness and repetition, strength and precision of counterargument, misdirection or confusion of argumentation, persuasiveness of language in the counterargument

To facilitate the complex task of automating the creation of reliable, fact-checked, persuasive interventions at scale and to also test the 2 hypotheses mentioned above among others, the research will be split into 4 phases, where the results of the earlier phases are used to inform those following.

Phase 1: How is Political Language Made Persuasive? A highly detailed semantic analysis of engagement with massive numbers of political social media postings will be undertaken, producing a statistical representation of the most efficient language-elements of political persuasion. A by-product of the analysis will be a collection of the ostensibly most persuasive postings for every single combination of measured humour-type, emotion-type and convincement level. These postings will be employed in an appropriately powered, crowd-sourced experiment to assess the persuasiveness of the postings (plus non-persuasive control texts). The experimental results will be correlated with the statistical representation to establish causality.

Phase 2: Most Effective Counterargument Strategies. Beginning with an initial set of known misinformational texts, an argument miner will be implemented, which interfaces with the APIs[8] of several (at least 3, to facilitate a fault tolerant[9] approach) reputable, fact-checking services. The services are queried based on the misinformation content and should return information implying the correctness of the assertion and, for non-factual information, associated facts which refute the initial statement. The argument miner will be refined to intelligently summarise longer texts and query for semantically similar misinformation (necessary, since misinformation typically mutates over time, although the same ground truths are required for debunking). In this phase a further (longitudinal) experiment will determine the level of debunking success over time. An experiment employing counterarguments of differing correlations to the misinformation topics will test whether introducing other topics is more effective when counterarguments match closely the subject matter of the misleading statement rather than introducing truths regarding other topics (i.e. whataboutism[10], paltering[11]).

Phase 3: Optimization of Countermeasures against Misinformation. The raw counterarguments from the argument miner (from phase2) will be programmatically enhanced using the verified statistical results from phase 1 to achieve controlled text generation of multiple counterarguments. These will be filtered using the classifiers developed in phase 1 to produce highly optimized counterarguments. The generated counterarguments (raw argumentation, improved, language-enhanced etc.) would finally be analysed and discussed by one or more focus groups of

[8] API: Application programming interface
[9] Fault Tolerant Software: https://users.ece.cmu.edu/~koopman/des_s99/sw_fault_tolerance/
[10] Whataboutism: Responding to an accusation by raising a different issue
[11] Paltering: Using truthful statements to mislead others

political science experts to determine the accuracy, persuasiveness, quality and thus expected efficacy of the automated debunking mechanisms. Hypotheses H1 and H2, are tested in phase 3.

Phase 4: Responses to Countermeasures. In a final phase, field experiments will be undertaken to determine to what extent debunking influences recipients of dis- and misinformation. To establish a baseline of expected responses to misinformation, a number of Twitter accounts belonging to current, probable serial purveyors of disinformation will be identified. In order to make large-scale determinations of whether opinions have been changed by counterarguments (which are informed and improved upon by the outcome of the focus groups described below), an additional NLM classifier will be implemented (with crowdsourced verification as in phase 1) to measure "view changing".

The Focus Group Approach

In the rest of this paper I will go into more detail of the third phase of the research project, exploring the use of focus groups to test the generated (persuasively improved) counter arguments as well as exploring the ethical considerations when employing this method as well as for the entire thesis.

The third phase (briefly elaborated upon above) is concentrated upon examining the generated counterarguments in a focus group setting. The main goal is for researcher (and by extension the readers of the thesis) to understand the effectiveness of the content of counterarguments and to show how the interaction between the group members may have an impact on the interpretation of misinformation and the generated counterarguments.

After reviewing more than forty years of research regarding misinformation E, Sakura and Li (2021) recognize that a majority of the research that has been conducted previously has focussed on quantitative experiments. The approach elaborated here would be able to provide a qualitative view of misinformation research. Furthermore, the advice, if acted upon, could lead to an improvement in the general experience of social media users and the further understanding of people's experiences and perspectives, as is the goal of qualitative research in general as discussed by Freeman (2006).

A positive side-effect of focus groups is, that they are often considered as rewarding by the participants (unlike single interviews or online questionnaires which may be perceived as somewhat exploitative), since participants can learn from each other.

Operationalisation of the Focus Groups.

At the beginning of the session participants are introduced to the moderator and the purpose of the study. They are informed about the terms of participation and are given the consent form. The moderator should encourage the participants to ask if they have any questions or concerns about the study and address them with the group. Before the group discussion starts, participants are requested

to fill out an online self-appraisal questionnaire, where general socio-demographic information is systematically collected such as their education, professional experience and background, which may prove relevant to the research. Other demographic variables such as age and sex may also be collected to provide a general overview of the participants but are not of special interest to this research. Participants – who are selected as political science experts – are additionally asked to assess their knowledge of misinformation and the approximate frequency with which they believe themselves to encounter misinformation.

Homogeneity / Homophily of Groups. Since the general aim of the focus group is to discuss counterarguments to misinformation, the recruited participants are people with the expertise in assessing such arguments, i.e. political science experts.

McLafferty (2004) discusses the composition of groups and provides different evidence of the homogeneity versus heterogeneity of groups. For the research proposed in this paper there should be a certain degree of homogeneity so as not to influence the research. A certain degree of closeness is, though, still guaranteed as the participants are not strangers but most likely will know each other professionally through conferences and collaborations as they are selected to be political science experts, whilst not well enough to impede any potential discussion points. Hydén and Bülow (2003) mention that there is also a discussion about the ideal group size of 6 to 10 people. There are advantages in having smaller groups, as members can be sensitive to the group dynamic leading to a more stimulating discussion. Freeman (2006) points out, that the group size should not be too small as to have too similar views and opinions but not be too large so that no one can contribute.

Group Size. Although there is sufficient literature on the use of focus groups, there is – according to Carlsen and Glenton (2011) – little on how focus groups are to be implemented and specifically about how the optimal number of focus groups is decided upon. The literature mentions the terms “data” and “theoretical saturation” but the definition is not enough to operationalise the concept. Saturation is defined by Strauss and Corbin (1991), to be when researchers have collected so much data that any new or relevant data would provide no further insight. They achieve this by applying so-called theoretical sampling in adjusting the theory and hypothesis as data is collected in an iterative process of analysis and reflection, deciding whether new data provides further information or not.

Carlsen and Glenton (2011) criticize that what is meant specifically by “new data” is inadequately defined. Some rules of thumb exist, suggesting two to five groups. Whilst, this is clearly dependent on the complexity of the research question and the composition, it is worth mentioning that as the number of participants in each group is not stated, one could shrink the group size in order to create another group of its own. A critical point is made by Charmaz (2005) that small group samples are justified through the point of saturation. Carlsen and Glenton (2011) add that, simply a lack of

sufficient funds or time is cited in some papers that they have analysed, to account for small or arbitrarily small sample sizes. Regarding funding, it is noted that when a study is pre-funded, the size and number of focus groups may have to be specified in advance, precluding any notion of an iterative approach to optimizing sample size. Guest et al (2006) strongly recommend using probabilistic sampling methodology, which is exceptionally true for populations which are not easily obtained. Purposive samples, the most commonly used type, are defined according to the authors as samples which select their participants according to predefined criteria. This, as opposed to theoretical samples, where participants are selected according to progressive insights. Guest et al (2006) highlight that research which is not adamant on statistical generalizability use non-probabilistic methodologies.

Data Collection. In agreement with Strauss (1987), the grounded theory approach, employing the constant comparison method – a process of iterative comparison – is introduced in this project, whereby further data-collection is decided upon only after analysis of the previous results. Termination of the process occurs when the emergence of new themes ceases and saturation is achieved. Using this technique of constant comparison, it is expected that data saturation will be recognised through identifying overarching concepts which the subject experts employ in each cycle of data-collection. A purposive sampling technique is chosen since it is believed that people with expert knowledge of political science and misinformation are considered to be a difficult sample to achieve. It is, however, expected that during the course of the project, further criteria might arise.

Conducting the Focus Groups

Introduction / Icebreaker. The introductory procedure is modelled according to the guide by Greenbaum (2000) and should be no longer than a quarter of the time allowed for the entire discussion. First, the moderator introduces him-/herself to the group and should point out that the main people who should talk will be the participants. The moderator explains the setting of the group and that the attendees are here to discuss their opinions on artificially generated counterarguments to misinformation, thus preparing for the discussion about to take place. The moderator also points out that the discussion will be recorded so that note taking on the part of the moderator can be kept to a minimum and the flow of the conversation is not disrupted. Next, the participants are asked to introduce themselves to the other group members. As an icebreaker they are asked to talk about what their current knowledge of misinformation and disinformation is. Additional material is provided in the form of common examples of misinformation (e.g. dolphins are swimming in Venice[12]). Printed out Tweets of misinformation examples may help guide the participants into talking more about the belief of people in misinformation, the dangers of misinformation and what possible interventions may help.

[12] https://www.nationalgeographic.com/animals/article/coronavirus-pandemic-fake-animal-viral-social-media-posts

Focussing Exercise. In order to concentrate the group's attention and interaction on the topic, a so-called focussing exercise is employed. This focussing exercise will take up most of the time of the focus group session and should take up about half of the time allocated for the entirety. In this case it is modelled after a Vignette study as conducted by Bloor et. al. (2001).

A ranking exercise is suggested where the group is offered a group of statements, and participants are asked to agree amongst themselves on the ranking of the statements in order of importance. The exercise as employed in this study would be of misinformation statements with various degrees of importance. The moderator asks the participants, after having read the text and having ranked them in order of importance, when they would deem it important to debunk the misinformation and when they would consider it unnecessary. In a next step, the counterarguments, along with the original misinformation statements, are presented to the participants and the experts are asked to assess the persuasiveness of the counterarguments. The moderator probes the participants on the elements that make the texts persuasive and conversely what makes the original misinformation statements persuasive. Throughout, the moderator continues to encourage all present to participate equally in the conversation by asking them, for instance, if they agree with statements made by other participants.

Conclusion of Discussion. The discussion is drawn to a close with the moderator asking the group to write down which improvements can be made to the misinformation statements and which improvements could be made to reduce the spread of misinformation if they were to directly voice their concerns to the CEOs of the social media companies. It is attempted to reach a consensus within the group by asking which the most important aspects of countering misinformation are. As Greenbaum (2000) notes, it is also recommended to highlight, that there is also the option of the group member to propose to various CEOs that the current social media algorithms are ok as they are and that no specific changes are required. Greenbaum (2000) also mentions that the aspect of writing down the suggestions communally may help the participants to convey their real feeling and let themselves be influenced by the other group members. The final stage of the focus group session should not last longer than a quarter of the time allocated.

Data Analysis.

Recording. A traditional offline focus group will be conducted, where people are seated in a circle whilst the moderator sits among the participants as the whole session is recorded (and later transcribed). This suggests, that there is a possibility to move the focus online in order to encourage a more active discussion. In an online setting the discussion could be conducted via a video conference allowing members to also use the written chat. The proposed vignette exercise which would be complemented with a video stream able to provide additional data of, for instance, facial reactions. This additional level of data would require an additional consent from the participants for the

purposes of this research, which would be added in the consent form the participants are asked to sign at the beginning of the discussion. If this level of consent is denied then only the written chat will be used for analysis. Recording of an online session for later analysis is technically unproblematic.

Coding. Since a grounded theory approach is undertaken according to Strauss (1987), an iterative process is applied. Instead of generating a theory, though, the results of the focus group discussion are combined to make recommendations to: on the one hand improve the counterarguments generated by the researcher specifically and on the other hand to learn how to increase the persuasiveness of debunking in general.

After the first round of focus group interviews has been performed the suggested approach is to do open coding, where directed reading is applied and a first formulation of key-concepts is generated. A concept-indicator model[13] is generated according to Strauss (1987). Over several rounds of coding it is expected that the data will summarize, distil and condense in the codes, as is described by Saldaña (2016) and thus more focused coding is applied. The goal is to look for patterns as documented in the transcribed texts of the discussion and relationships between topics, concepts and variables mentioned by the participants. The concept-indicator model is aimed to provide a visual representation of the patterns, the model is expected to result out of the codes using ATLAS.ti, a tool for qualitative research. With the choice of computer-aided coding it is also simple to perform simultaneous coding (i.e. applying two or more codes in the same text) as it is possible to find embedded codes that are part of a larger social scheme.

After several rounds of focussed coding, it is hypothesized that a change of perspective would be beneficial, which is why it is proposed to change the view from the level of the quotation to the level of a respondent. Saldaña (2016) explains that a change of filter through which the data is viewed may influence the types of questions that are asked during codes, thus creating more varied codes. The responses could then also be visualised on selected key concepts to gain a better understanding of the data.

Ethical Considerations.

The Role of the Moderator in the Focus Groups. In conducting the focus groups, the moderator has a vital role to play. Only when focus group members view themselves in a coherent group are they able to interact as such. It is subsequently up to the moderator to provide a conducive environment so that group members may view themselves as part of the group. The moderator must anticipate the flow of the discussion to guide the group to topics which are beneficial, as opposed to

[13] Strauss (1987) describes grounded theory to be based on a concept-indicator model, which directs the conceptual coding of a set of empirical indicators. Indicators are behavioural actions and events that are able to be observed or described in this case in the words of the focus group members. These data are indicators of a concept the analyst derives from them, at first provisionally but later with more certainty when a concept become more clearly defined through coding.

dead-ends. Additionally the moderator must deal with typical participant behaviour such as shy group members, encouraging them to talk. Rambling participants and dominant people may need to be subdued to allow each group member to contribute fully to the discussion. As such, the moderator plays a prominent role in guiding and directing the group and has a significant influence on the topic of interest (Freeman, 2006). The moderator must, thus, embrace their role in the group.

The moderator also has the job of assuring the participants that there would be privacy in gathering, storing and handling data as well as explaining the set-up of the discussion and giving the participants the opportunity to ask any questions before the discussion begins.

Millward (1995) identifies four types of content- and process-related moderator-style which dictate the relative structure of interviews. She recommends that a moderator low in control and high in process is the most appropriate for a focus group setting. The control over the content is held to a minimum but all relevant topics are discussed in depth.

The Role of Truth in Focus Groups. Although participant interaction is a central element of the focus group method, in recent years there has been a discussion about the discrepancy of the role of participant interaction in theory and in practice. The debate centres around two approaches to interactivity. The rationalist perspective highlights the individual and implicitly assumes that the participants already have fully formed ideas of so-called “truths” which will individually become apparent from the group discussion in the same way as they would via multiple bilateral interviews with all participants. Simultaneously, there is the notion that the group setting may contaminate the ideas of the individual. On the other hand, the social perspective of the dialogic perspective opposes that of the rationalist. Ideas and opinions are co-constructed by the social interaction between individuals in specific social contexts, which give them meaning.

The debate regarding participant interaction revolves around the social perspective accusing the rationalist perspective of viewing social interaction as secondary to individual points of view (Belzile & Öberg, 2012). This key difference between researchers subscribing to the one perspective or the other (determining whether social interaction is central or secondary) may also have an effect on the interpretation of this group-focus discussion, as the topic of concern is that of misinformation where social interaction has a potentially large role to play. There needs to be a distinction made between an individual’s understanding and that of the group’s. It may be that the individuals see a text as the truth or fact but, only after discussion in the group, may the individual find that their own understanding is not that of others. The reverse may also be true, where, only through the external knowledge of others, an individual may be able to evaluate a given text as not a falsehood but as the truth.

Broader Ethical Considerations in the Thesis.

There are three human interfaces that especially need to be considered.

Data gathered through:

1) Social media posts in phases 1 and 4

2) Political science experts participating in the focus groups of phase 3 and

3) Participants in the crowdsourced experiments of phase 1 and 2.

Personal data collected through social media such as the Twitter handles will be anonymized though they are publicly available data. As per the EUI Guidelines on Data Protection[14] the intended use of the data will comply with the terms and conditions established by the Data Controller. Sensitive data such as, for example, medical data are not intended to be collected and political opinions (which also fall under the category of sensitive data) are only collected insofar as expressed in the public Tweets posted by individuals.

According to SERISS[15] social media guidelines, as the data to be gathered is publicly available (and no protected Tweets or Private Messages are collected) there is no consent form needed. Additionally, the Tweets collected are from known groups (primarily parliamentarians) which gives a greater possibility of adjustment to exclude vulnerable populations within Twitter who might be potentially unaware of who they are sharing their information with online (Wheeler, 2018). Wheeler (2018) also argues that, in contrast to other platforms, the broad privacy settings of Twitter mean users have reasonable expectations of their privacy supporting analytic usage of the platform. Regarding data security, the experimental data will be stored on a private GitHub Repository, whose compliance regulations state that "GitHub has achieved both the AICPA Service Organisation Controls (SOC) 1 and SOC 2 Type 2 compliance for GitHub Enterprise Cloud"[16]. The stored data is anonymized, the Twitter handles being allocated numbers.

The focus group participants will be provided with an informed consent form at the beginning of the discussion, the EUI has a template which will be adapted for the situation at hand. They will be briefed at the beginning of the discussion on the purpose of the focus group (i.e. to provide recommendations and assess their opinion on the quality of generated counterarguments). It will be made clear to the subjects that they have a right to withdraw their consent at any time.

The crowdsourced experiments will be conducted anonymously. Some individual data will be gathered in the experiment to assess the opinions of the individuals. As with the focus groups the

[14] https://www.eui.eu/documents/servicesadmin/deanofstudies/researchethics/guide-data-protection-research.pdf
[15] Synergies for Europe's Research Infrastructures in the Social Sciences (https://seriss.eu/wp-content/uploads/2019/07/Social-media-guidelines_quick-ref.pdf)
[16] https://github.blog/2019-08-27-github-has-soc-1-and-soc-2-type-2-reports/

participants will be asked to read and agree to sign the informed consent where they will be briefed on the purpose of the study (i.e. to assess their opinion on the effectiveness of the generated counterarguments).

Truth in the Thesis. When dealing with disinformation, counter measures and, in essence, truths and untruths, we need to, at least briefly, consider the meaning of truth in the context of this work. Truth has been, for thousands of years, a central subject for philosophers. There is a wide spectrum of theories of truth, amongst the most notable are the correspondence (Crivelli, 2004) and coherence theories (Bradley, 1914; Putnam, 1981), sematic theories (e.g. Kripke 1975), Axiomatic Theories of Truth, the Identity Theory of Truth, Pluralist Theories of Truth, the Pragmatic Theory of Truth and the Revision Theory of Truth (Glanzberg, 2021). Behind each group of theories lies a wealth of philosophical reasoning, much of which contradicting or rationalizing the other theories. Additionally, much of the subject of philosophy relates to truth either by referring to it directly or by implying it.

Since it is far from clear that truth is sharply definable, and that truth itself is not intrinsic to this work, I shall proceed from the established position that – in general – publicised untruths are in different ways harmful, destructive or dangerous and that countering untruths with facts and grounded truths is a sometimes difficult but ultimately good thing. I posit the unremarkable idea that there do exist truths and lies which I shall refer to respectively – and with no further analysis – as T and T'. I evaluate not specific instances of truths and lies, but actors' beliefs in both and especially on how best to convince those believing lies to accept and believe in the truthful counterarguments.

Bibliography

Belzile, J. A., & Öberg, G. (2012). Where to begin? Grappling with how to use participant interaction in focus group design. *Qualitative Research*, *12*(4), 459–472. https://doi.org/10.1177/1468794111433089

Bloor, M., Frankland, J., Thomas, M., & Robson, K. (2001). Preparation and Conduct. In *Focus groups in social research*. SAGE Publications.

Bradley, F. H. (1914). *Essays on truth and reality*.

Brashier, N. M., Pennycook, G., Berinsky, A. J., & Rand, D. G. (2021). Timing matters when correcting fake news. *Proceedings of the National Academy of Sciences*, *118*(5), e2020043118. https://doi.org/10.1073/pnas.2020043118

Carlsen, B., & Glenton, C. (2011). What about N? A methodological study of sample-size reporting in focus group studies. *BMC Medical Research Methodology*, *11*(1). https://doi.org/10.1186/1471-2288-11-26

Charmaz, K. (2005). Grounded Theory in the 21st Century: Applications for Advancing Social Justice Studies. In N. Denzin & Y. Lincoln (Eds.), *The SAGE Encyclopedia of Qualitative Research Methods* (pp. 507–535). SAGE Publications, Inc. https://doi.org/10.4135/9781412963909.n189

Crivelli, P. (2004). *Aristotle on truth*. Cambridge University Press.

E, Q., Sakura, O., & Li, G. (2021). Mapping the field of misinformation correction and its effects: A review of four decades of research. *Social Science Information*, *60*(4), 522–547. https://doi.org/10.1177/05390184211053759

Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, *18*(3), 570–578. https://doi.org/10.3758/s13423-011-0065-1

Fein, S., McCloskey, A. L., & Tomlinson, T. M. (1997). Can the Jury Disregard that Information? The Use of Suspicion to Reduce the Prejudicial Effects of Pretrial Publicity and Inadmissible Testimony. *Personality and Social Psychology Bulletin*, *23*(11), 1215–1226. https://doi.org/10.1177/01461672972311008

Freelon, D., & Wells, C. (2020). Disinformation as Political Communication. *Political Communication*, *37*(2), 145–156. https://doi.org/10.1080/10584609.2020.1723755

Freeman, T. (2006). 'Best practice' in focus group research: Making sense of different views. *Journal of Advanced Nursing*, *56*(5), 491–497. https://doi.org/10.1111/j.1365-2648.2006.04043.x

Funke, D., & Flamini, D. (2021, August 13). *A guide to anti-misinformation actions around the world*. Poynter. https://www.poynter.org/ifcn/anti-misinformation-actions/

Glanzberg, M. (2021). Truth. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2021/entries/truth/

Greenbaum, T. L. (2000). The Discussion Guide. In *Moderating focus groups: A practical guide for group facilitation*. Sage Publications.

Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Hydén, L.-C., & Bülow, P. (2003). Who's talking: Drawing conclusions from focus groups—some methodological considerations. *International Journal of Social Research Methodology*, *6*(4), 305–321. https://doi.org/10.1080/13645570210124865

Kan, I. P., Pizzonia, K. L., Drummey, A. B., & Mikkelsen, E. J. V. (2021). Exploring factors that mitigate the continued influence of misinformation. *Cognitive Research: Principles and Implications*, *6*(1), 76. https://doi.org/10.1186/s41235-021-00335-9

Kripke, S. (1975). Outline of a Theory of Truth. *The Journal of Philosophy*, *72*(19), 690. https://doi.org/10.2307/2024634

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest*, *13*(3), 106–131. https://doi.org/10.1177/1529100612451018

McLafferty, I. (2004). Focus group interviews as a data collecting strategy. *Journal of Advanced Nursing*, *48*(2), 187–194. https://doi.org/10.1111/j.1365-2648.2004.03186.x

Millward, L. (1995). Focus Groups. In G. M. Breakwell, S. Hammond, C. Fife-Schaw, & J. A. Smith (Eds.), *Research methods in psychology* (pp. 274–292). Sage Publications.

Okuhara, T., Ishikawa, H., Okada, H., Ueno, H., & Kiuchi, T. (2020). Dual-process theories to counter the anti-vaccination movement. *Preventive Medicine Reports*, *20*, 101205. https://doi.org/10.1016/j.pmedr.2020.101205

Putnam, H. (1981). *Reason, Truth and History* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511625398

Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, *5*(1), 65. https://doi.org/10.1057/s41599-019-0279-9

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3E [Third edition]). SAGE.

Sangalang, A., Ophir, Y., & Cappella, J. N. (2019). The Potential for Narrative Correctives to Combat Misinformation†. *Journal of Communication*, *69*(3), 298–319. https://doi.org/10.1093/joc/jqz014

Scannell, D., Desens, L., Guadagno, M., Tra, Y., Acker, E., Sheridan, K., Rosner, M., Mathieu, J., & Fulk, M. (2021). COVID-19 Vaccine Discourse on Twitter: A Content Analysis of Persuasion Techniques, Sentiment and Mis/Disinformation. *Journal of Health Communication*, *26*(7), 443–459. https://doi.org/10.1080/10810730.2021.1955050

Shaar, S., Martino, G. D. S., Babulkov, N., & Nakov, P. (2020). That is a Known Lie: Detecting Previously Fact-Checked Claims. *ArXiv:2005.06058 [Cs]*. http://arxiv.org/abs/2005.06058

Shi, W., Wang, X., Oh, Y. J., Zhang, J., Sahay, S., & Yu, Z. (2020). Effects of Persuasive Dialogues: Testing Bot Identities and Inquiry Strategies. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376843

Stanovich, K. E. (2005). *The robot's rebellion: Finding meaning in the age of Darwin*. University of Chicago Press.

Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge University Press.

Strauss, A. L., & Corbin, J. M. (1991). *Basics of qualitative research: Grounded theory procedures and techniques* (3. printing). Sage.

The Royal Society. (2022). *The online information environment: Understanding how the internet shapes people's engagement with scientific information*. https://royalsociety.org/topics-policy/projects/online-information-environment/

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

Wheeler, J. (2018). Mining the first 100 days: Human and data ethics in Twitter research. *Journal of Librarianship and Scholarly Communication*, *6*(2). https://doi.org/10.7710/2162-3309.2235

Wilkes, A. L., & Leatherbarrow, M. (1988). Editing Episodic Memory following the Identification of Error. *The Quarterly Journal of Experimental Psychology Section A*, *40*(2), 361–387. https://doi.org/10.1080/02724988843000168