What Ethical Solutions Are There to Peer Distributed Misinformation?

Introduction

This essay attempts to find solutions to peer distributed misinformation based on ethical frameworks in Kant's *Groundwork of the Metaphysics of Morals*.[1] While misinformation is created as '[deliberately] manipulated information intended to deceive and mislead audiences'; the distribution of misinformation on peer-to-peer networks is complex and at times non-malicious. Finding effective solutions involves considering online behaviour as well as ethics. For this reason, the first part of this essay assesses who is responsible for peer distributed misinformation. Additionally, misinformation is explored as a transgression of ethics in Kant's *Groundwork of the Metaphysics of Morals*. These teachings provide strong moral frameworks which relate to both the roles of companies and users in peer distributed misinformation. The second part of the essay analyses the morality of solutions based on principles established in the *Groundwork of the Metaphysics of Morals*, such as autonomy and respect for human reasoning.

Part One

Who is responsible for peer distributed misinformation?

Firstly, we consider why social network companies are responsible. Some algorithms designed and used by networking companies foster misinformation, making misinformation more likely to be distributed by peers. Therefore, networking companies fail to uphold Kant's principle that the individual reasoning capabilities of each person should be respected.[2] This means the architecture of social media is not designed to safeguard people who are susceptible to misinformation. Notably, social networking sites filter users into likeminded groups in which echo chambers form.[3] Within these chambers unconventional or false beliefs are perceptively unanimous.[4] These chambers are epistemically dangerous, as they flaw how knowledge is acquired. Since humans naturally give credibility to sufficiently accepted beliefs, these chambers provide an environment for false information to be considered true because enough people believe it.[5] False beliefs can be held so strong in echo chambers that members become unresponsive to counterevidence.[6]

---

[1] Iosifidis, Nicoli, (2020) 64-65. (Digital, Culture, Media and Sports Committee of the House of Commons, 2019: 7) Iosifidis and Nicoli list five major distributors of misinformation: bodies close to the Russian government, right-wing groups, groups profiting such as those based in Macedonia, formal campaigns using marketing tools and peer-to-peer distribution networks. This essay focusses on the latter.
[2] Formosa, (2017) 75.
[3] Fantyl, (2021) 645.
[4] Edenberg, (2021), 259-261.
[5] Edenberg, (2021) 261.
[6] Fantl, (2021), 645.

Furthermore, algorithms decide which news stories users will find appealing. Likes, clicks, and shares factor into what news an algorithm presents. Consequently, algorithms potentially favour popularity over credibility.[7] For instance, a 2016 Twitter algorithm filtered timeline tweets based on popularity not chronology.[8] Therefore, social networking algorithms do not always present trustworthy sources. This flaws our epistemic relationship with information by partially taking away choice in the acquisition of knowledge. Consequently, users cannot be blamed for distributing and believing misinformation.

Despite the compounding role of algorithms in peer distributed misinformation, they have positive financial impacts for social networking companies. These algorithms boost time users spend on social media by exposing users to personalised content.[9] For this reason, algorithms increase the visibility of misinformation and help it become ingrained further making companies primarily responsible for peer distributed misinformation.

Additionally, companies are responsible as business decisions effect the reach of harm misinformation causes. Monetary business decisions often result in people being harmed and treated as a *mere means*.[10] This means decisions made by social media companies are not made from necessity or respect for the 'priori in every human being'. Instead, they are made because they are merely possible for financial purposes.[11] For instance, Twitter largely accepted President Donald Trump's misinformation which minimised the lethality of COVID-19.[12] In the midst of the pandemic Trump made tweets falsifying COVID-19. It was not until January 2021 when Trump was banned from Twitter.[13] As of the 29th of December 2020, there were 1034.78 deaths per million confirmed to be caused by COVID-19.[14] Estimating how misinformation factored into the deaths of so many people is complex. However, Twitter should have acknowledged the users who believed Trump's false claims by stopping Trump posting misinformation. Anthropologist, Alberto Acerbi explains, as a social species, humans heavily rely on following figures of prestige and popularity, such as a president.[15] For this reason Trump's twitter account was devastatingly harmful during the pandemic, yet his presence on Twitter was generally hugely lucrative. This instance demonstrates barring users from the community is a last resort as even harmful users contribute towards use, popularity, and sales.[16] This is morally wrong as Twitter should have always prioritised human life over business.

Now we consider why peer networks are responsible for misinformation. Peer networks share misinformation maliciously or carelessly, transgressing Kant's teaching that '[reason] recognizes as its highest practical function the grounding of a good will'.[17] People deliberately using misinformation maliciously may act pathologically – hoping to incite online arguments, convey hateful ideologies, and undermine social and political structures.[18] Likely to deliberately spread harmful misinformation are the individuals and groups termed by psychologist John Suler as *'rabble rousers', 'political paranoids'* and *'tenacious debaters.'* It is possible cyberspace contributes towards some people's

[7] Acerbi, (2020) 152.
[8] Acerbi, (2020) 127.
[9] Rhodes, (2021) 2.
[10] Formosa, (2017) 121.
[11] Kant, *Groundwork of the Metaphysics of Morals,* Zweyter Abschnitt *4.415.35* 'Man darf ihn nicht bloß als nothwendig zu einer ungewissen, bloß möglichen Absicht, vortragen, sondern zu einer Absicht, die man sicher und a priori bey jedem Menschen voraussetzen kweil sie zu seinem Wesen gehört.'
[12] *All the President's Lies about the Coronovirus*. TheAtlanic. Website.
[13] *Permanent suspension of @realDonaldTrump.* TwitterInc. Website.
[14] *Coronavirus Pandemic (COVID-19) – the data.* OurWorldInData. Website.
[15] Acerbi, (2020) 96.
[16] Suler, (2016) 330.
[17] Kant, *Groundwork of the Metaphysics of Morals,* Erster Abschnitt 4.396.30 'Weil die Vernunft, die ihre höchste practische Bestimmung in der Gründung eines guten Willens erkennt.'
[18] Suler, (2016) 331.

tendency to misbehave online. Humans tend to experience less inhibitions online in comparison to real life.[19] According to Suler, the uninhibition affect allows frustrated individuals to feel they can misbehave on social media as a mask of anonymity protects them from facing real consequences.[20]

This uninhibited affect may also influence people without malicious intentions to share misinformation. A substantial quantity of misinformation is likely shared by people acting more assertive than usual – broadcasting their internet findings while forgetting they are addressing real people not a computer.[21] Leslie Lawrenson's case exemplifies non-malicious distribution of misinformation. News articles report Lawrenson used social media to share anti-vaccine misinformation, being genuinely concerned about the vaccine's safety. Lawrenson himself died of COVID-19 and most likely did not wish others the same harm.[22] Nevertheless, all sharers of misinformation are autonomous. In Kant's ideology autonomy is a characteristic of our will, which rules itself – meaning we have capacity to rule ourselves and be our own moral agent.[23] Therefore, although there are accepted changes in online behaviour, it can be argued that different aspects of our nature are all part of our underlying self-direction.[24] Consequently, each person – whether they have malicious or non-malicious intentions – is fully responsible for their actions online. For this reason, networking companies are not responsible for the actions of self-governing individuals.

Reasons why both networkers and companies are individually responsible have thus been presented. Yet, illuminating the complexities of this argument, it is possible both parties simultaneously transgress Kant's moral teachings and contribute towards misinformation. Kant explains beneficent actions are often based on acting dutifully. Moreover, to act dutifully we must understand accepted laws or beliefs – '[sole] respect for the law is the incentive that can give an action a moral worth.'[25] In terms of misinformation, there is no established law that sharing information on social media indicates belief or disbelief in the information.[26] Having no law explain the epistemic relationship between the sharer and the information can have harmful consequences. For instance, Donald Trump deferred accountability for spreading misinformation by highlighting the difference between a *tweet* and a *retweet*: 'I didn't tweet, I retweeted'.[27] The share preserved racist ideologies towards Afro-Americans by falsely magnifying rates at which black people kill white people. This misinformation was shared in 2015 and was thus not responsible for Trump's ban from Twitter in 2021.[28] Yet, as neither networking companies nor users have clear boundaries on sharing information this example highlights how both parties are equally responsible for the acceptance of misinformation. Consequently, Twitter's directors and users did not act beneficently and dutifully.

---

[19] Kiesler, et al. (1984) 1125.
[20] Suler, (2016) 330.
[21] Kiesler, et al. (1984) 1125.
[22] *Anti-vaxxer dies of Covid nine days after saying the virus is 'nothing to be afraid of*'. TheIndependant. Website.
[23] Kant, *Groundwork of the Metaphysics of Morals,* Zweyter Abschnitt, 4.440 15 'Autonomie des Willens ist die Beschaffenheit des Willens, dadurch derselbe ihm selbst.'
[24] Formosa, (2017) 75.
[25] Kant, *Groundwork of the Metaphysics of Morals,* Zweyter Abschnitt, 4.440.5 'Sondern lediglich Achtung fürs Gesetz, diejenige Triebfeder sey, die der Handlung einen moralischen Werth geben kann.'
[26] Rini (2017) 47.
[27] *Donald Trump actually admitted that he doesn't check his facts. Seriously*. WashingtonPost, Website.
[28] *Permanent suspension of @realDonaldTrump.* TwitterInc. Website.

Part Two

What ethical solutions are there?

Both companies and users are responsible for peer-to-peer distributed misinformation, therefore resolving the problem will require change from both parties.

Reducing the harm caused by algorithms which foster misinformation should be the primary goal of social networking companies. The extent of harm algorithms cause is unanticipated and can depend on how they interact with human behaviour.[29] For instance, echo chambers form from a combination of filtering algorithms and a human tendency to trust sources that favour existing beliefs.[30] To combat this, computer scientist Randall Hill argues it is the moral duty of companies to publish how their algorithms utilise data and logic.[31] This may scale down the harm algorithms cause. Understanding these algorithms may allow users to critically view how algorithms manipulate engagement with social media. However publicising algorithms may have unanticipated risks. With unmonitored access, there is no way to predict how else these algorithms could be used. Kant teaches logical interest should guide us in presupposing the effects of actions.[32] Therefore, due to the unpredictable risks involved in this action, this is a dangerous solution.

Instead, companies should adopt editorial processes like mainstream news publishers. As algorithms demote and promote news content, social networking platforms are akin to publishers. Therefore, it is the moral duty of social media companies to revise content like mainstream news publishers.[33] Robust editorial measures may safeguard people against misinformation. For this reason, editorial measures constitute as a good solution as they comply with Kant's teaching that the individual reasoning capabilities of each person should be respected.

Currently, editorial processes on social networks are weak. For instance, Instagram involves users in the editorial process through a report function which flags misinformation. This is an ineffective editorial measure. The 'editors' may be unreliable as they can also report accurate news as misinformation.[34] Instead, like mainstream publishers, the aim of editorial processes should be to remove harmful misinformation before it reaches consumers. This may be possible with political misinformation which often has a direct source. Before political misinformation spreads to the newsfeeds of general users, it originates from political groups. From these groups the content is hyped by bots and partisans, resulting in the spread of misinformation.[35] Editorial measures should reduce the visibility and prevent the circulation of misinformation by ensuring content from partisan political groups is not re-posted or favoured by algorithms.

Controlling hashtags is an effective editorial feature presently in use. This process ensures people cannot search for misinformation and stops transmission via hashtag. However, this feature could be improved by informing users they have searched for misinformation and providing a definition of misinformation. This would confront believers in misinformation and potentially spur them to re-consider the epistemology of their beliefs.

---

[29] Hill, (2019) 40.
[30] Rhodes, (2021) 6. Brown, (2021) 219.
[31] Hill, (2019) 41.
[32] Kant, *Groundwork of the Metaphysics of Morals,* Dritter Abschnitt, 4.460.25 'Das logische Interesse der Vernunft (ihre Einsichten zu befördern,) ist niemals unmittelbar, sondern setzt Absichten ihres Gebrauchs voraus.'
[33] Brown, (2021) 220.
[34] Brown, (2021) 221.
[35] Rhodes, (2021) 2.

Unfortunately, large amounts of misinformation will still likely be visible even with robust editorial features. On one hand, visibility misinformation is non ideal; however, if users expect all misinformation to be removed, any remaining misinformation may appear more trustworthy than previously. This could create an 'implied truth effect.[36] For this reason, it is better for networking companies to highlight the transparency of misinformation sources in addition to removing it. Platforms can expose unreliable information sources by redirecting users to legitimate information before or while they engage with misinformation. Twitter effectively demonstrated this solution. While misinformation circulated that oregano oil could be used as a treatment for COVID, Twitter directed users searching for coronavirus-related hashtags to the website Center for Disease Control and Prevention.[37] Additionally, it may be possible to expose sources and themes of misinformation through announcements. Like adverts are displayed, announcements warning users of the latest and most harmful misinformation in circulation could be displayed. Overall, robust editorial measures constitute as a good solution as they factor the individual reasoning capabilities of each person into the structure of social networks.

While companies can make editorial changes to reduce peer distributed misinformation, social media users can embrace educational changes. It must be noted that the idea that users should take any responsibility is controversial. As algorithms interfere with knowledge acquisition on social media; some researchers believe misinformation is a structural problem beyond the control of users.[38] Brown states solutions requiring critical thinking are not robust.[39] In agreement, Croce and Piazza argue educational approaches are poorly motivated and structural intervention is the only way.[40] Despite this algorithmic problem, I suggest education approaches are not poorly founded.

Notably, epistemic nudging may increase awareness to misinformation. This method attempts to enlighten people online, helping them make informed judgements through peer influence.[41] Importantly, for nudging to be effective it must come from impartial sources.[42] Posts made by celebrities documenting their COVID-19 vaccinations exemplify this solution in practice.[43] Endorsing influential people to post misinformation counterevidence can alter perceived social norms on social media, positively influencing people's beliefs and behaviour.[44] Epistemic nudging may be effective for persons more receptive to social conditioning; such as paranoid believers who perceive authoritative removals of misinformation as further evidence of their suspicions.[45] As nudging is resistible, it influences our reasoning without taking away freedom of choice. Neither does it make users to adopt beliefs through force or indoctrination.[46] For this reason, epistemic nudging is a good solution as it values autonomy. Therefore, it complies with Kant's teaching that a person's autonomy should be respected as it is the ground for human dignity.[47]

A second argument supporting educational approaches is that there have already been instances when online users thought critically about online communication and behaviour. Therefore, people can potentially be encouraged to think critically about how they share information online. When the internet was first available to general users, the idea of *netiquette* was formed. This notion advocated cooperation, helpfulness, and trust in cyberspace and communication arenas such as the

[36] Brown, (2021) 220-221.
[37] Brown, (2021) 221.
[38] Croce, Piazza, (2021) 2.
[39] Brown, (2021) 219.
[40] Croce, Piazza, (2021) 2.
[41] Grundmann, (2021) 2.
[42] Grundmann, (2021) 5.
[43] Celebrities are Endorsing Covid Vaccines. Does it Help? TheNewYorkTimes. Website.
[44] Cookson, (2021) 2.
[45] Grundmann, (2021) 5.
[46] Grundman, (2021) 2.
[47] Kant, *Groundwork of the Metaphysics of Morals*, Zweyter Abschnitt, 4.436 'Autonomie ist also der Grund der Würde der menschlichen und jeder vernünftigen Natur.'

Usenet and mailing lists.[48] Attempts were made to formulate a set of online rules based on the ten commandments with 'thou shalt not use a computer to harm other people' as the first commandment.[49] The idea of commandments is notably strange at present; however, networking platforms could guide users in communicating online. For example, on Instagram there are currently no guidelines educating users of what is acceptable to share. I propose social networking companies can create interfaces which advise users not to share misinformation before they make a post. This may immediately reduce misinformation shares. Moreover, this solution could help educate users by increasing their epistemic humility and critical thinking thus making them less susceptible to misinformation.[50] Communication guidelines are a strong ethical solution as they allow users to make their own informed moral decisions. This follows Kant's teaching that a person's capacity to be their own moral agent should be respected.[51]

Conclusion

Finally, I must acknowledge this argument is limited in its deontological stance – moral conclusions are drawn based purely on a set of principles established in Kant's *Groundwork of the Metaphysics of Morals.* Notably, most academics do not suggest educational approaches are a moral solution. Perhaps other ethical frameworks could highlight flaws in my conclusion that educational should be implemented as well as structural changes. My moral conclusion favours the principles of autonomy and respect for human reasoning. Had the principles honesty and transparency factored more in Kant's ethical framework, I may agree that algorithms need to be published and structural change is the only way. However, social media algorithms may cause unpredictable harm if published. Moreover, for a structural change to be effective, the architecture of social media would need to prevent echo chambers forming and harmful influencers such as Trump from gaining following. This would be a significantly complex task as it is in human nature to form groups of similar beliefs and follow figures of prestige. Therefore, despite the limitations in my conclusion, this essay contributes towards finding a solution to peer distributed misinformation by highlighting the importance of educational approaches. Crucially, educational approaches such as displaying announcements, creating guidelines, and endorsing misinformation counterevidence posts may be easier to initialise than robust editorial measures. An editorial process may not be rapidly effective due to the scale of work involved in filtering out misinformation. For this reason, initialising both educational and editorial approaches is the most moral solution. If both companies and users embraced editorial and educational approaches, principles in a *Groundwork of the Metaphysics of Morals* could be upheld. Companies would cease treating users as *mere means* by prioritising human safety over financial gain. While users of social media could be their own moral agent – having the information to act and communicate beneficiary.

Word count 3012

Bibliography

Translations:

Timmermann, J. Gregor, M. (2011), tr., Immanuel Kant: Groundwork of the Metaphysics of Morals, Cambridge

Books and book chapters:

[48] Suler, (2016) 328.
[49] Scheuermann (1997) 270.
[50] Fantl, (2021), 646.
[51] Formosa, (2017) 75.

Acerbi, A. (2020) Cultural Evolution in the Digital Age, Oxford

Audi, R. (2015) Means, Ends and Persons: The Meaning and Psychological Dimensions of Kant's Humanity Formula, Oxford

Barak, A., Suler, J. (2008) Reflections on the Psychology and Social Science of Cyberspace in A. Barak Psychological aspects of Cyberspace: Theory, Research, Applications, Cambridge.

Brown, E.(2021) 'Regulating the Spread of Online Misinformation' in M, Hannon. J, de Ridder, *The Routledge Handbook of Political Epistemology*, New York

Edenberg, E., Hannon, M. (2021) Political Epistemology, Oxford

Formosa, P. (2017) Kantian Ethics, Dignity and Perfection, Cambridge

Gackenbach, J. (2007), Psychology and the Internet: Intrapersonal, Interpersonal, and Transpersonal Implications, 2nd ed, Amsterdam.

Henschke, A. (2017) Ethics in an Age of Surveillance: Personal Information and Virtual Identities, Cambridge

Hill, R (2019) 'Ethics of Immersive Technologies' in A. Abbas (ed.), *Next Generation Ethics: Engineering a Better Society*, Cambridge

Morahan-Martin, J (2016) 'Internet Abuse: Emerging Trends and Lingering Questions' in A. Barak (ed.*), Psychological Aspects of Cyberspace: Theory, Research, Applications, Cambridge.*

Sokal, A (ed.) (2010) (ed,) Beyond the Hoax: Science, Philosophy and Culture, Oxford

Southwell, B., Thorson, A. Sheble, L (2018) Misinformation and Mass Audiences, Texas

Suler, J. (2016), Psychology of the Digital Age: Humans Become Electric, Cambridge.

Sunstein, C. (2018) #Republic: Divided Democracy in the Age of Social Media, Princeton

Journal Articles:

Cookson, D (2021), 'A social norms approach intervention to address misperceptions of anti-vaccine conspiracy beliefs amongst UK parents', Plos One 16(11), e0258985 - e0258985.

Croce, M. Piazza, T (2021), 'Consuming Fake News: Can We Do Any Better?', Social Epistemology, 1-10

Engstrom, S (1992) 'The Concept of the Highest Good in Kant's Moral Theory', Philosophy and Phenomenological Research 52 (4) 747-780

Fanyl, J. (2021) 'Fake News vs Echo Chambers', Social Epistemology: A Journal of Knowledge, Culture and Policy 35(6), 645-659.

Grundmann, T. (2021) 'The Possibility of Epistemic Nudging', Social Epistemology 1-22.

Halsbury, (1959), Epistemology and Communication Theory, The Journal of the Royal Institute of Philosophy, 34(131), 289-307.

Iosifidis, P., Nicoli, N. (2020) 'The Battle to End Fake News: A qualitative content analysis of Facebook announcements on how it combats disinformation' International Communication Gazette 82(1) 60–81

Kiesler, S. J, Siegel. T, McGuire, (1984) 'Social Psychological Aspects of Computer-Mediated Communication', American Psychological Association 39 (10), 1123-1134

Niemiec, E. (2020) 'COVID-19 and Misinformation', EMBO reports 21(11), 51420-51420.

Parks, M. (1996), 'Making Friends in Cyberspace', Journal of Communication 46(1), 80 - 97.

Rhodes, S. (2021) 'Filter Bubbles, Echo Chambers, and Fake News: How Social Media Conditions Individuals to Be Less Critical of Political Misinformation' , Political Communication 1-22.

Rini, R. (2017) 'Fake News and Partisan Epistemology', Kennedy Institute of Ethics Journal, 27(2), 43—64.

Vosoughi, S., Roy, D. and Aral, S. (2017) 'The Spread of True and Fake News Online', MIT Initiative on the Digital Economy Research Brief, 1-5.

Walsh, J (2020) 'Social Media and Moral Panics: assessing the affects of technological change on societal reaction' International Journal of Cultural Studies 23(6) 840–859

Websites:

Alibhai, Z(2021) 'Anti-vaxxer dies of Covid nine days after saying the virus is noting to be afraid of'<https://www.independent.co.uk/news/uk/home-news/covid-anti-vaxxer-bournemouth-lawrenson-b1898012.html >[Last accessed 3 December 2021]

BBC (2021)< https://www.bbc.co.uk/news/uk-england-dorset-58080116> [Last accessed 3 December 2021]

Bickert, M (2020) 'Removing Holocaust Denial Content' <https://about.fb.com/news/2020/10/removing-holocaust-denial-content/> [Last accessed 27 November 2021]

Borchers, C. (2015), 'Donald Trump actually admitted that he doesn't check his facts. Seriously', *The Washington Post* < https://www.washingtonpost.com/news/the-fix/wp/2015/11/24/donald-trump-actually-admitted-that-he-doesnt-check-his-facts-seriously/ > [Last accessed 25 November 2021]

Gillbert, B(2021) 'Facebook says it doesn't read WhatsApp messages, but an investigation found it actually does' < https://www.businessinsider.com/facebook-reads-whatsapp-messages-encryption-2021-9?r=US&IR=T> [Last accessed 2 December 2021]

Goldman, A. *et al*. (2021), 'Social Epistemology', *The Stanford Encyclopedia of Philosophy* <https://plato.stanford.edu/archives/win2021/entries/epistemology-social/>[Last accessed 6 December 2021]

Ives, M (2021) 'Celebrities are Endorsing Covid Vaccines. Does it Help?'<https://www.nytimes.com/2021/05/01/health/vaccinated-celebrities.html > [Last accessed 7 December 2021]

Paz, C (2020) 'All the President's Lies about the Coronovirus' < https://www.theatlantic.com/politics/archive/2020/11/trumps-lies-about-coronavirus/608647/> [Last accessed 3 December]

Ritchie, H. et al. (2021) 'Coronavirus Pandemic (COVID-19) – the data' < https://ourworldindata.org/coronavirus-data> [Last accessed 3 December 2021]

Steup, M. *et al.* (2020)'Epistemology', *The Stanford Encyclopedia of Philosophy* <https://plato.stanford.edu/archives/fall2020/entries/epistemology/> [Last accessed 6 December 2021]

Twitter inc, (2021), 'Permanent Suspension of @realDonaldTrump', *Twitter (Blog)* < https://blog.twitter.com/en_us/topics/company/2020/suspension> [Last accessed 25 November 2021]