

IS J.S. MILL'S ACCOUNT OF FREE SPEECH SUSTAINABLE IN THE AGE OF SOCIAL MEDIA?



NEVIN CHELLAPPAH

ABSTRACT

In this paper, I examine whether John Stuart Mill's account of free speech can survive three main challenges posed by social media. First, I consider the problem of social media failing to distinguish between emotive and factual language. Second, I look at the problem of algorithms creating moralism. I then turn to a potential objection to my first two challenges. The objection elucidates the benefits of social media's emotional and algorithmic character, amplifying arguments and increasing public engagement. However, I take issue with this objection on consequentialist terms. I finally return to the third challenge, where I focus on how anonymity removes the consequences to our words; I contend that this final failure is the ultimate reason why Mill's account cannot persist in the modern age. In conclusion, I argue that Mill's account cannot withstand the problems posed by social media.



I. INTRODUCTION

Free speech is arguably the foundational value in liberal democracies. Contemporary liberals' endorsement of free speech traces back to John Stuart Mill's reasoning found in "Chapter II" of *On Liberty*.¹ Mill advocates for free speech due to its epistemic and social benefits of intellectual development. Social media provides a platform founded upon Mill's account, with minimal restrictions on free speech. Yet, the benefits that Mill said would come from free speech do not materialize in the context of online discourse. Hence, this essay argues that the classical liberal definition of free speech, as espoused by Mill, is no longer compatible with the digital age.

Social media presents three major challenges to Mill's account. First, social media allows for emotive discussion to take precedence over an exchange of ideas without distinguishing the two. Mill's inability to adequately categorize emotional propositions means that his account fails. The second is that moralism on social media, derived from algorithms, creates a false sense of objectivity. Mill's faith in different opinions being voiced in a civil manner means that his conception collapses. However, an objection can be waged against the first two challenges. Given that Mill's account is based on consequentialism, if the net impact of free speech on social media is more beneficial than harmful, then Mill's account may be preserved. This objection notes that both the emotional and algorithmic character of social media dialogue leads to more productive conversation, through amplifying arguments with truth-values and increasing engagement in civic discourse. However, this objection can be rendered ineffective. On the emotional character of online dialogue, the objection presupposes that only Mill's account creates the benefits that social media generates. In reality, an account that allows for censorship provides the same positive consequences, whilst limiting the harm caused. On the algorithmic character of online dialogue, the objection fails to recognize that algorithms do not lead to public engagement in discussion in the way Mill intended because algorithms are more compatible with profit than educational or democratic aims. Finally, the third challenge, that online anonymity removes the consequences of language, cannot be resolved by Mill, leaving his account inapplicable. Thus, given that Mill's account cannot withstand these three challenges, social media has seemingly eroded his conception of free speech.

II. MILL'S ACCOUNT OF FREE SPEECH

To further support the analysis of his framework in the social media age, it is important to outline Mill's account of free speech. Mill's primary concern is with the suppression of opinions by an authority. For him, the "evil of silencing...an opinion is that it is robbing the human race... [because] if the opinion is right, they are deprived...the opportunity of exchanging error for truth."² In response to censorship being presented as a trusted system to filter out true expressions from false ones, Mill posits that there is no perfect censor. The presumption of a perfect censor is proven false by history, with past authorities suppressing ideas that are currently accepted to be true—the Roman Catholic Church's censorship of Galileo's ideas comes to mind. Even if the opinion is wrong, Mill believes humanity loses "what is almost as great a benefit, the clearer perception...of truth produced by its collision with error."³ Emphasizing that "facts and arguments, to produce any effect on the mind, must be brought before it"⁴ illustrates Mill's central belief of the importance of viewpoint diversity and the value of individual thought.

Mill's argument for free speech can be formally presented as follows:

Premise 1: The truth is valuable, and people should be allowed to arrive at true beliefs.

Premise 2: Freedom of speech enables people to arrive at true beliefs.

Conclusion: Therefore, freedom of speech is valuable and ought to be promoted and protected.

Mill's argument narrows its focus on one specific aspect of free speech: free discussion. Free speech, according to Mill, means the freedom to express an opinion—further simplified to the assertion of the truth-value of a proposition. Mill does not consider speech as any utterance; rather, he understands speech as an action to seek knowledge. While appreciating speech as a means of doing is important, Mill cannot sidestep the sizable objections that social media highlights in cases where speech performs a different action. As the popularity of social media increases, platforms do not primarily aim to strive for knowledge. With social media no longer reflecting Mill's vision of a "marketplace of ideas," it bears asking whether a Millian definition of free speech still stands in this digital space.

1 John Stuart Mill, *On Liberty, Utilitarianism and Other Essays* (Oxford: Oxford University Press, 2015), 18-54.

2 Mill, *On Liberty*, 19.

3 Mill, *On Liberty*, 19.

4 Mill, *On Liberty*, 22.



III. THE FIRST CHALLENGE

A significant challenge posed by social media is that it does not differentiate between an exchange of ideas and an exchange of emotions. This is because social media does not attract attention through the truth-value of its propositions; instead, content is rewarded based on its popularity. Emotional content, which is easier to understand and connect with, leads to a snowball effect by which such content dominates on digital platforms. The fact that social media is often used as a form of entertainment means that non-propositional truth-value statements are more appealing. Such non-propositional statements online range from “I like most Tarantino movies I just think he should...die,”⁵ to “I would like to throw a jellyfish at your forehead,”⁶ to “Joe Biden, if I [do not] get stimulated, your son is getting eliminated.”⁷

The issue of discourse on social media is therefore an emotivist problem. Value judgements are associated as truth-value propositions rather than statements formed by sentiments. Generating an emotional reaction results in debate being marginalized. In its place, emotional claims become perceived as factual propositions. As Adam D. Moore identifies, “The gatekeep[ing mechanisms] of quality...are irrelevant,” if legitimacy takes form based on the number of likes or shares.⁸ This differs significantly from Mill’s advocated framework. This disparity illustrates that Mill’s conception of free speech is incoherent in the digital age, but understandably such incoherence derives from Mill’s idealism. Envisaging a Darwinian educational battle, Mill presupposes that free expression leads to intellectual discussion. For Mill, speech was purely an act to seek knowledge and develop this knowledge into action. Mill’s defense of free speech begins to look rather like, as Piers Norris Turner notes, “A defense of *not restricting viewpoints in frank and fair-minded public discussion*.”⁹ Hence, this idealistic belief that speech is only used in the context of academia or politics limits the applicability of Mill’s account.

The underlying issue that this challenge stems from is the sheer size of social media’s “marketplace of ideas.” A global platform is too large for productive discussion because it is affected by what Moore terms “content

pollution.”¹⁰ The overabundance of content to consume forces speech on a platform to prioritize emotional engagement over intellectual curiosity, as audiences prefer more “bingeable” content. The Darwinian battle no longer seeks the truth but instead attention. Since Mill’s argument was primarily epistemic in nature—attempting to establish the connection between free discussion and human flourishing—the same reasoning can now be used to suggest that such a connection has been digitally severed. Much of digital speech has little to no positive Millian value and yet such speech is deemed valuable on social media. This indicates that Mill’s account of speech as an act of seeking knowledge has been damaged in this context. Thus, Mill’s account is unable to survive the first challenge due to its failure to separate factual and emotional online discussion.

IV. THE SECOND CHALLENGE

Another challenge to Mill’s account posed by social media is that it does not facilitate a true “marketplace of ideas.” Viewpoint diversity and individuality are not applauded but are instead suppressed. This is because social media is funded through trading personal data to advertisers. Algorithms seek to categorize opinions and that personal data is then used, as Richard Sorabji explicates, “To target [users] with information, or disinformation, tailored as persuasive to [their] different personalities,” which is centered on information that they already agree with.¹¹ Ideas are reaffirmed rather than challenged, and subsequently online groups develop a perceived superiority of their own ideas, believing them to have objective truth. This leads to increased moralism and hostility when such groups are exposed to those who do not share their view.

Such a problem is further exacerbated by reasoned debate being substituted with emotional rhetoric online. This is indicative in the dilution of moral terms, most prominently seen in the frequent use of words like “Nazi”—a manifestation of Godwin’s law (i.e., as an online argument progresses, the chances of the Nazis being mentioned increases exponentially). Again, the problem of emotivism affects free speech on social media to a great extent because moral values lose their real meaning, instead turning into hostile attitudes towards anyone with a different opinion. Consequently, social media collapses into a marketplace of intellectual thuggery rather than ideas. This further suggests how Mill’s account is unsustainable in the digital age.

5 @Pyschofilmcritic, “I like most Tarantino...,” Instagram photo, April 19, 2021, <https://www.instagram.com/p/CN0zHM4ly0-/>.

6 @Trashcanpaul, “I would like to throw...,” Instagram photo, August 11, 2021, <https://www.instagram.com/p/CVvktKCPzpW/>.

7 @Sidetalknyc, “Joe Biden, if I...,” Instagram photo, March 26, 2021, https://www.instagram.com/p/CM495p_DnBP/.

8 Adam D. Moore, “Free Speech, Privacy, and Autonomy,” *Social Philosophy & Policy* 37, no. 2 (2020): 36, 10.1017/S0265052521000030.

9 Piers Norris Turner, “Introduction: Updating Mill on Free Speech,” *Utilitas* 33, no. 2 (2021): 126, 10.1017/s0953820821000029.

10 Moore, “Free Speech,” 49.

11 Richard Sorabji, “Free Speech on Social Media: How to protect our freedoms from social media that are funded by trade in our personal data,” *Social Philosophy & Policy* 37, no. 2 (2020): 209, 10.1017/s0265052521000121.



V. TWO OBJECTIONS

Before moving on to the third challenge, it is worth considering a potential objection to the first two challenges. This objection mitigates the two challenges by recontextualizing them, viewing the emotional and algorithmic character of social media not as problems, but as features that have net positive consequences for free speech. Given that Mill's argument for free speech is consequentialist, if this objection demonstrates social media's net positive influence on public discourse, then Mill's account may still be sustainable in the age of social media.

On the emotional character of online dialogue, this objection attacks the first challenge, suggesting that the lack of distinction between emotion and fact does not cause social media to drift away from Mill's account. Instead, emotional content can be used to amplify arguments with a truth-value—so it is compatible with Mill's principles. This is because the popularity contest of social media allows for emotional content to come into the mainstream. Accordingly, truth-value propositions are amplified by the emotional nature of their dialogue and find a public audience. Using emotional content to engage with people has a net benefit, as more ideas and information enter the public conversation. Given that Mill's argument for free speech is consequentialist, social media's clear benefit in promoting discussion suggests that Mill's account may still be vindicated.

A central tenet of Mill's argument is that free speech ensures that minority voices enter the public domain. The recent work of Jack Monroe is indicative of social media's ability to achieve this outcome. In a series of tweets, Monroe highlighted the rapid increase of the price of basic goods in UK supermarkets and how this disproportionately affects the poorest members of society. For example, Monroe tweeted that the price of canned spaghetti was once "13p, then 35p,...a price increase of 169%."¹² This is a factual statement, yet, in previous eras, it may not have been heard because it reflects the experiences of a minority group. Yet on social media, the emotional nature of her argument, that supermarkets contribute to the UK's cost-of-living crisis, meant that it became amplified. The positive effect of this amplification is demonstrated by the consequences of her tweets, with major supermarket Asda reducing the prices of their basic items in response. Under Mill's account, this consequence is beneficial because truth-value propositions from minority and marginalized backgrounds, that would previously not enter the public consciousness, can now do so through social media. Thus, if emotional content amplifies what Mill would consider valuable content in a way

that would not be possible without social media, then Mill's account remains sustainable in the digital age.

On the algorithmic character of social media, the objection rejects the second challenge, arguing that algorithms increase civic engagement and discourse. If there is such an increase, then Mill's account may remain sustainable. Such an objection is centered on social media's ability to provide targeted information, through algorithms, that users find appealing. All content is being seen more, yet emotional content is far more accessible than factual content. This is because emotional content is framed in an approachable manner that allows for a basic understanding of an issue. In short, emotional content creates interest.

Emotional connection, used as a marketing tool, can better lead someone to engage with public discourse. At least in this way, people are engaging with the emotional aspect of discourse, rather than not engaging at all. For example, climate change activists often share shocking images, from wildfires to meat production, which generally promotes increased dialogue. Mill's consequentialist reasoning exemplifies a collective duty to engage in constructive dialogue. Algorithms play a crucial role in this call to duty with their ability to categorize opinions and target advertising, offering new opinions and building upon ones that are beginning to form. Algorithms, being compatible with Mill's consequentialist principles, may ensure that social media has an educative benefit and support the classical liberal conception of free speech by reaffirming the emphasis of intellectual development.

This objection suggests that social media's emotional and algorithmic character results in more truth-value propositions entering public discourse and an increase in civic engagement with debate. From a consequentialist perspective, the objection might indicate that the positive effects of free speech on social media outweigh the negatives. Thus, Mill's account of free speech appears justified in the modern age.

VI. COUNTERARGUMENT TO THE OBJECTION

However, there are counterarguments that diffuse this objection. Even if the emotional character of free speech on social media leads to positive consequences for public discourse, this does not necessarily mean that these consequences could not be maintained if more restrictions were placed on speech. If social media is censored and its merits, of championing minority voices and increasing engagement, remain but its drawbacks are reduced, then a censored account of speech is preferable on consequentialist terms. This would reinforce the point that Mill's account is unnecessary in the digital age and that a new conception of free speech would be more favorable. There is no reason to

¹² Jack Monroe, @BootstrapCook, "Canned spaghetti," Twitter post, January 19, 2022. <https://twitter.com/BootstrapCook/status/1483778782964166662>.



accept the negative effects that come with free speech. For every Jack Monroe there are hundreds of people producing emotional content without truth-value, distracting people from the factual. If social media were to distinguish emotional truth-value statements from emotional non-truth-value statements allowing the former to take precedence and reducing the impact of the latter, then this suggests that a form of censorship, which Mill would not approve of, is beneficial.

A similar counterargument can be raised to the consequences of social media's algorithmic nature. While it is true that algorithms can increase engagement, which is valuable, this engagement is rarely at the level to which Mill aims. Algorithms create fevered debate, which conflicts with Mill's request for calm discussion. For Mill, tranquility within conversation is a necessary means of achieving the educative and democratic aims of public discourse; hence, he supports "giving merited honour to everyone...who has calmness."¹³ Yet, this concern for calmness and constructive dialogue is given much less significance by social media algorithms. The ease by which abuse slips into online dialogue highlights how algorithms utilize emotion not to engage people into new debates, but to segment audiences and pit them against each other. With arguments on social media turning into outbursts of emotion and insults, it is clear to see how digital conversation has been distanced from the Darwinian test Mill hoped for. Accordingly, Mill's belief that calmness leads to a positive outcome in discourse is severely challenged by social media's algorithmic character.

The reason why algorithms create frenzied discourse is that their intended consequences are fundamentally different from Mill's. Social media uses algorithms not for the purpose of engaging people with a wide spectrum of views, but to segment audiences for the purpose of profit. This leads to social media selling the data of individuals to companies and political organizations in order to maximize their revenue.¹⁴ If monetization is prioritized over balance, this demonstrates how the algorithmic character of social media uses speech to produce different consequences than Mill's educative aims. This poses an insurmountable challenge to Mill's account. If algorithms were to prioritize exposing different perspectives of debate to audiences, and limit the perspectives that are so often reinforced, this would prevent the negative impact of segmentation that comes in the form of moralistic conflict between opposing groups. This would create a "marketplace of ideas" more in Mill's image but would involve some element of censorship. Therefore, Mill's account remains unsustainable.

¹³ Mill, *On Liberty*, 54

¹⁴ Sorabji, "Free Speech Social Media," 237.

VII. THE THIRD CHALLENGE

Not only do these first two challenges show that Mill's account fails, but the third and final challenge posed by social media is perhaps the most damaging. This problem is that social media promotes anonymity as a surrogate for autonomy. Autonomy is central to Mill's argument for free speech. Mill argues that people only flourish when their actions are not mandated. Mill justifies this through utilitarianism, whereby actions are judged by the extent to which they maximize happiness or flourishing. His account of autonomy is based on the premise that freedom results in enriched flourishing. Mill uses this to justify his harm principle that, "the only purpose for which power can be rightfully exercised over any member...against [their] will is to prevent harm to others."¹⁵ This principle, and the central role that autonomy plays within it, is critical to Mill's defense of free speech.

Following the development of online platforms, anonymity is now seen as a necessary factor in ensuring autonomy. This was deemed a more democratic account because it meant that only opinions were being judged, not the person. To put it simply, such reasoning believes anonymity to be a precondition for autonomy. The ability to be anonymous exists as a selling point for social media. As Robert C. Post further illustrates, "The possibility that your digital character has more truth than your reality" suggests that social media allows for greater self-knowledge.¹⁶ These considerations mirror Joel Feinberg's four concepts of autonomy, one of them being the "the *sovereign authority* to govern oneself...within one's own moral 'boundaries.'"¹⁷ Maintaining that self-actualization must be an independent process falls nicely under the autonomy-building aspect of free speech and expression on social media. Instead of being displaced from government authorities, individuals are displaced from themselves and are given instead a new identity to discover ideas. On the surface, this seems to reinforce Mill's account. For Mill, free speech enables the flourishing of ideas which in turn leads to the flourishing of individuals' sense of self. As such, it is tempting to view anonymity as reinforcing Mill's consequentialist belief that speech has eudaimonic ends in terms of increasing self-fulfillment.

However, autonomy arguments lack significance in light of content pollution. This is because quality and relevance have no meaning when jeopardized by trivial and false personas. While most speech serves a

¹⁵ Mill, *On Liberty*, 13.

¹⁶ Robert C. Post, "Privacy, Speech, and the Digital Imagination," in *Free Speech in the Digital Age*, ed. Susan J. Brissson and Katharine Gelber (Oxford: Oxford University Press, 2019), 108, 10.1093/oso/9780190883591.001.0001.

¹⁷ Joel Feinberg, "Autonomy, Sovereignty, and Privacy: Moral Ideals in the Constitution," *The Notre Dame Law Review* 58, no. 3 (1983): 447, <https://scholarship.law.nd.edu/ndlr/vol58/iss3/1>.



function of either autonomy-building or autonomy-protecting, it is also true that many expressions are almost entirely fungible, destructive, or pointless. Stanley Fish maintains that “speech, in short, is never a value in and of itself but is always produced within the precincts of some assumed conception of the good.”¹⁸ Thus, speech is doing something with words; it is an action with purpose. The fact that anonymized content on social media is not end-seeking seems incompatible with Mill’s account of free speech.

This leads to the fundamental problem that Mill’s account cannot address in the digital age. Free speech does not mean consequence-free speech. Mill did not directly speak of consequences because for him the real consequence was whether an opinion survived the test of the marketplace. So, Mill’s argument of autonomy cannot withstand this problem. Speech seems inherently connected to an individual, but the ability to distance an individual from what they say to the point that it has no effect on them poses a grave challenge to free speech. In essence, anonymity neglects the act of speech, instead offering merely trivial words. Social media removes the good that speech should tend towards and accordingly changes the value of speech. Therefore, Mill’s definition of free speech cannot endure this problem and is unsustainable in the digital age

VIII. CONCLUSION

The technological effects on communication are echoed in the Phaedrus, in which Socrates cites the story of ancient Egyptian king Thamus’s reaction to Theuth’s invention of letters. Thamus fears that written communications will have a negative impact. Neil Postman urges us to take to heart the cautionary tale in this legend but offers a corrective to Thamus’s judgment. Postman observes that “technology is both a burden and a blessing, not either-or, but this-and-that.”¹⁹ It is clear that the medium of social media has influenced speech, allowing emotive discussion to take precedence over an exchange of ideas and allowing moralism, derived from algorithms, to create a false sense of objectivity. Mill’s classical liberal model cannot endure these challenges. However, it is the final challenge of anonymity that poses the greatest problem because it changes the fundamental nature of speech, and Mill’s account cannot withstand that issue. Speech has always been an act tending towards an end, so it cannot be helped but to believe that the possible demise of classical, liberal free speech is not a fatal flaw of the digital age. Instead, it is an unintended consequence of technological development that will have to be overcome just as any other.

¹⁸ Stanley Fish, *There’s No Such Thing as Free Speech...and it’s a good thing too* (Oxford: Oxford University Press, 1995), 104.

¹⁹ Neil Postman, *Technopoly: The Surrender of Culture to Technology* (New York: Vintage Books, 2010), 4-5.



Nevin Chellappah is looking forward to reading philosophy at the University of Durham in September. His primary interests lie in business ethics—specifically the potential intersection between environmental, social, and corporate governance and virtue ethics—philosophy of language, and aesthetics. Currently on his gap year, Nevin spends his time reading some philosophy and even better, trying to understand it.

