Journal of Applied Ethics and Philosophy

Printed in Japan

ISSN 1884 0590 (Online)

All queries should be directed to:
The Editor-in-Chief, Center for Applied Ethics and Philosophy
Faculty of Humanities and Human Sciences
Hokkaido University
N10 W7, Kita-ku, Sapporo 060-0810, Japan
jaep@let.hokudai.ac.jp

CONTENTS

Editorial Note

The *Journal of Applied Ethics and Philosophy* is an interdisciplinary journal that covers a wide range of areas in applied ethics and philosophy. It is the official journal of the Center for Applied Ethics and Philosophy (CAEP) at Hokkaido University. The aim of the *Journal of Applied Ethics and Philosophy* is to contribute to a better understanding of ethical and philosophical issues by promoting research in various areas of applied ethics and philosophy, and by providing researchers, scholars and students with a forum for dialogue and discussion on ethical and philosophical issues raised in contemporary society. The journal welcomes original and unpublished regular articles and discussion papers on issues in applied ethics and philosophy.

Nobuo Kurata
Editor-in-Chief

Many thanks to *Journal of Applied Ethics and Philosophy* reviewers

The *Journal of Applied Ethics and Philosophy* would like to thank the following individuals for generously reviewing manuscripts for us between June 2025 and January 2026. The support and expertise of these professionals promote and maintain the high quality of the journal's content. Thank you very much.

Martijn Boot	Taro Okuda	Tetsuji Iseda
Keiichiro Yamamoto	Kengo Miyazono	Mariko Nakano-Okuno

Nozick's Entitlement Thesis, Political Corruption and Distributive Injustice in Post-colonial African States

Olusola V. Olanipekun & Michael O. Aina
Department of Philosophy, Obafemi Awolowo University

Abstract

With corruption and injustice plaguing contemporary societies like incurable diseases, this paper examines the conjoined problems of political corruption and injustice in Post-colonial African States through the philosophical lens of Robert Nozick's entitlement theory of distributive justice. According to Nozick, holdings are just if they arise from just acquisition, just transfer and the rectification of the past injustices. However, in many African States, political corruption which manifests in different ways violates conditions for just acquisition and transfer. Consequently, all forms of unjust acquisitions invalidate the moral legitimacy of the current properties and wealth distributions among politicians and public office holders. This paper argues that applying Nozick's principles to the issue of political corruption in Africa challenges the assumption that current properties and wealth distributions among politicians and political office holders can be defended without addressing the deeply embedded foundational injustices. Furthermore, the paper examines the strengths and the limits of Nozick's entitlement thesis. It argues that African anti-corruption measures and corrective justice are not only compatible with Nozickian framework, but are also necessary to restore legitimate entitlements. Finally, it concludes that Nozick's entitlement theory offers a valuable framework for addressing ongoing social injustices by promoting fair principles that ultimately contribute to a more just and well-ordered society.

Keywords: Distributive (In)justice, Entitlement Thesis, Political Corruption, Rectification, Post-colonial African States.

Introduction

One of the perennial problems undermining good governance, economic development, and social cohesion in post-colonial African states is political corruption. According to Kwame Gyekye, "Political corruption could be defined as illegal, unethical and unauthorized exploitation of one's political or official position for personal gain or advantage." (Gyekye, 1997:193). Also, Inge Amundsen identifies political corruption as an unethical behaviour which violates the norms of the system of political order (Amundsen, 2006:3). Meanwhile, Transparency International (2022) broadly defines corruption as the abuse of entrusted power for private gain. While Transparency International defines corruption generally as "the abuse of entrusted power for private gain," this paper focuses explicitly on political corruption, which refers to the unjust acquisition or retention of political power and public resources by state officials through coercion, bribery, electoral manipulation, embezzlement, or patronage networks. This distinction is thoughtfully important because, within Nozick's entitlement theory, corruption becomes a problem of justice not merely because it is immoral, but because it violates the principles of just acquisition and just transfer. For instance, when political actors buy votes, divert state funds, or manipulate electoral processes to obtain offices they were not freely granted by citizens, they engage in an unjust acquisition of political power. Similarly, embezzlement of public revenue constitutes an unjust transfer of communal resources into private hands. As a result, political corruption is interpreted here not simply as inefficiency or moral decay, but as a systematic violation of entitlement norms, making it subject to Nozick's rectification principle (Nozick, 1974, pp. 152-153).

Just like in different parts of the world, corruption in Africa manifests in forms of embezzlement of public funds, bribery, electoral manipulation, nepotism, and so on. Scholars such as Jean Francois Bayart and Jean

Francois Madard have connected the persistence of corruption to historical legacies of colonial rule, weak institutional frameworks, and the centralization of power within post-independence governments (Bayart, 1993; Médard, 2002). Despite various anti-corruption initiatives, the problem remains deeply entrenched, raising critical questions about justice, legitimacy, and entitlement to political and economic resources.

While many existing literatures examine corruption through economic perspective (Mauro, 1995), institutional perspective (North, 1990), or cultural perspective (Smith, 2007), this paper, on the other hand, examines corruption within the purview of political philosophy. Even though, scholars such as Kwame Gyekye (1997), and Olanipekun Olusola (2021) also examined corruption within the purview of moral and political philosophy, this paper deviated from them by specifically seeking to apply Robert Nozick's Entitlement Theory of distributive Justice to understand corruption as a violation of rightful acquisition and transfer. *In Anarchy, State, and Utopia* (1974), Nozick posits that holdings are just if acquired and transferred according to principles of justice, and all injustices in acquisitions or transfers require rectification. By interpreting political corruption as "unjust acquisition", the paper aims to show how corrupt practices distort legitimate entitlements and thereby perpetuate broader patterns of injustice in African societies. This paper, therefore, critically examines both the strengths and limits of adopting Nozick's framework in explaining African realities.

Focusing on Nozick's entitlement thesis, the paper emphasizes the three main components of the theory, namely, justice in acquisition, justice in transfer, and rectification of injustice. Outlining how corrupt acts violate Nozickian principles, the paper treats political corruption as a form of unjust acquisition. The paper as well draws practical example from Nigeria. This is done with the mind of illustrating the practical dynamics of political corruption and its distributive consequences. It discusses efforts at corrective justice, drawing from Nozick's rectification principle and real-world anti-corruption initiatives. Finally, it concludes by critically reflecting on the main implications of Nozick's theory for governance in Africa.

Nozick's Entitlement Theory of Justice

One of the most prominent scholars who contributed immensely to the issue of distributive justice after John Rawls is Robert Nozick. Nozick's entitlement theory of justice presents a historical rights-based account of distributive justice. A question may be asked that "Why Nozick despite several alternative accounts?". This is necessary because justification is required for choosing Nozick's entitlement theory rather than more egalitarian models of justice such as global egalitarianism, sufficientarianism, luck egalitarianism, or republican non-domination. Egalitarian model generally begins with the question of how goods should be distributed, whereas the Nozickian approach begins with how goods were acquired and whether those acquisitions respect individual rights or the principles of distribution. (Nozick, 1974, pp.149-182). In many African countries, corruption manifests in different ways such as bribery, fraudulent privatization, embezzlement of public funds, unjust land seizures and so on. These are not failures of the principles of distributive but failures of just acquisition and transfer. Consequently, Nozick's historical and rights-based conceptions of justice is more appropriate for analyzing African corruption than patterned theories. This does not mean egalitarian theories are irrelevant, rather, they address different moral questions. The central claim of this paper is that before asking how resources should be redistributed, we must first ask whether current holders of wealth and power are entitled to them in the first place.

By extension, contrary to John Rawls's theory of justice as fairness, Nozick argues that justice is fundamentally historical and concerns the legitimacy of acquisitions and transfers rather than the final pattern of distribution. This distinction is crucial especially in relation to the African context. Rawls's account defends the existence of well-structured and collectively owned assets within a functioning liberal democratic society. However, these conditions are strange to post-colonial African states that were just emerging from colonial rule, coupled with the problems of new political authority and still in the process of reclaiming land and resources.

Now, contrary to Rawlsian theory which emphasizes a distribution of tangible assets among citizens, Nozick's theory is best suited for historical and socio-political realities of post-colonial African nations. As a matter of fact, while many African countries were in the process of acquiring or reclaiming assets upon gaining independence, Nozick's view on entitlement and the rights of individuals to their holdings provides a more appropriate lens. It is through this lens that one could examine how domestic populations in post-colonial African states began to assert their claims to resources in the face of both internal and external challenges. Meanwhile, Rawls's distributive principle requires already established assets for equitable distribution.

In line with Nozick's entitlement theory, inequality becomes a matter of justice only when it results from violations of the principles of just acquisition, just transfer, or rectification (Nozick, 1974, pp. 150-153). Therefore, this paper is not advocating a patterned or egalitarian view of justice. Instead, it argues that much of the inequality seen in African societies is the product of historical injustices such as colonial expropriation and contemporary political corruption, where state officials acquire resources through coercion, fraud, or theft. These forms of inequality are unjust not because they deviate from a preferred social pattern, but because they arise from unjust acquisition and non-consensual transfer and these are direct violation of Nozickian principles.

The first component of Nozick's entitlement theory of distributive justice is the principle of 'Justice in Acquisition'. This component addresses the original acquisition of unowned resources. Resting on John Locke's theory of property rights, Nozick proposes that individuals are entitled to holdings if they acquire them without violating the rights of others. Although, Nozick adapts Locke's notion of "mixing labor" with resources, he notably modifies the Lockean proviso−the condition that appropriation must leave "enough and as good" for others. He interprets the proviso in a way that allows significant inequalities, provided that the acquisition does not worsen the position of others relative to a no-ownership baseline (Nozick,1974:177-178).

The second component of Nozick's theory is 'Justice in Transfer'. This component concerns with the voluntary exchange or transfer of holdings. According to Nozick, as long as transactions are free from force, fraud, or coercion, any transfer of holdings is just. Individuals are entitled to give, sell, or donate property as they wish. There is no need for a central authority to redistribute holdings according to a favored pattern or end-state principle. The emphasis here is on consent and voluntariness. He rejects any form of patterned redistribution because it necessarily involves continuous interference in individuals' lives- violating their rights to liberty and property. For instance, he famously critiques redistribution as requiring "continuous redistribution to maintain a pattern" (Nozick, 1974:163). However, the conditions of voluntariness are often systematically undermined in a corrupt political environment. In other words, when public resources are captured through coercion or bribery, transfers cannot be considered just according to Nozickian standards.

The third component of Nozick's entitlement theory of distributive justice is the 'Rectification of Injustice'. This component focuses on how to respond to past injustices. If holdings have been acquired or transferred unjustly, through theft, fraud, coercion, or other rights violations− justice requires rectification. However, he provides only a broad sketch rather than a comprehensive formula for rectification. This suggests that distributive principles must depend on the historical details of the injustices involved (Nozick, 1974:153). Also, rectification may involve returning holdings to their rightful owners, compensating victims, or otherwise adjusting holdings to approximate what would have occurred under just conditions. Nozick acknowledges that historical injustices can dramatically distort current holdings, but he also recognizes the practical difficulties of tracing exact injustices and calculating precise remedies. This principle is so significant for African states, where histories of colonization, land dispossession, and systemic corruption have profoundly shaped current distributions of wealth and political power.

Despite Nozick's brilliant philosophical presentation and analysis, his entitlement theory has attracted several criticisms. For instance, Jeremy Waldron challenges his inadequate account of the original acquisition in addressing historical injustices. According to Waldron, societies emerging from conquest, colonization, or systemic theft cannot easily assume that current holdings are legitimate. (Waldron, 1992:4). In Africa, colonial powers often seize lands and resources without consent and any claim to "just acquisition" is always problematic. Also, Nozick's excessive dependence on voluntary transfers suggests relatively equal bargaining conditions and this is rare in non-liberal societies. Critics such as Gerald A. Cohen argues that structural inequalities, including poverty, illiteracy, and lack of political power, can distort so-called "voluntary" agreements, rendering many transfers unjust in practice (Cohen, 1995). Similarly, Nozick's approach could be charged of underestimating the role of social structures and public goods. Nozick's individualistic framework pays little attention to the fact that rights and holdings are sustained within broader structures. But the implication of the above is that, in states where institutions and political structures are weak, there will be no clear distinction between just and unjust transfers.

Admittedly, while several well-known objections have been raised against Nozick's entitlement theory, it is important to know that these criticisms do not entirely invalidate the core Nozickian commitments used in this paper for several reasons. First, Nozick's theory does not deny that inequality can be morally troubling; rather, it holds that inequality is unjust only if it results from violations of just acquisition or transfer. This strength allows the theory to address corruption by enhancing just redistribution. Second, in response to Cohen's criticism that self-ownership leads to exploitation, Nozick argues that voluntary exchanges and respect for bodily autonomy are morally superior to forced redistribution. (Nozick, 1974, p.174). Third, while Rawlsians argue that Nozick lacks a principle for defending the worst-off, Nozick's rectification principle explicitly requires that past injustices be corrected before holdings can be morally legitimate (Nozick, 1974, p.152). What the above view suggests is that Nozick's theory is considered to be philosophically capable of addressing contemporary political corruption and historical injustice in Africa when given proper interpretation through the lens of rectification and entitlement.

Political Corruption as an Unjust Acquisition

Susan Rose-Ackerman conceives political corruption as the violation of the public trust, where officials prioritize personal or group interests over the collective good (Rose-Ackerman, 1999). Given our general understanding of the concept, it is no longer a news that political corruption encompasses a wide range of activities and appears in different forms in different societies. As a matter of fact, political corruption is a universal phenomenon. However, in the African context, J.F. Bayart argues that corruption often assumes systemic dimensions, becoming embedded within political institutions and social networks (Bayart, 1993). In Bayart's view, the implication of corruption is that corrupt practices often form part of a broader "politics of the belly," where informal distribution of the state resources is done within a few sets of people for personal benefits.

For this paper, political corruption could also be interpreted as a violation of property rights and entitlements. The reason is because, public officials who divert state assets for personal benefit are engaging in the illegitimate acquisition or transfer of holdings that rightfully belong to the citizenry (Nozick, 1974). Thus, from a Nozickian perspective, political corruption constitutes a fundamental breach of the principles of justice in acquisition and transfer. For Nozick, individuals are entitled to holdings only if these holdings are acquired justly and transferred voluntarily (Nozick, 1974:150-153).

Arising from Nozickian litmus test, political corruption disrupts distributive justice in two major ways: First, corrupt acquisitions violate the principle of justice in acquisition. State resources- revenues, land, and public contracts, are held on behalf of the citizenry. When officials appropriate such resources for personal gain, they acquire holdings unjustly, without rightful entitlement (Nozick, 1974:171). Second, corruption undermines justice in transfer. In many corrupt transactions, such as bribery or electoral manipulation, consent is compromised. Citizens and firms often feel coerced into participation, distorting the voluntary nature of exchanges (Rose-Ackerman, 1999:7). Transparency International reports about some African states reveals that elections are tainted by fraud or vote-buying, and also violate the principle of voluntary political consent (Transparency International, 2022). This is a clear indication that political corruption in Africa involves both unjust acquisition and unjust transfer. Moreover, corruption initiates a chain of injustice, distorting not only the original transaction but subsequent economic and political interactions, a dynamic Nozick recognizes as critically undermining just distributions (Nozick, 1974:152).

Furthermore, the moral drawn from the above view is that voluntariness is a key condition for legitimate transfer in Nozick's Entitlement Theory. However, in a corrupt system, voluntariness is systematically undermined due to bribery, unfair competitions and other factors. Also, electoral corruption disrupts the voluntary transfer of political authority. Elections manipulated through fraud, intimidation, or vote-buying negate the idea that governance is based on genuine citizen consent (Bayart, 1993:233). Consequently, in African context, political offices are held unjustly on many occasions just like property is held unjustly in Nozick's approach due to unjust distributions that divorce holdings from rightful entitlement.

Corruption and Distributive Injustice in African States

One of the main claims around the history and development of political corruption and distributive injustice in post-colonial African states is that they are closely tied or linked to colonialism. As a matter of fact, Walter Rodney clearly stated that during the colonial era, European powers seized land, labor, and resources through violent and coercive means, establishing deeply unjust initial holdings (Rodney, 1972). Combining Nozickian perspective with Rodney's claim, this initial acquisition (i.e, European powers seized land, labor, and resources through violent and coercive means) was fundamentally illegitimate. Why? The main reason is because it involved force, fraud, denial of consent and devoid of voluntariness.

For instance, in Nigeria, Toyin Falola reveals how British colonial policies restructured indigenous land tenure systems to enable the extraction of agricultural products like palm oil and cocoa (Falola, 2009). By implication, land became concentrated in the hands of colonial agents and local elites favored by the British, setting the foundation for post-colonial inequalities. Similarly, Caroline Elkins also points out that in Kenya, British settlers seized large tracts of fertile land, particularly in the "White Highlands", displacing indigenous communities through legal and military force (Elkins, 2005). The above view suggests that unjust acquisition renders many African communities landless till today. The same unjust acquisition manifested in South Africa as emphasized by A.J. Christopher. Christopher reveals that the Natives Land Act of 1913 systematically stripped black South Africans of land rights, concentrating ownership among the white minority (Christopher, 1994). The above view is the aftermath of Apartheid experience which further created a highly unequal distribution of wealth and opportunity among the native South Africans.

Admittedly, by making reference to Nigeria, South Africa and Kenya, we do not claim that these three countries represent all African states uniformly because we do not erase significant differences in historical experience, institutional capacity, and legal frameworks. However, they were chosen due to their political relevance and postcolonial trajectories. They mainly function as illustrative case studies that illuminate how unjust acquisition and political corruption can be interpreted through Nozick's entitlement theory.

Following Nozick's theory, the initial injustices arising from unjust transfer in African societies require rectification. However, post-independence African states largely failed to address these colonial wrongs, and to make the matter worse, political corruption often perpetuated or worsened them due to the failure and insensitivity of some post-independent African leaders.

Shortly after independence, many African states experienced a transformation due to changes of power from the hands of the colonizers to the hands of the colonised. Consequently, many African elites became involved in practices that are clear violation of Nozick's principles of justice in transfer. They frequently engaged in corruption, misallocating resources for personal enrichment, manipulating elections, and engaging in nepotism.

As a matter of fact, Nigeria, Africa's largest economy, has been plagued by systemic corruption, particularly related to its vast oil revenues. Successive governments have diverted oil wealth into private hands through inflated contracts, "ghost workers," and outright embezzlement (Smith, 2007). Institutions like the Nigerian National Petroleum Corporation (NNPC) have been at the center of massive scandals involving billions of dollars in missing revenues (Transparency International, 2022).

Furthermore, Hakeem Onapajo, opines that electoral corruption remains rife in Africa. For vote-buying, ballot-box stuffing, and judicial manipulation are recurrent problems (Onapajo, 2014). By implication, political authority is often secured not through voluntary, legitimate consent but through fraud and coercion, undermining justice in transfer in many African countries.

Arising from the above, it is evident that corruption's impact on distribution is so profound. Beyond distorting acquisition and transfer mechanisms, corruption engenders inequality, undermines meritocracy, and erodes trust in institutions. In Nigeria, for instance, despite decades of oil wealth, poverty rates remain alarmingly high in the country. In fact, the World Bank (2020) estimation shows that around 40% of Nigerians live below the poverty line. Whereas, little fraction of the political officeholders controls vast wealth that is largely acquired through corrupt means. Interestingly, this is a clear violation of the conditions for just acquisition and transfer from Nozickian point of view.

Just like in Nigeria, in Kenya, Roger Southhall pointed out that land ownership remains highly unequal such as a direct result of colonial seizures compounded by post-independence corruption. Political elites have used their influence to acquire land, often displacing marginalized communities (Southall, 2005). Similarly, in South Africa, while political liberation was achieved, economic liberation remains elusive. The coefficient- a measure of inequality remains one of the highest in the world (World Bank, 2018).

Relying on Nozick's theory, holdings acquired through unjust means or unjust transfers require rectification. However, in some post-independent African states, corruption has deepened distributive injustices rather than correcting historical injustices. By implication, the persistence of systemic corruption, coupled with the failure to rectify colonial wrongs, breed injustice and leaves African societies far from the Nozick's analysis of a just distribution. By just distribution, Nozick mean all form of distributions that are rooted in legitimate acquisition and voluntary transfer.

Corrective Justice and Anti-Corruption Efforts

Admittedly, Nozick acknowledges that past injustices (such as theft, fraud, or coercion) vitiate the legitimacy of current holdings (Nozick, 1974:152). Now, how do we address this problem historical injustice? Nozick proposes a 'Rectification Principle'. As hinted earlier, this principle states that if past injustices have occurred, steps must be taken to rectify them to restore justice in holdings. However, it is obvious that Nozick was silent about providing a detailed formula for rectification. This is not unconnected to the complexity of historical injustices. Interestingly, despite the omission, Nozick suggests that an appropriate rectification would depend on the specific historical circumstances, including the magnitude of the injustice and the degree to which it has distorted present holdings (Nozick, 1974:231). This could be applied to Postcolonial African states by focusing how anti-corruption efforts must not only punish wrongdoers, but must also restore unjust acquisitions and wrong distributions of public resources. This will not only repair the systemic distortions in political and economic systems, it will also promote fair principles that ultimately contribute to a more just and well-ordered society.

However, we are not ignorant of the fact that the practical implementation of corrective justice is not free from certain significant challenges. In a society like Africa where corruption is systemic rather than episodic, it should be noted that simply punishing a corrupt public office holder or sentencing them to certain number of years of imprisonment may be insufficient. The reason is because doing so does not amount to rectification. A full Nozickian rectification would require, first, identifying victims and perpetrators, second, restoring stolen or misappropriated assets, and lastly, rebuilding institutions to ensure future voluntary and fair transfers. However, this is not always the case.

Interestingly, some Post-colonial African states have attempted partial rectification through institutional reforms, asset recovery initiatives, and the establishment of anti-corruption commissions. However, the effectiveness of these mechanisms is mixed. For example, Nigeria established the Economic and Financial Crimes Commission (EFCC) in 2003 to investigate and prosecute financial crimes, including corruption. The EFCC has been involved in recovering significant sums of stolen public funds. For example, between 2015 and 2020, the EFCC reportedly recovered over $3 billion in assets (EFCC, 2020). High-profile convictions, such as those of former state governors and politicians, suggest some progress. However, critics argue that the EFCC is vulnerable to political manipulation and selective enforcement, undermining its legitimacy (Amundsen, 2010). From a Nozickian standpoint, the EFCC's efforts represent partial attempts at rectification. Yet the deep structural distortions caused by decades of corruption, such as entrenched inequality and weakened institutions, are far harder to remedy than individual asset recovery.

Despite these efforts, significant obstacles to full corrective justice exist. As Nozick himself notes, rectifying historical injustices is extraordinarily complex (Nozick, 1974:231). In Africa, corruption is layered upon colonial expropriation, making clear identification of victims and rightful owners difficult. Secondly, corruption often becomes embedded within the political economy itself, creating vested interests resistant to reform (Bayart, 1993). To avoid merely asserting that contemporary corruption is "layered upon colonial expropriation," it is important to briefly indicate the mechanism linking past dispossession to present injustice. Colonial administrations in states such as Kenya and Nigeria centralized land and mineral rights under the authority of the state, displacing traditional ownership structures (Mamdani, 1996, pp. 62-67). After independence, many postcolonial governments retained these centralized powers, enabling political elites to appropriate land, oil revenues, and public resources in ways that mirror colonial extractive practices. Thus, present-day corruption does not simply follow colonialism chronologically; it often relies on the same institutional mechanisms of resource control, making Nozick's rectificatory justice directly relevant. Thirdly, effective rectification requires strong, independent institutions, a condition lacking in many African states (Bratton & van de Walle, 1997). Finally, without sustained political will, anti-corruption bodies often lack the power to pursue elites or implement broader structural reforms (Hope, 2016).

However, this paper does not reduce contemporary corruption to colonialism alone by exonerating present political leaders of responsibility. While colonial structures created initial conditions of unjust acquisition, Nozick's theory insists that moral responsibility lies with current holders of power who continue to violate principles of just transfer and acquisition. Thus, both historical and contemporary agents are accountable: colonialism explains how entitlement was first violated, but current political actors perpetuate injustice through embezzlement, electoral fraud, and resource misappropriation. Nozick's rectification principle allows us to distinguish between inherited injustice and actively committed injustice. Thus, while Nozick's rectification principle offers a moral and philosophical justification for anti-corruption efforts.

Nozick's Entitlement Theory and the Challenge of Ongoing Injustice in Africa

Admittedly, it may be a daunting task to fully rectify historical injustices such as colonization, slavery, and the likes, but this does not imply that all forms of injustice are irreparable. In fact, such injustices that continue to undermine justice and equity across many African nations today still need to be confronted and rectified. For instance, land grabbing remains a major issue in countries like Ethiopia (Human Rights Watch, 2011) and Mozambique (Milgroom & Spoor, 2013), where vast tracts of fertile land have been leased or sold to foreign investors at the expense of local farmers and indigenous communities. The negotiation is often done without the consent of the rightful owners and there may be no provision for compensation. Thus, unjust transfer of this nature undermines Nozick's principle of justice in transfer because they are products of manipulation and rather than voluntary, informed agreement.

Basically, adapting Nozick's entitlement thesis to African context involves considering the historical and cultural factors that shape entitlements and recognizing the need for rectification of past injustices. For instance, policies aimed at redistributing land or resources could be informed by Nozick's principles, ensuring that such redistribution is based on fair and just principles.

Interestingly, adapting Nozick's theory requires balancing individual and communal rights. This can be achieved by recognizing the importance of communal interests while also protecting individual liberties. For example, community-based natural resource management refers to systems in which local communities participate in the ownership, decision-making, and supervision of resources such as land, forests, minerals, and water. The purpose is to balance individual benefits—such as wages, royalties, small-scale business opportunities, or direct income from resource extraction—with communal interests, which include environmental sustainability, equal access, and the welfare of both current and future members of the community.

While Nozick emphasizes individual rights and self-ownership, many traditional African moral systems- such as *Ubuntu ethics* (Molefe, 2024) in South Africa, as well as *Omoluabi ethics* (Olanipekun, 2017), among the Yorubas in Nigeria- stress communal responsibility and relational personhood. These views are not entirely incompatible. Nozick's entitlement theory can accommodate communal claims if communities, rather than individuals alone, are seen as legitimate holders of property or rights based on historical ownership and collective acquisition. Thus, African communal traditions can complement Nozick's view by expanding the concept of entitlement beyond the solitary individual to include legitimate collective ownership.

Furthermore, there is an additional point that worth considering. While Nozick offers a valuable critique of arbitrary redistribution in African states, his redistribution strategy can be further enriched by incorporating other perspectives that are unique to the African context. For instance, Olusola Olanipekun suggested that... one of the most practical approaches to the question of just distribution and acquisition may involve a return to embracing foundational African values such as honesty, goodwill, good reputation, and contentment. (Olanipekun, 2021, p:15). As Olanipekun suggests, when a society's historical record is marred by deep injustices, the process of restructuring that society may be hindered unless these core values- whose erosion has contributed to the persistence of injustice are restored.

Conclusion

This paper examined the conjoined problems of political corruption and injustice in Post-colonial African States through the philosophical lens of Robert Nozick's entitlement theory of distributive justice. According to Nozick, holdings are just if they arise from just acquisition, just transfer and the rectification of past injustices. However, in many African states, political corruption which manifests in different ways negates the conditions for just acquisition and transfer. This paper argued that the unjust acquisitions consequently invalidate the moral legitimacy of the distributions. Thus, applying Nozick's principles to the issue of political corruption in Africa challenges the assumption that current properties and wealth distributions among politicians and political office holders can be defended without addressing the deeply embedded foundational injustices. Furthermore, the paper examined the strengths and the limits of Nozick's entitlement thesis. It suggested that anti-corruption measures and corrective justice are not only compatible with Nozickian framework, but are also necessary to restore legitimate entitlements. Finally, it concluded that Nozick's entitlement theory offers a valuable framework for addressing ongoing social injustices by promoting fair principles that ultimately contribute to a more just and well-ordered society.

References

Amundsen, I. (2006). "Political Corruption" in *Anti-Corruption Resource Centre CMI*, vol. 6(3)1-36

Amundsen, I. (2010). *Political Corruption in Africa: Lessons* from Country Case Studies. Bergen: Chr. Michelsen Institute.

Bayart, J.F. (1993). *The State in Africa: The Politics of the Belly*. London: Longman.

Bratton, M., & van de Walle, N. (1997). *Democratic Experiments in Africa: Regime Transitions in Comparative Perspective*. Cambridge: Cambridge University Press.

Christopher, A. J. (1994). *The Atlas of Apartheid*. London: Routledge.

Cohen, G. A. (1995). *Self-Ownership, Freedom, and Equality*. Cambridge University Press.

Elkins, C. (2005). *Imperial Reckoning: The Untold Story of Britain's Gulag in Kenya*. New York: Henry Holt.

Falola, T. (2009). *Colonialism and Violence in Nigeria*. Bloomington: Indiana University Press.

Githongo, J. (2006). *Inefficiency and Corruption in Public Sector Activities: A Review of Kenya's Experience*. Nairobi: Kenya Anti-Corruption Commission.

Gyekye, K. (1997). Tradition and Modernity: Philosophical Reflections on African Experience:
Accra: Sankofa.

Hope, K. R. (2016). *Corruption and Governance in Africa: Swaziland, Kenya, Nigeria*. Cham: Palgrave Macmillan.

Mamdani, M. (1996). *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton: Princeton University Press.

Mauro, P. (1995). "Corruption and Growth". *Quarterly Journal of Economics*. Vol.110, (1):681-712

Médard, J.F. (2002). "Corruption in the Neo-Patrimonial States of Sub-Saharan Africa," in *The Political Economy of Corruption*. London: Routledge. pp. 379-403.

Milgroom, J., & Spoor, M. (2013). "Gender and Generational Impacts of Large-scale Land Acquisitions in Mozambique: Insights from the ProSAVANA Program". *Conference Paper for the International Conference on Global Land Grabbing II*, 17-19, Cornell University.

Molefe, M. (2024). *Ubuntu Ethics: Human Dignity, Moral*

Perfectionism and Needs. New York: Routledge.

Moyo, S. (2011). *Land Reform Under Structural Adjustment in Zimbabwe: Land Use Change in the Mashonaland Provinces*. Uppsala: Nordiska Afrikainstitutet.

Ndlovu-Gatsheni, S. J. (2013). "Coloniality of Power in Postcolonial Africa: Myths of Decolonization". Dakar: CODESRIA.

North, D. C. (1990). *Institutions, Institutional Change, and Economic Performance*. CUP

Nozick, R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books.

Olanipekun, O.V. (2017) "Omoluabi: Re-thinking the Concept of Virtue in Yoruba Culture and Moral System" in *Africology: The Journal of Pan African Studies*, 10 (9):217-231.

Olanipekun, O.V. (2021) "Political Corruption in Africa: Revisiting Kwame Gyekye's Moral Solution" in *The African Review*, 48 (1):1-18.

Onapajo, H. (2014). "The Dominance of Incumbents in the Use of Violence to Rig Elections", *African Spectrum*, Vol.49, (2): 27-51

Rawls, J. (1971). *A Theory of Justice*. Cambridge: Harvard University Press.

Rodney, W. (1972). *How Europe Underdeveloped Africa*. Dar es Salaam: Tanzania Publishing House.

Rodney, W. (1972). *How Europe Underdeveloped Africa*. London: Bogle-L'Ouverture Publications.

Rose-Ackerman, S. (1999). *Corruption and Government: Causes, Consequences, and Reform*. Cambridge: Cambridge University Press.

Southall, R. (2005). "The Ndungu Report: Land and Graft in Kenya" *Review of African Political Economy*, 32(103):142

Smith, D. J. (2007). *A Culture of Corruption: Everyday Deception and Popular Discontent in Nigeria*. Princeton: Princeton University Press.

Waldron, J. (1992). "Superseding Historic Injustice". *Ethics*. Vol. 103, (1):4-28.

Zondo Commission Report. (2022). *State Capture Inquiry Final Report*. Pretoria: Government of South Africa.

Human Rights Watch. (2011). "Waiting here for Death: Forced Displacement and Villagization in Ethiopia's Gambella Region".

Economic and Financial Crimes Commission (EFCC) (2020). "Annual Report 2020". Abuja: EFCC.

Transparency International. (2022). "Corruption Perceptions Index 2022". Retrieved from [https://www.transparency.org/en/cpi/2022] (https://www.transparency.org/en/cpi/2022)

World Bank. (2018). "Overcoming Poverty and Inequality in South Africa: An Assessment of Drivers, Constraints and Opportunities". Retrieved from [https://www.worldbank.org/]

World Bank. (2020). "Nigeria Development Update: Rising to the Challenge". Retrieved from [https://www.worldbank.org/] (https://www.worldbank.org/).

Social-Media Companies Have a Duty to Remove Misinformation

Douglas R. Campbell
Department of Philosophy, Alma College

Abstract

The goal of this article is to argue that social-media companies ought to remove misinformation from their platforms. I begin by outlining four kinds of harms that misinformation can produce. In the following section, I clarify the conceptual foundations of this duty, including the difference between purposeful and general social-media sites, influenced by Lavinia Marin's recent research, and the correct characterization of online debates about topics such as climate change and vaccination.

Keywords: social media; misinformation; choice architecture; epistemic harm; disinformation; infodemic; information environment

I argue that social-media companies have a duty to remove misinformation. This duty exists in order to mitigate four kinds of harm: harm to a person's well-being, harm to democratic institutions, harm to an information environment, and the harms that result from a pernicious choice architecture. For the purposes of this article, it is sufficient to define misinformation as content that is either false or misleading.[1] This capacious understanding of misinformation includes content that asserts something false, content that asserts something true but omits important facts in tendentious ways, and deepfakes, which are not, strictly speaking, true or false but which are misleading. I argue that such content is harmful in the first section. However, most of this article is not about why such content is harmful; instead, I spend most of this article walking through important conceptual foundations that underlie the duty to remove misinformation. For instance, Lavinia Marin's distinction between purposeful and general social-media sites is key here, and so is Ludwig Wittgenstein's understanding of a language-game.

Social-media companies occupy such an important role in 21st-century society that they have a duty to remove misinformation. They should prioritize removing content in proportion to how harmful it is. Accordingly, content that undermines people's health and well-being and content that undermines democratic institutions should be removed first. Content that, in contrast, attributes incorrect titles to Taylor Swift's albums is not much of a priority, even though the false beliefs generated by that misinformation do harm our information environment.

As I argue below, the common defense that removing misinformation is tantamount to censorship and violates free-speech rights should not make us doubt the duty to remove misinformation. In the first place, such a defense implausibly generates a duty to not remove misinformation. More importantly, it ignores the place that social-media companies occupy in today's society, and it is based on a characterization of online discourse that I argue is false.

1. Harms of misinformation

A duty to remove misinformation amounts to a duty to mitigate the harms caused by social-media sites. In this section, I shall identify four harms that misinformation causes and that the duty to remove misinformation would mitigate. The harmful nature of misinformation has been subject to fruitful and thorough research in recent years, and I shall sometimes use the footnotes to direct readers

1 See Carey 2023 for a more sophisticated treatment of misinformation as content that has a disposition to worsen a person's overall epistemic position.

to thoughtful discussions elsewhere. It is not the goal of this article to contribute to this research exactly, but to justify the claim that there exists a duty to address the harms caused by misinformation. For this reason, I shall move briskly through some of the harms, saying enough to establish that the harms exist but not much more.

The first kind of harm generated by misinformation is the simplest: the way that false beliefs make us worse off as individuals. This is most evident in the case of misinformation that concerns our health and well-being. For example, the misinformation that proliferated online during the COVID-19 pandemic harmed people by persuading them that getting vaccinated was dangerous. Social-media platforms give airtime to people such as the so-called Disinformation Dozen, a small group of online influencers who were responsible for roughly two-thirds of all anti-vaccine misinformation during the pandemic.[2] One member of the Disinformation Dozen, Sherri Tenpenny, gained such a high level of prominence on account of her social-media presence that she eventually testified in front of the Ohio House of Representatives' health committee. Her testimony reflected the beliefs that she spread online, namely, that the COVID-19 vaccine would magnetize people and that keys and other household items would stick to people who were vaccinated. Early in 2022, she predicted that "by the end of 2022, every fully vaccinated person over the age of 30 may have the equivalent of full-blown vaccine-induced immune suppressed AIDS."[3] Such misinformation can convince people to avoid important medical interventions.

There is nothing special about COVID-19 misinformation in this regard. Social-media sites have been hotbeds for pro-anorexia communities for a long while.[4] People in these communities convince others not to get treatment for their psychiatric condition, which is terribly unfortunate in light of the fact that anorexia is more lethal than any other mental illness. This misinformation can have obvious and measurable negative effects on a person's life, even to the point of death.

Health-related misinformation stands out here as the most obvious culprit, but other forms of misinformation can harm a person's well-being, too. People in so-called incel, or involuntarily celibate, communities can become so entrenched in misogynistic and sexist misinformation that they miss out on meaningful relationships with women. Members of QAnon and flat-Earther communities can similarly become so obsessed with the misinformation and conspiracy theories that circulate around them that they become estranged from people in their lives. That too is harmful.

The second kind of harm is the damage done to democratic institutions. The democratic process works in the following manner: I, a voter, have values; competing politicians express values and support policies based on those values; I vote for the politician who most closely reflects my own values and consequent policy preferences; the elected official then works in the government alongside other elected officials to draft policies that reflect those values. The democratic process breaks down under the weight of misinformation. Former president of the U.S.A., Barack Obama, got at this point well when he said:

> If we do not have the capacity to distinguish what's true from what's false, then by definition the marketplace of ideas doesn't work. And by definition our democracy doesn't work. We are entering into an epistemological crisis.[5]

To understand why this is the case, imagine a candidate for political office spreading the misinformation that climate change is a hoax perpetrated by the Chinese government.[6] There are many voters who care about the environment and who deeply value ensuring that future generations inherit a healthy planet. Those voters might be persuaded by this candidate that climate change simply does not exist, and then they vote accordingly. The policies that result from this elected official will reflect the belief that climate change is not occurring, even though it is, and will therefore fail to align with the values of the electorate (even though it aligns with their false beliefs). Misinformation threatens the integrity of democratic processes by thwarting the representation of voters in policy-making.

The third kind of harm is the poisoning of our information environment.[7] The more false or misleading content there is online, the worse our position with respect to the formation of beliefs becomes. We will tend to encounter more false or misleading assertions, in general. The quality of the average piece of information declines because there is a greater quantity of low-quality information available. High-quality sources of information might be falsely slandered, which in turn will cause us to miss out on true beliefs that those sources of information might have provided us with. We could very well even abandon true beliefs that we had already formed through testimony provided by high-quality sources; the slander against them could incorrectly convince us to drop those beliefs. Online influencers trade in misinformation all too frequently: the Disinformation Dozen, mentioned earlier, is one

2 See Srikanth 2021.
3 See Gage 2022.
4 Parsell 2008 has an interesting discussion about why pro-anorexia and similar communities thrive online.
5 Goldberg 2020.
6 See Emery 2016.
7 This is a complicated kind of harm, which I unpack only in a cursory manner here. See Carey 2023, de Ridder 2021, and Marin 2021b for more thorough discussions.

example. In this information environment, going online is akin to wading through vast quantities of misinformation. Even the vigilance required of us as doxastic agents is exhausting, which is a kind of harm in itself.

The last kind of harm is the harm produced by a choice architecture that is built on misinformation. This kind of harm is rarely, if ever, mentioned in discussions of misinformation, and it is subtle. The term 'choice architecture' originated in libertarian-paternalist research.[8] The choice architecture is the context in which a choice is made, and libertarian paternalists view modifications of the choice architecture that they call *nudges* as permissible ways of promoting the likelihood that some particular choice will be made, without the use of coercion. A clear offline example is the design of grocery stores: customers are being nudged towards the purchase of candies by placing them prominently near cash registers, where customers generally spend a few minutes waiting. The likelihood that customers will buy candies increases, but no coercion is used. Social-media developers design the choice architecture of their sites in ways that maximize the amount of time spent on them.[9] For instance, notification icons are designed to be a shade of attention-catching red, and many video-streaming sites play the next video automatically after the completion of the current one.

Misinformation plays a key role in the choice architecture of a social-media site. One way to get people to spend more time on the site is to show them content that will grab their attention. One prominent kind of nudge is the recommendation that a site will offer by saying that some content is trending or worth checking out. The problem is that sometimes this content is misinformation because the algorithm or the developers have predicted that the content that will be most effective at maximizing user-engagement metrics is misinformation. Certainly, one of the most heinous examples of this was Meta's knowing promotion of pro-eating-disorder content on Instagram in order to nudge young girls to stay on the platform longer.[10] The removal of misinformation would clean up the choice architecture of a site, no longer allowing developers to promote clickbait or other misleading or false content in order to boost engagement metrics.

If user-engagement metrics fall as a result of the removal of misinformation, that might well be a good thing. There are randomized controlled trials showing that reducing social-media usage improves anxiety, fear of missing out, loneliness, and depression.[11] If changes to the choice architecture entail that we spend less time online, that might be better for us. Even if this psychological research turns out to be false upon further investigation, many philosophers have argued that we have a right to our attention that the constant distractions posed by social-media sites violate.[12] The duty to remove misinformation would not stop people from distracting us with the truth, but it would remove many of the usual weapons, such as clickbait headlines, that sites use to distract us from whatever else we would prefer to spend our time on.

There is much to be said about the way that social-media platforms' algorithm plays a role in aggravating the harms of misinformation. Since their algorithms are proprietary and thus not available for our scrutiny, it is hard to speak about them in concrete terms. Yet, we can speak generally about this problem. The algorithmic boosting of some content for the sake of user-engagement metrics amplifies misinformation and increases the harm done. In a profound sense, social-media sites do the opposite of what this article argues that they ought to do: instead of removing, they promote misinformation. Even if they changed their algorithms to downplay misinformation, it would still be wrong of them to fail to remove misinformation because the harms outlined in this section would still occur, even if at a reduced rate. For this reason, social-media sites ought to outright remove misinformation.

Ultimately, these are four kinds of harm that the duty to remove misinformation would mitigate. A careful bit of reflection reveals that not all misinformation is equally harmful. Misinformation that threatens our well-being and democratic institutions is the most harmful. Health-related misinformation can cause the most visible harm, whereas content that undermines democracies is most likely the most consequential per unit. However, even apparently harmless misinformation, such as incorrectly stating that Taylor Swift was born in 1988, worsens our information environment and is thus responsible for at least one of the four harms. In what follows, I shall present the conceptual foundations of the duty to remove misinformation, and I maintain that social-media companies should remove content while prioritizing what is most harmful. Misinformation about Taylor Swift, and similar subjects, should be removed last, and the most harmful misinformation should be removed first. Ultimately, all misinformation should be removed because even the least harmful kind still produces harm to our information environment, but some misinformation produces even more harm and, consequently, should be prioritized.

8 See, e.g., Thaler and Sunstein 2008 and Campbell 2022.

9 Campbell 2022 discusses many examples of such nudges as well as their moral status.

10 See Wells, Horwitz, and Seetharaman 2021.

11 E.g., Hunt et al. 2018.

12 E.g., Puri 2021, Tran 2015, and Chomanski 2023.

2. The conceptual foundations of the duty to remove misinformation

We should remove misinformation in order to prevent or mitigate these harms. However, there is a problem. The most prominent objection to removing misinformation is that doing so constitutes censorship and violates people's free-speech rights. This is built on a way of setting up content-moderation dilemmas as a competition between two values: respecting freedom of speech and mitigating harm. This way of setting up the problem is common both inside and outside academia.[13] There is an interpretation according to which this objection is simply false. People have free-speech rights to protect them from government overreach. To say that a company's decision to remove your misinformation violates your free-speech rights is as false as saying that your mother's decision to tell you to stop talking violates your free-speech rights. Private individuals are morally permitted to remove content from their platforms. This is how we might reply to this interpretation of the free-speech objection. However, there are two much more compelling and intellectually fruitful ways of construing this objection.

First, we can establish a more compelling version of this objection by reframing it as being less about people's free-speech rights and more about the integrity of public discourse on social-media sites. Sites such as X and Instagram occupy such an important role in democratic societies that they might be plausibly conceived of as the spaces in which citizens work out ideas together. The elimination of an important idea, even a false one, such as that climate change is a hoax perpetrated by a foreign government, from social-media sites effectively eliminates it from public discourse altogether. It is not fair to point out that people in free societies could still circulate such ideas in print, by means of brochures and pamphlets: ideas circulated on YouTube, for example, reach an audience with a speed and scope of distribution that is not paralleled by any other form of communication that humanity has ever developed. The removal of even false ideas harms public discourse by taking out of the public's consideration possibly important things to reflect on and debate.

This objection to the duty to remove misinformation is powerful, but I disagree with it. It is a helpful objection to consider because it provides us with an opportunity to reflect on the conceptual foundations of the duty to remove misinformation.

We should begin by observing exactly what the objection entails. Specifically, it entails that there is a duty to *not* remove content on social-media sites. This is not merely the negation of the central claim of this article, which is that there is a duty to remove misinformation. The objection affirms that it is morally impermissible to remove content posted online. The duty to not remove content is implausibly strict. Imagine Asian owners of a social-media site. Now imagine that on this site, many users post misinformation that claims that Asian people uniquely spread some disease and that anti-Asian racism is, therefore, justified. This is not unlike what happened at times throughout the COVID-19 pandemic amid speculation concerning the origins of the virus. If the owners of the site had a duty to not remove this misinformation, then they would not be permitted to remove any of this content. It would not matter how much it violated their own values or how uncomfortable they felt owning a site that was responsible for airing so much racist content. People would have to tolerate misinformation on their own platforms even when they had the tools and desire to remove it. If the value of their company started to decrease rapidly, they would still have to tolerate it.

Consider how counterproductive this would be. One important tool for social change is the pressure that business owners, and other similarly positioned people, can exert on people with socially undesirable behaviors, such as spreading racist misinformation. If we have a duty to not remove misinformation, then we lose this important tool for social change.

Moreover, social-media sites *already* remove all sorts of content, and it is obviously good that they do so. The removal of child pornography is a clear example of content that we already recognize ought to be removed.

13 For a recent academic example, see Kozyreva 2023. See also Stroud 2019's discussion of what he calls trade-offs in the attempt to remove misinformation. Here are a few examples from outside academia. First, Monbiot 2021's way of setting up the problem: "we have a right to speak freely. We also have a right to life. When malicious disinformation — claims that are known to be both false and dangerous — can spread without restraint, these two values collide head-on." Second, see Zuckerberg 2019's discussion about the practices of his own company: "we make a lot of decisions that affect people's ability to speak. […] Frankly, I don't think we should be making so many important decisions about speech on our own either." This is reflected in Mudde 2020's headline from The Guardian during the pandemic: "don't let free speech be a casualty of coronavirus. We need it more than ever." Howard 2021: 168 sums up this debate thus: "that something should be done to confront this misinformation seems undeniable. The controversy arises when we consider what, exactly, the appropriate solution is. One possibility is to reduce the circulation of misinformation by insisting that social media networks remove it from their platforms, or at least reduce its dissemination. Yet in the public debate, this suggestion continually inspires a chorus of opposition from those who think that such a move would run afoul of one of the most cherished values of liberal democracy: freedom of speech" (emphasis in original).

So, clearly, there is no duty to *not* remove content in general. As for misinformation in particular, sites already have unobjectionable removal practices. For instance, imagine if I post a false review of a restaurant on a review site such as Google Reviews or Yelp. Imagine that I do this specifically in order to skew perceptions of the restaurant by remarking on dishes that do not exist there: I talk about chicken tikka masala and naan in a review for an Italian restaurant. Google and Yelp would delete my review according to their present removal guidelines, and nobody would disagree with this removal. It is clear that such misinformation's pollution of our information environment cannot and should not be tolerated. The whole purpose of the review site falls apart if it tolerates false reviews.

I said earlier that there were two intellectually fruitful ways of construing the free-speech objection. The first, we have seen, entails that there is a duty to *not* remove content, and I have argued that, on the basis of this entailment, this objection does not work. The second denies that there is a duty to *not* remove content; instead, it maintains that it is permissible to leave content untouched. It is permissible to remove, but it is permissible to *not* remove, too. This objection is plausible to the extent that it avoids what was unintuitive about a duty to not remove, but it overlooks the tremendous harm that misinformation can cause, which I discussed in the first section. The duty to remove content online *is* a duty to mitigate harm. Since this objection *concedes* that there is no duty to *not* remove content and that removing content does not trample on free-speech rights of users, it is mysterious why failing to mitigate the harms done by misinformation is permissible.

The need to remove misinformation is even greater when we consider cases that bear on public discourse about weighty matters. Imagine if someone spreads misinformation concerning the backgrounds and political views of some candidates for political office on Wikipedia. Wikipedia is so important to many people's research into candidates that the need to remove this information increases in urgency. This is partly why the objection that highlights the place of social-media platforms in 21st-century society fails: *the place of social-media platforms is so important to public discourse that the duty to remove misinformation follows naturally*. The objection drew the wrong conclusion from the importance of social-media platforms: it concluded that we should let discussions unfold without intervention, but, in fact, interventions are needed because of how important they are.

I have cited hypothetical cases on Wikipedia, Google Reviews, and Yelp. At first, these might seem disanalogous to cases on Facebook, Instagram, TikTok, and so on. That is because the former are *purposeful* social-media sites, whereas the latter are *general* social-media sites. The distinction between purposeful and general sites was invented by Lavinia Marin in a discussion that clarified the norms of sharing information online.[14] Purposeful social-media sites are those that have a clear purpose that unifies the behaviors of users on the site, and this purpose gives us norms that we can use to assess whether someone's behaviors are objectionable. For instance, LinkedIn is dedicated to discussions about users' careers. If someone is spreading misinformation about his or her employment history, then it is obvious that such content should be removed. We can arrive at that judgment by citing the site's purpose. This is what is going on when we agree that false, misleading reviews on Yelp should be removed.

It is trickier to assess cases on general social-media sites because they lack a clear, unifying purpose. That is what it means to be a *general* site: users use such sites as Facebook and Instagram for a variety of purposes, perhaps even changing purposes of use daily. The fact that Facebook, Instagram, Reddit, and so on, allow such a variety of uses that users find acceptable makes it harder for us to justify the removal of content. Some Reddit communities might be dedicated to finding a romantic partner, in which case spreading misinformation about oneself is wrong and deceptive, whereas there are other Reddit communities that might be satirical, in which it would be uncharitable to view posts as deceptive and misinformation.

Nevertheless, I maintain that we can still justify the removal of misinformation from general social-media sites. Firstly, we could still point to the harms of misinformation, enumerated in the first section of this article, to justify the duty to remove misinformation on simple consequentialist grounds. However, I think that there is an important philosophical concept that clarifies the matter. This is the concept of a *language-game*, which was invented by Ludwig Wittgenstein and introduced into the ethics of social media by Lavinia Marin.[15]

Wittgenstein proposed the concept of a language-game in order to explain that our words and utterances get their meanings from the contexts in which they are embedded. Here is an example that he gives:

> The language is meant to serve for communication between a builder A and an assistant B. A is building with building stones: there are blocks, pillars, slabs and beams. B has to pass the stones, and that in the order in which A needs them. For this purpose they use a language consisting of the words "block," "pillar," "slab," "beam." A calls them out; — B brings the stone which he has learnt to bring at such-and-such a call. — Conceive this as a complete primitive

14 See Marin 2021a.

15 See Wittgenstein 1958 and Marin 2021a.

language.[16]

This simple language-game suffices to illustrate the point. After all, shouting 'block!' could mean different things in different language-games. Wittgenstein's concept is rich and deep, but for our purposes here, we can focus on the way that utterances are considered acceptable or illegitimate depending on the context in which they are uttered. For instance, poetry is an entirely different language-game than a broadcast on CNN, and different utterances are acceptable in each. I emphasize the centrality of Wittgenstein's concept here because it will prove indispensable to our analysis of misinformation on social-media sites.

Now let us return to the distinction between purposeful and general social-media sites. Each purposeful social-media site is a different language-game. There are different norms for communicating about ourselves on LinkedIn than there are on Tinder or Bumble. There are things that we might say about ourselves on Bumble that would raise eyebrows on LinkedIn, and a conversation focused on our professional history is par for the course on LinkedIn but would be entirely out of place on Tinder. While each purposeful site has its own distinct language-game, the rules are easily discerned. To figure out what is acceptable to share on LinkedIn is, in the vast majority of cases, as easy as it is to figure out that if someone at a construction site says the word 'block!' to you, he or she is asking for a block.

General social-media sites look different because of the variety of uses. There is *some* language-game happening on Facebook, but it is not always clear what the language-game is. Consequently, the rules are not so obvious; neither is what is acceptable. However, we are armed with some psychological and sociological research that illuminates the answer to this question. People generally post on social-media sites for entertainment, as part of status-seeking, for self-expression, or for socialization.[17] Some researchers say that the diversity of motivations can be reduced to two basic needs: the need to belong and the need to present oneself to others.[18] One group of researchers summarizes their findings by saying that "catchiness rather than truthfulness often drives information (and misinformation) diffusion on social media."[19] The language-game on general social-media sites is something like the language-game on a playground: the kids talk about a variety of subjects, but they are not really interested in the truth, by and large.[20] They speak, spread rumors, and share gossip in order to fit in, show what their values are, have fun, exclude people whom they want to exclude, and promote social cohesion among people that they want to include.

In light of this research, we should not have any objections to removing misinformation. The objection to the duty to remove misinformation was founded on the view that conversations on social-media sites, such as X, were part of an important public debate that it would be wrong to intervene in. On the one hand, when such a debate is happening, we *ought* to intervene for the same reason that we think that Wikipedia's content should be carefully scrutinized: the cost of negligence and laxity is too high. On the other hand, research by psychologists and sociologists suggests that conversations on social-media sites should not generally be characterized this way. We should be as ready to jump in and intervene in these conversations as we would be if we overheard kids on a playground spreading rumors and gossip. After all, as bad as the children's rumors and gossip might be, the misinformation spread on social-media sites can be much more dangerous.

It is because of this increased danger that we should respond differently to misinformation online than we would to rumor-spreading on a playground. Whereas we might respond to the spread of a rumor by privately correcting the offending child, the harms caused by social-media platforms instead call for the removal of the content. The harms I cataloged in the first section of this article generate this duty for us, but to differing degrees. The one most salient, for the topic at hand, is the harm done to our information environment. The *only* way to clean up our information environment is to remove the content. Otherwise, we will still have a polluted information environment.

There is an objection to this view that conversations on social-media sites are like conversations on a playground. The objection is that although research has revealed that this is how people *post* online, we have reason to believe that people *consume* content differently. Around half of Americans get news from social-media sites.[21] The spread of misinformation on social-media sites is, to some extent at least, due to an unfortunate asymmetry: *people post as if they were in a playground, but they consume as if there were a serious debate occurring between knowledgeable people*.

Let us pursue a bit more the possibility that conversations on social-media sites are important and high-minded debates that are a worthy source of news information. This view is not plausible in light of what we know about the motivations of those who post, but it is worth taking seriously since many people consume news on social-media sites as if it were true. It might be the case that this view entails that we should be reluctant

16 Philosophical Investigations §2. Translated by G.E.M. Anscombe.

17 See Chen et al. 2015, Diddi and LaRose 2006, and Lee and Ma 2012.

18 See Nadkarni and Hofman 2012.

19 Chen et al. 2015: 111.

20 Marin 2021a similarly infers that conversations on general social-media sites are like the gossip and rumors spread by kids on a playground.

21 See Walker and Matsa 2021.

to remove ideas from these debates. However, I think that we should ultimately reject this possibility because these sites are simply *bad venues for hosting the sorts of debates that are characterized this way*. Consider some of the misinformation that gets spread online regarding climate change or vaccination. Recall Sherri Tenpenny's misinformation, mentioned earlier, that the COVID-19 vaccine magnetizes people. These are important scientific claims. They should be tested in whatever experimental setting is appropriate: perhaps randomized controlled trials, or something else, where appropriate. The proper institutional venues for putting forward these claims are academic ones, with the infrastructure provided by peer review. There might be flaws to institutional peer-review systems, but none of these flaws are remedied by Instagram or YouTube. Social-media sites, in turn, offer precisely *zero* advantages with respect to reliability or accuracy.

In other words, psychological and sociological research makes it clear to us that people are *not* posting for reasons that would make us refrain from removing content online. People do generally consume content as if there were important debates happening on social-media sites, but to the extent that people sincerely believe that this is happening, they are mistaken. The important debates concerning the topics that misinformation is often about are happening elsewhere, in settings that have reliable, even if imperfect, institutions guiding them. Moreover, *if* there were truly important debates happening on social-media sites, the importance of those debates would entail that we ought to intervene in order to ensure the integrity of the debates, just as we intuitively recognize the importance of ensuring that Wikipedia's information is accurate.

Without valuable institutional safeguards, debates on social-media sites are insensitive to the truth in ways that were evident as soon as we learned the motivations that people generally have when posting. Some people might still insist that, even without institutional peer-review processes, high-quality content will still rise to the top. This objection goes like this: perhaps peer review has merits, but we can get by without it in other settings and still arrive at high-quality information. I present one experiment in order to rebut this objection. Researchers performed an experiment on Reddit in which they made a comment and then either upvoted it immediately, downvoted it immediately, or left it alone.[22] If high-quality content rises to the top, then we would expect that, over time, the initial upvote or downvote would be canceled out by corrections from other users. In fact, that is not what happened at all. Both the upvote and the downvote had the effect of increasing overall turnout: the initial popularity of the comment led to sustained popularity in the long term. The upvote predicted further upvotes down the road. In other words, people like content that other people like, even when the initial liking is manufactured and artificial. Interestingly, in the experiment, while the downvote increased turnout, it did not predict further downvotes. Overall, then, we should conclude from this research that we have no reason to believe that the success of content online tracks anything related to the truth. Perhaps the success of content is due to luck or the money that some influencers spend buying followers or employing bots, or some other factor.[23] This insensitivity to the truth provides us with another reason to abandon any remaining reluctance to remove misinformation from social-media sites.

Lastly, I will concede that it can be hard to tell what counts as misinformation. I do not mean that the definition of misinformation is unclear. I mean that it is sometimes difficult to determine whether some piece of online content is an instance of misinformation. For instance, in the early days of the COVID-19 pandemic, content that concerned the efficacy of masking or the possibility of asymptomatic transmission of COVID-19 was rife on social-media sites, but it was not obvious what the truth was. This point is well-taken, but this should not make us doubt the existence of a duty to remove misinformation. It should make us carefully inquire into whether some content is misinformation. We should not lose sight of the fact that misinformation is very often easily identified as such, even if not always.

Moreover, it would be wrong to doubt the existence of a duty on account of the fact that there are sometimes cases in which it is not obvious whether the duty applies. Think about the duty to rescue someone from drowning when you are easily able to do so. Now imagine that you are on a beach and that you see someone who *might* be drowning. Some of the time, it is obvious whether a person is struggling in the ocean, but this is one of those cases in which it is not so obvious. There are many reasonable ways to respond to this: you can try to get a better view of the person, while bearing in mind that time might be sensitive, or you can rush out there and err on the side of caution. Either way, nothing about the difficulty of determining whether the duty applies to this specific case should make us doubt the existence of the duty. The same goes for the duty to remove misinformation and the difficulty of determining whether

22 See Muchnik et al. 2013 for the original research; see also Sunstein 2017: 112 for a discussion.

23 For this reason, I in general think that we should reject the claim that if we subject ideas to competition in a so-called marketplace of ideas, the winner of that competition will be the truth. There is a discussion of this claim in Kozyreva et al. 2022. As Sunstein 2021: 20 puts it, "far from being the best test of truth, the marketplace ensures that many people accept falsehoods."

something *is* misinformation. I recognize that this case of the drowning person is not identical in every respect to the case of misinformation, but the general point is that the existence of a duty should not be doubted on the grounds that it is not always obvious when that duty applies.

A final objection: perhaps misinformation should not be removed; instead, social-media companies should simply flag misinformation as false or misleading. If this is sufficient, then we can preserve the freedom of speech of the users who post the misinformation, while also disarming their content's ability to harm people. There are a few problems with this suggestion, and they are all related to people's cognition. First, the presence of misinformation is polluting the information environment, which in turn harms users' ability to reliably form beliefs well. Therefore, since people are not able to reason well, they are not able to deliberate properly about whether to believe the misinformation or the flag alerting them to the fact that it is misinformation. The flag-but-do-not-remove proposal misunderstands the problem, which, in reality, concerns people's doxastic abilities being undermined. Second, the proposal overlooks the role of cognitive biases, such as confirmation biases, in determining what content people online take seriously. Perhaps some users will be persuaded by the flag, but not all will be. Accordingly, the misinformation should be removed for the sake of those unconvinced users. Third, the flag is no different in principle from everyone else on social-media sites who speaks out against misinformation. If those other people will not make a difference, then neither will the flag. Perhaps users who debunk misinformation are not so visible to users who consume misinformation; so, the hope is that the flag will be seen by the users who are consuming misinformation because it is adjacent to the content. However, the first two problems with this proposal explain why we should not be so optimistic.

The flag-but-do-not-remove proposal claims to resolve the harms of misinformation by merely *adding more speech* to our information environment, but the addition of more speech will not clear up the information environment. It merely contributes to it. The fact that social-media sites already contain accessible, high-quality information does not reduce the damage done to our information environment. So, adding more speech, in the form of flags, will not, either.

For these reasons, I conclude that social-media companies have a duty to remove misinformation.

Bibliography

Campbell, Douglas R. 2022. "Nudging and Social Media: The Choice Architecture of Online Life," G*iornale Critico di Storia Delle Idee* 2002 (2): 93-114.

Carey, Brandon. 2023. "Misinformation and Epistemic Harm," *Social Philosophy Today* 39: 89-100.

Chen, X., Sin, S. C. J., Theng, Y. L., Lee, C. S. 2015. "Why do social media users share misinformation?" In PL Bogen, S Allard, H Mercer, M Beck, SJ Cunningham, D Goh, G Henry (eds.) *Proceedings of the 15th ACM/IEEE-CE on Joint Conference on Digital Libraries - JCDL '15*, New York,111-114.

Chomanski, Bartlomiej. 2023. "Mental Integrity in the Attention Economy: in Search of the Right to Attention," *Neuroethics* 16: 8-19.

De Ridder, Jerome. 2021. "What's So Bad about Misinformation?" *Inquiry* 1-23 https://doi.org/10.1080/0020174X.2021.2002187

Diddi, A., & LaRose, R. 2006. "Getting hooked on news: Uses and gratifications and the formation of news habits among college students in an internet environment." *Journal of Broadcasting and Electronic Media* 50 (2): 193–210.

Emery, David. 2016. "Did Donald Trump Claim Global Warming Is a Hoax?" September 27, 2016 https://www.snopes.com/fact-check/donald-trump-global-warming-hoax/

Gage, Brandon. 2022. "Anti-vax activist claims COVID-19 vaccines cause AIDS," *Salon* April 3, 2022. https://www.salon.com/2022/04/03/anti-vax-activist-claims-19-vaccines-cause-aids_partner/

Goldberg, Jeffrey. 2020. "Why Obama Fears for Our Democracy." *The Atlantic*, November 16, 2020 .https://www.theatlantic.com/ideas/archive/2020/11/why-obama-fears-for-our-democracy/617087/

Howard, Jeffrey. 2021. "Coronavirus Misinformation, Social Media, and Freedom of Speech," F. Niker and A. Bhattacharya, *Political Philosophy in a Pandemic: Routes to a More Just Future*. New York: Bloomsbury, 167-176.

Hunt, Melissa G., Rachel Marx, Courtney Lipson and Jordyn Young. 2018. "No More FOMO: Limiting Social Media Decreases Loneliness and Depression," *Journal of Social and Clinical Psychology,* 37 10: 751-768.

Kozyreva, Anastasia, Herzog, Stefan M., Lewandowsky, Stephan, Hertwig, Ralph, Lorenz-Spreen, Philipp, Leiser, Mark, and Reifler, Jason. 2023. "Resolving content moderation dilemmas between free speech and harmful misinformation," *Proceedings of the National Academy of the Sciences*. 120: 7 doi: 10.1073/pnas.2210666120.

Lee, C. S., and Ma, L. 2012. "News sharing in social media: The effect of gratifications and prior experience." *Computers in Human Behavior* 28 (2): 331-339.

Marin, Lavinia. 2021a. "Sharing (mis) information on social networking sites. An exploration of the norms for distributing content authored by others," *Ethics and Information Technology* 23 (3): 363-372.

—. 2021b. "Three contextual dimensions of information on social media: lessons learned from the COVID-19 infodemic," *Ethics and Information Technology* 23: 79-86.

Monbiot, George. 2021. "Covid lies cost lives – we have a duty to clamp down on them" *The Guardian* January 27, 2021 https://www.theguardian.com/commentisfree/2021/jan/27/covid-lies-cost-lives-right-clamp-down-misinformation

Mudde, Cas. 2020. "Don't let free speech be a casualty of coronavirus. We need it more than ever," *The Guardian* , April 6, 2020. https:/www.theguardian.com/commentisfree/2020/apr/06/coronavirus-free-speechhungary- fake-news

Lev Muchnik, Sinan Aral, and Sean J. Taylor. 2013. "Social Influence Bias: A Randomized Experiment," *Science* 341 (6146): 647–51.

Nadkarni, A., and Hofmann, S. G. 2012. "Why do people use Facebook?" *Personality and individual differences*, 52 3: 243–249.

Parsell, Mitch. 2008. "Pernicious Virtual Communities: Identity, Polarisation, and the Web 2.0," *Ethics and Information Technology* 10: 41-56.

Puri, Anuj. 2021. "The Right to Attentional Privacy," *Rutgers Law Record* 48 (1): 206–221.

Srikanth, Anagha. 2021. "12 prominent people opposed to vaccines are responsible for two-thirds of anti-vaccine content online: report" *The Hill* March 21, 2021. https://thehill.com/changing-america/well-being/prevention-cures/544712-twelve-anti-vaxxers-are-responsible-for-two/

Stroud, Scott. 2019. "Pragmatist Media Ethics and the Challenges of Fake News," *Journal of Media Ethics* 34 (4): 178-192.

Sunstein, Cass. 2017. *#Republic: Divided Democracy in the Age of Social Media*. Princeton, New Jersey: Princeton University Press.

—. 2021. *Liars: Falsehoods and Free Speech in an Age of Deception*. Oxford: Oxford University Press.

Thaler, R. and Sunstein, C. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven CT: Yale University Press.

Tran, Jasper L. 2015. "The Right to Attention," *Indiana Law Journal* 91: 1023–1064.

Walker, Mason, and Matsa, Katerina Eva. 2021. "News Consumption Across Social Media in 2021," *Pew Research Center* September 20, 2021 https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/

Wells, Georgia, Jeff Horwitz, and Deepa Seetharaman, 2021. "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show," *Wall Street Journal* September 14, 2021. https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739?mod=hp_lead_pos7&mod=article_inline

Wittgenstein, Ludwig. 1958. *Philosophical Investigations*. Oxford: Blackwell Publishing. Translation by G.E.M. Anscombe.

Zuckerberg, Mark. 2019. "Standing for voice and free expression," *Facebook* October 17, 2019 https://about.fb.com/news/2019/10/mark-zuckerberg-stands-for-voice-and-free-expression/

Notes to Contributors

1. All submitted papers are subject to anonymous peer-review, and will be evaluated on the basis of their originality, quality of scholarship and contribution to advancing the understanding of applied ethics and philosophy.

2. Papers should be at most 8,000 words, including references.

3. An abstract of 150-300 words and a list of up to 5 keywords should be included at the beginning of the paper.

4. Papers can be submitted any time of the year by e-mail to jaep@let.hokudai.ac.jp . If the authors wish their papers to be included in the next volume (to be published in March 2027), however, they are advised to submit their papers by September 15th 2026.

5. Submitted manuscripts may be published either as original papers or as discussion papers. If you wish your manuscript to be published as a discussion paper, please indicate this at the time of submission. Manuscripts submitted as original papers may, at the discretion of the editorial board, be published as discussion papers.

6. It is the journal's policy to request no more than one round of revisions after initial submission. Beyond that point, any submission will normally be either accepted or rejected outright.

7. In-text references should be cited in standard author-date form: (Walzer 1977; Kutz 2004), including specific page numbers after a direct quotation, (Walzer 1977, 23-6).

8. A complete alphabetical list of references cited should be included at the end of the paper in the following style:

   Cohen, G.A. (1989), 'On the Currency of Egalitarian Justice', Ethics, 99 (4): 906-44.
   Kutz, C. (2004), 'Chapter 14: Responsibility', in J. Coleman and S. Shapiro (eds.), Jurisprudence and Philosophy of Law, Oxford, UK: Oxford University Press, 548-87.
   Walzer, M. (1977), Just and Unjust Wars: A Moral Argument with Historical Illustrations, New York: Basic Book.

9. Accepted papers will be published in electronic formats.

10. The Editorial Board reserves the right to make a final decision on publication.