

Consumer Trust Under Siege: Fake Reviews as Digital Deviance in Platform Economies

*Andrew Goodwin*¹, *Tina H. Deshotels*² and *Craig J. Forsyth*³

¹*University of Louisiana, Lafayette*

²*Jacksonville State University*

³*University of Louisiana, Lafayette. E-mail: craig.forsyth@louisiana.edu*

TO CITE THIS ARTICLE

Julie S. Costopoulos & Marshall A. Jones (2026). Mitigating Forensic Risk for Practitioners. *Journal of Crime and Criminal Behavior*, 6: 1, pp. 63-79.

Abstract: Consumer trust is a foundational element of the digital marketplace, where reviews serve as informal yet powerful systems of reputation and decision-making. However, the proliferation of fake reviews is strategically crafted, commercially motivated, and often algorithmically disguised, where algorithms are used to shape visibility, credibility, or authenticity, often without the user's awareness, thereby threatening consumer trust. This paper investigates fake reviews as a form of digital deviance, drawing on theories of social capital, deviance, and the commodification of public opinion. Using the New York Attorney General's 2013 *Operation Clean Turf* as a central case study, this research examines how fake reviews are produced, legitimised, and institutionalised within platform capitalism. To strengthen the analysis and address broader patterns in enforcement, we also draw on three additional regulatory cases—FTC v. Fashion Nova, FTC v. Sitejabber, and the FTC's 2024 nationwide fake-review rule—to situate *Operation Clean Turf* within a wider landscape of contemporary review fraud and regulatory response. Through qualitative analysis of legal documents, news coverage, and regulatory responses, we identify key themes: the industrialisation of deception, platform exploitation, economic rationality, and the normalisation of unethical behaviour. We also explore recent cases from 2024 involving companies like Vevor and Sitejabber, alongside the emerging threat of AI-generated reviews. Findings reveal that fake reviews are not isolated anomalies but systemic artefacts of a digital economy that incentivises manipulation over authenticity.

Keywords: Digital Deviance, Social Capital, Consumer Trust

Introduction

The digital marketplace has revolutionised how consumers evaluate products, engage with businesses, and make purchasing decisions. At the heart of this new economic environment is a crucial yet fragile asset: consumer trust. Reviews on platforms like Yelp, Google, and Amazon operate as informal reputation systems that guide consumer behaviour and shape market dynamics. Yet, the rise of fake reviews—intentionally deceptive endorsements or critiques—has placed this trust under siege. Indeed, the statistics indicating that “75% of consumers are concerned about fake reviews” (Bazaarvoice, 2023) and “39% of U.S. consumers who read reviews trust them less than they did five years ago” (CivicScience, 2024) underscore the growing crisis of confidence in online platforms.

This study explores the sociological dimensions of fake reviews, positioning them as a form of deviant behaviour facilitated by technological innovation, corporate power, and shifting norms of consumerism. We situate this phenomenon within the broader context of platform capitalism, an economic system in which digital platforms—such as Amazon, Yelp, and Google—act as central intermediaries that extract, monetise, and control data, labour, and visibility. In platform capitalism, user-generated content (like reviews) becomes a commodified asset, and algorithms serve as gatekeepers of reputation, trust, and economic opportunity. These algorithms not only determine which reviews are displayed most prominently but also assess the credibility of reviewers, prioritise certain businesses over others, and flag or remove content deemed suspicious. In doing so, algorithms mediate access to visibility and legitimacy, effectively shaping what consumers see, believe, and act upon. Their influence means that businesses and review writers often tailor content to appeal to algorithmic preferences, reinforcing the system’s logic and making it vulnerable to manipulation. Compounding the issue, Search Engine Optimisation (SEO) firms—companies that specialise in increasing online visibility through tactics such as keyword optimisation, content placement, and reputation management—often play a key role in this commodification process. In some cases, SEO firms engage in manipulative practices like generating or purchasing fake reviews to artificially boost a business’s digital reputation, blurring the line between legitimate marketing and deceptive amplification.

Drawing on theories of social capital, deviance, and the commodification of public opinion, we argue that fake reviews represent a form of digital deception

that undermines both interpersonal trust and the integrity of online communities. While *Operation Clean Turf* (2013) serves as our foundational case, we analyse it alongside several recent enforcement actions—including *FTC v. Fashion Nova* (2022), *FTC v. Sitejabber* (2024–2025), and the FTC’s 2024 nationwide rule banning fake reviews—to demonstrate how the underlying dynamics identified in the 2013 case persist and evolve across different organisational contexts and regulatory interventions. For instance, in 2024, Vevor, a \$500 million e-commerce tool company, was reported to be manipulating Amazon reviews and listings with little consequence (New York Post, 2024). That same year, the FTC charged Sitejabber, an online review platform, with deceptively inflating product ratings by incorporating pre-purchase opinions (The Verge, 2024). Compounding the issue is the rise of AI-generated fake reviews, which watchdogs report are increasingly difficult to detect and are flooding platforms across industries (AP News, 2024).

Our research centres on *Operation Clean Turf* not as an isolated incident, but as a lens through which to understand a persistent and evolving challenge in the digital economy. This paper seeks to answer the following research questions:

- How are fake reviews produced, distributed, and legitimised within the digital marketplace?
- What are the social, economic, and ethical consequences of review manipulation?
- How do formal and informal controls attempt to regulate this form of deviance?

We have organised the remainder of our paper as follows. First, we have reviewed the literature on trust, deviance, and platform capitalism. We then present our theoretical framework, integrating Social Capital Theory, theories of commodification, and deviance theory. After outlining our qualitative methodology, we have analysed the *Operation Clean Turf* case along five thematic lines: techniques of deception, algorithmic exploitation, rational deviance, corporate complicity, and regulatory response. Our findings reveal a pattern of strategic deception in which businesses, SEO firms, and freelancers collaborate to manipulate public perception by fabricating online reviews. The erosion of trust undermines the social capital that makes online communities function effectively (Coleman, 1988; Putnam, 2000), diminishing the legitimacy of peer-based systems. We have then discussed the

implications of our findings for theory, consumer behaviour, and policy. We finally conclude by offering directions for future research and structural reform.

Addressing these questions is important not only for scholars of technology and deviance but also for policymakers, platform designers, and consumers. As digital interactions increasingly mediate trust and reputation, understanding how these systems are subverted is vital to preserving social and economic integrity.

Literature Review

Trust, Technology, and Social Capital

Trust is foundational to both social relationships and market transactions. In the digital context, trust is institutionalised through user-generated content—reviews, ratings, and testimonials—which consumers use to navigate choices in the absence of face-to-face interactions (Resnick & Zeckhauser, 2002). These systems support what Fukuyama (1995) calls “high-trust environments,” where economic and social interactions flourish based on the assumption of honesty and reciprocity. However, the infiltration of fake reviews destabilises this trust infrastructure, undermining both platform integrity and consumer autonomy.

Social Capital Theory helps explain the significance of this erosion. Coleman (1988) and Putnam (2000) emphasised that social capital arises from norms, shared understandings, and mutual obligations within communities. Digital marketplaces, though virtual, operate as trust-dependent communities. When platforms are flooded with fabricated endorsements—whether by firms like Vevor, which reportedly manipulated Amazon listings in 2024, or by sites like Sitejabber, recently charged by the FTC for deceptive review practices—consumer trust and social capital decrease. These disruptions reflect a breakdown of the informal social controls that can reinforce authentic behaviour in online spaces.

Social capital theory provides a useful framework for understanding the relational dimensions of digital trust. Coleman (1988) and Putnam (2000) emphasised how norms of reciprocity, shared expectations, and community sanctions maintain trust within social networks. Fake reviews, however, violate these norms and exploit the symbolic currency of authenticity, diluting the value of online reputation and weakening collective confidence (Del Vicario et al., 2016).

Deviance, Anonymity, and Platform Capitalism

From a sociological perspective, fake reviews constitute a form of white-collar deviance (Sutherland, 1940)—a strategic, profit-driven manipulation of informational norms for organisational gain. Theorists such as Chambliss (1967) and Clinard & Quinney (1973) argued that corporate crime is often normalised and structurally embedded, facilitated by legal ambiguity and weak enforcement mechanisms.

Moreover, digital deception aligns with what some scholars have described as platform-enabled deviance (Marwick & Boyd, 2011), where technologies amplify both the opportunities and motivations for unethical behaviour. As digital commerce intensifies, businesses may perceive deceptive reviews as necessary tactics for visibility and growth, especially in highly saturated markets. These behaviours are normalised within corporate ecosystems and frequently go unpunished, as seen in the Vevor case, where manipulation persisted despite awareness from the host platform (New York Post, 2024).

Labelling Theory (Becker, 1963) further illuminates how certain acts are defined and sanctioned as deviant based not solely on their nature, but on power dynamics. Smaller sellers engaging in review manipulation may face swift penalties, while larger, high-revenue companies often receive leniency or avoidance of scrutiny. Deviance, in this sense, is differentially labelled depending on the actor's market position.

Moreover, the anonymity of online spaces enables what Marwick and Boyd (2011) term platform-enabled deviance, in which technology facilitates deceptive acts without fear of reputational or legal consequences. The growing prevalence of AI-generated fake reviews, now widespread across e-commerce and service platforms (AP News, 2024), highlights how automation accelerates and obscures deviance, making detection increasingly difficult. This, combined with the anonymity afforded by online environments, diminishes risk to reputations and weakens informal controls, enabling behaviours such as astroturfing—the artificial generation of grassroots support or criticism (Kang, 2010).

The Commodification of Public Opinion

The increasing commodification of opinion—where consumer feedback becomes a tradable, monetised form of data—reshapes the boundaries between genuine

expression and market strategy. In a digital economy where attention is profitable, and visibility is power, corporations are incentivised to manipulate reviews, blur the line between advertising and opinion, and outsource manipulation to low-wage, global labour markets (Graham et al., 2017; Roberts, 2019).

The Sitejabber case illustrates how reviews can be repackaged—incorporating pre-purchase sentiment as part of official ratings, thus commodifying opinion. Similarly, AI-generated reviews complicate authorship and expression, reducing the boundary between human feedback and algorithmic content generation (Graham et al., 2017), where content, such as reviews, comments, or product descriptions, is generated by artificial intelligence systems trained to mimic human language and sentiment. These systems can rapidly generate large volumes of persuasive, tailored content that appears organic but is strategically optimised to influence platform algorithms and consumer perception. In the context of platform capitalism, where digital visibility and reputation are commodified, this automation turns influence itself into a product—mass-produced, scalable, and detached from authentic consumer experience. From the perspective of deviance theory, such practices represent a new form of digital norm violation, where manipulation becomes normalised and embedded within platform logics. As businesses exploit these tools to fabricate credibility, the boundary between legitimate marketing and deceptive behaviour becomes increasingly blurred, challenging both regulatory frameworks and public trust.

This commodification process also echoes Marxist critiques of capitalism's tendency to subsume authentic human interactions into market logics. As Hearn (2010) notes, in the “publicity economy,” even sincerity becomes a performative commodity. In this context, fake reviews are not just lies—they are a structural feature of a system where truth is secondary to marketability.

Regulation and Control: Formal and Informal Mechanisms

The regulation of fake reviews relies on both formal legal frameworks and informal platform moderation. *Operation Clean Turf* exposed the limitations of traditional law enforcement in policing decentralised acts of deception. Indeed, these acts are deliberately concealed using algorithms, making them difficult to detect or distinguish from legitimate activity. By mimicking the patterns of authentic user behaviour—such as timing, tone, location diversity, or engagement—fake reviews can evade automated detection systems and human moderators alike. This allows

deceptive content to circulate widely while maintaining the appearance of legitimacy. Although the investigation led to \$350,000 in penalties for businesses and SEO firms, enforcement was reactive and symbolic.

Platforms like Yelp and Google deploy algorithmic filters and machine learning tools to detect fake reviews. Yet these systems are imperfect and often struggle to distinguish genuine from inauthentic engagement, especially when sophisticated actors use IP spoofing, AI-generated text, or review farming (Luca & Zervas, 2016), the coordinated production of large volumes of fake reviews by hired individuals or firms, often operating across borders and using multiple accounts and devices to simulate authentic user activity. These networks may consist of freelancers, bots, or outsourced labour that can rapidly generate ratings and comments tailored to exploit platform algorithms. The reliance on automated detection also raises concerns about false positives, platform bias, and labour exploitation, as moderation tasks are frequently outsourced to invisible, precariously employed content workers (Roberts, 2019).

Theoretical Framework

This study draws on three interrelated theoretical perspectives:

1. Social Capital Theory (Coleman, 1988; Putnam, 2000): Explains how trust is generated and maintained in social networks and how its erosion affects online communities.
2. Theories of Deviance:
 - Labelling Theory (Becker, 1963): Helps explain how certain behaviours, like fake reviews, are defined and sanctioned as deviant based on social context and power dynamics.
 - Corporate and Organisational Deviance (Clinard & Yeager, 1980): Provides a lens for understanding how businesses normalise unethical behaviour under pressure to compete in marketplaces driven by algorithms.
3. The Commodification of Public Opinion (Hearn, 2010; Fuchs, 2014): Highlights how digital capitalism transforms user expression into a marketable asset, incentivising manipulation.

Together, these theories frame fake reviews not simply as isolated acts of dishonesty, but as emergent phenomena shaped by social and technical infrastructures,

economic pressures, and shifting cultural norms around truth, visibility, and trust. These theoretical perspectives guide our analysis across multiple enforcement cases, enabling comparison between early forms of coordinated deception (e.g., Operation Clean Turf) and more recent platform-level or corporate review manipulation uncovered by the FTC.

Methods

In this study, we conducted a qualitative multi-case comparative analysis of fake online reviews, centring on Operation Clean Turf (2013) while examining three additional FTC enforcement cases to broaden the scope and demonstrate recurring patterns across time, platforms, and regulatory agencies.

Data Collection

1. **Document Analysis:** We collected and analysed official press releases, legal documents, and news articles related to Operation Clean Turf. Key sources include:
 - The New York Attorney General's press release detailing the investigation's findings and settlements.
 - News coverage from reputable outlets such as *The Guardian* and *Reuters*, providing insights into the operation's execution and outcomes.
2. **Content Analysis:** We examined the specific tactics employed by businesses and SEO firms to generate fake reviews, as reported in the collected documents. This involved identifying patterns in the methods used to create and disseminate deceptive content.
3. **Thematic Analysis:** Applying sociological theories—Social Capital Theory, Commodification of Public Opinion, and Theories of Deviance—we interpreted the broader social and economic implications of fake reviews. This theoretical lens allowed us to contextualise the findings within frameworks of trust, corporate behaviour, and consumerism.

Analytical Framework

- **Social Capital Theory (Coleman, 1988):** To assess how fake reviews impact trust and reciprocity within online communities.

- Theories of Deviance (Becker, 1963): To understand the normalisation and regulation of deceptive practices in digital marketplaces.
- Commodification of Public Opinion (Hearn, 2010): To explore how consumer feedback is transformed into a marketable asset, influencing corporate strategies.

Findings

The case study of Operation Clean Turf, led by the New York Attorney General's Office in 2013, offers a revealing look into the industrialisation of fake review production and its sociological implications. While Operation Clean Turf anchors our analysis historically, themes identified in this case also appear across the additional FTC actions analysed for this study, highlighting their persistence and evolution in the decade that followed. Drawing on Social Capital Theory (Coleman, 1988), we examine how these practices erode trust and reciprocity in online communities. The Commodification of Public Opinion (Hearn, 2010) provides a lens to understand how consumer feedback is transformed into a tradable asset that influences corporate reputation and market strategy. Finally, the theories of Deviance (Becker, 1963) help us interrogate the normalisation of deceptive online behaviours as rationalised and often institutionalised business practices. The findings are organised into five thematic areas:

Techniques of Deception: Astroturfing and Digital Masking

Using Becker's (1963) theory of deviance, the techniques employed in Operation Clean Turf illustrate how deceptive behaviours became normalised within digital marketing culture. Astroturfing—a method of fabricating grassroots support—was central. These reviews were algorithmically disguised as legitimate user feedback, violating norms of fair advertising.

- IP Spoofing and fake profiles were used to bypass algorithmic filters on platforms such as Yelp, Google Places, and CitySearch. For instance, Zamdel, Inc. generated over 1,500 fraudulent reviews using multiple IP addresses.
- Freelancer recruitment practices through Craigslist and oDesk (now Upwork) show the outsourcing of digital deviance to low-wage global labour markets.
- Sample advertisements reveal the routine nature of the work:
 - “Must have multiple IPs... Need to understand how Yelp filters work.”
 - “I will pay \$10 per review... must have at least 10 friends on Yelp.”

- “We are offering \$1.00 per post.”

These strategies represent a routinisation of deviance—where deception becomes embedded in digital marketing as a normalised and even professionalised task.

Platform Exploitation and Algorithmic Evasion

Informed by Hearn’s (2010) theory of commodification, this theme reveals how firms exploited platform algorithms to commodify visibility and reputation. Companies manipulated algorithmic systems to promote favourable content while evading detection.

- SEO firms like XVIO, Inc. designed review campaigns to bypass Yelp’s filters, including “secret shopper” programs offering free goods in exchange for positive reviews.
- The reviews were posted across multiple platforms—Yelp, Google Maps, CitySearch, Judy’s Book, InsiderPages, and ScoreLive—demonstrating a strategic, multi-platform approach to maximising reputational capital.

These actions illustrate the transformation of user trust into a marketable commodity, where visibility and reputation are bought and sold in digital economies.

Economic Incentives and Rational Deviance

This theme engages both Becker’s deviance theory and Coleman’s concept of social capital. Fake reviews were understood by offenders as rational economic behaviours, embedded in cost-benefit calculations. At the same time, these behaviours deteriorated trust—an essential form of social capital—within online communities.

- Legal penalties ranged from \$2,500 to \$100,000—relatively small compared to potential gains in customer acquisition and brand reputation.
- Examples of incentive-based deception include:
 - US Coachways, Inc.: Offered \$50 gift cards for undisclosed positive reviews.
 - Laser Cosmetics: Mobilised friends, employees, and SEO firms for coordinated review campaigns.

These cases illustrate how rational deviance was reinforced through economic incentives and minimal consequences, further undermining the trust that platforms rely upon to function effectively.

Corporate Complicity and Scope of Offenders

From the perspective of deviance normalisation, the investigation uncovered widespread corporate complicity. A diverse array of businesses—from medical services to nightlife—were implicated, suggesting that deceptive practices have become structural features of digital reputational economies.

Notable offenders:

- Zamdel, Inc. – Over 1,500 fake reviews.
- XVIO, Inc. – Paid for strategic shopper-based reviews.
- US Coachways, Inc. – Engaged employees and freelancers for fabricated posts.
- Laser Cosmetica – Used SEO firms and internal networks.
- Swam Media Group/Scores Media Group – Created 175 fake reviews for an NYC adult club.

Other offenders included: A&E Wig Fashions, Inc.; Platinum Dental (A.H. Dental P.C.); Body Laser Spa Inc.; Bread and Butter NY (La Pomme Nightclub); Metamorphosis Day Spa Inc.; Magic Smile, Inc.; The Web Empire, LLC; Webtools Internet Solutions; Stillwater Media Group; Medical Message Clinic and HerballYours.com; iSEOiSEO; Main Street Host; Envision MT Corp.

This widespread involvement across sectors underscores how corporate actors commodify consumer trust and rationalise deviant strategies in pursuit of market dominance.

Regulatory Response: A Limited Deterrent

Although the crackdown sent an important symbolic message, its effectiveness in restoring social capital and deterring future deviance was minimal. The relatively low fines and limited jurisdiction of enforcement allowed deviant practices to persist and evolve.

- Only New York-based firms were targeted, and enforcement was reactive, not systemic.
- Yelp and others have improved detection algorithms, but offenders adapt quickly and circumvent filters.
- The low cost of digital deception, combined with global labour access, keeps the fake review economy viable.

In sociological terms, formal sanctions failed to rebuild social capital or shift the normative framework that currently tolerates reputational manipulation.

Discussion

This study has examined the production, circulation, and consequences of fake reviews in the digital marketplace, focusing on the landmark case of Operation Clean Turf. By incorporating multiple cases—from a state-level 2013 sting to federal enforcement actions between 2022 and 2025—we demonstrate that the mechanisms identified in Operation Clean Turf are not unique but instead recur across different industries, platforms, and regulatory jurisdictions. Our findings reveal a pattern of strategic deception in which businesses, SEO firms, and freelancers collaborate to manipulate public perception by fabricating online reviews. These practices not only distort consumer experiences but also raise broader sociological questions around deviance, trust, and digital capitalism. Fake reviews are strategically manufactured and industrially produced, often through globally outsourced labour, and these deceptive practices are embedded in corporate operations across a wide range of industries. Technological savvy and platform loopholes are exploited to maximise impact and avoid detection, while economic rationality drives deviance, as the benefits of a positive reputation often exceed the risks of enforcement. These findings support the use of sociological theories to understand digital deception not as isolated misconduct, but as a systemic feature of platform capitalism.

The case illustrates how consumer feedback—once considered a form of peer-to-peer knowledge sharing—has become commodified. Platforms like Yelp, Google, and CitySearch operate in an attention economy where visibility is currency and reputations are assets to be bought, sold, and managed. In some cases, SEO firms engage in manipulative practices like generating or purchasing fake reviews to artificially boost a business's digital reputation, blurring the line between legitimate marketing and deceptive amplification. Platforms like Yelp, Google, and CitySearch operate in an attention economy where visibility is currency and reputations are assets to be bought, sold, and managed. Following Hearn (2010) and Fuchs (2014), we understand this commodification as part of a broader trend in which emotional expression, authenticity, and opinion are transformed into monetisable data.

This commodification not only incentivises manipulation but also challenges traditional boundaries between advertising and genuine consumer engagement.

As public opinion becomes a strategic resource, businesses are encouraged to manufacture it through artificial means, eroding the value of real consumer experiences and distorting the public sphere.

The Normalisation of Deviance

Deceptive review practices are not the result of rogue actors or isolated incidents. Rather, they reflect a normalisation of deviance within digital capitalism. As Becker (1963) and Clinard & Yeager (1980) suggest, deviance becomes normalised when it is rationalised, routinised, and embedded in organisational culture. The companies implicated in Operation Clean Turf often viewed fake reviews as standard tools for “reputation management” and as rational responses to competitive pressure.

The use of fake reviews became a techno-cultural script—a normalised behaviour that conforms to platform logic, economic incentives, and market expectations. In such a context, ethics are subordinated to metrics, and deceit becomes not only tolerated but strategically advantageous. This supports Becker’s labelling theory, which helps explain how actions that violate norms can be redefined as acceptable under certain economic and institutional conditions. It also highlights how social capital—the trust that underpins user communities—becomes both a target and a casualty of commodification. As digital trust is systematically undermined, the very basis for peer-to-peer interaction is eroded.

Regulatory and Platform Responses

Despite its symbolic impact, Operation Clean Turf reveals the limitations of current regulatory frameworks. The penalties issued were modest, enforcement was localised to New York, and the investigation was reactive rather than systemic. This aligns with Clinard and Quinney’s (1973) observation that corporate deviance often escapes meaningful punishment due to the complexity of enforcement and the ambiguity of legal standards.

The penalty was symbolic for several reasons. First, the fines—ranging from \$2,500 to \$100,000—were minimal compared to the economic benefits companies gained through enhanced visibility and increased consumer trust. Second, enforcement failed to produce structural reform, such as algorithmic transparency or regulation of review practices across platforms. Third, the operation was reactive rather than preventive, serving more as a public demonstration of enforcement than

a sustainable deterrent. Finally, within the framework of platform capitalism, where reputation is commodified and visibility drives profit, the low penalties and rare enforcement actually reinforced the logic of rationalised deviance.

Platforms like Yelp and Google deploy algorithmic filters and machine learning tools to detect fake reviews. Yet these systems are imperfect and often struggle to distinguish genuine from inauthentic engagement, especially when sophisticated actors use IP spoofing, AI-generated text, or review farming (Luca & Zervas, 2016). The reliance on automated detection also raises concerns about false positives, platform bias, and labour exploitation, as moderation tasks are frequently outsourced to invisible, precariously employed content workers (Roberts, 2019).

Many of these deceptive acts are algorithmically obfuscated, deliberately concealed within the logic of platform algorithms, making them difficult to detect or distinguish from legitimate activity. By mimicking the patterns of authentic user behaviour—such as timing, tone, location diversity, or engagement—fake reviews can evade automated detection systems and human moderators alike. This algorithmic camouflage allows deceptive content to circulate widely while maintaining the appearance of legitimacy, further complicating enforcement and reinforcing consumer uncertainty. Postmodern and Marxist theories help us understand how these forms of digital manipulation blur the line between authenticity and performance, revealing how truth itself becomes a malleable, marketable construct in the publicity economy.

Implications

This research has theoretical and practical implications. First, theoretically, this case underscores the utility of integrating theories of deviance, social capital, and critical media theory in analysing digital phenomena. First, Labelling Theory helps us understand how certain behaviours are criminalised while others are normalised depending on power and context. Secondly, Social Capital Theory highlights how trust is not only an interpersonal phenomenon but also a systemic asset in the digital economy. Finally, Postmodern and Marxist theories contribute to our understanding of how authenticity is constructed, packaged, and sold under late capitalism.

The continued relevance of these theories in analysing digitally mediated trust and deception signals a fertile ground for future sociological inquiry.

Practically, for consumers, fake reviews represent a breach of trust with real-world consequences. If reviews cannot be trusted, the foundation of online shopping—peer evaluation—collapses. This may result in increased consumer cynicism, where users discount all reviews as potentially fake, greater reliance on brand recognition or influencer endorsement, reinforcing power imbalances between large corporations and small businesses, and reduced platform loyalty, as users migrate to spaces they perceive as more trustworthy. The erosion of trust undermines the social capital that makes online communities function effectively (Coleman, 1988; Putnam, 2000), diminishing the legitimacy of peer-based systems.

Finally, there are also policy implications. First, our study calls for stronger enforcement mechanisms. Governments should expand the legal definition of false advertising to explicitly include fake reviews and increase penalties to create a meaningful deterrent. Secondly, regulatory agencies should partner with digital platforms to share data and enhance transparency. A centralised review, authenticity certification or disclosure standard may also be beneficial. Thirdly, there should be labour protections. The gig workers tasked with creating or moderating fake reviews operate in regulatory grey zones. Labour protections and ethical outsourcing policies should be part of any response to the fake review economy. Finally, there needs to be public education initiatives to help consumers recognise signs of fake reviews and make informed choices.

Limitations and Future Research

While this study offers a detailed examination of one major regulatory case, it also presents several limitations. Operation Clean Turf is a snapshot from 2013 and is geographically limited to New York State. Future research could explore whether similar enforcement efforts have been attempted in other jurisdictions or whether deceptive review practices have evolved since then. Additionally, we focused on platforms like Yelp and Google, but fake review economies also flourish on Amazon, Facebook, TikTok, and other social commerce platforms. Future studies should explore cross-platform dynamics and the role of platform design in facilitating or deterring deception. There is also a lack of the consumer perspective. Future research could investigate how consumers identify, interpret, and respond to fake reviews, and whether demographic factors (e.g., age, digital literacy) affect their vulnerability. Finally, there is a need for deeper ethnographic research into the lived

experiences of review writers and gig workers engaged in deceptive labour, especially in the Global South, where much of this work is outsourced. Although this study now includes multiple cases, they remain limited to U.S.-based regulatory actions, leaving open questions about comparative international enforcement or platform-specific variation outside government interventions.

Conclusion

Fake reviews are not just a nuisance or ethical lapse—they are a structural feature of platform capitalism. As companies chase visibility and algorithmic favour, deception becomes a rational strategy, and public opinion becomes a battleground of data manipulation. Tackling this problem requires more than technical fixes or occasional lawsuits; it demands a deeper reckoning with how digital economies commodify truth and erode the foundations of trust.

References

- Bazaarvoice. (2023). *Shopper experience index 2023*. <https://www.bazaarvoice.com/resources/shopper-experience-index-2023/>
- Becker, H. S. (1963). *Outsiders: Studies in the sociology of deviance*. Free Press.
- Chambliss, W. J. (1967). Types of deviance and the effectiveness of legal sanctions. *Wisconsin Law Review*, 1967(4), 703–719.
- CivicScience. (2024). *Consumer trust in online reviews declines*. <https://civicscience.com/consumer-trust-in-online-reviews-declines/>
- Clinard, M. B., & Quinney, R. (1973). *Criminal behaviour systems: A typology*. Holt, Rinehart and Winston.
- Clinard, M. B., & Yeager, P. C. (1980). *Corporate crime*. Free Press.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, S95–S120.
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., & Quattrocioni, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3), 554–559.
- Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. Free Press.
- Graham, M., Lehdonvirta, V., & Barnard, H. (2017). Digital labour and development: Impacts of global digital labour platforms and the gig economy on worker livelihoods.

- Transfer: European Review of Labour and Research*, 23(2), 135–162. <https://doi.org/10.1177/1024258916687250>
- Hearn, A. (2010). Structuring feeling: Web 2.0, online ranking and rating, and the digital ‘reputation’ economy. *Ephemera: Theory & Politics in Organisation*, 10(3/4), 421–438.
- Kang, C. (2010). Astroturfing and manipulation of online opinion. *The Washington Post*.
- Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, 62(12), 3412–3427.
- Marwick, A., & Boyd, D. (2011, September 21–24). The drama! Teen conflict, gossip, and bullying in networked publics [Conference presentation]. *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, University of Oxford, Oxford, England.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.
- Resnick, P., & Zeckhauser, R. (2002). Trust among strangers in internet transactions: Empirical analysis of eBay’s reputation system. *Advances in Applied Microeconomics*, 11, 127–157.
- Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press.
- Sutherland, E. H. (1940). White-collar criminality. *American Sociological Review*, 5(1), 1–12.