COMMENT

FACING CONTRACTING ISSUES: THE PSYCHOLOGICAL AND FINANCIAL IMPACTS OF FACEBOOK OUTSOURCING CONTENT MODERATION

CHLOÉ NURIK†

INTRODUCTION ........................................................... 1552
I. REGULATIONS AND FACEBOOK’S MODERATION SYSTEM .... 1554
   A. The Structure of Facebook’s Moderation System ................ 1556
   B. The Role of Moderators ................................................ 1557
II. DETRIMENTAL IMPACTS OF LABOR ARRANGEMENTS: HARMS TO MODERATORS .................................................. 1560
III. DETRIMENTAL IMPACTS OF CONTRACTUAL LABOR ARRANGEMENTS: FINANCIAL & LEGAL RISKS TO PLATFORMS 1567
   A. Worker Classification in the Gig Economy ....................... 1567
   B. Prevailing Worker Classification Tests ............................. 1571
   C. Applying Prevailing Classification Tests to Moderators ..... 1575
   D. Risks of Misclassification: Lawsuits & Fines .................. 1578
IV. REFORMS ..................................................................... 1581
   A. Rationales & Stakes of Reform ....................................... 1581
   B. State Legislators ............................................................ 1583
   C. Platforms ....................................................................... 1585
   D. Moderators & Civil Society ............................................ 1586
CONCLUSION .................................................................... 1587

† I would like to offer my sincerest thanks to Professor Amanda Shanor for her guidance on this Comment as well as her continuous support and encouragement. Additionally, thank you to the talented staff and editors of the University of Pennsylvania Law Review for their hard work on this Comment. Lastly, I am grateful for interviewees who shared their stories and provided valuable insights into the processes and labor dynamics of content moderation.
INTRODUCTION

Historically, the editors of print media served as gatekeepers, controlling content and news dissemination. Communication generally flowed one-way, as readers obtained material, but could not directly create their own content or publish it in channels of mass distribution. New technology altered these dynamics, and Web 2.0 ushered in an era of user-generated content. In particular, the rise of social media platforms supposedly heralded a time of unmediated communication when individuals could wrest control away from traditional gatekeepers. However, amid the “[l]ong [h]angover of Web 2.0,” it has become increasingly apparent that social media sites are not unfiltered and unregulated spaces. Instead, “a variety of covert corporate and institutional constraints . . . shape what appears online and what is concealed from public view” through platform content moderation.

Changes in the scope and scale of modern communications have necessitated an evolution in gatekeeping strategies. For example, users share an average of nearly five billion items on Facebook each day. Further, more than three million items are reported daily by Facebook users or flagged by artificial intelligence (AI) screening systems for moderators to review. The enormous volume and instantaneity of user-generated content have given rise to different labor practices than those deployed in traditional media. Rather than directly hiring moderators, platforms such as Facebook rely on

1 See David Manning White, The “Gate Keeper”: A Case Study in the Selection of News, 27 JOURNALISM Q. 383, 384 (1950) (analyzing a wire editor’s selection or rejection of news stories).
3 See Edward Lee, Warming Up to User-Generated Content, U. ILL. L. REV. 1459, 1460 (2008) (“From blogs to wikis to podcasting to ‘mashup’ videos and social networking sites like Facebook and MySpace, the Web 2.0 culture encourages users to engage, create, and share content online.”).
4 See JASON GAINOUS & KEVIN M. WAGNER, TWEETING TO POWER: THE SOCIAL MEDIA REVOLUTION IN AMERICAN POLITICS 11 (2013) (“Social media has no obvious gatekeeper. Anyone can join and participate given sufficient knowledge and resources.”).
Facing Contracting Issues

moderators provided by third-party vendors.9 While scholarship and public discourse have addressed a variety of issues related to content moderation,10 the labor arrangements involved in this process have not yet been foregrounded, but are nevertheless essential to understanding the dynamics and stakes of platform governance.

This Comment explores the labor arrangements of content moderators who work for vendors that partner with Facebook. The analysis is grounded in original interview data with former and current moderators.11 Facebook was selected as this Comment’s focus due to its size, number of users, hyper-curated nature, and outsized influence on the social media industry. Examining the work arrangements of this site’s moderators provides a window into platform employment issues more generally. In this Comment, I argue that Facebook's system of contracting out content moderation to third-party vendors—rather than hiring moderators directly—engenders psychological impacts to workers and presents legal and financial risks for the platform. On a broader level, contracting out moderation underscores the deep inequities baked into the gig economy and digital work,12 the growing complexity and precarity of modern labor arrangements,13 and the need for greater regulation of tech industries.14

This Comment is divided into several parts. Part I examines Facebook's moderation system. I outline the regulatory underpinnings of this system and describe the platform's labor arrangements with vendors and moderators. Part II discusses the detrimental impacts of Facebook's employment structure on moderators with respect to labor conditions and workplace support. Part III

11 To protect interviewees, I use pseudonyms (indicated by the * symbol) for the names of moderators and former moderators quoted in this Comment.
12 See, e.g., Arianne Renan Barzilay & Anat Ben-David, Platform Inequality: Gender in the Gig-Economy, 47 SETON HALL L. REV. 393, 431 (2017) (“G]ender inequality is re-configured and reproduced through platform-facilitated labor . . . .”).
14 See Nikos Koutsimpogiorgos, Jaap van Slageren, Andrea M. Herrmann & Koen Frenken, Conceptualizing the Gig Economy and Its Regulatory Problems, 12 POL’Y & INTERNET 525, 527 (2020) (presenting four key regulatory pillars related to the gig economy).
explores how the outsourcing of moderation poses financial, legal, and reputational risks to Facebook. Lastly, Part IV considers leading reform proposals and stresses the urgent need for regulatory change and social activism.

I. REGULATIONS AND FACEBOOK’S MODERATION SYSTEM

Social media sites face few legal restrictions in their regulation of user content. In the United States, Section 230 of the Communications Decency Act of 1996 immunizes platforms from civil liability for both user-generated content and the moderation of such material. Although there are several carveouts to Section 230 immunity, courts have interpreted this statute broadly, treating it as a “sacred cow.” In contrast, several nations outside the United States have more extensive regulations. For example, the European Union’s Code of Conduct on Countering Illegal Hate Speech Online requires companies to review and, if necessary, remove illegal hate speech within twenty-four hours of its posting. Beyond this transnational code, individual countries have taken up the mantle of oversight in a process akin to a regulatory arms race. Germany’s Network Enforcement Act requires social media sites to remove “manifestly unlawful” material, including defamatory and terrorist content, within twenty-four hours of posting or face steep penalties.

---

15 47 U.S.C. § 230(c)(2)(A). Notably, Section 230 immunity does not require platforms to moderate online content, but rather immunizes sites if they choose to moderate. See, e.g., Green v. Am. Online (AOL), 318 F.3d 465, 472 (3d Cir. 2003) (“Section 230(c)(2) does not require AOL to restrict speech; rather it allows AOL to establish standards of decency without risking liability for doing so.” (footnote omitted)).

16 Section 230 immunity is subject to several (small) exceptions. Companies must follow a notice-and-takedown procedure when instances of copyright infringement occur. See 17 U.S.C. § 512. Additionally, immunity does not apply to federal criminal laws, state laws “consistent with” Section 230, charges for sex trafficking, or violations of the Electronic Communications Privacy Act of 1986. See 47 U.S.C. § 230(e).

17 To this point, courts have “treated § 230(c) immunity as quite robust, adopting a relatively expansive definition of ‘interactive computer service’ and a relatively restrictive definition of ‘information content provider.’” Carafano v. Metrosplash.com, Inc., 339 F.3d 1119, 1123 (9th Cir. 2003) (footnotes omitted); see also Universal Commc’n Sys., Inc. v. Lycos, Inc., 478 F.3d 413, 419 (1st Cir. 2007) (“Section 230 immunity should be broadly construed.”). Such immunity has seemingly transcended issues of third-party speech by immunizing matters concerning platform design, infrastructure, and business practices. See, e.g., Langdon v. Google, Inc., 474 F. Supp. 2d 622, 631 (D. Del. 2007) (extending Section 230 immunity to cover Google’s refusal to run particular ads).


fines.\textsuperscript{20} Similarly, Australia’s \textit{Sharing of Abhorrent Violent Material Act} stipulates that platforms must remove or stop hosting audio-visual content depicting terrorist acts, murders, torture, kidnaps, or rapes “expeditiously” or be fined up to ten percent of their annual turnover.\textsuperscript{21} Lastly, the United Kingdom is in the process of creating a statutory duty of care to protect users (especially children) and to tackle illegal or harmful content and activities online.\textsuperscript{22} These initiatives may provide inroads to the problem of self-regulated content moderation by installing external enforcement mechanisms. However, certain platforms have already attempted to bypass these requirements, perhaps suggesting that some companies view the payment of regulatory fines as more cost-effective than full regulatory compliance.\textsuperscript{23}

The paucity of external oversight has empowered platforms to enact private governance of speech.\textsuperscript{24} Platforms such as Facebook design their own community guidelines and enforce the same standards across the globe.\textsuperscript{25} There is no meaningful system of checks and balances or due process on these sites.\textsuperscript{26} Instead, “some nontransparent form of private governance or

\textsuperscript{22} Online Safety Bill 2022–23, HL Bill [87] (Eng.).
\textsuperscript{24} See Laura DeNardis & Andrea M. Hackl, Internet Governance by Social Media Platforms, 39 TELECOMMS. POL’Y 761, 762 (2015) (“This inquiry . . . . addresses governance by social media rather than governance of social media.”).
\textsuperscript{25} See Chloé Nurik, "Men Are Scum": Self-Regulation, Hate Speech, and Gender-Based Censorship on Facebook, 13 INT'L J. COMM. 2878, 2880 (2019) (critiquing Facebook's policy of applying the same content standards globally).
\textsuperscript{26} See Nicolas Suzor, Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms, SOC. MEDIA & SOC’Y, Sept. 2018, at 8 (“[U]sers have no realistic chance of challenging decisions made by platforms in any formal legal process.”). Facebook’s Oversight Board represents one recent effort to provide users with a formal appeals process. The Oversight Board began hearing a limited number of high-interest cases of content removal in 2020. See Brian Fung, Facebook’s Oversight Board Is Finally Hearing Cases, Two Years After it was First Announced, CNN (Oct. 22, 2020, 12:45 PM), https://www.cnn.com/2020/10/22/tech/facebook-oversight-board/index.html [https://perma.cc/GD8B-WGUY] (noting that Facebook’s Oversight Board was
bureaucracy serves as prosecutor, judge, jury, and executioner.” Scholarship has largely focused on the implications of these “Constitution-free zones” for users but has neglected how the self-regulatory speech control of platforms impacts another key population: moderators.

A. The Structure of Facebook’s Moderation System

To enforce its standards, Facebook relies on users, algorithms, and an international team of content moderators. Facebook initially deployed Harvard students as volunteer moderators, but five years later created a content moderation division. The platform subsequently developed a sophisticated and sprawling moderation process, revealing that “[f]or large-scale platforms, moderation is industrial, not artisanal.”

Facebook’s first line of defense is comprised of users who are deputized as “a volunteer corps of regulators” to report content on the platform that allegedly violates communal guidelines. Users also serve as group administrators, managing and moderating content in Facebook groups. Beyond this initial line of defense, AI proactively detects certain types of content violations such as terrorist content and child exploitation material. While Facebook plans to increasingly deploy AI for content moderation, existing technology is limited and introduces a host of problems, including errors, decontextualization, and biases, which “tend to disproportionately affect already marginalized groups and speakers.” These issues became more pronounced during the COVID-19 pandemic when platforms such as

prepared to hear cases for the first time). However, there are a host of factors that limit the effectiveness of this governance structure. For a discussion of the problems with Facebook’s Oversight Board, see Nuriik, Censored, Commodified, and Surveilled, supra note 6, at 108.


28 Benjamin F. Jackson, Censorship and Freedom of Expression in the Age of Facebook, 44 N.M. L. REV. 121, 166 (2014).


30 GILLESPIE, CUSTODIANS OF THE INTERNET, supra note 5, at 77.


32 See Joseph Seering, Tony Wang, Jina Yoon & Geoff Kaufman, Moderator Engagement and Community Development In the Age of Algorithms, 21 NEW MEDIA & SOC’Y 1417, 1418 (2019) (conceptualizing how volunteer moderators engage with their communities).


Facebook increasingly relied on AI to manage content. This reliance caused the number of content removals for child exploitation and self-harm material to plummet, revealing that “[e]ven the most sophisticated algorithms and machine-learning tools still can’t replicate the computing power of an army of human content moderators.”

**B. The Role of Moderators**

At the heart of Facebook’s moderation system is a global army of moderators. But this army is grossly inadequate in number: Only 15,000 moderators worldwide police content for nearly two and a half billion users (a ratio of one to 160,000). This small group of moderators are routinely overworked and exposed to harmful content, toiling under highly stressful working conditions without adequate protection or support.

Although moderators play an essential role in the social media ecosystem, platforms are intentionally opaque about their labor practices. Social media sites “purposefully mask the processes of content moderation to downplay their curatorial function, to maintain their ideological positioning as venues of open expression, and to avoid being held accountable (and even liable) for the millions of subjective (and at times discriminatory) decisions adjudicated each day.” As a result, the public receives little information about the “silent,

---

35 See Issie Lapowsky, *How COVID-19 Helped—and Hurt—Facebook’s Fight Against Bad Content*, PROTOCOL (Aug. 11, 2020), https://www.protocol.com/covid-facebook-content-moderation [https://perma.cc/DVE4-KMZ2] (describing the impacts of Facebook’s decision to deploy more AI for content moderation during the pandemic and noting that with increased reliance on AI, the company was “less equipped” to “catch” bad content).


37 BARRETT, WHO MODERATES THE SOCIAL MEDIA GIANTS?, supra note 8, at 4, 8.

38 See Elizabeth Dwoskin, *Facebook Content Moderator Details Trauma That Prompted Fight for $52 Million PTSD Settlement*, WASH. POST (May 13, 2020, 4:00 PM) [hereinafter Dwoskin, Facebook Content Moderator Details Trauma], https://www.washingtonpost.com/technology/2020/05/12/facebook-content-moderator-ptsd [https://perma.cc/J52H-U9ZD] (“Workers have spoken with The Washington Post regarding suffering from PTSD and other forms of psychological trauma, such as paranoid ruminations, frequent nightmares and an inability to sleep.”).


secret workforce” who decide which content is removed from the world’s largest platforms. By treating moderation as “ghost work,” platforms such as Facebook naturalize the presence of problematic content (making it seem organic rather than curated and calculated) and disguise the treatment of moderators, blocking investigations and reforms.

Furthermore, the hidden labor of content moderation is inherently feminized, both literally and metaphorically. Although platforms such as Facebook do not publish demographic information about moderators, it is suspected by both interviewees and researchers that women and members of minority groups comprise a sizable portion of the content moderation workforce, similar to gig workers generally. The demographics of moderators sharply contrast with full-time social media employees who largely embody “the mythical white, male ‘brogrammer.’” In this manner, while individuals from historically marginalized groups may work as moderators, a group of mostly nondiverse Silicon Valley employees set content policies and moderation standards, highlighting the “digital ceiling” for women and minorities in tech industries. Additionally, moderation is feminized because of its “invisibility, lower pay, and marginal status within the technology sector,” traits that are often associated with women’s labor.

---

43 See Sarah T. Roberts, Commercial Content Moderation: Digital Laborers’ Dirty Work, in The Intersectional Internet: Race, Sex, Class, and Culture Online 147, 149 (Safiyya Umoja Noble & Brendesha M. Tynes eds., 2016) (“When those in-house processes are kept from users, as they almost always are, the logic for the existence of the material on a site becomes opaque and the content becomes normalized.”).
44 See Interview with Hazel*, Former Content Moderator (Jan. 31, 2021) [on file with author] [hereinafter Hazel* Interview] (describing the demographics of Facebook moderators); see also Melissa Gregg & Rutvica Andrijasevic, Virtually Absent: The Gendered Histories and Economies of Digital Labour, 123 FEMINIST REV. 1, 2 (2019) [hereinafter Gregg & Andrijasevic, Virtually Absent] (contextualizing gendered online content review as part of a longer trajectory of women’s hidden labor).
46 Gregg & Andrijasevic, Virtually Absent, supra note 44, at 3.
47 Sarah T. Roberts, Behind the Screen: Content Moderation in the Shadows of Social Media 93–94 (2019) (“[I]ts policies regarding user-generated content were developed in the specific and rarefied sociocultural context of educated, economically elite, politically libertarian, and racially monochromatic Silicon Valley, USA.”).
Content moderation also heavily relies on outsourcing, a labor dynamic that has historically perpetuated the repression of workers who identify as women (especially women of color). Thus, the labor of moderation both reflects and reinforces the minoritizing of social media employment and the gig economy, entrenching inequalities rather than dismantling societal hierarchies.

The precariousness of content moderation is further attributable to the contingent and contractual nature of this work. Facebook mostly farms out content moderation to third-party vendors (e.g., Upwork, Sutherland, Deloitte, Accenture, Cognizant, Genpact, and Pro Unlimited) that hire, train, and manage moderators. Although the number of moderators in the United States is growing, many vendors outsource moderation to countries such as the Philippines, India, Ireland, Portugal, Spain, Germany, Latvia, Kenya, Mexico, and Turkey. Most of these international workers are classified as “tier 3 moderators,” who review the bulk of reported content and escalate concerns to “tier 2 moderators,” who act as supervisors either remotely or at local centers. In contrast, “tier 1 moderators,” who are usually lawyers and policymakers, create and revise content moderation policies. These employees are primarily located at Facebook’s headquarters in the United States. Moderation is thus outsourced both geographically and structurally, as moderators are distanced from Facebook through a web of contractual relationships across the globe. Moderators, journalists, and scholars have considered the following:

50 See generally GENDERED COMMODITY CHAINS: SEEING WOMEN’S WORK AND HOUSEHOLDS IN GLOBAL PRODUCTION (Wilma A. Dunaway ed., 2014) (presenting a collection of works that provide a historical and international analysis of women’s involvement in outsourcing).


52 See, e.g., Klonick, The New Governors, supra note 29, at 1639-41, at 1639-41 (listing the companies that handled outsourced moderation from Facebook); BARRETT, WHO MODERATES THE SOCIAL MEDIA GIANTS?, supra note 8, at 12 (same); Satariano & Isaac, The Silent Partner, supra note 9 (describing Facebook’s outsourcing arrangement with Accenture); Casey Newton, The Trauma Floor: The Secret Lives of Facebook Moderators In America, VERGE (Feb. 25, 2019, 8:00 AM) [hereinafter Newton, The Trauma Floor] https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona [https://perma.cc/HTN5-T39F] (noting how a former moderator sued the vendor Pro Unlimited).

53 See Newton, The Trauma Floor, supra note 52 (“Until recently, most Facebook content moderation has been done outside the United States. But as Facebook’s demand for labor has grown, it has expanded its domestic operations to include sites in California, Arizona, Texas, and Florida.”).

54 See Klonick, The New Governors, supra note 29, at 1640 (describing the geographical reach of content moderators); see also BARRETT, WHO MODERATES THE SOCIAL MEDIA GIANTS?, supra note 8, at 3 (same).


56 Id. at 1641.

57 Id.
increasingly drawn attention to the irony at the core of Facebook’s content moderation system: “[C]ontent moderation seems so important to running Facebook that it ought to be regarded as falling within the company’s core activities, not as an ancillary chore to be handled by contractors.”

Nevertheless, Facebook heavily relies on a contractual model to save money and cloak problematic labor relations.

II. DETERIMENTAL IMPACTS OF LABOR ARRANGEMENTS: HARMs TO MODERATORS

The Byzantine structure of Facebook’s content moderation system intentionally minimizes accountability, liability, and visibility. In this manner, the layered nature of this labor arrangement insulates Facebook and vendors, allowing them to establish problematic work conditions. In this Part, the Comment discusses issues that moderators face on the job, such as time constraints, high accuracy requirements, exposure to disturbing content without adequate psychological support, and the culture of secrecy.

Moderators experience job instability and the never-ending pressure to meet stringent quotas and deadlines. Dillon*, a moderator, explained that the moderators at his location are all hired on a one-year contract, with the threat of unemployment acting as a “looming sword over your head.” To remain employed, moderators must meet specified quotas, reviewing hundreds of posts, profiles, and videos per day. While Facebook refutes that vendors enforce quotas, interviewees, journalists, and scholars have confirmed the use of quotas and reported that moderators are questioned by their supervisors if they do not reach target numbers. Put on “a virtual

---

58 BARRETT, WHO MODERATES THE SOCIAL MEDIA GIANTS?, supra note 8, at 18.
59 See Interview with Dillon*, Content Moderator (Sept. 19, 2020) (on file with author) [hereinafter Dillon* Interview] (discussing the motivations behind Facebook’s contractual system); see also Hazel* Interview, supra note 44 (same); Interview with Vincent*, Former Engineer, Facebook (Feb. 2, 2021) (on file with author) [hereinafter Vincent* Interview] (same); Interview with Eric*, Former Content Moderator (Feb. 2, 2021) (on file with author) [hereinafter Eric* Interview] (same).
60 Quotas have proven problematic in other fields as well, such as policing. See Shaan Ossei-Owusu, Police Quotas, 96 N.Y.U. L. REV. 529, 535 (2021) (“[P]olice quotas shape the enforcement of criminal laws by introducing a host of perverse incentives into an already insecure body of criminal procedure.”).
61 Dillon* Interview, supra note 59.
62 See Newton, The Trauma Floor, supra note 52 (“[A moderator] will spend less than 30 seconds on each item, and … will do this up to 400 times a day.”).
64 See, e.g., Dillon* Interview, supra note 59 (“Our quota is around 200 a day.”); Vincent* Interview, supra note 59 (discussing quotas); Eric* Interview, supra note 59 (noting that he met his
assembly line,” 65 moderators are given strict time limits per piece of content, which can be as brief as ten to thirty seconds. 66 As a result, “[e]ach complaint is . . . getting just a sliver of human attention.” 67 This rapid pace is made even more challenging by the decontextualized nature of content review, wherein moderators are given few data points and are asked to “interpret things in the way almost like a robot would.” 68

Further compounding this problem, moderators must achieve a certain level of accuracy (usually around ninety-five percent to ninety-eight percent), which is measured by comparing a subset of a moderator’s decisions with those of a quality assurance (QA) analyst who serves in a supervisory capacity. 69 Although vendors such as Accenture give moderators time to improve their performance, interviewees expressed frustration that QA analysts are often misinformed about content policies, leading to discrepancies in moderation decisions that erroneously lower their accuracy scores. 70 The requirement to maintain a high level of accuracy has even led to moderators lobbying QAs to change their decisions “in a ritual known as ‘getting your points back.’” 71 In one case, a QA analyst reported that he “would sometimes return to his car at the end of a [workday] to find moderators waiting for him . . . . to intimidate him into changing his ruling,” 72 revealing the intense pressure moderators face to maintain a high level of accuracy.

daily quotas but that it was sometimes difficult for other moderators to do the same); Newton, The Trauma Floor, supra note 52 (“[A particular moderator] does not have a quota of posts to review, [but] managers monitor his productivity, and ask him to explain himself when the number slips into the 200s.”)

65 Roberts, Digital Detritus, supra note 39.

66 Interview with Bridget*, Former Content Moderator (Oct. 14, 2020) (on file with author) [hereinafter Bridget* Interview]; see also Olivia Solon, To Censor or Sanction Extreme Content? Either Way, Facebook Can’t Win, GUARDIAN (May 23, 2017, 1:00 PM), https://www.theguardian.com/news/2017/may/22/facebook-moderator-guidelines-extreme-content-analysis [https://perma.cc/N83N-TGYZ] (noting that moderators have about ten seconds to review each piece of content); Newton, The Trauma Floor, supra note 52 (noting that moderators spend less than thirty seconds per piece of content).

67 GILLESPIE, CUSTODIANS OF THE INTERNET, supra note 5, at 121.

68 Eric* Interview, supra note 59.

69 Vincent* Interview, supra note 59; see also Newton, The Trauma Floor, supra note 52 (“The company has set an accuracy target of 95 percent, a number that always seems just out of reach.”); Jason Koebler & Joseph Cox, The Impossible Job: Inside Facebook’s Struggle to Moderate Two Billion People, VICE (Aug. 23, 2018, 5:15 PM), https://www.vice.com/en/article/xwk9zd/how-facebook-content-moderation-works [https://perma.cc/A2J2-BMF2] (“Moderators face consequences if they fall below a 98 percent accuracy rate . . . .”)

70 See, e.g., Bridget* Interview, supra note 66 (expressing frustration with the QA system); Hazel* Interview, supra note 44 (same); Eric* Interview, supra note 59 (same); BARRETT, WHO MODERATES THE SOCIAL MEDIA GIANTS?, supra note 8, at 17 (“The QAs ‘calls often amount to highly subjective second-guessing.’”)

71 BARRETT, WHO MODERATES THE SOCIAL MEDIA GIANTS?, supra note 8, at 17.

72 Newton, The Trauma Floor, supra note 52.
Moderators have critiqued the relentless drive for efficiency and accuracy in content moderation. Everlee*, a former content moderator, explained: “The entire call center setup is just not conducive to doing content moderation because it’s got a huge emphasis on numbers, less focus on . . . the people actually doing the work.”73 In this manner, the frantic tempo of moderation and the micromanaged nature of the work (where each second, even down to bathroom breaks, is logged) are at odds with the subjective, emotional labor fueling content review.74

Due to these pressures, turnover and burnout rates are high in this field.75 In this metrics-obsessed environment, moderators have reported feeling replaceable and devalued.76 Former content moderator Eric* described how vendor Accenture treated moderators in a dehumanizing way, focusing exclusively on their productivity: “They don’t see us as people anymore when all they’re looking at are the metrics that are involved and whether they get a bonus or whether their contract gets renewed.”77 It is important to note that although moderators are hired by third-party vendors (and not directly by platforms), sites such as Facebook are complicit in these arrangements, financially benefitting from this exploitative system.

In addition to experiencing stressful working conditions, content moderators undergo psychological stress as a result of their daily exposure to harmful and lurid content on the platform. Facebook moderators routinely

---

73 Interview with Everlee*, Former Content Moderator (Nov. 5, 2020) (on file with author) [hereinafter Everlee* Interview].

74 See David Gilbert, Facebook Is Forcing Its Moderators to Log Every Second of Their Days—Even in the Bathroom, VICE (Jan. 9, 2020, 10:20 AM) [hereinafter Gilbert, Facebook Forcing Moderators to Log Their Days], https://www.vice.com/en/article/23bee8/facebook-moderators-lawsuit-ptsd-trauma-tracking-bathroom-breaks [https://perma.cc/YEK5-GNWV] (describing how moderators are required to clock out for bathroom breaks). The time of moderators is so micromanaged that there have been reports of moderators throwing up into trash cans by their desks instead of taking bathroom breaks. See Casey Newton, Bodies in Seats, VERGE (June 19, 2019, 8:00 AM) [hereinafter Newton, Bodies in Seats], https://www.theverge.com/2019/6/19/18681845/facebook-moderator-interviews-video-trauma-ptsd-cognizant-tampa [https://perma.cc/3KKH-G5RY] (“And so to avoid receiving an ‘occurrence,’ as the company calls unapproved absences, contractors who have exhausted their break time come to work sick—and occasionally vomit in trash cans on the production floor.”).

75 See Interview with Waverly*, Former Content Moderator (Sept. 25, 2020) (on file with author) [hereinafter Waverly* Interview] (“Folks generally aren’t in that job for more than two years.”); see also Everlee* Interview, supra note 73 (“[B]ecause the turnover in the field is so high, you’re considered a long-term moderator if you stay for six months.”); Lauren Weber & Deepa Seetharaman, The Worst Job in Technology: Staring at Human Depravity to Keep It Off Facebook, WALL ST. J. (Dec. 27, 2017, 10:42 PM), https://www.wsj.com/articles/the-worst-job-in-technology-staring-at-human-depravity-to-keep-it-off-facebook-1514398398 [https://perma.cc/R4CN-8AKj] (“Turnover is high, with most content moderators working anywhere from a few months to about a year before they quit or their assignments ended . . . .”).

76 See Newton, Bodies in Seats, supra note 74 (“More than anything else, the contractors described an environment in which they are never allowed to forget how quickly they can be replaced.”).

77 Eric* Interview, supra note 59.
view content including suicides, murders, beheadings, graphic pornography, child abuse and grooming, bestiality, animal cruelty, and other materials which are “not normal things for humans to process.”

Consequently, interviewees described the work of moderation as “drowning,” “traumatizing,” and “basically a death sentence.” Moderators’ repeated exposure to such material has been linked to depression, anxiety, nightmares, and in some cases, post-traumatic stress disorder (PTSD). Aware of the emotional toll that this work involves, Accenture requires content moderators to sign a form acknowledging that their job may trigger PTSD in an attempt to absolve the company of responsibility.

The lack of psychological support for moderators further reveals the exploitative nature of this work. For example, after former moderator Hazel* reported mistreatment by a supervisor, she was reassigned to work on the child exploitation queue, an emotionally taxing job for which she was provided little mental health support and few resources. Interviewees have also reported that they have access to only a handful of counseling sessions and are required to pay for their own health insurance, despite their starting wage being only fourteen percent of the median salary for Facebook employees. In contrast to Facebook employees who receive an array of health benefits, moderators have difficulty receiving and paying for the help they need.

---

78 Waverly* Interview, supra note 75; see also Newton, The Trauma Floor, supra note 52 (explaining that moderators are routinely exposed to graphic material including murders and bestiality); Dwoskin, Facebook Content Moderator Details Trauma, supra note 38 (noting that moderators are exposed to images of gang rape, child abuse, and suicide).

79 Waverly* Interview, supra note 75.

80 Eric* Interview, supra note 59.

81 Hazel* Interview, supra note 44.

82 See Dwoskin, Facebook Content Moderator Details Trauma, supra note 38 (discussing how moderators have experienced PTSD); see also Miriah Steiger, Timir J. Bharucha, Sukrit Venkatagiri, Martin J. Riedl & Matthew Lease, The Psychological Well-Being of Content Moderators: The Emotional Labor of Content Moderation and Avenues for Improving Support, PROC. OF 2021 CHI CONF. ON HUM. FACTORS IN COMPUTING SYS., May 2021, at 4 (explaining that moderators may be at risk for developing anxiety and depression).

83 See Madhumita Murgia, Facebook Content Moderators Required to Sign PTSD Forms, FIN. TIMES (Jan. 26, 2020), https://www.ft.com/content/98aad2fo-3e9f-11ea-a012-bae547046f75 [https://perma.cc/7CRW-J6T3] (noting that moderators at a European Accenture site were required to sign forms acknowledging the risks of developing PTSD); see also Casey Newton, YouTube Moderators Are Being Forced To Sign A Statement Acknowledging The Job Can Give Them PTSD, VERGE (Jan. 24, 2020, 10:15 AM), https://www.theverge.com/2020/1/24/21075830/youtube-moderators-ptsd-accenture-statement-lawsuits-mental-health [https://perma.cc/S6XZ-FQCG] (discussing Accenture’s use of these acknowledgement forms as “part of an apparent legal defense”).

84 Hazel* Interview, supra note 44.

85 Id.; see also Eric* Interview, supra note 59; Elizabeth Dwoskin, Inside Facebook, the Second-Class Workers Who Do the Hardest Job Are Waging a Quiet Battle, WASH. POST (May 8, 2019, 3:11 PM), https://www.washingtonpost.com/technology/2019/05/08/inside-facebook-second-class-workers-who-do-hardest-job-are-waging-quiet-battle [https://perma.cc/5G3U-QNGM] (“[C]ontent moderators . . . earn a starting wage that is 14 percent of the median Facebook salary.”).
they need, highlighting the second-class status of these workers. Illustrating this disparity, when Hazel* required medication to alleviate her stress, she was unable to pay for an appointment with a psychiatrist based on her paltry salary as a moderator. She ultimately had to turn to Facebook employees who pooled together funds to pay for her appointment. Describing the incident, Hazel* noted: “There is something seriously wrong with this system if I have to rely on Facebook employees so that I can go to the doctor.” Furthermore, moderators have trouble trusting on-site counselors since there have been instances of managers pressuring providers to divulge confidential information from sessions with moderators.

The culture of secrecy surrounding content moderation exacerbates the psychological stress moderators experience. Moderators must sign “draconian nondisclosure agreements that forbid them to discuss their work in even the most rudimentary terms,” including which platform their vendor contracts with and where their workplace is located. This “insane level of secrecy” increases the burdens on moderators who are instructed to refrain from discussing their work with loved ones or speaking to an outside therapist who has not been approved. Nondisclosure agreements (NDAs) are also used to

86 See, e.g., Vincent* Interview, supra note 59 (comparing the treatment of content moderators to that of Facebook employees); Eric* Interview, supra note 59 (noting the lack of employee benefits for moderators); BARRETT, WHO MODERATES THE SOCIAL MEDIA GIANTS?, supra note 8, at 4 (“Despite the crucial role they perform, content moderators are treated, at best, as second-class citizens.”).
87 Hazel* Interview, supra note 44.
88 Id.
89 Id.
90 See Eric* Interview, supra note 59 (discussing the lack of confidentiality in sessions with on-site counselors); see also Open Letter from Content Moderators Re: Pandemic, FOXGLOVE (Nov. 18, 2020) [hereinafter Open Letter 1], https://www.foxglove.org.uk/2020/11/18/open-letter-from-content moderators-re-pandemic [https://perma.cc/35EQ-U8J9] (presenting a letter on behalf of approximately three hundred moderators and noting that “Facebook management” as well as management at vendors “ask ‘coaches’ to reveal confidential details of counselling sessions”).
92 See Bridget* Interview, supra note 66 (mentioning that she was not allowed to say which platform’s content she moderated); see also Eric* Interview, supra note 59 (stating that a non-disclosure agreement kept moderators at his site from revealing the location where they worked).
93 Everlee* Interview, supra note 73.
94 See Elizabeth Dwoskin, A Facebook Contractor Posted Bruce Springsteen Lyrics to Protest Their Working Conditions. He was Fired Two Weeks Later, WASH. POST (June 27, 2019, 1:02 PM) [hereinafter Dwoskin, Facebook Contractor was Fired], https://www.washingtonpost.com/technology/2019/06/27/facebook-contractor-posted-video-bruce-springsteen-lyrics-his-profile-protest-working-conditions-he-was-fired-two-weeks-later [https://perma.cc/7DKS-QyWR] (“[Moderators] also asked for exceptions to the confidentiality agreements that would enable them to talk to outside psychotherapists and clinicians about the traumatic impact of the work . . . .”).
suppress dissent, as moderators worry that speaking to reporters or other third parties may cause them to lose their jobs. This concern is understandable considering that moderators have previously faced retaliation for protesting.\(^95\) In this lopsided arrangement, moderators are concerned about the possibility of having their confidentiality breached by company-based counselors, while vendors have devised a system for ensuring that their secrets (and by extension, the secrets of platforms) will be protected. Former moderator Everlee\(^*\) described the long-term implications of this asymmetric relationship: “It’s . . . taking advantage of people from lower income communities, communities of color, kind of destroying a little part of them, and then turning them back within a year to try to figure out how to recover from that.”\(^96\)

Thus, moderation imposes a heavy toll on workers who routinely view distributing material, manage without adequate support, and are expected to prioritize confidentiality above their own wellbeing. These issues stem from the fact that the work of content moderation “exists in a regulatory gray zone”\(^97\) since moderators are hired by vendors and do not receive the same benefits and protections as Facebook employees.\(^98\) Independent contractors are not granted unemployment insurance benefits, paid sick days, or family and medical leave protections.\(^99\) Additionally, workers’ compensation programs do not usually extend coverage to contractors, nor do they cover psychological damages.\(^100\)

Further compounding this problem, both Facebook and the third-party vendors it contracts with have minimal legal oversight,\(^101\) and the platform “largely allows the companies that hire the moderators to police themselves.”\(^102\) Without the threat of liability, Facebook and its vendors pass

\(^95\) See Hazel* Interview, supra note 44 (discussing moderators being punished for speaking out); see also Dwoskin, Facebook Contractor was Fired, supra note 94 (explaining how a moderator was fired after protesting work conditions).

\(^96\) Everlee* Interview, supra note 73.


\(^98\) See generally GRAY & SURI, GHOST WORK, supra note 42 (describing the lack of protections or benefits for contractors).


\(^100\) See Dwoskin, Facebook Content Moderator Details Trauma, supra note 38 (“Governmental workers’ compensation systems generally do not apply to third-party contractors and do not cover psychological damages . . . ”).

\(^101\) See Waverly* Interview, supra note 75 (addressing the limited external oversight of Facebook and vendors).

the buck to one another, creating a vacuum of accountability. To this point, Everlee* described the site’s lax attitude toward addressing problems: “Facebook can just choose to ignore [moderators’ issues] if [it] want[s] to, because [it] can just close [its] ears and be like, ‘Look, we understand that there’s some issues here, but that’s your problem to figure out.’” Since platforms and vendors have evaded legal responsibility and effectively used NDAs to limit public exposure, there have been few reasons for social media sites to deviate from the highly profitable system of content moderation.

In some ways, the industrial nature of Facebook’s moderation labor is necessary given the insufficiency of algorithmic moderation and the massive scale of user-generated content on the platform. Nevertheless, Facebook has increasingly acknowledged the fundamental problems with its outsourced model of moderation. In the past few years, the site has received greater pushback from moderators, full-time employees, and the press, all of whom have made this issue salient. Moderator protests and public letters have been international in scope and have been supported by allies within the company who have signed demand letters and formed coalitions to advocate for change. These efforts have resulted in some positive developments, such as resiliency programs to improve moderator self-care, raises in moderator salaries, audits, and vendor partner summits. Additionally, as part of a settlement in 2020, Facebook agreed to pay $52 million to former and

---

103 See Dwoskin, Facebook Content Moderator Details Trauma, supra note 38 (noting that a moderator was caught between Facebook and a vendor, both of which claimed that providing more support was the other’s responsibility).
104 Everlee* Interview, supra note 73.
105 See Vincent* Interview, supra note 59 (noting the lack of incentive for platforms to alter contractors’ working conditions).
107 Open Letter 1, supra note 90.
current moderators who experienced PTSD as well as to increase psychological support and to implement technical tools for moderators generally.\(^\text{110}\)

However, such reforms have not gone far enough to meaningfully improve moderators’ working conditions. To this point, the $52 million payout—which represents approximately half of a percentage point of Facebook’s 2020 revenue—fails to address the structural underpinnings of content moderation or the contractual arrangements undergirding this system.\(^\text{111}\) Interviewees stressed that Facebook can only achieve genuine reform by abandoning its reliance on outsourcing.\(^\text{112}\) In the absence of such systemic change, moderators (especially those from historically marginalized communities) will continue to provide hidden emotional labor that is disrespected and devalued.

### III. Detrimental Impacts of Contractual Labor Arrangements: Financial & Legal Risks to Platforms

In addition to harming moderators, outsourcing moderation presents financial, legal, and reputational risks for Facebook. This Part discusses the importance of worker classification, addresses the issues presented by worker classification in the gig economy, outlines the tests used for distinguishing employees and independent contractors, applies the tests to moderators, and describes the risks of misclassification for employers.

#### A. Worker Classification in the Gig Economy

The issue of worker classification is centrally important to both workers and employers. While companies may use independent contractors for legitimate business purposes (such as versatility, specialization, and expertise),\(^\text{113}\) designating workers as independent contractors can also be a

---


\(^{112}\) See, e.g., Hazel* Interview, *supra* note 44 (discussing how Facebook needs to designate moderators as employees); Eric* Interview, *supra* note 59 (same); Everlee* Interview, *supra* note 73 (same).

\(^{113}\) See Katharine Dunn, *The Pros and Cons of Hiring Independent Contractors in an Era of Regulatory Change*, 11 INT’L IN-HOUSE COUNS. J. 1, 2 (2017) [hereinafter Dunn, *Pros and Cons of Hiring Independent Contractors*] (listing benefits to hiring independent contractors, such as staffing versatility and specialized expertise).
financial boon. It is estimated that companies can save between twenty to forty percent of costs by hiring independent contractors as this enables them to evade paying for benefits (such as health insurance or worker’s compensation insurance), office space, minimum wage, and overtime protections.\(^{114}\) However, in obtaining this “competitive advantage,” companies open themselves up to the possibility of government audits, misclassification fines, and lawsuits.\(^{115}\)

For workers, operating as independent contractors can also be a mixed bag. Looking to the positive, being an independent contractor can bring independence, flexibility, autonomy, and better work/life balance.\(^{117}\) However, such freedom and balance may prove illusory in practice for some independent contractors.\(^{118}\) Additionally, designating workers as independent contractors may lead to limited bargaining power, lower wages, and less protection on both an individual and collective level.\(^{119}\) There has been limited legal redress for online contract workers in the United States so far, due in large part to the novelty of these claims and the lack of applicable}

\(^{114}\) See Gray & Suri, Ghost Work, supra note 42, at 32 (noting that companies can save forty percent of costs by not paying benefits or office rentals); see also Noam Scheiber, A Middle Ground Between Contract Worker and Employee, N.Y. TIMES (Dec. 10, 2015), https://www.nytimes.com/2015/12/11/business/a-middle-ground-between-contract-worker-and-employee.html?searchResultPosition=1 [https://perma.cc/WGZ9-4TPH] (explaining that using an employment model may cost companies an additional twenty to thirty percent); Tran & Sokas, The Gig Economy and Contingent Work, supra note 45, at e63 (noting that most employers that hire independent contractors “do not provide benefits such as health or workers’ compensation insurance”); Alana Semuels, What Happens When Gig-Economy Workers Become Employees, ATLANTIC (Sept. 14, 2018) [hereinafter Semuels, What Happens When Gig Workers Become Employees], https://www.theatlantic.com/technology/archive/2018/09/gig-economy-independent-contractors/570307 [https://perma.cc/q2GC-U6WR] (“On the other are labor advocates, who argue that gig-economy companies push much of the cost of their business onto workers, who don’t receive worker protections that were once standard, such as minimum wage and overtime protections.”).


\(^{116}\) See Dunn, Pros and Cons of Hiring Independent Contractors, supra note 113, at 3 (describing the risk of government audits and fines).


\(^{119}\) See David Weil, Lots of Employees Get Misclassified as Contractors. Here’s Why it Matters, HARV. BUS. REV. ONLINE (July 5, 2017), https://hbr.org/2017/07/lots-of-employees-get-misclassified-as-contractors-heres-why-it-matters [https://perma.cc/6LQ3-G3ZS] (“Though its form varies, the impacts of misclassification are almost always the same: the underpayment of wages, absence of benefits, and increased exposure to a variety of risks.”).
The dramatic consequences of independent contractor classification for companies and workers underscore the stakes at issue. This classification is especially critical for a precarious labor force such as content moderators who are not provided with benefits or protections when they are designated as independent contractors.

In recent years, the gig economy, technological advancements, and organizational changes have put pressure on traditional classifications of employee/independent contractor. Part of the problem in categorization is the emergence of “two-sided platform[s],” wherein an online intermediary (such as Uber) connects customers with service providers (such as drivers). While providers control the amount of business that they accept (making them seem like independent contractors), they operate within the confines established by the platform (making them seem like employees). The liminal nature of this work arrangement has resulted in factfinders being “handed a square peg and asked to choose between two round holes.”

The issue of classification in the gig economy has played out most visibly with the ride-sharing companies Uber and Lyft. While some have argued that the employment status of Uber and Lyft drivers has been “unsettlingly settled,” the problem remains unsettlingly unsettled in practice. Since 2016, there have been legal and political conflicts concerning how to classify independent contractors on these applications. In cases where Uber and Lyft workers have attempted to be classified as employees, there have been a range of outcomes: motions to dismiss have been denied, motions for summary

---

120 See Miriam A. Cherry, The Sharing Economy and the Edges of Contract Law: Comparing U.S. and U.K. Approaches, 85 GEO. WASH. L. REV. 1804, 1845 (2017) (“To date, plaintiffs in these new online work cases in the United States have had their cases dismissed in part because they have yet to find a solid doctrinal ground for recovery.”).
121 See Keith Cunningham-Parmeter, From Amazon to Uber: Defining Employment in the Modern Economy, 96 B.U. L. REV. 1673, 1676 (2016) (“Since the end of the Great Recession, U.S. businesses have aggressively engaged in a series of organizational changes—from classifying workers as independent contractors, to hiring subcontractors, to utilizing staffing agencies—to delegate employment-related responsibilities to outsiders.” (footnote omitted)).
125 See Jillian Kaltner, Employment Status of Uber and Lyft Drivers: Unsettlingly Settled, 29 HASTINGS WOMEN’S L.J. 29, 54 (2018) (concluding that despite evidence that Uber and Lyft drivers are misclassified as independent contractors, their status is “unsettlingly settled”).
settlements have been reached, preliminary injunctions have been denied, motions to compel arbitration have been granted, ballot measures have been struck down, settlements have been reached, and deals have been cut to block substantial state regulation. This dizzying array of outcomes underscores the complexity and legal uncertainty of this issue.

However, the categorization of content moderators is arguably even more complicated than other gig workers since there is not a triangular relationship among platforms, workers, and users (as in the case of Uber and Lyft), but rather one among platforms, workers, and vendors. The relationship among these parties likely does not rise to the level of joint employment, wherein an individual is hired by more than one employer, due to the stringent requirements of this doctrine.

Thus, vendors and platforms are not jointly liable for workers. Nevertheless, courts can still apply the traditional categorization tests to a second employer (such as a platform), even when the workers in question are

---

129 See, e.g., Cunningham v. Lyft, Inc., 17 F.4th 244, 255 (1st Cir. 2021) (reversing the district court’s decision to deny Lyft’s motion to compel arbitration).
133 See Julia S. Van de Walle, Doe v. Wal-Mart: Revisiting the Scope of Joint Employment, 30 BERKELEY J. EMP. & LAB. L. 589, 592 (2009) (“Joint employment doctrine is meant to determine when more than one employer employs an individual such that both employers are jointly liable to the employee and to [third parties] injured by the employee acting within the scope of employment.”). The factors considered in joint employment doctrine include: (1) whether the putative joint employers share power to control workers; (2) whether the putative joint employers share control in hiring/firing workers; (3) the degree of permanency in the relationship between the putative joint employers; (4) whether one of the employers controls or is controlled by the other employer; (5) whether the work is performed on one of the employer’s premises; and, (6) whether the employers share responsibility for traditional employer functions, such as managing payroll. See Salinas v. Com. Interiors, Inc., 848 F.3d 125, 141-42 (4th Cir. 2017) (outlining these factors). Many of these factors are not met in the case of Facebook and vendors.
already employees of another company (such as a vendor). To examine the status of content moderators, it is first necessary to consider the prevailing tests that are used to classify workers as independent contractors or employees.

B. Prevailing Worker Classification Tests

Agencies and jurisdictions use different tests to classify workers as either independent contractors or employees, which this Section will explore. These distinct tests may lead to different outcomes, which is particularly relevant in the case of content moderation that crosses geographic and jurisdictional borders. In this Section, I outline several major tests, including the right-to-control test, the economic reality test, tests deployed by agencies, and tests utilized in the states of Arizona, Florida, Texas, and California.

The most widespread test is the common law right-to-control test, used by the Employee Retirement Income Security Act (ERISA) and anti-discrimination statutes. Derived from the common law of agency, this test evaluates the employer’s “right to control the manner and means by which the product is accomplished.” The test considers several factors: the level of skill required, the source of tools and equipment, the location of the work, the duration of the arrangement, the worker’s discretion over “when and how long to work,” the method of payment (i.e., by time or job completed), whether the worker is engaged in a distinct line of work, and whether the work is part of the regular business of the employer. These factors are non-exhaustive and non-determinative, and factfinders balance various considerations to reach an ultimate conclusion.

---

134 See, e.g., Rutherford Food Corp. v. McComb, 331 U.S. 722, 730 (1947) (determining that meat boners who were supplied by a contractor were also employees of the slaughterhouses in which they worked).

135 See Jenna Amato Moran, Independent Contractor or Employee? Misclassification of Workers and Its Effect on the State, 28 BUFF. PUB. INT. L.J. 105, 107 (2009) [hereinafter Moran, Independent Contractor or Employee?] (“Courts, government agencies, and states use various standards to determine the classification of workers.”).


138 Id. (footnote omitted).

139 See RESTATEMENT (SECOND) OF AGENCY § 220(2)(a)-(j) (AM. LAW INST. 1958) (listing relevant factors for determining whether a worker is an independent contractor).

140 See Cmty. for Creative Non-Violence, 490 U.S. at 752 (“No one of these factors is determinative.” (citing Ward v. Atl. Coast Line R. Co., 362 U.S. 396, 400 (1960))).

141 See Moran, Independent Contractor or Employee?, supra note 135, at 109 (“These factors are weighed and balanced.”).
weighing exercise, factfinders attend to the degree of control an employer possesses, even if that employer does not exercise such control in practice. The right-to-control test is known not only for its ubiquity, but also for its stringency, as the test “generally gives employers more latitude to classify workers as independent contractors than do other legal approaches.”

Another frequently used common law test—the economic reality test—classifies workers under several major statutes, including the Fair Labor Standards Act of 1938 and the Family and Medical Leave Act of 1993. Taking a broader and “more inclusive approach” than the right-to-control test, the economic reality test examines the degree of the employer’s control as well as the worker’s level of dependency on the employer. This test inquires as to “whether the putative employee is economically dependent upon the principal or is instead in business for himself.” Multiple factors are also considered under this test, including the employer’s degree of control, the duration of the arrangement, the source of tools and equipment, opportunities for profit or loss, the amount of personal initiative required for success in open market competition, and the extent to which the services are integral to the employer’s business. The economic reality test focuses on the purposes of the underlying legislation and attends to social and structural inequalities, such as differences in bargaining power.

---

142 See Cotter v. Lyft, 60 F. Supp. 3d 1067, 1075 (N.D. Cal. 2015) (“The company need not exercise its full right of control for a worker to be deemed an employee.”).


144 See id. (listing the statutory uses of the economic reality test). In 2022, the Department of Labor issued a notice of proposed rulemaking pertaining to independent contractor classification under the Fair Labor Standards Act. See Employee or Independent Contractor Classification Under the Fair Labor Standards Act, 87 Fed. Reg. 62218 (proposed Oct. 13, 2022) (to be codified at 29 C.F.R. pts. 780, 788, 795) (“The Department is . . . proposing to return the consideration of investment to a standalone factor, provide additional analysis of the control factor . . . and return to the longstanding interpretation of the integral factor . . . ”).

145 See Myra H. Barron, Who’s an Independent Contractor? Who’s an Employee?, 14 LAB. LAW. 457, 460 (1999) [hereinafter Barron, Who’s an Independent Contractor?] (“Generally, if the individual is not only controlled by, but also depends on, the employer, that person is an employee entitled to the coverage under that law rather than an independent contractor.”)

146 Lilley v. BTM Corp., 958 F.2d 746, 750 (6th Cir. 1992) (citing Dole v. Snell, 875 F.2d 802, 804 (10th Cir. 1989)).

147 See Bales & Woo, The Uber Million Dollar Question, supra note 136, at 472 (describing the factors used to assess whether someone is an employee under the economic reality test).

148 See Barron, Who’s an Independent Contractor?, supra note 145, at 460 (“The economic reality test focuses on the purpose of the legislation under which the test is applied.”). For a historical example of this test, see NLRB v. Hearst Publ’ns, Inc., 322 U.S. 111, 127 (1944) (classifying newsboys as employees because they had limited bargaining power compared with Hearst Publications and
While the two common law tests share several factors, aspects of control such as hours and location are not generally considered in the economic reality test. As a result, individuals may be classified as employees under the economic reality test, but as independent contractors under the right-to-control test,149 making the test applied pivotal.150 Of note, some factfinders utilize a hybrid approach between the two common law tests (i.e., considering economic realities but prioritizing an employer’s right to control).151

Various government agencies have derived their own tests. For example, the Internal Revenue Service (IRS) deploys a modified version of the right-to-control test.152 Collapsing its previous twenty-factor test,153 the IRS’s test now considers three categories: behavioral control, financial control, and the type of relationship.154 Behavioral control considers whether an employer has a right to control the worker’s task through instruction (such as when, where, and how to work) or training.155 Financial control evaluates whether business expenses are reimbursed, the degree of the worker’s investment in tools or facilities, the extent to which the worker’s services are generally available in the relevant market, how the worker is paid, and the extent to which the worker can realize profit or loss.156 Under the type of relationship prong, the IRS considers the relationship description in written contracts, whether the worker is provided employee-type benefits, the duration of the relationship,


149 See Barron, Who’s an Independent Contractor?, supra note 145, at 458 (“[B]ecause the economic reality test takes a broad approach to the employment relationship, an individual may be an employee under that test, but an independent contractor under the narrower right to control test.”).

150 See Charles J. Muhl, What Is an Employee? The Answer Depends on the Federal Law, MONTHLY LAB. REV., Jan. 2002, at 5 (“[T]he same person can be classified as an employee under one test and the relevant Federal laws to which that test is applied, but as an independent contractor under another test and its relevant Federal laws.”).

151 See id. at 9 (describing how the hybrid approach combines elements of the right-to-control test and the economic reality test to determine who is classified as an employee). The hybrid approach has fallen out of favor after the Supreme Court’s decision in Nationwide Mutual Insurance Co. v. Darden, 503 U.S. 318, 324-25 (1992), which held that Congress, through its amendments, signaled to the judiciary that it wanted it to apply the common law factors, of which none are determinative, but instead must be weighed to conclude whether a person is an employee.

152 See Bales & Woo, The Uber Million Dollar Question, supra note 136, at 469 (“[T]he IRS looks to several unweighted and non-dispositive common-law factors while using all the ‘information that provides evidence of the degree of control and the degree of independence . . . .’” (quoting I.R.S. Publication 15-A 1, 7)).

153 See Harned, Kryda & Milito, Creating a Workable Legal Standard, supra note 143, at 103 (explaining that the IRS previously used a twenty factor test).


155 Id.

156 Id.
and the extent to which the services are part of the employer’s regular business.\textsuperscript{157} The IRS instructs that businesses must weigh these factors in a context-dependent manner as no single factor or “no ‘magic’ or set number of factors” designates a worker as an employee or independent contractor.\textsuperscript{158}

Beyond agencies, states also have their own tests for classifying a worker as either an employee or an independent contractor. This Section describes worker classification standards in the four states that have major Facebook moderation hubs: Arizona, Florida, Texas, and California.\textsuperscript{159} Arizona utilizes a right-to-control test, considering whether the employer can control the worker’s activities and whether the work is part of the business of the employer.\textsuperscript{160} Similarly, Florida uses the common law right-to-control test, considering factors such as skills, supervision, source of tools and equipment, place and duration of work, whether the worker is engaged in a distinct occupation, and whether the work is part of the employer’s regular business.\textsuperscript{161}

Texas uses a twenty-factor version of the common law right-to-control test.\textsuperscript{162} The state recently amended its test to clarify that gig workers are independent contractors as long as: (1) payment is made on a per-job basis; (2) the platform does not prescribe specific work hours; (3) the platform does not prohibit the contractor from working for another platform; (4) the platform does not restrict the contractor from engaging in other business; (5) the worker can determine the time and location of work; (6) the worker bears their own expenses; (7) the worker provides their own tools and equipment; (8) the worker does not have to follow specific instructions from the employer; and (9) the platform does not require mandatory meetings or training.\textsuperscript{163}

Lastly, California applies the “ABC Test,” which designates a worker as an independent contractor if the worker (A) is free from the employer’s control (both in contract and in practice), (B) performs work that is outside the employer’s usual course of business, and (C) engages in an independent trade

\textsuperscript{157} Id.
\textsuperscript{159} Newton, \textit{The Trauma Floor}, supra note 52.
\textsuperscript{160} ARIZ. REV. STAT. ANN. § 23-902(B) (2018).
\textsuperscript{161} Florida, EVERSHEDES SUTHERLAND, https://www.workerclassification.com/State-Resources/Florida [https://perma.cc/M72Z-3MRV].
\textsuperscript{162} See 40 TEX. ADMIN. CODE § 821.5 (2007) (including factors such as whether there is a set number of working hours, whether the individual has a significant investment in the business, who supplies the tools, and the duration of the relationship, among others).
\textsuperscript{163} 40 TEX. ADMIN. CODE § 815.134 (2019).
or business. Significantly, the ABC Test “creates a presumption of employment.”

C. Applying Prevailing Classification Tests to Moderators

There is a strong case for classifying content moderators as platform employees rather than independent contractors under any of the aforementioned tests. However, the application of these tests is complicated by the lack of transparency regarding the work of content moderation. This obfuscation makes it difficult to categorize the labor of moderation and to ascertain the level of control that Facebook and vendors exert over moderators. Nevertheless, information gleaned from interviews, news reports, and public press releases indicates that moderators act more as employees than as independent contractors.

Under the right-to-control test, some factors lean toward classifying moderators as independent contractors while other factors lean toward classifying moderators as platform employees. Moderators engage in a distinct line of work, have short-term contracts, and are skilled laborers—three factors that are associated with independent contractor status. The location where work is performed could cut either way. Moderators have traditionally been required to work in person at the offices of vendors. While their work is not performed on Facebook’s campus—a fact that makes moderators seem more like independent contractors—it is also not the paradigmatic “home-based” work that many independent contractors perform. The remaining factors counsel in favor of classifying moderators as platform employees. Moderators do not supply their own tools or equipment and are not paid on a piecework basis. While vendors, rather than Facebook, determine work structure and hours, moderators have little discretion over their schedules. In fact, some moderators are assigned undesirable shifts (such

165 See Harned, Kryda & Milito, Creating a Workable Legal Standard, supra note 143, at 102 (“[I]t creates a presumption of employment, making it more difficult for unscrupulous employers to misclassify employees as independent contractors to avoid legal obligations.” (footnote omitted)).
166 See Everlee* Interview, supra note 73 (“It’s an insane level of secrecy.”); see also Eric* Interview, supra note 59 (discussing the NDAs that moderators are often required to sign).
167 Dillon* Interview, supra note 59; Hazel* Interview, supra note 44.
168 See Jennifer Pinsof, Note, A New Take on An Old Problem: Employee Misclassification in the Modern Gig-Economy, 22 Mich. TELCOMMS. & TECH. L. REV. 341, 363 (2016) (“Typically, if a worker works onsite[,] he is more likely classified as an employee, whereas if a worker primarily works outside his employer’s place of business, he is more likely an independent contractor.”).
170 Dillon* Interview, supra note 59.
as overnight shifts)\textsuperscript{171} and are heavily micromanaged while conducting their tasks.\textsuperscript{172}

Most significantly, the work of moderation is central to Facebook’s business. Despite Facebook’s self-categorization as a technology company rather than a media company,\textsuperscript{173} user-generated content and the moderation of such material are core to the platform’s viability and profitability. As stated by author Tarleton Gillespie, “moderation is, in many ways, the commodity that platforms offer.”\textsuperscript{174} Unlike repairmen who sporadically fix problems or specialists who are called in for limited projects,\textsuperscript{175} moderators review content that is part of the platform day in and day out. Moderators have even recognized their own centrality: “Without our work, Facebook is unusable . . . . Facebook needs us.”\textsuperscript{176} In other cases, factfinders have looked beyond companies’ opportunistic self-characterizations to truly evaluate whether the work individuals perform is essential to the company’s business, a move that would make it more likely for moderators to be classified as platform employees.\textsuperscript{177} Further, since the right-to-control factors are not exclusive, a factfinder may also consider the high degree of control Facebook has over moderators through detailed content guides, specified quotas, required accuracy levels, and stringent NDAs.\textsuperscript{178} When balanced together, the factors weigh in favor of classifying moderators as platform employees. However, there is reason for caution since the right-to-control test typically affords companies wide “latitude to classify workers as independent contractors.”\textsuperscript{179}

Compared to the right-to-control test, moderators are more likely to be considered platform employees under the economic reality test. As discussed

\textsuperscript{171} Bridget* Interview, supra note 66.
\textsuperscript{172} See Gilbert, Facebook Forcing Moderators to Log Their Days, supra note 74 (describing how moderators must frequently log their tasks and clock out even for bathroom breaks).
\textsuperscript{174} GILLESPIE, CUSTODIANS OF THE INTERNET, supra note 5, at 13.
\textsuperscript{175} See Dynamex Operations W., Inc. v. Cnty. of L.A., 416 P.3d 1, 37 (Cal. 2018) (noting that occasional repairs are not part of a company’s usual course of business, and thus such workers would not be deemed employees of the company).
\textsuperscript{176} Open Letter 1, supra note 90.
\textsuperscript{177} See, e.g., O’Connor v. Uber Techs., Inc., 82 F. Supp. 3d 1133, 1142 (N.D. Cal. 2015) (“Even more fundamentally, it is obvious drivers perform a service for Uber because Uber simply would not be a viable business entity without its drivers.” (citing Yellow Cab Coop., Inc. v. Worker’s Comp. Appeals Bd., 226 Cal. App. 3d 1288, 1293–94 (1991))).
\textsuperscript{178} See discussion infra Section II.
\textsuperscript{179} See Harned, Kryda & Milito, Creating a Workable Legal Standard, supra note 143, at 100 (noting that the right-to-control test gives employers more room to deem workers as independent contractors than do other legal approaches).
Facing Contracting Issues

above, factors such as skill and duration point toward classifying moderators as independent contractors, while the source of equipment, the essential nature of the work, and the degree of Facebook’s control over moderators point toward classifying them as employees. Additional factors specific to the economic reality test support the conclusion that moderators are platform employees. Moderators do not operate their own independent businesses, do not have an opportunity for personal profit or loss, and their success is not contingent on their initiatives in the open market.

Additionally, there is a fundamental imbalance in bargaining power between moderators and Facebook (as well as between moderators and vendors), indicating moderators’ dependence on the platform. In this vein, former moderator Waverly* reported feeling “so exceptionally powerless.” This sentiment is echoed by statements of other interviewees who felt disenfranchised due to the economic clout of Facebook and vendors, retaliation against protests, threats of NDAs being enforced, neglect from HR, the lack of job mobility, and the limited transferability of skills developed on the job. To this point, former moderator Everlee* explained that when moderators speak up, “it puts a target on [their] back[s],” and that if moderators go public with their complaints about Facebook and vendors, they run the risk of “get[ting] black balled from a job with them ever again.”

These statements underscore how moderators do not act autonomously and independently, but rather are subject to direct and indirect control by Facebook, weighing toward categorization as platform employees under the economic reality test.

Moderators may also be classified as platform employees under the IRS’s test. To reiterate, the IRS’s test considers three categories: behavioral control, financial control, and the type of relationship between worker and employer. The first prong leans toward classification as employees since moderators are subject to Facebook’s behavioral control through instruction (including detailed internal rulebooks) and training. In fact, several of the moderators I interviewed were trained directly by Facebook rather than by vendors. With regard to the second prong, Facebook exerts financial control over moderators since moderators do not purchase their own equipment, do

---

180 Waverly* Interview, supra note 75.
181 Everlee* Interview, supra note 73; Hazel* Interview, supra note 44.
182 Everlee* Interview, supra note 73.
183 See supra note 154 and accompanying text.
184 See Nick Hopkins, Revealed: Facebook’s Internal Rulebook on Sex, Terrorism and Violence, GUARDIAN (May 21, 2017, 1:00 PM), https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence [https://perma.cc/3PDV-T2W7] (noting that Facebook has more than 100 internal training manuals on content moderation).
185 See, e.g., Everlee* Interview, supra note 73; Hazel* Interview, supra note 44.
not provide services generally available in the open market, do not receive pay on a piecemeal basis, and do not have the opportunity to realize individual profit or loss. The final factor—the type of relationship between worker and employer—could weigh in either direction. The labor of moderators is essential to Facebook’s business, but moderators have short-term contracts and are not provided employee-type benefits by Facebook. Since the IRS test is highly context specific, it is harder to predict how moderators may be classified under this test—although the balance of factors seems to weigh in favor of employment status.

Finally, the likely outcomes of state classification tests correspond with the previously discussed application of the common law and IRS tests. Under Arizona’s modified right-to-control test and Florida’s reliance on the standard right-to-control test, moderators may be classified as platform employees, although this finding is not guaranteed given the wide latitude that these tests afford to employers.

The status of moderators remains uncertain under Texas’s test related to gig workers. Factors that favor independent contractor status include the fact that Facebook does not prescribe specific work hours (schedules are likely set by vendors) and does not determine the location of work (which takes place within the vendor’s physical premises). In contrast, factors that favor employment status include the payment of a salary, the fact that moderators do not supply their own equipment, the fact that moderators do not moderate content for more than one platform, the fact that moderators are required to follow specific instructions, and the fact that moderators must receive mandatory training.

California’s “ABC Test” would likely characterize moderators as platform employees. This outcome is based on the test’s presumption of employment status. Additionally, under this test, there are a host of factors that weigh in favor of employment—including that moderators are subject to Facebook’s control, perform work within Facebook’s usual course of business, and do not operate their own independent businesses. It is worth noting that moderators could be classified as independent contractors in one state and employees in another, further complicating this issue and highlighting the importance of uniformly making moderators Facebook employees.

D. Risks of Misclassification: Lawsuits & Fines

Since there is a good chance that content moderators may be classified as platform employees under at least some of the aforementioned tests,

---

186 Everlee* Interview, supra note 73; Hazel* Interview, supra note 44; Eric* Interview, supra note 59.
187 See supra note 179 and accompanying text.
Facebook’s decision to contract out moderation exposes the company to legal and financial risks, especially in light of increased government attention to this problem. Scrutiny of worker misclassification has increased in recent years; one study found that up to thirty percent of audited employers misclassified their employees as independent contractors. As a result, states and the federal government have ramped up efforts to detect and punish misclassification, and audits for misclassification have become more common. In the short span between 2004 and 2012, twenty-two states revised their statutory definitions of independent contractors or modified enforcement structures for misclassification. At the federal level, several bills have been introduced (but have not been passed) setting forth procedures for determining employment status or increasing the penalties for misclassification. Additionally, in 2014, the Department of Labor provided funds to nineteen states for worker misclassification detection and enforcement. Alongside these government efforts, there have been a growing number of lawsuits (particularly in the technology sector) against individual companies for worker misclassification. These actions reveal heightened scrutiny for misclassification in the current moment as well as resultant legal risks for companies.

Facebook’s outsourcing of moderation also may harm the company financially. If misclassification is proven, Facebook may have to pay back taxes

---


189 See, e.g., Bran Noonan, The Campaign Against Employee Misclassification, N.Y. STATE BAR ASS’N, Oct. 2010, at 42 (“Over the past few years, New York State and the federal government have increasingly extended their efforts to eliminate the misclassification of workers as independent contractors.”).

190 See Harned, Kryda & Milito, Creating a Workable Legal Standard, supra note 143, at 93-94 (“Audits for such misclassifications are becoming more frequent . . . ”).


194 See, e.g., supra notes 126–129 and accompanying text.
and contributions.\textsuperscript{195} In cases of unintentional misclassification, employers must pay fines of $50 for each W-2 form they failed to file, penalties (with interest) totaling 1.5% of the wages for failure to withhold incomes tax, 40% of Social Security and Medicare taxes that were not withheld from the employee, and 100% of the matching Social Security and Medicare taxes the employer should have paid.\textsuperscript{196} Employers are also on the hook for a failure to pay taxes penalty, which is equal to 0.5% of the unpaid tax liability for each month taxes went unpaid, up to 25% of the total tax liability.\textsuperscript{197} However, if an employer qualifies for a safe harbor exemption under IRS § 530, which considers reporting and substantive consistency as well as the employer’s reasonable basis for categorization, it will not be subject to “federal employment taxes, penalties, and interest for such misclassification.”\textsuperscript{198} In the case of intentional misclassification, the employer may be subject to criminal penalties of up to $10,000 for each misclassified worker and up to one year in prison.\textsuperscript{199} There may also be costs on the state level. For example, worker misclassification is a felony in Florida,\textsuperscript{200} and in California, employers who have engaged in a pattern or practice of willful misclassification are subject to penalties up to $25,000 for each violation.\textsuperscript{201} Employers may also have to pay the costs of legal proceedings and settlements, such as Uber’s 2019 classification settlement for $20 million.\textsuperscript{202} Finally, a company may experience reputational damage from its misclassification, which can translate into future financial losses.\textsuperscript{203} Avoiding negative press is paramount in an era of values-

\textsuperscript{195} \textit{See} Barron, \textit{Who’s an Independent Contractor?}, supra note 145, at 463 (noting that the employer may be on the hook for taxes and contributions).

\textsuperscript{196} \textit{See} Matt Chodosh, \textit{Employee Misclassification Too Big to Ignore}, BLOOMBERG (July 1, 2021, 4:00 AM) [hereinafter Chodosh, Misclassification Too Big to Ignore], https://news.bloombergtax.com/daily-tax-report/employee-misclassification-too-big-to-ignore [https://perma.cc/U264-BD2F] (describing the possible penalties for misclassification).

\textsuperscript{197} Id.


\textsuperscript{199} Chodosh, \textit{Misclassification Too Big to Ignore}, supra note 197.


\textsuperscript{201} \textit{See} CAL. LAB. CODE § 226.8(c) (West 2013) (setting forth the maximum penalty of $25,000).


based consumerism in which “customers expect [companies] to declare and act on their corporate values.”

Consequently, a company that misclassifies its workers may ultimately pay more than if it had properly classified its workers as employees from the start. These serious consequences highlight how Facebook’s outsourcing of moderation not only jeopardizes workers’ wellbeing but also jeopardizes the company’s bottom line.

IV. REFORMS

In this Part, I lay out the rationales and stakes of reforms, which are critically important to rectifying the issues discussed in prior Parts. I first discuss the general stakes of reforms before offering reform suggestions aimed at particular constituencies, such as state legislators, platforms, and moderators/civil society.

A. Rationales & Stakes of Reform

It is clear that the current status quo is unsustainable for moderators, Facebook’s reputation and finances, and the system of online content review generally. It is equally clear that individual lawsuits and settlements are inadequate to rectify this problem. For the past few years, moderators have pursued actions against Facebook and other platforms for claims such as negligent exercise of retained control, negligent provision of unsafe equipment, violation of California’s Unfair Competition Law, negligent infliction of emotional distress, fraudulent concealment, and violation of Florida’s Deceptive and Unfair Trade Practices Act. These claims have been settled (in a way many moderators have found to be unsatisfactory), voluntarily dismissed by plaintiffs, or dismissed by courts for lack of standing, failure to state a claim, or lack of personal jurisdiction. Such outcomes

LAES] (“Moreover, in an economy where 70% to 80% of market value comes from hard-to-assess intangible assets such as brand equity, intellectual capital, and goodwill, organizations are especially vulnerable to anything that damages their reputations.”).


205 See Barron, Who’s an Independent Contractor?, supra note 145, at 457 (“However, those who misclassify their employees as independent contractors potentially incur more liability for such misclassification than they would had they classified the worker as an employee in the first instance.”).

206 See discussion supra notes 109–111 and accompanying text.

207 See, e.g., Order Granting Motion to Dismiss, Young v. ByteDance Inc., No. 22-CV-01883-VC (N.D. Cal. Oct. 19, 2022) (dismissing a claim at the request of the plaintiff); Order Granting Motion to Dismiss with Leave to Amend, Doe v. YouTube, Inc., No. 20-CV-07493-YGR, 2021 LEXIS 131339, at *2 (N.D. Cal. July 14, 2021) (dismissing with leave to amend since the complaint
reveal that courts are not well suited to provide needed redress in this area due in part to the novelty of these claims, information and power asymmetries between Facebook and moderators, and the United States’s “adversarial legal system in which employers enjoy tremendous structural advantages.”

Looking beyond the courts, the job insecurity of moderators is a multifaceted problem that requires sustained involvement of various stakeholders. It is important to foreground the needs and desires of moderators when devising alternative labor arrangements. Moderators have consistently pushed for less restrictive NDAs in addition to designation as formal employees of Facebook. Robust reforms will also necessitate dealing with structural issues and emerging problems that call for societal debate and resolution.

The labor of moderation connects with broader questions of whether and how to regulate the tech industry and the gig economy, which have hastened the development of “workers without workplaces and employees without employers.” Additionally, this topic connects to concerns regarding the breakdown of employee protections established in the New Deal era. Employers have increasingly foregone traditional labor protections in the fragmented “gloves-off” economy, requiring a reenvisioning of socioeconomic policies to safeguard workers’ rights. Lastly, the status and experiences of moderators breathe new life into the debate of whether rights should only be granted to employees or whether certain rights should extend to all

See Ellen Berrey, Steve G. Hoffmann & Laura Beth Nielsen, Situated Justice: A Contextual Analysis of Fairness and Inequality in Employment Discrimination Litigation, 46 LAW & SOC. REV. 1, 4-5 (2012) (discussing how an institutional concept of fairness “obscures the employers’ disproportionate resources, power, and control in litigation”).

209 See, e.g., Hazel Interview, supra note 44; Eric Interview, supra note 59; Everlee Interview, supra note 73; Open Letter to Facebook from Content Moderators, GOOGLE DOCS (July 2021) [hereinafter Open Letter 2], https://docs.google.com/forms/d/1WBwyFy5wWB8ojYDbHOcZeEZoCThkkxZy8ozZzDno/viewform?edit_requested=true [https://perma.cc/TNQ6-RXSW] (urging Facebook to modify NDAs and to bring moderators in house).

210 See Katherine V.W. Stone, Legal Protections for Atypical Employees: Employment Law for Workers Without Workplaces and Employees Without Employers, 27 BERKELEY J. EMP. & LAB. L. 251, 251 (2006) (examining the legal consequences of the changing nature of employment relationships for workers without traditional employment relationships).

211 See Harned, Kryda & Milito, Creating a Workable Legal Standard, supra note 143, at 104 (describing broad employee protections that originated in the New Deal era).

workers. These issues raise the stakes of possible reform efforts and underscore the importance of reimagining worker protections to safeguard vulnerable communities, such as moderators.

In the following Sections, I discuss potential reforms for state legislators, platforms, and moderators/civil society.

B. State Legislators

There are several steps that state legislators should take to rectify these problems. First, legislators should consider modifying the independent contractor tests used in their jurisdictions. Rather than relying on tests that stress the worker’s degree of dependence on the company, legislators should instead focus on tests that emphasize the company’s degree of dependence on the worker. This change is especially important in the modern economy, wherein “Uber, the world’s largest taxi company, owns no vehicles,” “Facebook, the world’s most popular media owner, creates no content,” “Alibaba, the most valuable retailer, has no inventory,” and “Airbnb, the world’s largest accommodation provider, owns no real estate.” Such a pattern reveals the degree of dependence companies have on contractors who are nevertheless excluded from employee protections. When applied to Facebook, moderators’ labor is central to the platform, cultivating dependence on their work: “Moderation is not an ancillary aspect of what platforms do. It is essential, constitutional, definitional. Not only can platforms not survive without moderation, they are not platforms without it.”

Tests foregrounding the company’s degree of dependence on workers would better recognize emerging forms of labor arrangements and would help mitigate prevailing inequities. The necessity of this adjustment cannot be overstated on the heels of a pandemic that has furthered the “essential worker paradox,” wherein the labor of gig workers has become essential but is nevertheless treated as outside the employment relationship.

---


214 See Robert Sprague, Worker (Mis)Classification in the Sharing Economy: Trying to Fit Square Pegs into Round Holes, 31 ABA J. LAB. & EMP. L. 53, 54 (2015) ("[A]nalysis of work in the sharing economy should turn from the worker's dependence on the company to the company's dependence on the worker.”).


216 GILLIESPIE, CUSTODIANS OF THE INTERNET, supra note 5, at 21.

the centrality of workers in state classification tests would go a long way to providing moderators greater protections.

Alternatively, legislators could create a third category of worker, which sits between independent contractor and employee status. Countries such as Canada and Germany have recognized an intermediate class of laborer titled “dependent contractors,” who form an exclusive relationship with a client and are in a position of economic dependence. Workers in this category enjoy rights such as collective bargaining.

Similarly, scholars have advocated for the designation of “independent workers,” who would receive benefits such as collective bargaining and civil rights protections but not unemployment insurance or minimum wage safeguards. However, such proposals raise concerns about deliberate mischaracterization, transposing current issues with two categories (independent contractors and employees) into a system with three categories.

A final option for legislators is changing the categorization of moderators or providing benefits directly to this group. For example, California has required that drivers for cannabis companies be classified as employees instead of independent contractors. Rather than changing prevailing designations, other jurisdictions have supplied protections to vulnerable workers. To this effect, a district court in New York granted a preliminary injunction ordering app-based ride-share services to pay unemployment insurance benefits to drivers. Additionally, in 2015, the Seattle City Council passed a law allowing Uber and Lyft drivers to form unions. Lastly, 

218 See Catherine Tucciarello, Comment, The Square Peg Between Two Round Holes: Why California’s Traditional Right to Control Test Is Not Relevant for On-Demand Workers, 13 SETON HALL CIR. REV. 351, 369 (2017) (“For example, courts and labor and employment statutes in Canada and Germany recognize an intermediate class, called the ‘dependent contractor.’” (quoting Benjamin Sachs, A New Category of Worker for the On-Demand Economy?, ONLAW. (June 22, 2015))).

219 See Moran, Independent Contractor or Employee?, supra note 135, at 130 (outlining the rights of dependent workers).


221 See, e.g., MARTIN RISAK, FRIEDRICH-EBERT-STIFTUNG, FAIR WORKING CONDITIONS FOR PLATFORM WORKERS: POSSIBLE REGULATORY APPROACHES AT THE EU LEVEL 18 (2018) (“The introduction of an intermediate category could well lead to evasive strategies being opted for by employers . . . .”).

222 See Semuels, What Happens When Gig Workers Become Employees, supra note 114 (noting this change in classification).


California passed a bill specifying that workplace quotas cannot come at the expense of health and safety and mandating that warehouses provide detailed records about quotas and workers' actual rates of productivity. Each of these schemes would provide relief to moderators currently caught between vendors and platforms and therefore in need of state protection.

C. Platforms

To protect their bottom lines, reputations, and the workers who make their sites run, platforms such as Facebook should bring moderators in-house and make them employees. Platforms should also conduct more regular audits of vendors and expand access to moderation sites. Currently, platforms routinely deny requests from journalists and scholars to visit moderation sites, and vendors control moderators’ access to the press as well as attending meetings with platforms. In other words, despite widespread calls for greater social media transparency, visibility around the labor conditions of moderators has not been prioritized in the United States.

Platforms and vendors should provide information about the labor of moderation to further scholarship and advance reform efforts. Additionally, platforms should adopt more robust standards for vendors, fund research about the impacts of working as a content moderator, expand the support offered to moderators, and supply healthcare and other resources after moderators leave the industry.

---


227 See Hazel* Interview, supra note 44 (discussing how access to the press is controlled); see also Everlee* Interview, supra note 73 (“It’s an insane level of secrecy.”).

228 See Nicolas P. Suzor, Sarah Myers West, Andrew Quodling & Jillian York, What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation, 13 INT’L J. COMM. 1526, 1527 (2019) (“We argue that there is a pressing need for more specificity in identifying what information should be provided and to whom.”).


230 See BRARRETT, WHO MODERATES THE SOCIAL MEDIA GIANTS?, supra note 8, at 2, 25 (recommending funded research and continued healthcare benefits).
D. Moderators & Civil Society

Finally, moderators, allies, and civil society can also make inroads to address the current problems surrounding content moderation. Moderators should continue to enlist employees from platforms as critical allies. Moderators may draw upon the successes of other platform workers—most notably, drivers—who have channeled precarity into strategic opportunity by leveraging their liminal status. As an example of such a strategy, moderators may look to the Independent Driver’s Guild (IDG), an advocacy group representing more than 250,000 drivers nationally. IDG has worked on issues such as tipping, deactivation appeals, minimum wage, benefits, wellness, and pandemic safety—topics that are also relevant to moderators.

In replicating the success of IDG, moderators may turn to the Silicon Valley Rising Campaign (SVR), a coalition of “labor, faith leaders, community-based organizations and workers” dedicated to responsible contracting. While the SVR does not specifically focus on moderation, moderators may still be able to partner with this organization. Alternatively, moderators could partner with organizations such as the Trust & Safety Professional Association, which supports moderators but does not currently focus on issues of contracting. However, unionization of moderators may be difficult due to several factors, including the vast geographic spread of moderators.

Another option for moderators is to seek support from advertisers and users to put pressure on platforms and vendors. This strategy previously achieved measured success with Facebook. Namely, in 2013, activists who were upset by Facebook’s lack of response to sexist imagery and rape threats and jokes on the site turned to direct action. Activists sent companies screenshots of their ads, which were situated next to abusive imagery of

---

231 Hazel* Interview, supra note 44; Vincent* Interview, supra note 59.
232 See Jill Edenshade, Elizabeta Shifrin & Karina Rider, Leveraging Liminality: How San Diego Taxi Drivers Used Their Precarious Status to Win Reform, 60 LAB. HIST. 79, 88 (2019) (describing how drivers used their liminal status to their advantage).
235 About Us, SILICON VALLEY RISING, https://siliconvalleyrising.org/about [https://perma.cc/538N-92AD].
236 See Drootin, “Community Guidelines”, supra note 226, at 1239 (noting this limitation).
238 See Drootin, “Community Guidelines”, supra note 226, at 1241 (outlining the challenges of unionization).
239 See Nurik, Censored, Commodified, and Surveilled, supra note 6, at 133 (describing the 2013 campaign).
women on the site. As Facebook’s advertising revenue began to fall (from companies such as Nissan UK), the company negotiated with consumer advocacy groups, leading to greater involvement of activists in the moderation process, increased training of moderators, and more frequently updated community standards. A similar campaign, “Stop Hate for Profit,” was launched in 2020 to encourage advertisers to boycott Facebook and Instagram until the platforms meaningfully addressed the presence of racism on their sites. While problems have persisted on Facebook, the platform’s responsiveness to advertisers demonstrates the potential of deploying boycotts and strategic alliances with advertisers to achieve social change.

CONCLUSION

Although moderators toil day in and day out to review an unfathomable amount of content on platforms, they are treated as second-class workers. Despite providing economic value for sites, moderators are hired by vendors rather than by platforms. Moderators’ labor is critical for platforms but is kept under wraps through a maze of contractual relationships, restrictive NDAs, and blame shifting between platforms and vendors to avoid liability and accountability. In this manner, moderators work under stressful and exploitative conditions and find their labor hidden, devalued, and disregarded by Facebook and third-party vendors.

These problematic work arrangements reveal deeply rooted structural problems. In the words of scholar Tarleton Gillespie, “the fundamental arrangement [undergirding content moderation] itself is flawed,” since “the challenges [platforms] face are now so deep as to be nearly paradoxical.” To this point, platform content moderation has led to “the distancing and mistreatment of workers in a company dedicated to community,” secrecy on a site of public communication, and disenfranchisement in a place of supposed empowerment.

This Comment sheds light on the labor of content moderation, informed by interviews with former and current content moderators. As this Comment has demonstrated, contracting out content moderation harms both workers and platforms. Moderators lack the protections afforded to Facebook

240 Id.
241 Id. at 134.
243 GILLESPIE, CUSTODIANS OF THE INTERNET, supra note 5, at 198, 205.
244 Nurik, Censored, Commodified, and Surveilled, supra note 6, at 159.
employees and experience job precarity. They face relentless pressure from quotas, time limits, and accuracy scores in a micromanaged work environment that resembles an assembly line. Rates of burnout and turnover are high as the job routinely exposes moderators to graphic imagery without adequate psychological support and resources. The contractual nature of content moderation also poses risks for Facebook. By relying on outside vendors, Facebook opens itself up to government audits, misclassification fines, lawsuits, and negative press.

The treatment of moderators is not an isolated issue, but instead is reflective of broader inequities in the gig economy, social media regulation, and contemporary labor arrangements. These broader implications underscore the need for urgent societal debate and reform. Although Facebook and vendors have instituted some improvements in recent years, little substantive progress has been made in adjusting this problematic dynamic. Moving forward, Facebook and vendors should become more transparent about the hidden labor fueling content review and install meaningful reforms to better protect moderators. State legislators, platforms, moderators, and civil society should also become involved to rectify these issues. Thus, it is necessary to foreground the labor of moderation when considering social media research, reform, and regulation.