From Scalability to Subsidiarity in Addressing Online Harm

Amy A. Hasinoff[1] (iD) and Nathan Schneider[2] (iD)

Abstract

Large social media platforms are generally designed for scalability—the ambition to increase in size without a fundamental change in form. This means that to address harm among users, they favor automated moderation wherever possible and typically apply a uniform set of rules. This article contrasts scalability with restorative and transformative justice approaches to harm, which are usually context-sensitive, relational, and individualized. We argue that subsidiarity—the principle that local social units should have meaningful autonomy within larger systems—might foster the balance between context and scale that is needed for improving responses to harm.

Keywords

platform studies, restorative justice, scale, scalability, social media, subsidiarity, transformative justice

Content moderation on social media platforms has eluded simple solutions, even for some of the world's wealthiest corporations. When harm happens online, there seem to be few options for appropriate and legitimate responses. This is in part because platforms are generally designed to prioritize economic efficiency while holding a large number of users' attention; responding to harm takes the form of simple, semi-automated sanctions such as putting warning labels on posts or demoting or removing content. Empirical studies suggest that many users perceive moderation regimes as arbitrary and unfair (Myers West, 2018; Nurik, 2019). As Sarita Schoenebeck and Lindsey Blackwell put it, "In typical platform-driven moderation systems, all violators are treated equally, with users who unintentionally violate rules receiving the same sanctions as users who deliberately try to cause harm" (Schoenebeck & Blackwell, 2021, p. 136).

Online harm occurs in a variety of ways and generally takes two related forms: individually targeted harassment (Vogels, 2021) and harmful content that may be targeted at a group or at no one in particular (Scheuerman et al., 2021). The lack of context sensitivity in the responses to all forms of harm appears to be a widespread pattern among social platforms. In some respects, platforms try to address harm by mimicking criminal-legal systems. Schoenebeck and Blackwell (2021) find that social media governance "has largely been informed by Western models of criminal justice, which rely on sanctions (e.g., punishment) to encourage compliance with formal rules and laws" (p. 134). However, in pursuit of frictionless scale, important features for context

sensitivity in Western justice systems are absent; there are typically limited due process rights or systems of appeals, no democratic processes for creating and amending rules, and no juries of one's peers. Facebook's Oversight Board has been referred to as a "supreme court" for the platform (Douek, 2019; Klonick, 2019), and yet there is no corresponding system of lower courts. Instead, a single rulebook—an internal company document—governs moderation practices for the company's billions of global users. Reddit encourages more diverse community rules, but moderators frequently make decisions without community input (Seering et al., 2019), thanks in part to the platform's "automoderator" tool that automates decisions and enforcement (Jhaver et al., 2019). Platform software also provides little support for community involvement in governance or for holding moderators and administrators accountable (Schneider, 2022).

Scholars have long critiqued dreams of achieving large scale without commensurate cost. In his whimsical essay "On Being the Right Size," biologist J. B. S. Haldane (1926) stressed that for living things, "a large change in size inevitably carries with it a change in form." More recently,

[1]University of Colorado Denver, USA
[2]University of Colorado Boulder, USA

Corresponding Author:
Nathan Schneider, University of Colorado Boulder, 1511 University Avenue, 478 UCB, Boulder, CO 80309-0401, USA.
Email: nathan.schneider@colorado.edu
Twitter: @ntnsndr

anthropologist Anna Lowenhaupt Tsing (2012) identifies the aspiration for *scalability*: "the ability to expand—and expand, and expand—without rethinking basic elements." Tsing then develops a theory of *nonscalability* as she draws on examples from colonial plantations to the global trade in rare mushrooms to argue that "scalability never fulfills its own promises." Both Haldane and Tsing use examples in nature to question the widespread human ambition to achieve limitlessly large-scale and increasingly efficient social organization—for Haldane, state socialism, and for Tsing, computerized capitalism. In nature, they observe, differences of scale are differences of kind, involving distinct structures and constraints.

In the context of online platforms, business models designed for scalability are considered highly desirable and are often required to secure financing (Graham, 2012; Sullivan, 2016). Tarleton Gillespie (2018) documents how social platform companies have had to "become fundamentally different than they once were" when they achieve large user-bases, adopting scalable techniques for moderation when more hands-on methods for smaller scales "simply will not translate" (p. 76). Massive social platforms boast that their scalability gives them the ability to develop sophisticated and efficient automated content moderation tools. Facebook founder Mark Zuckerberg asserted that, due to his company's wealth and vast user-base, "we're able to do things that I think are just not possible for other folks to do" with content moderation (Feiner & Rodriguez, 2019). Yet as the dangers of using automation to achieve scalability through low-cost content moderation become more and more evident, Gillespie (2020) suggests, "Maybe we should not automate." Alex Hanna and Tina M. Park (2020) go further to present a provocation "against scale" and "ask readers to consider what potential resistances to scale thinking may look like." In large networks, a range of harms are situated in a variety of different social contexts; we argue that as networks seek the benefits of scale, the responses to harm must still reflect those differences.

Some scholars and journalists have begun asking whether the largely offline practices of restorative and transformative justice could provide an alternative model for online moderation and governance (Garfield, 2019; Hasinoff et al., 2020; Paul & Borton, 2021; Salehi, 2020; Schoenebeck & Blackwell, 2021; Schoenebeck, Haimson, & Nakamura, 2021; Schoenebeck, Scott, et al., 2021; Warzel, 2019; Xiao, 2021; Xiao et al., 2022). Instead of simply shunning or removing rule-violating content or users, a restorative justice approach to online conflict and abuse relies on community participation, centers the needs of people who have been harmed, and pursues the repair of that harm when possible. Transformative justice additionally stresses transforming incidents into opportunities for clarifying norms and enacting social change. Empirical research investigating the application of restorative and transformative justice to online harm has largely used interviews and surveys to determine

how moderators might implement these alternatives, what kinds of repair those who have been harmed want, and how users might react to participating in these kinds of justice processes (Schoenebeck, Haimson, & Nakamura, 2021; Schoenebeck, Scott, et al., 2021; Xiao, 2021; Xiao et al., 2022). The literature does not yet explain how the governance of large online platforms could facilitate key principles of restorative and transformative justice—such as attention to individual needs and community self-determination—while also fostering a balance with accountability to potentially planetary-sized networks.

This article argues that subsidiarity could facilitate context-sensitive processes in large-scale online networks, including for responses to harm. Subsidiarity is the principle that local social units should have meaningful autonomy wherever possible while maintaining their connections and responsibility to the larger systems in which they exist. Through an analysis of the context-specific and relational practices of restorative and transformative justice, we show the advantages of subsidiarity over scalability for the development and governance of large-scale, intricate social networks. Like the call that Janneke Adema and Samuel A. Moore (2021) make for "scaling small," we contend that local specificity can occur in systems with great reach and complexity.

First, we explore lessons from restorative and transformative justice; these point to a need for processes that are sensitive to context and attentive to the precise nature of harms and needs of people who have been harmed. Second, we turn to how the logic of subsidiarity can reorient the aspiration of scale, enabling highly adaptable processes that are nonetheless applicable across large networks. Third, we summarize some design considerations for applying subsidiarity to online networks and for developing tools that enable contextual responses to harm and conflict.

Restorative and Transformative Justice

Criminal legal systems and social media platforms both respond to harm primarily by enforcing rules and punishing the person who caused it. In contrast, restorative and transformative justice processes typically focus on one or more of the following goals: pursuing immediate harm reduction strategies, meeting some of the harmed person's needs, helping the person who caused harm take accountability, or confronting injustice in the community—including its conditions, norms, and rules—to prevent future harm.

While restorative justice mechanisms tend to focus on repairing harm for those who have experienced it, transformative justice processes put more of their energy in transforming the underlying conditions that led to the harm. For example, in a restorative justice approach, a prosecutor might decide to drop the charges against a teenager who stole a loaf of bread if they successfully complete a restorative justice program. The process might involve facilitators meeting one-on-one with each person involved and could culminate

in a community meeting where the teenager publicly apologizes to the store owner and agrees to compensate them by working some hours at the store. A transformative justice approach to the same incident would begin not with police involvement, but with the store owner contacting a community organization for help. The organization might similarly facilitate individual meetings followed by a group conference, but a transformative justice process would focus on identifying and addressing the underlying cause of the incident. The teenager, the store-owner, and other community members might, for example, work together to create a food bank for the community.

Kathleen Daly explains that the term "restorative justice" began to appear in the 1990s, emerging out of diverse experimental practices in North America and beyond, such as mediation and family counseling (Daly, 2016, p. 19). Some scholars stress how Indigenous legal traditions and restorative justice influenced each other (Chartrand & Horn, 2016), and many practitioners refer to Western restorative justice as "rooted" in Indigenous practices (e.g., Shah et al., 2017). However, Paora Moyle and Juan Marcellus Tauri view this as appropriation; they describe the restorative justice practice of family group conferencing as a crime control product incorrectly marketed around the world as a Māori practice— and they argue that settler governments use the practice, both in New Zealand and elsewhere, in ways that actually hinder Indigenous communities' self-determination (Moyle & Tauri, 2016; Tauri, 2020). In the United States, the United Kingdom, Australia, and a handful of other countries, restorative justice is most widely adopted as a diversion from criminal prosecution for nonviolent juvenile offenses, typically at the discretion of the prosecutor. Primary and secondary schools have also adopted restorative justice processes to mediate conflicts and repair harm without suspending or expelling students. A handful of universities and nonprofit organizations offer restorative justice training and certifications for facilitators.

In contrast, transformative justice programs are typically grassroots organizations that operate entirely outside of formal institutions or criminal legal systems. Many transformative justice practitioners want to abolish prisons and police, which they view as ultimately exacerbating violence rather than reducing it. As Mimi Kim explains, complex problems like "interpersonal violence [are] . . . embedded in systems of structural harm" (Kim, 2021, p. 7), including social and cultural norms, state violence, and the marginalization of communities of color. Unlike the formal relationships with criminal legal systems, certifying bodies, and other institutions that restorative justice often relies on, many transformative justice initiatives are ad hoc, local, and informal. The Project Nia and Interrupting Criminalization initiative "One Million Experiments" archives and highlights dozens of transformative justice projects and proposals in a podcast, website, and zine series (Project Nia & Interrupting Criminalization, n.d.). These efforts all address community problems without engaging with the criminal legal system— from a non-police emergency response network in Detroit to a safe-ride fund for Black trans women that first launched in New York.

Restorative and transformative justice both developed in response to critiques of criminal legal systems' focus on punishing people who have caused harm as both cruel and ineffective. Howard Zehr (2015), one of the founders of the restorative justice movement, explains that in Western criminal legal models of justice, crime is defined as "a violation of the law and the state" and justice "requires the state to determine blame (guilt) and impose pain (punishment)." Researcher-activist Mariame Kaba contends that criminal punishment "rarely, if ever, encourages people to take accountability for their actions. Instead, our adversarial court system discourages people from ever acknowledging, let alone taking responsibility for, the harm they have caused" (Kaba, 2021). Research demonstrates that punishment has limited power to deter rule-breaking (e.g., Nagin, 2013), though criminal law generally assumes otherwise. Restorative justice acts upon people who have caused harm though persuasion instead of fear of further punishment; according to John Braithwaite (2002), restorative justice involves a process of "reintegrative shaming," which coaxes people who have caused harm to move past denying and minimizing the effects of their actions, to take responsibility and ultimately to build a new sense of self as someone who is committed to community norms of morality.

Restorative and transformative justice also emerged in response to criminal legal systems' relative lack of concern for the needs and interests of people who have been harmed. Indeed, research demonstrates that people who have been harmed are not primarily interested in or satisfied with punishment (Alliance for Safety and Justice, 2016; Daly, 2014; Wemmers, 2002). Zehr (2015) explains that in restorative justice, crime is "a violation of people and relationships" which creates "obligations" and that justice involves "victims, offenders, and community members in an effort to put things right." Restorative justice pursues accountability though repair; typically the nature of that repair is determined through individualized, context-specific conversations.

Restorative justice programs have been extensively researched and evaluated. In a review of dozens of studies that compare restorative justice programs to conventional criminal-legal processes, Lawrence Sherman and Heather Strang (2007) find that restorative justice programs offer the following benefits: they often reduce recidivism more or at least as much as prison; they at least double the number of offenses ultimately brought to justice; they have higher rates of satisfaction for all participants; they reduce posttraumatic stress symptoms for people who have been harmed; and they function at a lower overall cost than courts and prisons. Each of these findings is validated by at least two randomized, controlled trials. This review also finds that restorative justice programs reduce re-offending more effectively for serious

crimes and crimes against persons than property crimes such as burglaries; however, mental health outcomes are still better for people who have experienced any of these crimes. In general: "The evidence on victims is far more consistent than it is on offenders. On average, in every test available, victims do better when they participate in RJ [restorative justice] than when they do not" (Sherman & Strang, 2007, p. 22).

Research on transformative justice is typically based on case studies rather than controlled trials (e.g., Kelly, 2011; Kim, 2011, 2021; Nocella & Anthony, 2011). Like restorative justice, there is a robust literature on transformative justice in handbooks, toolkits, white papers, and popular books (e.g., brown, 2020b; Chen et al., 2016; Creative Interventions, 2012; Dixon & Piepzna-Samarasinha, 2020b; Kershnar et al., 2007). Transformative justice activists and organizers often focus on their communities at local levels (e.g., the Bay Area Transformative Justice Collective and Philly Stands Up), but have also become leaders in influential social movements. For example, Kaba's (2021) book about prison abolition and transformative justice debuted on the *New York Times* bestseller list, likely reflecting her prominent role in national conversations about Black Lives Matter. This kind of impact is particularly relevant for an approach that does not restrict its ambitions to specific outcomes in specific cases. As activist-writer adrienne maree brown explains, transformative justice is focused on "practices that go all the way to the root of the problem and generate solutions and healing there, such that the conditions that create injustice are transformed" (brown, 2020a, p. 148).

Justice and Scale

Restorative and transformative justice practitioners generally agree that their processes vary across cases (Chen et al., 2016; Dixon & Piepzna-Samarasinha, 2020b; Johnstone & Van Ness, 2013; Kelly, 2011; Valandra & Yazzie, 2020). These processes often involve multiple community stakeholders, careful attention to each person's needs, one-on-one meetings, and sometimes group conferences. Each of these aspects can be time-consuming and resource-intensive. While transformative justice activists tend to avoid operating at large scales, some restorative justice programs pursue forms of scalability through standardization. Those that serve as diversion programs within the criminal legal system often need to have standardized processes, timelines, and outcomes. For example, prosecutors who refer people who have caused harm to these programs will typically only drop the charges if the process is complete within a certain timeframe.

At any scale, successful restorative justice programs depend on well-trained facilitators and context-appropriate processes. In a review of research documenting negative experiences with restorative justice, Jung Jin Choi, Gordon Bazemore, and Michael J. Gilbert warn that when practitioners are not properly trained in restorative justice, "processes

can produce adverse outcomes for some victims [if] they become offender focused or insensitive to the needs and concerns of victims" (Choi et al., 2012, p. 41). They explain,

> The primary means of preventing harms to victims is to ensure that RJ processes are, at all times, sensitive to the needs of victims. If not, the practices used become ethically questionable and at odds with the fundamental principles of RJ. . . . All practitioners including professional facilitators and volunteers [should] be adequately trained in restorative justice theory, values, principles and practices before they conduct RJ processes and provided periodic (at least annual) refresher or advanced training. (Choi et al., 2012, p. 39)

These researchers stress that with the correct training, practitioners can ensure that people who have been harmed are thoroughly prepared to participate in a process, including having a clear understanding of their role, realistic expectations, and a necessary awareness of the potential costs and benefits. Choi and coauthors note that "highly structured RJ processes" with "formalized artificial structure[s]" may lead people who have caused harm to undermine their apologies with deflections of shame and responsibility, which are perceived as insincere (Choi et al., 2012, p. 38). In other words, excessively standardized processes risk compounding the harm. Transformative justice practitioner Ejeris Dixon, likewise, highlights the need for creative facilitation and ongoing practice over rote procedures: "We must practice community safety as we would practice an instrument or a sport. By practicing in slow, measurable, and deliberate ways, we build the knowledge we need to diffuse and address conflict within our communities" (Dixon, 2020, p. 15).

Transformative justice advocates, in particular, express concern that efficient, standardized processes can limit the potential of their efforts to collectively change community norms. Practitioner Shira Hassan argues that transformative justice cannot scale without losing what distinguishes it from restorative approaches: "We know that TJ can't be scaled because we know what scaling looks like. It looks like RJ, we've got that" (Piepzna-Samarasinha, 2020, p. 168). For Hassan and others, when a practice becomes standardized and institutionalized for mass adoption it gains efficiency but loses the potential for genuine social justice. As Kaba explains,

> We have to embrace the messiness of [the] process. The messiness is inherent. It will always be there. And by messy, I mean that there are multiple U-turns that are happening all the time, that people are sometimes their best selves and sometimes not, that we move forward in some places and backwards in another, and that all this stuff is actually part of the work. (Dixon & Piepzna-Samarasinha, 2020a, p. 175)

For reasons like this, brown is skeptical about the potential of translating transformative justice to online spaces, particularly because of social media's apparent lack of tolerance for

slowness and deliberation: "Real time is slower than social media time, where everything feels urgent. Real time often includes periods of silence, reflection, growth, space, self-forgiveness, processing with loved ones, rest, and responsibility" (brown, 2020a, p. 150). In short, while scalability demands quick resolution to incidents, both restorative and transformative justice call for slower, more individualized care and negotiation.

While criminal legal systems tend to focus on determining a specific verdict or carrying out a sentence, transformative and restorative justice programs emphasize processes over outcomes. Choi and coauthors (2012) warn that focusing on achieving a specific outcome—such as getting an apology or closing the case—de-centers the person who has been harmed and risks re-traumatizing them. In contrast to outcome-focused practices, restorative justice and especially transformative justice processes often aim to use incidents of harm as opportunities to prevent the same kind of harm from happening again. For those who connect justice to social change, standardized processes and preset outcomes and expectations that aim to resolve cases only create more barriers. As historian Bench Ansfield and organizer Jenna Peters-Gordon explain: "If we reach for 'success,' we are undermining the work" in transformative justice; they aim instead "to make our organizing so nourishing and resilient that we can weather mistakes, build off of failures, and appreciate the revolutionary potential of even the subtlest gains" (Ansfield & Peters-Golden, 2013, p. 31). Applying this kind of justice to large online platforms will require rethinking the relationship between context and scale.

Subsidiarity

Subsidiarity is the principle that local social units should have meaningful autonomy within larger systems, and that such arrangements contribute to the health and accountability of the system as a whole. In this section, we explain how the logic of subsidiarity poses an alternative to scalability for online platforms, which could potentially support restorative and transformative approaches to justice.

The concept of subsidiarity first appeared in the writings of the seventeenth-century Calvinist legal philosopher Johannes Althaus, but during the past century it has been more widely associated with Catholic social teaching (Brennan, 2014; Follesdal, 1998). In the Catholic understanding of "social justice," justice cannot occur through only what a community receives from above but must involve meaningful self-determination in that community. Thus subsidiarity emerged as a theological principle—a way of understanding the social order in its relationship to the Church and to God—although it has also taken hold in secular contexts, such as in the organizing documents of the European Union. The Latin word *subsidium* refers to "help," and served as the basis of *subsidiarium*, the word for subsidiarity that Pope Pius XI used to describe the principle:

It is an injustice and at the same time a grave evil and disturbance of right order to assign to a greater and higher association what lesser and subordinate organizations can do. For every social activity ought of its very nature to furnish help to the members of the body social, and never destroy and absorb them. (Pius, 1931, sec. 79)

Nils Christie (1977) offers a similar critique of criminal justice institutions, arguing that courts and lawyers have "stolen" conflicts from the individuals involved and from local communities, which means that people who have been harmed lose the chance to participate meaningfully in the resolution of conflicts, and communities lose the opportunity to clarify their norms. He speculates about replacing national and state-level criminal legal systems with a proliferation of local neighborhood courts, without professional lawyers or judges, but populated instead with a rotation of all the members of that community.

Despite their abuses and failings, criminal legal systems tend to bear at least the intention of being attentive to local context—far more than social platforms do. For example, the common law systems found in many Anglophone countries produce layered and parallel bodies of case law. Local legislators can create their own laws, and local judges can interpret laws according to community norms, short of contradicting higher jurisdictions or local precedent. Sentencing, however unjust it might be, is at least intended to correspond to the precise nature of the offense. Context-sensitive legal systems are expensive to maintain and use, which disadvantages people who lack wealth. When jurisdictions try to avoid the costs of deliberate processes, such as by compelling people into plea bargains rather than trials, the result can appear to be a betrayal of justice. The injustice lies in something similar to what social media platforms do as a matter of course: applying uniform rules indiscriminately, without careful attention to the specifics of the case. Justice seems to demand subsidiarity.

Online social platforms already practice limited forms of subsidiarity. Reddit has user-managed Subreddits; YouTube has channels; Facebook has Groups; Wikipedia has distinct language-specific communities. Under the authority of user-moderators, these units operate with some independence from the corporate authorities that own and maintain the infrastructure and provide the moderation tools. The moderators of a Subreddit or a Facebook Group can set and enforce their own rules, and most moderators experience no direct interaction or training with platform companies (Seering et al., 2019). In this way, pseudo-subsidiarity enables platform companies to outsource significant portions of moderation labor to users they do not need to pay (Seering et al., 2019; Terranova, 2000), equipping them with only the blunt digital equivalents of blacklisting, censorship, and exile. Platforms provide a similarly limited range of options to volunteer community moderators (Schneider, 2022). While some communities and moderators address

conflicts and harm in ways that align with restorative and transformative justice (e.g., Xiao, 2021), these practices emerge despite the available software rather than because of it. As Joseph Seering et al. (2019) argues, meaningful community self-moderation has been relatively rare and inadequately studied. The most visible moderation decisions, such as those dealing with celebrity users, tend to occur through opaque corporate processes without clear accountability to user communities.

Subsidiarity is a framework that could help answer Schoenebeck and Blackwell's call for social media platforms "to transition away from paternalistic, top-down models of governance in favor of giving communities more control over their own experiences" (Schoenebeck & Blackwell, 2021, p. 150). Even the meaning of the term *community* differs widely from platform to platform and from group to group. Communities might be constituted by algorithmic matching, "follower" relationships, permissioned groups, and more. If communities are to hold more power over addressing harm among their participants, then participants must also be able to shape how their communities are defined in terms of both membership and conduct. Boundary definitions, in turn, shape how communities interact with each other and the larger systems of which they are part. In this and other ways, subsidiarity invites forms of repair that are sensitive to the context where the harm occurs while also contributing to the health of the larger system.

Subsidiarity and Scale

The concept of subsidiarity emerged historically not as a critique of large-scale systems so much as a strategy for building and organizing them. Applying the concept to online networks offers an opportunity to explore questions about how to design large-scale systems without the false promise of scalability. The need for subsidiarity has been a widespread recognition across diverse human societies; the extent to which online governance has neglected it is an aberration.

Consider, for instance, the cooperative movement—a form of business based on ownership and control by direct participants rather than outside investors (Schneider, 2018). Cooperatives aspire to a different kind of scale than investor-owned businesses by maintaining lean central offices and supporting diverse, culturally relevant practices among local member businesses. For example, while investor-owned retail chains may seek to drive local competitors out of business and replace them, a national-scale cooperative would more likely aid and enable local businesses through joint purchasing and shared services (Taylor, 2021). Under a logic of subsidiarity, the purpose of a central authority in cooperative networks is to support the local units in meeting local needs through their own processes. Whereas social media platforms typically accumulate the wealth and power their users create into centralized corporations, cooperative businesses tend to push their gains out to their members. Similarly, Divya Siddharth, Danielle Allen, and E. Glen Weyl (2022) argue

that subsidiary ownership and governance could enable emerging blockchain networks to achieve their often-stated aspiration of decentralization.

Following Shagun Jhaver, Seth Frey, and Amy Zhang's (2021) typology for multilevel governance, subsidiarity can occur at large scales in either federalist or polycentric forms. Federalism involves hierarchy (Bednar, 2009). Smaller units aggregate into larger units, which may form even larger units in turn. Federalist subsidiarity, then, prioritizes the vibrancy and autonomy of those local levels, except only when power from further up the hierarchy is necessary. The Catholic Church reflects a federalist structure, with priests leading semi-autonomous parishes under bishops led by a pope. In the Chinese political tradition, the concept of harmony explains how local autonomy can function beneath a strong central government through a federalist hierarchy, whether headed by an emperor or the Communist Party (Wang et al., 2016). This framework grants regional officials the freedom to diverge and experiment in contextually sensitive ways, while a commitment to harmony rules out direct challenges to the central government's authority. Subsidiarity likewise does not constitute a rejection of centralized power over large domains per se.

In contrast to federalism, the idea of polycentricity "connotes many centers of decision making that are formally independent of each other" (V. Ostrom et al., 1961). A system may include inequalities of authority, but these are not necessarily strict hierarchies with a single apex (Jhaver et al., 2021; E. Ostrom, 2010; V. Ostrom, 1972/1999). For instance, the authority structure of Calvinist subsidiarity—such as among the Congregationalism of Puritan colonizers in New England—is polycentric; power flows across a network of semi-autonomous congregations. Similarly, in Sunni Islam, authority flows through the *ulama*, the networked legal scholars belonging to various lineages and schools of thought (Zaman, 2007). In a given community, Muslims may be able to choose among jurists to consult for resolving a dispute, which creates competition among overlapping regimes. Authority derives from distinct juridical traditions simultaneously. Polycentric forms of subsidiarity can appear more anarchic than federalist ones, lacking any central authority, but they still enable a kind of shared identity among peer communities.

Polycentric designs might have a natural affinity with the structures of online networks, as well as with the vision of transformative justice to address harm without top-down coercive power. While transformative justice projects are loosely united by certain principles, handbooks, podcasts, and leading thinkers, which facilitates communication and skill-sharing across communities, they nonetheless resist central authority and forms of institutionalization. In contrast, restorative justice is often more federated in its relationship to criminal legal systems and state institutions. Federated structures have also thrived in online contexts, perhaps due to the legal requirement for central ownership over servers and branding, including in corporate-owned social media and nonprofit peer-production projects such as Wikipedia.

A genre of social media known as "federated networks" is designed for subsidiarity in ways that dominant platforms are not (Mansoux & Abbing, 2020; Zignani et al., 2018). Despite their name, however, federated networks are more polycentric than federalist. They rely on free and open source software, enabling users to host platforms on their own servers and connect with users on servers elsewhere. Examples include Mastodon (a microblogging platform), Matrix (a chat protocol), and PeerTube (a video-sharing platform). Individuals or groups who run servers on these networks can set tailored rules for content moderation, dispute resolution, and which other servers they "federate" with. In a few high-profile cases, this approach has demonstrated the capacity to confront far-right hate groups through peer-to-peer mechanisms—such as, for instance, coordinated blocking and pressuring software developers to implement better moderation features (Caelin, 2020). Federated networks manifest subsidiarity in that they facilitate local governance; the software is designed to maximize communities' contextual control while also connecting communities together into much larger systems.

Whether it occurs in federalist or polycentric forms, subsidiarity expects that processes for preventing and addressing harm are contextually appropriate and accountable to the people they most directly affect—and always in the context of a larger system. Unlike scalability, subsidiarity insists that there is no one-size-fits-all algorithm for different contexts. This difference makes space for relationship-based processes such as restorative and transformative justice. But subsidiarity is not a formula. It invites designers to imagine platform architectures that prioritize the craft of community self-governance rather than imposing universal solutions.

Design Considerations

Addressing online harm through subsidiarity, without relying on one-size-fits-all algorithms, requires a different approach to platform design. The governance processes built into the software for most online communities, including the ways they can address online harm, are crude and simplistic. Building with subsidiarity in mind could allow designers and user communities, if they so choose, to follow restorative and transformative justice in their departure from surveillance and punishment and turn toward practices of encounter, repair, and community accountability. In this section, we explore the unique affordances of social platforms for facilitating such processes, the potential to diversify responses to harm, and the need for inviting communities, rather than just individuals, to be accountable for harm.

Making Online Processes Appropriate to Online Contexts

Compared to offline abuses and conflicts, online contexts have affordances that can change the nature and experience of harm. Two particularly relevant affordances are the potential visibility and spreadability of communicative acts (boyd, 2014, p. 11), which can both amplify harm and enable people who cause harm to remain anonymous or pseudonymous. However, a subsidiarity-informed approach suggests that the technological affordances of social media may also offer advantages for restorative or transformative justice processes.

First, a recurring concern in institutionalized restorative justice programs is that people who have been harmed may feel coerced to accept apologies and repress their desires for retribution, while those who have caused harm might feel compelled to offer insincere apologies (Acorn, 2004; Menkel-Meadow, 2007; Suzuki & Wood, 2017). As such, many practitioners argue that community processes can only produce genuine personal or community change if they are truly voluntary—and indeed they are critical of restorative justice processes where unresolved cases return to prosecutors' offices. In other words, if joining and leaving a community is easier online, people can more easily opt out of an unwanted accountability process.

Second, the network effects of social media platforms facilitate the formation of groups of people around a common cause or interest. Group formation is vital for restorative and transformative justice processes, which often involve forming small, temporary support groups around the person who caused harm and the person who was harmed (Mingus, 2016). People who have been harmed can also find support from strangers on large networks; a range of studies has shown that social media can facilitate support for victims, involving people even outside a person's circle of friends and acquaintances (Powell, 2015; Rentschler, 2014, 2017; Vitis & Gilmour, 2017). One study found that many people who used hashtags such as #MeToo to share their experiences of sexual violence felt that signals of engagement, including from strangers, offered a meaningful acknowledgment of the harm they experienced (Mendes et al., 2018). Coordinating networks of support for a person who has been harmed can be especially valuable when those who caused harm are not engaging in a process.

Third, online interactions can be more accessible than offline ones. When the COVID-19 pandemic forced restorative and transformative justice practitioners to move their processes online, many raised concerns about the drawbacks of digital communication: inequalities of access to technology and adequate internet connection speeds, privacy and confidentiality concerns, the loss of eye contact and visible body language cues about participants' emotional states, superficial and simplistic interactions, and decreased intimacy and interpersonal connection (Paul & Borton, 2021). Yet conducting meetings and group conferences through video, audio, or text-only chat also presents some advantages, such as convenience and physical safety. While the physical distance in video chat or asynchronous communication is a drawback for many, for some participants, "it is precisely the emotional distance . . . [which allows them to] engage in a restorative process more fully" (Paul & Borton,

2021, p. 92). Long before the widespread availability of video-conferencing, practitioners have used letter-writing and other asynchronous media to allow people who have been harmed to communicate with people they are not comfortable interacting with in person.

Diversifying the Responses to Harm

Platform administrators are asked to adjudicate a wide range of incidents. Currently, social media platforms typically offer only a few basic options: refer to local police if the incident violates a criminal law; ban the user from the platform; remove the piece of content from the platform; limit the distribution of the content; do nothing (Caplan, 2018; Gillespie, 2018). Research indicates that many people who have been harmed by online harassment and abuse are not satisfied with these outcomes (e.g., Vogels, 2021). By fostering context-sensitive responses, including community-specific definitions of harm and the appropriate forms of redress, subsidiarity could broaden the range of options.

Systems oriented toward meeting the needs of people who have been harmed require flexible, diverse tools. In survey-based studies about hypothetical resolutions to online harassment, researchers found that individual preferences were associated with racial, gender, and class identities (Schoenebeck, Haimson, & Nakamura, 2021; Schoenebeck, Scott, et al., 2021). For example, "women and participants with higher parent education levels were more likely to prefer apologies in general" (Schoenebeck, Scott, et al., 2021, p. 8); the authors suggest that those with particular kinds of privilege might perceive apologies as more sincere. As such, facilitators need system designs that allow them to tailor the forms of repair and responses to harm to each individual case.

Moving to restorative or transformative frameworks for justice would push platforms away from their current approach to moderation that focuses almost exclusively on superficially removing or demoting objectionable content without addressing underlying causes. As Niloufar Salehi argues, this should represent an opportunity:

> What is needed is not more sophisticated ways to identify and remove offending content—just as we don't need better ways of policing and imprisoning people—but ways of supporting survivors and transforming the societies in which harm happens, including our online social worlds. (Salehi, 2020)

Likewise, Schoenebeck and Blackwell (2021) argue that a shift from content to addressing root causes "will allow social media platforms to . . . [implement] interventions that discourage harmful behaviors before they manifest on the platform" (Schoenebeck & Blackwell, 2021, p. 149). In one study, participants reflected on a specific incident of online harm and expressed desires both for immediate actions and for the "transformation of online environments in the long term" (Xiao et al., 2022, p. 12). Platforms could design social

spaces that enable a much wider range of options for decision-making and power-sharing (Schneider et al., 2021). If platforms do not allow communities to choose among diverse systems of organization and governance, transformation will be difficult.

Facilitating Accountability for Communities, Not Just Individuals

In both restorative justice and in the criminal-legal concept of a "jury of one's peers," communities are often assumed to be normalizing, positive moral forces. However, as criminologist Giuseppe Maglione (2017) argues, many communities are more contingent, unstable, and heterogeneous than people imagine. Furthermore, Judith Herman (2005) critiques restorative justice programs for failing to recognize that some forms of violence—particularly sexual violence—flourish because of community norms, not in spite of them. As such, communities often do not adequately support people who have experienced sexual violence:

> The so-called community cannot be counted on to do justice to victims because public attitudes toward these crimes are conflicted and ambivalent at best. . . . Because these crimes, by design, shame and stigmatize the victim, a restorative justice model, which relies on traditional community standards, will inevitably fail, for the same reason that the conventional justice system fails. Community standards are the standards of patriarchy. (Herman, 2005, p. 598)

Furthermore, in a study of three decades of restorative justice policy, Maglione finds that "the community" is regularly portrayed as "innocent," neglecting the fact that "social networks can be sites whereby deviant values are learned, rationalised and practiced" (Maglione, 2017, p. 459). Certainly, online communities are not "innocent" either, whether their internal norms facilitate harm to members or whether they encourage harm to other groups.

In structures based on subsidiarity, communities—not just individuals—can be accountable for the behavior they enable and encourage. In hierarchical systems, one way to mitigate the problem of communities that normalize harm is to train facilitators to be attentive to systems of privilege and oppression. Christina Parker points out that without a well-developed awareness of power, practitioners working in schools risk excluding and silencing some students:

> Any restorative dialogue requires thoughtful consideration of historical oppressions and how certain cultural groups, values, and beliefs are marginalized. Yet even in these dialogues, the dominant cultural group's worldview remains the default. (C. Parker, 2020, p. 69)

From a restorative justice perspective, careful training and attention to detail in group conferences enables facilitators to keep dynamics of power and patterns of marginalization in mind and work toward avoiding the reproduction of existing

forms of oppression. At the same time, when a community normalizes or even encourages harm, the impetus to change may need to come from outside of it. In a federalist system, higher levels of authority might set standards for facilitator training and carry out audits of the local units. In polycentric systems, a group of communities might initiate a process to confront harmful patterns in another community in their network. The case of how an Ojibway First Nations community in Manitoba, Canada addressed endemic sexual abuse, as depicted in the film *Hollow Water* (Dickie, 2000), illustrates a combination of federalist government resources and polycentric support from outside tribal leaders. While the federal criminal justice system prosecuted some people, punishing individuals had failed to end the cycles of abuse. The film chronicles the culturally specific practices they chose to engage in to both foster individual accountability and to work toward transforming community norms, which involved inviting leaders from other places to help advance the process.

The challenge and strength of subsidiarity as a framework is that it demands the definition and management of relationships between the local units and the larger systems. Context-specific efforts to meet the needs of people who have been harmed cannot stand alone; there must also be value-based mechanisms for challenging the internal norms and outward practices of wider communities when necessary. Schoenebeck and Blackwell stress that "social media governance must be principled rather than neutral, and that a principled approach requires platforms to reckon with their role in enabling, or magnifying, structural injustices" (Schoenebeck & Blackwell, 2021, p. 144). The same goes for the communities that platforms host. While governance based on scalability may offer an appealing clarity through standardized top-down control, subsidiarity invites designs based on negotiation—between the local and the global, between habits and innovations, between resolving the conflicts at hand and transforming their underlying causes.

Conclusion

This article has offered an alternative to the logic of scalability for addressing online harm by drawing on diverse legacies, from centuries-old religious doctrines to recent experiments in non-carceral justice. Practices of restorative and transformative justice have shown that to address harm without violence or coercion, it is necessary to engage in highly context-sensitive work in communities where harm has occurred. The concept of subsidiarity offers a frame for doing just that with locally autonomous processes inside large-scale networks.

Transformative justice practitioner Shira Hassan argues that it is not possible to standardize transformative justice while keeping its core intentions and aims intact. Hassan says, "What we need is for as many organic pockets of people who are practicing the work to start documenting it. So that we can understand all the different kinds of intersections. And learn from all those intersections to create a

better practice together" (Piepzna-Samarasinha, 2020, p. 168). We argue that subsidiarity is a framework that can help platform designers create spaces, tools, and networks to support those "organic pockets of people."

The imperative for rapid, prolonged growth that social media makes possible—and that venture-capital investors' business models demand (Graham, 2012)—may be fundamentally incompatible with taking justice seriously, particularly in ways that people perceive as fair, accountable, and non-coercive. While scalability may enable social platforms to grow exponentially at low marginal cost (G. G. Parker et al., 2016), genuine subsidiarity resists such economies of scale. Just as any new city requires its own schools, libraries, and plumbing, each new community on an online platform cannot rely solely on processes established elsewhere for resolving its conflicts. As Gillespie (2020) writes, the inherent complexities of human conflict online "should be understood as a limiting factor on the 'growth at all costs' mentality" that has become so widespread in the dominant internet economy. Without enabling the human beings in the loop to engage in human-scale processes, platforms can expect the persistence of the symptoms that have become commonplace: anxieties about "cancel culture," failures to protect vulnerable people from abuse, and a feeling that rules are enforced arbitrarily. Having more accountable responses to harm online may require departing from the dominant engines of growth on which the online economy has relied. If what we argue for is radical that is only because platforms' drive to scalability has become such an unquestioned norm. Past scholarship suggests that scalability is precisely what enables platforms to wield so much economic and social power; challenging scalability means challenging that power.

While offline restorative programs are less expensive than courts and incarceration, embracing subsidiarity as an approach to online harm would likely cost more for a large platform, because they would need to replace some of their semi-automated moderation with trained human facilitators. Much of the savings might occur outside the platform at first, such as in reduced need for public resources, hotlines, and private individual counseling to help people deal with online abuse and harassment. Recognizing this, governments may see fit to require that platforms provide support for more holistic processes and support systems when harm occurs, rather than allowing platforms to simply hide the problem from view and externalize the true cost. Platforms might even find that, over the long term, providing resources for communities to build their capacity to address harm and resolve conflicts might be more cost-effective than self-perpetuating cycles of harm and punishment, with related expenses of litigation, lobbying, and public relations.

We contend that the framework of subsidiarity provides a pathway out of the dilemma platform studies scholars have identified between large-scale systems and human-scaled responses to harm. This opens opportunities for new work on how to achieve context-sensitive problem solving across large networks. First, ethnographers might better understand

online practices that align with restorative and transformative justice, identifying the forms they take and the resources they require. Second, engineers might strive to create software that is suited to supporting practices of subsidiarity, such as restorative and transformative practices and community-level control. Third, policymakers might explore, at the levels of both government regulation and corporate governance, how to facilitate and protect community-based processes. Because subsidiarity may require new social media business models—or at least more resources for platforms' trust and safety departments—further research on such strategies is vital.

Digital technologies have altered many aspects of social and economic life, making once-difficult tasks appear simpler and easier. Resolving conflict and addressing harm, however, are challenges that have only deepened with the advent of online social media. They neither can nor should be automated away. Emerging online social infrastructures must not take the business imperatives of speed, growth, and scalability for granted at the expense of the demanding but vital work of justice.

Declaration of Conflicting Interests

Funding

ORCID iD

Amy A. Hasinoff ![ORCID] https://orcid.org/0000-0001-7530-8278
Nathan Schneider ![ORCID] https://orcid.org/0000-0002-2745-0533

References

Acorn, A. E. (2004). *Compulsory compassion a critique of restorative justice*. UBC Press. http://www.deslibris.ca/ID/404141

Adema, J., & Moore, S. A. (2021). Scaling small; or how to envision new relationalities for knowledge production. *Westminster Papers in Communication and Culture*, *16*(1), Article 1. https://doi.org/10.16997/wpcc.918

Alliance for Safety and Justice. (2016). *Crime survivors speak: The first-ever national survey of victims' views on safety and justice*. Alliance for Safety and Justice. https://allianceforsafetyandjustice.org/wp-content/uploads/documents/Crime%20Survivors%20Speak%20Report.pdf

Ansfield, B., & Peters-Golden, J. (2013). How we learned not to succeed in transformative justice. *Make/Shift: Feminisms in Motion*, *12*, Article 31.

Bednar, J. (2009). *The robust federation: Principles of design*. Cambridge University Press.

boyd, d. (2014). *It's complicated: The social lives of networked teens* (1st ed.). Yale University Press.

Braithwaite, J. (2002). *Restorative justice & responsive regulation*. Oxford University Press.

Brennan, P. M. (2014). Subsidiarity in the tradition of catholic social doctrine. In M. Evans & A. Zimmermann (Eds.), *Global perspectives on subsidiarity* (pp. 29–47). Springer Netherlands. https://doi.org/10.1007/978-94-017-8810-6_3.

brown, adrienne maree. (2020a). What is/isn't transformative justice? In E. Dixon & L. L. Piepzna-Samarasinha (Eds.), *Beyond survival: Strategies and stories from the transformative justice movement* (pp. 148–150). AK Press.

brown, adrienne maree. (2020b). *We will not cancel us: And other dreams of transformative justice*. AK Press.

Caelin, D. (2020). *Decentralized social networks vs. the trolls*. ActivityPub Conference. https://conf.tube/videos/watch/d8c8ed69-79f0-4987-bafe-84c01f38f966

Caplan, R. (2018). *Content or context moderation? Artisanal, community-reliant, and industrial approaches*. Data & Society Research Institute. https://datasociety.net/wp-content/uploads/2018/11/DS_Content_or_Context_Moderation.pdf

Chartrand, L. N., & Horn, K. (2016). *A report on the relationship between restorative justice and Indigenous legal traditions in Canada*. Department of Justice Canada. https://www.justice.gc.ca/eng/rp-pr/jr/rjilt-jrtja/index.html

Chen, C.-I., Dulani, J., & Piepzna-Samarasinha, L. L. (2016). *The revolution starts at home: Confronting intimate violence within activist communities* (2nd ed.). AK Press. https://www.akpress.org/revolutionstartsathome.html

Choi, J. J., Bazemore, G., & Gilbert, M. J. (2012). Review of research on victims' experiences in restorative justice: Implications for youth justice. *Children and Youth Services Review*, *34*(1), 35–42. https://doi.org/10.1016/j.childyouth.2011.08.011

Christie, N. (1977). Conflicts as property. *British Journal of Criminology*, *17*(1), 1–15. https://doi.org/10.1093/oxfordjournals.bjc.a046783

Creative Interventions. (2012). *Creative interventions toolkit: A practical guide to stop interpersonal violence*.

Daly, K. (2014). Reconceptualizing sexual victimization and justice. In I. Vanfraechem, A. Pemberton, & F. M. Ndahinda (Eds.), *Justice for victims: Perspectives on rights, transition and reconciliation* (pp. 378–396). Routledge.

Daly, K. (2016). What is restorative justice? Fresh answers to a vexed question. *Victims & Offenders*, *11*(1), 9–29. https://doi.org/10.1080/15564886.2015.1107797

Dickie, B. (2000). *Hollow water*. https://www.nfb.ca/film/hollow_water/

Dixon, E. (2020). Building community safety: Practical steps toward liberatory transformation. In E. Dixon & L. L. Piepzna-Samarasinha (Eds.), *Beyond survival: Strategies and stories from the transformative justice movement* (pp. 11–17). AK Press.

Dixon, E., & Piepzna-Samarasinha, L. L. (2020a). Be humble: An interview with Mariame Kaba. In E. Dixon & L. L. Piepzna-Samarasinha (Eds.), *Beyond survival: Strategies and stories from the transformative justice movement* (pp. 170–178). AK Press.

Dixon, E., & Piepzna-Samarasinha, L. L. (Eds.). (2020b). *Beyond survival: Strategies and stories from the transformative justice movement*. AK Press.

Douek, E. (2019). Facebook's oversight board: Move fast with stable infrastructure and humility. *North Carolina Journal of Law & Technology*, *21*, Article 1.

Feiner, L., & Rodriguez, S. (2019, May 23). Mark Zuckerberg: Facebook spends more on safety than Twitter's whole revenue for the year. *CNBC*. https://www.cnbc.com/2019/05/23/facebook-fake-account-takedowns-doubled-q4-2018-vs-q1-2019.html

Follesdal, A. (1998). Survey article: Subsidiarity. *Journal of Political Philosophy*, *6*(2), 190–218. https://doi.org/10.1111/1467-9760.00052

Garfield, B. (2019, August 2). Repairing justice: How to fix the internet. *On the Media*. https://www.wnycstudios.org/podcasts/otm/episodes/on-the-media-fix-the-internet

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Big Data & Society*, *7*(2). https://doi.org/10.1177/2053951720943234

Graham, P. (2012, September). *Startup = Growth*. http://paulgraham.com/growth.html

Haldane, J. B. S. (1926, March 26). On being the right size. *Harper's Magazine*. http://www.phys.ufl.edu/courses/phy3221/spring10/HaldaneRightSize.pdf

Hanna, A., & Park, T. M. (2020). Against scale: Provocations and resistances to scale thinking. In *Proceedings of the CSCW 2020 Workshop Reconsidering Scale and Scaling in CSCW Research*. http://arxiv.org/abs/2010.08850

Hasinoff, A. A., Gibson, A. D., & Salehi, N. (2020, July 27). *The promise of restorative justice in addressing online harm. Brookings: Tech Stream*. https://www.brookings.edu/techstream/the-promise-of-restorative-justice-in-addressing-online-harm/

Herman, J. L. (2005). Justice from the victim's perspective. *Violence against Women*, *11*(5), 571–602. https://doi.org/10.1177/1077801205274450

Jhaver, S., Birman, I., Gilbert, E., & Bruckman, A. (2019). Human-machine collaboration for content regulation: The case of Reddit automoderator. *ACM Transactions on Computer-Human Interaction*, *26*(5), 1–35. https://doi.org/10.1145/3338243

Jhaver, S., Frey, S., & Zhang, A. (2021). *Designing for multiple centers of power: A Taxonomy of multi-level governance in online social platforms*. ArXiv:2108.12529 [cs]. http://arxiv.org/abs/2108.12529

Johnstone, G., & Van Ness, D. (Eds.). (2013). *Handbook of restorative justice*. Routledge.

Kaba, M. (2021). *We do this 'til we free us: Abolitionist organizing and transforming justice*. Haymarket.

Kelly, E. L. (2011). Philly stands up: Inside the politics and poetics of transformative justice and community accountability in sexual assault situations. *Social Justice*, *37*(4(122)), 44–57.

Kershnar, S., Haines, S., Harkins, G., Greig, A., Wiesner, C., Levy, M., Shah, P., Kim, M., & Carr, J. (2007). *Toward transformative justice: A liberatory approach to child sexual abuse and other forms of intimate and community violence*. generationFIVE.

Kim, M. E. (2011). Moving beyond critique: Creative interventions and reconstructions of community accountability. *Social Justice, 37*(4(122)), 14–35. https://doi.org/10.2307/41478931

Kim, M. E. (2021). Transformative justice and restorative justice: Gender-based violence and alternative visions of justice in the United States. *International Review of Victimology*, *27*, 162–172. https://doi.org/10.1177/0269758020970414

Klonick, K. (2019). The Facebook oversight board: Creating an independent institution to adjudicate online free expression. *Yale Law Journal*, *129*, Article 2418.

Maglione, G. (2017). Communities at large: An archaeological analysis of the "community" within restorative justice policy and laws. *Critical Criminology*, *25*(3), 453–469. https://doi.org/10.1007/s10612-017-9349-8

Mansoux, A., & Abbing, R. R. (2020). Seven theses on the fediverse and the becoming of FLOSS. In K. Gansing & I. Luchs (Eds.), *The eternal network: The ends and becomings of network culture*. Institute of Network Cultures. https://monoskop.org/images/c/cc/Mansoux_Aymeric_Abbing_Roel_Roscam_2020_Seven_Theses_on_the_Fediverse_and_the_Becoming_of_FLOSS.pdf

Mendes, K., Ringrose, J., & Keller, J. (2018). #MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies*, *25*(2), 236–246. https://doi.org/10.1177/1350506818765318

Menkel-Meadow, C. (2007). Restorative justice: What is it and does it work? *Annual Review of Law and Social Science*, *3*(1), 161–187. https://doi.org/10.1146/annurev.lawsocsci.2.081805.110005

Mingus, M. (2016, June 2). Pods and pod mapping worksheet. *Bay Area Transformative Justice Collective*. https://batjc.wordpress.com/resources/pods-and-pod-mapping-worksheet/

Moyle, P., & Tauri, J. M. (2016). Māori, family group conferencing and the mystifications of restorative justice. *Victims & Offenders*, *11*(1), 87–106. https://doi.org/10.1080/15564886.2015.1135496

Myers West, S. (2018). Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society*, *20*(11), 4366–4383. https://doi.org/10.1177/1461444818773059

Nagin, D. S. (2013). Deterrence in the twenty-first century. *Crime and Justice*, *42*(1), 199–263. https://doi.org/10.1086/670398

Nocella, A. J., & Anthony, J. (2011). An overview of the history and theory of transformative justice. *Peace & Conflict Review*, *6*(1), 1–10.

Nurik, C. (2019). "Men are scum": Self-regulation, hate speech, and gender-based censorship on Facebook. *International Journal of Communication*, *13*, 2878–2898.

Ostrom, E. (2010). Beyond markets and states: Polycentric governance of complex economic systems. *The American Economic Review*, *100*(3), 641–672.

Ostrom, V. (1999). Polycentricity (Part 1). In M. D. McGinnis (Ed.), *Polycentricity and local public economies: Readings from the workshop in political theory and policy analysis*. University of Michigan Press. (Original work published 1972)

Ostrom, V., Tiebout, C. M., & Warren, R. (1961). The organization of government in metropolitan areas: A theoretical inquiry. *American Political Science Review*, *55*(4), 831–842. https://doi.org/10.1017/S0003055400125973

Parker, C. (2020). Who's in and who's out? Problematizing peacemaking circles in diverse classrooms. In E. C. Valandra & J. R. Yazzie (Eds.), *Colorizing restorative justice: Voicing our realities* (pp. 65–86). Living Justice Press.

Parker, G. G., Alstyne, M. W. V., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. W.W. Norton.

Paul, G. D., & Borton, I. M. (2021). *Creating restorative justice: A communication perspective of justice, restoration, and community*. https://rowman.com/ISBN/9781498576451/Creating-Restorative-Justice-A-Communication-Perspective-of-Justice-Restoration-and-Community

Piepzna-Samarasinha, L. L. (2020). Every mistake I've ever made: An Interview with Shira Hassan. In E. Dixon & L. L. Piepzna-Samarasinha (Eds.), *Beyond survival: Strategies and stories from the transformative justice movement* (pp. 163–169). AK Press.

Pius, X. I. (1931). *Quadragesimo Anno: On reconstruction of the social order*. https://w2.vatican.va/content/pius-xi/en/encyclicals/documents/hf_p-xi_enc_19310515_quadragesimo-anno.html

Powell, A. (2015). Seeking rape justice: Formal and informal responses to sexual violence through technosocial counterpublics. *Theoretical Criminology*, *19*(4), 571–588.

Project Nia & Interrupting Criminalization. (n.d.). *One million experiments*. https://millionexperiments.com

Rentschler, C. (2014). Rape culture and the feminist politics of social media. *Girlhood Studies*, *7*(1), 65–82.

Rentschler, C. (2017). Bystander intervention, feminist hashtag activism, and the anti-carceral politics of care. *Feminist Media Studies*, *17*(4), 565–584. https://doi.org/10.1080/14680777.2017.1326556

Salehi, N. (2020, August 31). Do no harm. *Logic Magazine*. https://logicmag.io/care/do-no-harm/

Scheuerman, M. K., Jiang, J. A., Fiesler, C., & Brubaker, J. R. (2021). A framework of severity for harmful content online. *Proceedings of the ACM on Human-computer Interaction, 5(CSCW2)*, 368:1–368:33. https://doi.org/10.1145/3479512

Schneider, N. (2018). *Everything for everyone: The radical tradition that is shaping the next economy*. Nation Books.

Schneider, N. (2022). Admins, mods, and benevolent dictators for life: The implicit feudalism of online communities. *New Media & Society*, *24*, 1965–1985. https://doi.org/10.1177/1461444820986553

Schneider, N., De Filippi, P., Frey, S., Tan, J. Z., & Zhang, A. X. (2021). Modular politics: Toward a governance layer for online communities. *Proceedings of the ACM on Human-Computer Interaction*. https://dl.acm.org/doi/10.1145/3449090

Schoenebeck, S., & Blackwell, L. (2021). Reimagining social media governance: Harm, accountability, and repair. *Yale Journal of Law and Technology*, *23*(1), 113–152.

Schoenebeck, S., Haimson, O. L., & Nakamura, L. (2021). Drawing from justice theories to support targets of online harassment. *New Media & Society*, *23*, 1278–1300. https://doi.org/10.1177/1461444820913122

Schoenebeck, S., Scott, C. F., Hurley, E. G., Chang, T., & Selkie, E. (2021). Youth trust in social media companies and expectations of justice: Accountability and repair after online harassment. *Proceedings of the ACM on Human-Computer Interaction*, *5*, 1–18. https://doi.org/10.1145/3449076

Seering, J., Wang, T., Yoon, J., & Kaufman, G. (2019). Moderator engagement and community development in the age of algorithms. *New Media & Society*, *21*, 1417–1443. https://doi.org/10.1177/1461444818821316

Shah, S., Stauffer, C., & King, S. (2017). *Restorative justice listening project*. Zehr Institute for Restorative Justice. https://3b191c22-5460-4333-a26b-100c5a142cec.filesusr.com/ugd/c28cc3_1c445161c5ff4db9a15d7b59282283dc.pdf

Sherman, L., & Strang, H. (2007). *Restorative justice: The evidence*. Smith Institute.

Siddarth, D., Allen, D., & Weyl, E. G. (2022, May 12). The Web3 decentralization debate is focused on the wrong question. *Wired*. https://www.wired.com/story/web3-blockchain-decentralization-governance/

Sullivan, T. (2016, April 1). Blitzscaling. *Harvard Business Review*. https://hbr.org/2016/04/blitzscaling

Suzuki, M., & Wood, W. R. (2017). Co-option, coercion and compromise: Challenges of restorative justice in Victoria, Australia. *Contemporary Justice Review*, *20*(2), 274–292. https://doi.org/10.1080/10282580.2017.1311194

Tauri, J. M. (2020). Indigenous peoples and the globalization of restorative justice. *Social Justice*, *43*(3), 46–67.

Taylor, K. (2021). An analysis of the entrepreneurial institutional ecosystems supporting the development of hybrid organizations: The development of cooperatives in the U.S. *Journal of Environmental Management*, *286*, 112244.

Terranova, T. (2000). Free labor: Producing culture for the digital economy. *Social Text*, *18*(2(63)), 33–58. https://doi.org/10.1215/01642472-18-2_63-33

Tsing, A. L. (2012). On nonscalability: The living world is not amenable to precision-nested scales. *Common Knowledge*, *18*(3), 505–524. https://doi.org/10.1215/0961754X-1630424

Valandra, E. C., & Yazzie, J. R. (2020). *Colorizing restorative justice: Voicing our realities*. Living Justice Press.

Vitis, L., & Gilmour, F. (2017). Dick pics on blast: A woman's resistance to online sexual harassment using humour, art and Instagram. *Crime, Media, Culture*, *13*(3), 335–355. https://doi.org/10.1177/1741659016652445

Vogels, E. A. (2021). *The state of online harassment* (p. 54). Pew Research Center.

Wang, X., Juffermans, K., & Du, C. (2016). Harmony as language policy in China: An internet perspective. *Language Policy*, *15*(3), 299–321. https://doi.org/10.1007/s10993-015-9374-y

Warzel, C. (2019, August 20). Could restorative justice fix the internet? *New York Times*. https://www.nytimes.com/2019/08/20/opinion/internet-harassment-restorative-justice.html

Wemmers, J.-A. (2002). Restorative justice for victims of crime: A victim-oriented approach to restorative justice. *International Review of Victimology*, *9*(1), 43–59. https://doi.org/10.1177/026975800200900104

Xiao, S. (2021). *A restorative justice approach to addressing harm in online gaming communities* [Unpublished manuscript].

Xiao, S., Cheshire, C., & Salehi, N. (2022, April). Sensemaking, support, safety, retribution, transformation: A *restorative justice approach to understanding adolescents' needs for addressing online harm*, CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3491102.3517614

Zaman, M. Q. (2007). *The Ulama in contemporary Islam: Custodians of change*. Princeton University Press.

Zehr, H. (2015). *The little book of restorative justice* (Revised and updated). Good Books.

Zignani, M., Gaito, S., & Rossi, G. P. (2018, June 15). *Follow the "mastodon": Structure and evolution of a decentralized online social network*. Twelfth International AAAI Conference on Web and Social Media. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17862

Author Biographies

Amy A. Hasinoff (PhD, University of Illinois at Urbana-Champaign) is an Associate Professor of Communication at the University of Colorado Denver. Her research interests include gender and sexuality in the context of new media.

Nathan Schneider (MA, University of California, Santa Barbara) is an Assistant Professor of Media Studies at the University of Colorado Boulder. His research interests include social movements, religion, and the governance of online life.