

ORIGINAL ARTICLE

Does Drill Rap Cause Violence, and, Even if it Does, Should it Be Censored?

Tareeq Jalloh

The University of Oxford, Oxford, UK

Correspondence: Tareeq Jalloh (tareeqjalloh@gmail.com)

Received: 16 April 2025 | Revised: 19 December 2025 | Accepted: 24 December 2025

Keywords: black art | censorship | drill music | feminist philosophy of pornography | police restrictions | rap music

ABSTRACT

The critique that rap causes violence (causal claim) goes back as far as the 1980s, but today drill rappers are having their freedoms restricted based on instances of a causal claim. Police in the UK regularly remove drill videos from social media and streaming platforms, and drillers are being restricted from making music and associating with people via criminal behavior orders based on the claim that drill music causes serious violence. If this causal claim is true, it provides potential grounds for such restrictions on drillers. This paper argues against censoring drill on the assumption that it causes violence for three reasons. First, even according to the most charitable rendition of the causal claim, there is insufficient evidence to support it. Second, there are less speech-restrictive and more effective approaches to violence reduction than censorship-oriented approaches. Third, critics should not be advancing the causal claim due to its racist public meaning.

1 | Introduction

There has been recent philosophical interest in the way drill rap music—a subgenre of hip hop, originating in Chicago during the early 2010s—has been treated as incriminating evidence by the police and courts.¹ Jalloh [2] and Nowak [3], for instance, argue that when law enforcement treats rap as autobiographical, this wrongs rappers as knowers, constituting various epistemic injustices. Magnus and Malone [4] argue that when rap is treated in this way, rappers suffer art-interpretive injustices, harmed in their creative capacities as artists. Whilst these claims pay attention to whether rap music *evidences* violence (or what I call the *evidencing claim*), much less philosophical attention has been paid to the critique that rap *causes* violence (the *causal claim*). The critique that rap causes violence goes back as far as the 1980s. Today drill rappers (drillers) are having their freedoms restricted based on instances of the causal claim. For instance, police regularly remove drill videos from social media and streaming platforms, and drillers are being restricted from

making music and associating with people via criminal behavior orders based on the assumption that drill music causes serious violence [5]. It is an important question whether the causal claim is true, and whether it justifies such restrictions.

In this paper, I argue against censoring drill on the assumption that it causes violence for three reasons. First, even according to the most charitable rendition of the causal claim, there is insufficient evidence to support it. Second, there are less speech-restrictive and more effective approaches to violence reduction than censorship-oriented approaches. Third, critics should be worried about the racist public meaning of advancing the causal claim and the censorship-oriented approach to drill. The paper proceeds as follows: Section 1 motivates the need for an investigation of the causal claim. Section 2 uses a framework found in the feminist literature on pornography to characterize a sensible articulation of the causal claim [6], and shows that the current evidence claiming to support the view that rap causes violence is unlike the evidence necessary to support this view. Section 3

argues that restricting drillers on the basis of the causal claim is unjustified. Section 4 concludes by arguing that critics should not be advancing the causal claim due to its racist public meaning.

2 | Motivating an Investigation of the Causal Claim

According to police recorded data and crime survey data, 2018 saw a significant increase in violent crime (including knife crime) in England and Wales [7]. This was followed by a media panic, which often blamed drill rap for this rise [8, 9]. This incentivized the head of the Metropolitan Police—Cressida Dick—to petition online video platform YouTube to remove drill content, with around 30 music videos removed and some drill artists blocked from uploading [9]. The policy of removing YouTube videos has continued to this day, with at least 1636 videos removed by the Metropolitan Police (the Met) to date [10]. One of the Met's justifications for this practice of removal is that it removes content containing "information on gang violence and murders" which "could lead to further retaliatory violence" [11].² They are concerned with how drill lyrics and videos, describing stabbings and taunting rivals, can "raise the risk of [gang] violence" [12]. In response to a freedom of information request asking the Met to provide details about flagging drill videos to social media and music streaming companies for removal, the Met claimed that "The MPS (the Metropolitan Police Service) ask social media platforms to take robust and rapid approach to removing content which is flagged as being harmful that is, likely to provoke/cause real-life violence" ([8], emphasis added). The worry that drill music causes violence is also reflected in a response from a spokesperson of the Home Office, who responded to concerns about the censorship of drill music by claiming, "We are committed to protecting freedom of expression but are clear *causing* or inciting serious violence and other serious crimes is not acceptable" [13] (emphasis added).³

Academics have questioned the implications of such policies. Klienberg and McFarlane assert, The current policy of removing drill music videos from social media platforms such as YouTube remains controversial because "it risks conflating the co-occurrence of drill rap and violence with a causal chain of the two" [15]. Lynes et al. argue that to a popular readership, censoring drill lends misplaced credence "to the suggestion that drill music was causing violence or at least that certain drill music artists used the music genre to incite violent crime" [9], p. 1.

In order to establish whether restrictions based on the claim that rap music causes violence are justified, critics have at least two tasks. First, they must provide evidence that rap music causes violence—call this the *evidentiary burden*. The question whether drill has some causal role in instigating violence of some sort is a big question. My aim is narrower: to challenge the Met on their own grounds by querying whether it can at least be shown that it causes retaliatory gang violence. This makes for a more charitable reconstruction of their claims, as one might think that it would be easier to establish a causal link to gang violence. Second, critics must show that restricting rappers' freedoms is justified on the basis of this evidence—call this the *justificatory burden*. In the following sections, I argue that neither burden has been met.

3 | The Evidentiary Burden

When investigating the causal claim—that rap music causes violence—it may seem all that is needed to decide is empirical data, and philosophical enquiry would be useless in our endeavors. However, to avoid imprecision and to ensure charitable interpretations of those advancing the causal claim, we need to be precise on the terms in which the critique is framed—"cause" and "violence"—if we are to assess whether imposing significant restrictions of freedoms on people is justified. This is because it might be unclear what kind of violence critics of drill are talking about. Moreover, we need to make precise the kind of causal relationship critics are looking to prove.

Other forms of media have also been accused of causing violence. A popular media form that is often criticized for its alleged connection to violent harm is pornography. Pornography has been criticized on many fronts, and the critique that pornography causes harm—violent and otherwise—is known as the "harm hypothesis." There is an extensive literature on the proposed harms of pornography, and within this literature, we find frameworks that make sense of what "cause" and "harm" mean. One especially influential framework developed to address this issue is offered by Anne Eaton [6]. Eaton critiques the conception of "cause" and "harm" typically employed in pornography debates. I will set out Eaton's claims and show how we can employ an analogous framework to address debates about rap causing violence.

Eaton theorizes causation in a non-deterministic way. She draws on a probabilistic account of causation used in epidemiology.⁴ The process by which pornography may shape attitudes and/or may lead to harmful behavior can be specified by drawing an analogy with smoking [6]. Smoking does not cause cancer in a deterministic sense: many smokers do not develop lung cancer. Smoking is neither necessary nor sufficient to develop lung cancer. Still, medical experts and the general public accept as true the claim that "smoking causes cancer" since it increases the likelihood of developing cancer. This is a probabilistic notion of causation, but it is still enough to vindicate the claim that smoking causes cancer.⁵ More specifically, according to this notion of causation:

x is a cause of y if and only if (i) x occurs earlier than y and (ii) the probability of the occurrence of y is greater, given the occurrence of x, than the probability of the occurrence of y given not-x [6], p. 696.

If pornography increases the probability of, say, sexualized violence against women, then we would have vindicated a relationship between pornography and harm to women.

With this notion of causation in mind, Eaton makes the following suggestions of what a sensible rendition of the harm hypothesis might involve:

1. Because the causal relation between pornography and its purported harms is probabilistic, exposure to pornography is neither necessary nor sufficient for its putative harms, like sexual assault, but raises the chances of sexual assault

depending on context (its relationship to other factors, see 2).

2. Proponents of the harm hypothesis should emphasize that violence against women is the cumulative effect of several factors—such as addiction, poverty, background misogyny and adverse childhood experiences—in a complex causal mechanism, and inequalities in pornography is just one of these factors.
3. Antipornography feminists should specify the type of pornography that is the target of their critique, which is inequalities in pornography and no other types of pornography which may be liberatory
4. Inequalities in pornography's role in sexism should be considered as a feedback loop: inequalities in pornography results from gender inequality and also facilitates this inequality.
5. An improved harm hypothesis should articulate a “careful and nuanced delineation of pornography's alleged harms, recognizing a wide range of potential injuries that differ in terms of character and severity” [6], p. 714.
6. A sensible rendition of the harm hypothesis should be cautious and not necessarily in favor of state regulation. This rendition is sensitive to the entire range of factors potentially contributing to pornography's alleged harms (like sexual assault) when proposing measures of prevention and redress.

Before I extend Eaton's framework to the case of drill rap, it is important to note some clarifications. With regards to suggestion 3—specifying the exact type of pornography/rap that is the target of their critique—those advancing the causal claim and basing restrictions on it (the Metropolitan police) are primarily worried about drill rap. Drill rap is a subgenre of rap—heavily influenced by “gangsta rap”—which often references specific incidents of violence in taunting ways. The police are not worried about rap music that depicts violence or music that just has violent themes, but rather they are worried about how drill rap glorifies violence and antagonizes rivals in ways that might cause gang members who are taunted to commit acts of retaliatory violence. It should not be a critique of rap as a genre since there is a lot of non-violent rap, and there is no reason to think that non-violent rap would cause violent attitudes or behaviors.

Relatedly, regarding suggestion 5, Eaton is worried about pornography raising the risk of harmful attitudes or behaviors in the general population, whether that is an acceptance of rape myths, underestimating women's intellects, or sexual assault. Similarly, a sensible rendition of the causal claim about drill would be aware of the potential harms that might come about from listening to drill music, and see what harm results with what probability. For instance, a sensible rendition of the causal claim might outline harms that range from increased violent attitudes in the general population to increased violent behaviors, including retaliatory gang violence. However, given that this paper is particularly concerned with whether the policing of drill is justified, the primary harm I investigate is the one most salient to the police, which is that drill rap causes retaliatory gang violence.⁶ Whilst this is disanalogous to Eaton's fifth

suggestion, it is a departure that enables us to zoom in on the kinds of harm the police claim licenses censorship. In this sense, it is a more specific version of a broader rendition of a sensible causal claim.

With this notion of causation in mind, I construct an analogous sensible rendition of the claim that (R)ap causes violence:

R1. Because the causal relation between rap and offline retaliatory gang violence is probabilistic, exposure to rap is neither necessary nor sufficient for offline retaliatory gang violence, but raises the chances of offline retaliatory gang violence depending on context (its relationship to other factors, see R2).

R2. Proponents of the causal claim should emphasize that offline retaliatory gang violence is the cumulative effect of several factors—such as, poverty, inequality, lack of opportunity and funding, lack of access to a good education, drug and alcohol abuse, adverse childhood experiences and poor mental health [17]—in a complex causal mechanism, and drill rap is just one of these factors.

R3. The causal claim about rap should be seen as pertaining to rap with antagonizing messages towards rival gangs—drill rap. The police believe drill music constitutes a form of “coded communication” between gangs and antagonisms raise the probability of retaliatory offline violence.

R4. Drill rap's role in offline retaliatory gang violence should be considered as a feedback loop: drill rap results from retaliatory gang violence and also facilitates retaliatory gang violence.

R5. The causal claim should articulate drill rap's alleged harms, whilst recognizing a range of potential harms resulting from drill rap. This paper, whilst cognisant of other harms that might arise from drill, is primarily concerned with the police's worry about how antagonisms in drill rap music are raising the probability of offline retaliatory violence in those who are taunted, which they think in turn leads to rising knife crime and youth violence (see Section 1).

R6. A sensible rendition of the claim about drill rap causing violence is cautious and not necessarily in favor of state intervention, since it is sensitive to the many other purported harms when coming up with preventative measures.

3.1 | What Must Be Done for the Causal Claim to Be Upheld?

How might we verify the claims that pornography causes harm, or rap causes violence, in the senses specified above (R1-R6)? I think there are important recommendations in Eaton's framework, which can illuminate how anti-rap critics might verify this sensible rendition of the causal claim. Eaton suggests antipornography feminists would do well to verify the harm hypothesis by employing epidemiological methods [6]. Epidemiologists use clinical trials to establish an association between an action and a disease—that is, smoking and cancer. However, the investigation does not stop there. An association does not tell us (a) whether the association between smoking and cancer is special (more significant than

other causes), or (b) the direction of causality—perhaps lung cancer creates urges to smoke, nor (c) whether lung cancer and smoking are symptoms of a common cause [6], p. 704. To verify an association found in clinical trials, epidemiologists “compare disease rates among large groups with differing levels of exposure to the suspected cause” [6], p. 704. In attempting to prove a connection between smoking and lung cancer, ecological studies are employed to investigate the incidence of exposure to smoking and prevalence of lung cancer. For instance, studies comparing the United States, Japan, Norway, Poland and France found that as smoking increases, lung cancer also increases [6], p. 704. This was a crucial first step in establishing a positive association between smoking and lung cancer.

There have not been ecological studies of this kind investigating the causal claim about drill rap music. For instance, there are no studies, which look to find increased reports in violent crime and increased violent rap consumption, specifically studies which look to find increased reports of retaliatory violence in gang members and increased consumption of rap music including specific taunts and threats. Even if there *were* ecological studies that demonstrated a correlation between incidence of exposure to antagonizing drill rap and prevalence of retaliatory gang violence, there would need to be many, large studies to demonstrate a positive relationship between antagonizing drill rap and offline retaliatory violence. Moreover, even if rap critics produce a coherent body of studies demonstrating a positive association between drill rap and offline retaliatory gang violence, it would only provide data for groups, but importantly does not give us exposure data for individuals within populations.⁷ With regards to smoking, for instance, these ecological studies do not tell us whether the individuals who developed lung cancer are the same individuals that smoked. To establish a positive correlation, we also need exposure and outcome data for individuals in a population.

To remedy this missing data, epidemiologists often turn to studies of individual characteristics—case-control and cohort studies—in search of data about individual populations. Case-control studies are retrospective; they start with the outcome (like a disease) and look back to see if past exposures were different between those who have the disease (cases) and those who do not (controls). For instance, regarding smoking and cancer, a case-control study compares the smoking histories of two groups: one with lung cancer and one without lung cancer. Cohort studies are generally prospective. They start with an exposure and follow a group of people to see if they develop a disease. In the case of smoking and cancer, cohort studies compare smokers with nonsmokers, determining the rate of lung cancer in each group [6], p. 706.

No studies looking to establish the causal claim about rap have employed case-control and cohort studies. For instance, there have not been studies that have compared the rap consumption histories of violent criminals (including violent gang members) with non-violent criminals, or those not known to have committed any crimes. Nor have there been any studies comparing the frequency of violent crimes committed by gang-associated offenders who listen to violent antagonizing (particularly antagonizing rap music aimed at them) rap music with those who do not listen to violent antagonizing rap.

In fact, even if we find through ecological, case-control and cohort studies a meaningful correlation between drill and violence, we *still* need to do more to establish a causal relation. Epidemiologists standardly appeal to six criteria to determine whether a correlation is causal (with respect to a disease).

- C1. Temporality: the suspected causal factor must precede the onset of the disease, and the interval between exposure to the causal factor and disease must be considered
- C2. Strength/Replication: Firmer causal relationships are provided by stronger correlations. Further, findings should replicate
- C3. Quantal-dose relationship: increasing duration, total exposure, intensity and level of suspected causal factor leads to a progressive increase in the risk of disease.⁸
- C4. Plausibility: the correlation should fit within our current state of knowledge
- C5. Consideration of alternative explanations: have investigators considered alternative explanations?
- C6. Cessation data: “if a factor is a cause of a disease, the risk of the disease should decline upon reduction or elimination of exposure to the factor” [6], p. 709–710.

What do these six criteria look like regarding rap? First, consuming rap must precede the violent effects, as the first condition of the sensible rendition of the causal claim makes clear. This would mean that a video taunting another gang would have been released before the alleged offline retaliatory response. Second, studies investigating the correlation between rap music and violent attitudes or behaviors in gang-associated youth must provide strong correlations that replicate. Third, increasing the duration, total exposure, intensity and level of violent taunts, threats and antagonisms increase the likelihood of violent attitudes or behaviors in gang-associated youth. Fourth, should a correlation between violent rap music and violent effects be found, this should be explainable by our best theories and hypotheses. Fifth, investigators must take alternative causes of violent effects into account. For instance, have investigators looked at causes other than drill rap as explanations for violent attitudes or behaviors (such as a location being revealed in music videos [18, 19]) Finally, if rap music is a causal factor of violent effects, the risk of violent effects should decrease with the decrease in the consumption of rap. For instance, if rap music is a causal factor of offline retaliatory violence, the risk of offline retaliatory violence should decrease with the decrease in gang-associated youth’s consumption of rap. Eaton admits that anti-pornography feminists are far from providing evidence of the harm hypothesis, which meets these criteria [6]. Similarly, as I show in the next subsection, the evidence claiming to support the causal claim about rap has yet to meet these criteria, and therefore has yet to verify the sensible rendition of the causal claim (R1-R6).⁹

3.2 | Existing Empirical Literature

What is the state of the empirical literature linking drill rap to violence? To date, the only empirical study investigating the link

between drill rap and violence was conducted by Klienbergh and McFarlane [15]. They conducted a sentiment trajectory analysis of 549 drill songs—between December 2013 and November 2018—and violent crime data (such as homicide and violence with injury) from the Met within the same time frame. Their study came up empty in locating a causal link between drill music and violent crime.¹⁰ Given this, what other empirical literature could the police point to in order to show that drill rap music causes violent attitudes and behaviors? I think there are at least two sets of literature the police might employ for evidence of the causal claim. One set of studies could look at how rap produces copy-cat violence by looking at the general effects of rap music on general audiences [20]. The second set of studies could look at how adherence to the “code of the street” causes retaliatory violence in gang-associated youth [18–22]. I look at each in turn.

3.2.1 | General Effects of Rap Music on Violent Attitudes and Behaviors

The most rigorous analysis of the literature on rap music's alleged violent effects was conducted by Charis E. Kubrin and Ronald Weitzer in 2010 [20]. They found that many of the studies of rap's effects are nonempirical [20]. Many authors theorize about rap music's effects on one's identity, political inclinations, attitudes and behaviors based on mere anecdotal evidence and selective quotations [20]. Few authors, however, base their claims about rap on data or observations. Kubrin and Weitzer further identify some writings on rap as “quasiempirical”. These cases differ from mere anecdote insofar as there is (some) data to support the points of the authors, but the methodological shortcomings of studies limit the ability to draw conclusions from them [23]. Kubrin and Weitzer identify additional weaknesses with the quasiempirical literature, such as small sample sizes in surveys and studies; selective use of examples—in the form of rap lyrics, videos, writings—to “prove” a point; and even when studies have a good sample size and regression analyses are involved, they still suffer from methodological weaknesses and lack external validity [24, 25].¹¹ Kubrin and Weitzer identify some studies that use empirical data to back up their conclusions [20]. However, these studies are rare and often conclude that the relationship between rap and violent behaviors and attitudes is complex, rather than concluding that listening to rap promotes violence [20, 26]. Mahiri and Conner's study, for instance, revealed that youths who listen to rap had more progressive values than the values represented in the rap discourse to which they are exposed [26]. They also found that some youths had cogent critiques of negative characterizations of men and women in rap, particularly representations of men as “gangstas” and women as “hoes.” Their study challenges “the simple connections that the dominant public discourse and media so often draw between rap music and pervasive negative influences on black youth” [26], p. 135, as cited in Reference [20].

There have been experimental and ethnographic studies into the effects of rap music, but these also have significant limitations. Christy Barongan and Gordon Hall ([25], as cited in Reference [20]), for instance, attempted to measure the behavioral effects of subjects exposed to misogynistic rap. Participants were split into two groups; one group listened to misogynistic

rap music and the other listened to neutral rap. Experimenters then showed all subjects neutral, sexually violent (depicting the rape of a woman) and assaultive (depicting a man physically assaulting a woman) film vignettes. Participants were then asked to choose one vignette that they would like to show one of two female confederates. They found that subjects exposed to misogynistic rap songs were more likely than those exposed to neutral rap to show a female confederate a sexually violent or assaultive film vignette (doing so was interpreted by the researchers as an aggressive act). However, this study suffers from methodological limitations. For instance, findings of attitudinal and behavioral change in lab settings might suffer from demand effects—“Subjects believing that the experimenter condones what is presented”—or findings might be the result of immediate stimuli in the context of the lab, and these attitudinal, behavioral effects are thus short-lived [20], p. 134. In both cases, the findings lack external validity and may be “artefacts of the experiment” [20], p. 134. Experimental studies like Barongan and Hall's must be satisfied with acts like showing a female confederate a film vignette depicting an assault or rape of a woman for ethical reasons, but it is dubious whether such artificial conditions reveal anything about the kind of real-world conduct rap music is criticized for. Many of the studies seem to be worried about how rap might lead to violent attitudes, but this tells us nothing about how consuming rap might translate into the types of harms the police are worried about. If drill rap music has a tendency to make violence attractive to its consumers, this will not necessarily translate into the violence—youth crime, gang murders and knife crime—that the police are concerned with. This is because the alleged cultivation of violent attitudes might be counterbalanced by the consumers' other attitudes. For instance, if taunting a rival by making fun of their dead friends on a song causes a rival to have violent attitudes, whether this manifests in the rival stabbing the taunter will involve the balance of other commitments, such as a commitment to his family, career, and social and emotional wellbeing. These commitments will be weighed against the alleged violent attitudes you get from hearing a song with pursuing acts that might lead to your death or land you in prison. Therefore, these other considerations might inhibit the behaviors the police seem concerned with.

In the real world, there are often multiple influences on a person's behavior and attitudes, but in a laboratory setting, exposure to violent messages is decontextualized and unreflective of the numerous influences affecting behavior and/or attitudes. We might think that a “natural experiment” would do better by considering a subject's real-world experience. However, “even natural experiments in the real world—either comparing different contexts cross-sectionally or longitudinal changes in, say, violent programming—have not demonstrated a link between media violence and aggressive behavior of children or adults” [27] as cited in Reference [20], p. 135.

Aside from experimental studies, some ethnographic studies investigate the real-world effects of rap. These studies include those exploring how youths actively decode rap's messages. They look to explore the link between rap and listeners' perceptions. These types of studies are very scarce in the rap literature [20]. One example of this kind of ethnographic study is found in Mahiri and Conner's 2003 study (see above), which found that youths interpreted rap in various critical ways, such as resisting

and critiquing stereotypical images in rap. They also found youths understood and critiqued images of crime and violence in rap in ways that increased their desire to circumvent and improve violent situations in their lives and communities.

Recent studies have found that exposure to violently-themed music does not cause real-world violence through increasing violent behaviors and attitudes [28]. Others have found that violent themes in music can give fans a range of positive psychological functions [29]. These studies coupled with the inadequacy of the studies looking at general effects of rap on general audiences show that the causal claim is far from being established. Moreover, even if these studies conclusively showed that rap causes violent effects, we might question whether they are helpful for establishing the sensible rendition of the causal claim. As the sensible rendition of the causal claim highlights (R5), the types of violence the police are concerned with are offline retaliatory violence, where drill music constitutes a form of “coded communication” between gangs and antagonisms raise the probability of retaliatory offline violence. They are not concerned with how listening to rap music with violent themes and messages might precipitate copy-cat crime. That is, these studies are too imprecise, focusing on general audiences, instead of the specific producers and listeners that are targeted for interventions by the police. The police are concerned with whether antagonisms in rap music cause offline retaliatory gang violence. We cannot know whether antagonisms in drill raise the probability of retaliatory violence in gang-associated youth based on studies looking at the general effects of rap. Given this, proponents of the causal claim might appeal to a different set of evidence. I turn to this next.

3.2.2 | Adherence to the “Code of the Street” and Retaliatory Violence

A proponent of the causal claim might proceed in a different way. To substantiate the sensible causal link between drill and violence, they could instead argue that drill rap music raises the probability of retaliatory gang violence by appealing to the literature supporting the view that gang-associated youth use social media to taunt rivals and trade insults in ways that cause offline gang retaliation [18, 19]. By using the “code of the street” as theory, researchers claim that insults occurring in online spaces raise the probability of offline retaliatory violence. The basic premise of “the code,” as developed by Elijah Anderson, is that gang-associated youth adhere to a street code, which revolves around the display of a predisposition to violence and a capability for retaliatory violence [21]. For instance, if someone wrongs you, you must bring them down for what they did to you, and also challenge others so that your reputation is bolstered. According to Anderson, this code emerged out of extreme structural disadvantage: poverty, unemployment, racism, and mistreatment by and lack of faith in the police and judicial system. This structural disadvantage led urban males to rely on violence to achieve status in their community. Those who adhere to the code are socialized to have a disposition for aggression and violence to defend one’s reputation.

As a result of growing fears over digitally mediated gang violence, researchers have written exploratory articles about how

gang-associated youth transfer the code of the street directly to social media. Stuart argues that the existing scholarship is based in a theory of *parallelism*, where online activity mirrors offline activity since the “same gang violence mechanisms on the urban street unfold online” [30], p. 1012 as cited in Reference [19], p. 193. Crucially, by framing social media as a reflection and analogue for violent conflicts, the parallelism thesis predicts that social media necessarily increases both the intensity and frequency of offline retaliatory violence. For instance, scholars maintain that drill rap music videos posted on social media sites such as YouTube can raise the probability of offline gang violence, consistent with the code of the street. Keir Irwin-Rodgers, James Densley and Craig Pinkney, for instance, assert that “Intentionally or not, such videos ‘flame wars’ by challenging the status and reputation of particular gang members, and consistent with the ‘code of the street’ subsequently provoking some form of retaliation” [31] p. 5. Scott et al. write, “using social media through rap music and videos is tied to increased physical conflicts and violence, as they can intimidate and threaten gangs to retaliate against each other” [32], p. 9.

However, these articles based on the parallelism thesis profoundly overestimate the causal relationship between social media content and offline retaliatory gang violence. As Stuart writes, “Even a cursory review of recent crime statistics reveals that violent crime in urban areas has decreased to historic lows in the same years that gang-associated youth have increased their use of social media” [19] p. 193. Assuming a sort of parallelism, or direct positive or causal relationship between social media use and physical retaliation, overlooks important opportunities for youth to de-escalate or avoid offline gang violence [18, 19]. Further, arguing that online interactions mirror offline interactions neglects how offline conflicts are transformed by social media. While gangs indeed challenge rivals online, these challenges are characterized “by distinct rules, conventions, and consequences that afford new and different kinds of outcomes while constructing previous ones” [19] p. 193. Crucially, insults uttered online in drill music videos carry different meanings and demand different responses than if they were merely uttered in the street (more on this below).

Moreover, many of the studies upon which researchers claim that social media increases the probability of gang violence are methodologically limited. As Stuart claims, “Previous studies rely primarily on self-report surveys and/or content analyses of disembodied social media posts several years after their creation, often asking people other than gang-associated youth to interpret the meanings and motivations of a given piece of social media content” [19] p. 193. For instance, in a 2017 report, Irwin-Rodgers and Pinkney warn that music videos could raise the probability of offline violence based on discussion from participants, including a focus group of young people and a group discussion involving a police officer [33]. When an interviewer asked, “So what things happened after this music video was uploaded?” participants of the focus group of young people responded that “shootings and stabbings” occur after rap videos taunting and threatening rivals are uploaded [33], p. 20. Another responded that “it caused more stabbings for a time” [33], p. 20. This report relied primarily on interpretations of non-gang-associated youth—including the police, who have repeatedly shown they lack familiarity with the conventions of drill and are

therefore not well epistemically situated to decode what results after drill music is uploaded to social media [2, 14, 18, 19, 34]. For instance, according to the police there were a number of different stabbings and shootings that occurred as a result of drill videos being put up over the course of a week, but they offer no real support to show that the drill videos caused the shootings and stabbings [33], p. 20.¹²

In another study, Patton et al. [30] examine communication of a gang member on the social media platform X (formerly known as Twitter), using inductive textual analysis. They conclude that the same gang violence mechanisms occurring on the urban street unfold online [30].¹³ Their findings suggest that because gang-involved persons adhere to the code of the street, opportunities for violence are created via online behavior [30]. However, this study suffers from the limitations outlined by Stuart. For instance, it draws primarily on non-gang involved parties to interpret the meanings of online content. At most these studies (and report) show that, in some cases, a piece of online content may have preceded an increase of offline violence (according to some witnesses, even if not perhaps fully reliable). To say something meaningful about drill, we need more than bare assertions (or at best anecdotal evidence) of increase of violence, and a more in-depth analysis. A growing number of studies have provided this in-depth analysis and evidence contradicting the parallelism thesis [18, 19, 35–36]. These studies demonstrate that gang-associated youth's use of social media does not independently raise the probability of offline gang violence and even provides them with new ways of preventing, avoiding and lowering the risk of violence. Acknowledging deficiencies in studies supporting the parallelism thesis, Stuart conducted ethnographic research, including direct and real-time observations alongside producers of gang-related online content and participants in gang conflicts. He asserts that “lacking access to such on-the-ground, real-time data, existing studies are unable to determine the relationship between a given piece of social media content and the physical violence that may or may not result” [19], p. 193.

Stuart's work as director of an after-school youth violence prevention programme allowed him to gain access to gang-associated youth and their families. He recruited 60 young men, 29 of them being members of the gang faction CBE—Corner Boys Entertainment—and 31 others coming from allied gang factions.¹⁴ Stuart enlisted participants to review each day's social media activity with him. This included reviewing their posts and uploads and debrief sessions, which focused attention on providing key data, prompts and discussions for materials, as well as asking “participants about the origins, aims meanings, and consequences of particular online antagonisms” [19], p. 196. This meant Stuart could follow feuds as they developed in real time from online to offline spaces. He found that gang-associated individuals engage in “online antagonisms as part of an innovative and intentional effort to build violent reputations without the necessity of behaving violently offline” [18, 19], as cited in Reference [22], p. 310–311. Stuart calls upon researchers to do better by way of providing evidence of the role that online antagonisms play on offline retaliatory behavior. He writes that it is “incumbent upon researchers to increase the rigor of empirical evidence and claims,” since “there exist no systematic data or definitive analyses to determine the precise causal relationship between social media and offline violence” [22], p. 311.

Crucially, the most in-depth analysis of taunts in drill music to date—Stuart's ethnographic work—provides evidence against the parallelism thesis—used to justify the causal claim—specifically because he uses empirical methods to show how online threats are in fact different from offline threats.

As we have seen, the existing evidence drill-critics can appeal to in order to establish the causal claim is riddled with problems. The academic literature that has been used to support the claim that rap causes offline retaliatory violence looks nothing like the studies needed to show a causal relationship between drill rap music and the violence the Met are concerned with. If through ecologic, case-control studies and cohort studies, a meaningful relationship between rap music and violent retaliatory attitudes or behaviors can be found, proponents of the causal claim would do well to adopt the six criteria (C1–C6) to show that rap music causes these violent retaliatory effects. The evidence reviewed in this section, however, is insufficient to provide support for the claim that rap music causes violence (R1–R6).

4 | The Justificatory Burden

I have attempted to provide a charitable reconstruction of the causal claim critics of rap music appeal to, and have argued that existing evidence is insufficient to establish any such causal connection. Thus, the *evidential burden* of showing that rap causes violence is far from being met. I now turn to consider the *justificatory burden* that police must meet to justify their approach of censoring drill music and restricting drill artists' movement and freedoms. I argue that even if we assume drill is a cause of violence, the police intervention of censoring drill is not justified since other interventions are less speech-restrictive and better-suited to prevent violence; in other words, censorship is not necessary for the purposes of preventing serious youth violence.

Freedom of expression is a fundamental human right. It is enshrined in Article 19 of the Universal Declaration of Human Rights and Article 10 of the Human Rights Act 1998. It is of inherent value to the individual as it involves the external communication of an individual's “forum internum” or inner realm of thoughts, beliefs and convictions, and as such, is connected to foundational values associated with the *forum internum*, such as personal autonomy and human dignity [37], p. 93. Freedom of expression also has consequentialist value. It is valuable to liberal democracy, political participation and artistic and media freedom. Scholars also highlight the epistemic value of free speech [37], p. 93. For instance, John Stuart Mill argues that human fallibility justifies greater tolerance for freedom of expression and that conflicting doctrines often share truth between them [38]. Whilst freedom of expression is vital to our ability to convey art, opinions, convictions and beliefs, and to meaningfully participate in liberal democracy, there are limits to freedom of expression on certain grounds, such as to ensure national security, public order, public health and public morals [39].

The epistemic, consequentialist and inherent value of freedom of expression suggests that there is a substantial burden of justification when limiting speech. Given this, human rights law provides guidelines on when it is permissible to restrict free speech. For instance, the Universal Declaration of Human

Rights provides a three-part test for legitimate restrictions to freedom of expression. Any restriction on freedom of expression must be (1) provided for by law; (2) pursue a legitimate aim; (3) be necessary for a legitimate aim [40, 41]. The first condition requires that the law restricting speech has a written basis; complies with international standards on human rights, and is sufficiently precise to allow people to adapt their behaviors to the law [40, 41]; it must be “clear, publicly accessible and easily understood by everyone” [39]. The second condition requires that restrictions have valid grounds for restricting speech, which tends to include: respect of the rights and reputations of others; the protection of public health; the protection of public order; and the protection of morals [40, 42]. The third condition requires that restrictions be effective, proportionate and the least restrictive possible to achieve the legitimate pursuit [40, 43]. The UK primarily protects the right to free speech through the Human Rights Act 1998, which includes similar restrictions to free speech as the Universal Declaration of Human Rights does [44]. For instance, restrictions on free speech in the UK must be set out in laws, justified with a legitimate aim, and be necessary and proportionate.

Having shown why free speech requires such a high justificatory burden, and when restrictions on free speech are legitimate, we are better placed to see why it does not follow that restricting drill is justified, even if we suppose it causes violence in the sense specified in Section 1. For the sake of argument, suppose the first two conditions were met: it was clear, publicly accessible and easily understood by everyone that drill videos causing violence are to be censored, and censoring these videos pursues the legitimate aim of preventing crime. We might still be worried that censoring drill is not necessary for tackling retaliatory gang violence (as per condition three above). When asking whether a restriction on a human right—such as censorship of speech—is a necessary intervention, we must assess whether there are alternative measures which would tackle the legitimate aim—such as addressing retaliatory gang violence—in ways less intrusive to the human right in question and equally effective [43]. To illustrate this, consider Gerards’ hypothetical case of a prohibition on wearing headscarves during gym classes to protect the safety of pupils [45]. This is because some headscarves do not allow pupils to move freely, and there is risk that children may trip or fall over their headscarves. An alternative measure might be the prescription of elastic headscarves that do not hamper their movements and are not a danger to their safety. This would be an equally effective and preferable measure over a complete prohibition of headscarves from the perspective of the protection of freedom of religion [45].

In the case of drill, there are less restrictive and more effective alternatives to tackling retaliatory gang violence. Public health approaches are shown, by evidence, to be more effective than censorship-oriented approaches [17]. For instance, a public health approach is more effective than censorship for promoting well-being and fostering positive societal and structural change. It therefore offers a long term, preventative solution, over censorship, which is often reactive and individualizes the problem in a way that marginalizes structural causes [17], p. 325. Many have advocated for a public health approach to serious violence, including gang violence and knife crime [17], p. 325. A public health approach “suggests policy makers should

search for a ‘cure’ by using scientific evidence to identify what causes violence” and develop/testing interventions [46]. This approach is beneficial insofar as it shifts the emphasis away from police enforcement-led interventions, which often worsen the problem [17]. For instance, in the specific case of censoring and restricting drillers, artists are often barred from the various financial, emotional and social benefits their music brings them [18, 19]. It is thus counterproductive from a crime-reduction perspective. A public health approach promises a science-based (epidemiological) understanding of violence, whilst establishing more evidence-based, progressive and sustainable solutions.¹⁵ It is thus wholly consistent with the sensible rendition of the causal model I have outlined in Section 2.

The UK has already claimed great success for such an approach to reducing gang violence in Glasgow/Strathclyde, which had one of the highest rates of knife crime in the world before a public health approach was taken. For instance, Operation Reclaim—which focused on reopening sports and recreational areas to provide safe and supervised educational, recreational and diversionary activities—combined with Operation Phoenix, which brought together teenagers to participate in sports and recreation activities and foster relationships with the police, has been highly successful interventions. Over the course of these operations, gang violence was recorded as having diminished between 40% and 70% [17], p. 328.

Drillers’ testimonies align with a public health approach, showing that they are well aware of the underlying structural causes of serious violence. As London-based drill artist Abra Cadabra argues.

Targeting musicians is a distraction [...] The cuts that affect schools, youth clubs, social housing, benefits, are making life harder for the average person living on or below the poverty line in this city. There are people doing mad tings (committing acts of violence) not because they want to, but because the situation has forced them to Reference [48]. Whilst there has been great success of taking a public health-based approach to gang violence and knife crime in Glasgow, in places like London, the public health approach is seriously lacking. Williams and Squires outline that this is due to disregarding research and evidence-based principles that are supposed to underpin a public health approach. Documents surrounding the approach to serious youth violence allude “to a public health based strategy, but it is not yet a public health based strategy,” since “the Government’s rhetoric on public health has not yet been reflected in its actions” (Home Affairs Select Committee Deliberations, as cited in Reference [17], p. 333). Instead, the World Health Organization noted that strategies to reduce violence in places like London are not based in research but primarily consist of police-led (and non-evidence-based) interventionist approaches [17], p. 334. This is reflected in censorious interventions targeting drill. Crucially, evidence indicates the public-health approach is more effective than censorship since it focuses on prevention, empowers communities by emphasizing community-based solutions, tackles root causes, and offers a longer-term solution to tackling violence. Censorship, at best, offers a short-term solution, and risks being counterproductive and shifting attention away from structural causes of violence. Moreover, and importantly for justifying restrictions on free speech, the public health approach

is a less speech-restrictive means than censorship of addressing serious youth violence. It is therefore more favorable from the perspective of protecting freedom of speech.

It should also be clear why we cannot both censor drillers and take a public health approach. Given the importance of free expression, censorship should only ever be a last resort in crime control. Insofar as there exists a less restrictive approach that is more effective than censorship, critics would have to first show why a public health approach fails before proposing censorship, even in conjunction with a public health approach. Further, censoring and restricting drillers' freedoms requires considerable resources in the form of policing, surveillance and court proceedings. This takes resources away from community-building approaches, which we know to be effective, such as opportunities for recreation and youth clubs.

In sum, censoring drillers requires a high justificatory burden, which I have argued has not been met. Supposing that drill causes violence, this is insufficient to show that censorship is justified. For censorship to be justified requires more than showing it could alleviate a problem: it typically requires showing that the problem cannot otherwise be addressed in a more effective and/or less-restrictive way. The public health approach is a more effective and less speech-restrictive means of addressing the alleged problem: retaliatory gang violence. So censorship advocates, such as the police, must show first that such an approach is unpromising, and if they cannot do so, then policies of censorship cannot be justified. It seems highly prejudicial that law enforcement would zoom in on music censorship as an intervention, as opposed to, say, tackling poverty, increasing opportunities for those in high-risk areas, and dedicating money to activities for the youth, such as youth clubs. Moreover, this justificatory burden becomes even more untenable when we think about the racist social meaning of such interventions, as I will discuss in my concluding section.

5 | The Racist Social/Public Meaning of the Causal Claim

The claim that rap music causes violence, as we saw, lacks robust empirical support, yet it is a prevalent critique of rap music. Since the causal claim is empirically unsubstantiated, and robust empirical evidence supporting it is required, we might be left questioning why it is made so often. It could be argued that these claims are caused by racist attitudes themselves. There are some empirical studies that might suggest that anti-rap attitudes are indeed caused by anti-black attitudes [49–54]. These scholars suggest that attitudes about rap music, specifically about the criminality of rap, provide an important example of anti-black attitudes that can be expressed without mentioning race. I am not advancing here a claim about what *causes* anti-rap critiques, however. Any such claim would have to meet the rigorous empirical standards I set out in Section 1. However, leaving open what might cause such critiques for further investigation, I instead argue that such critiques have a racist public meaning.

Archer and Matheson define *public meaning* as “the meaning that others can justifiably attribute to our acts given the context in which we perform them” [55], p. 6. A key issue with

prejudicial public meanings is that they send prejudicial messages or reinforce prejudicial stereotypes. For instance, Archer and Matheson give the following example. Say a man was to interrupt a woman in an important business meeting. In a business culture that is patriarchal, it seems reasonable to understand this interruption as expressing sexism, denigration of women, regardless of whether we think (or can verify that) the interrupter has sexist attitudes and intentions. The problem here is that his actions reinforce sexism.

There is a rich literature in philosophy highlighting the importance of social/public meaning in a range of contexts (I use the two terms interchangeably here). For example, Charles Mills, in assessing arguments in favor of black men's duty to marry black women, concludes that even if a black man's marriage to a white woman is (a) not motivated by racism, and (b) they are aware that it is not motivated by racism, there is always the fact that regardless of his motivations, black people will perceive him as marrying out of racial self-hatred and reinforcing white supremacy [56]. In other words, he writes that such interracial unions with white women might send the message that “black women just ain't good enough” [56], p. 149. Here, it seems Mills acknowledges that despite such a union not being motivated or caused by racist beliefs, attitudes or intentions about white women's beauty, there still might be a racist public meaning that is communicated by such a union. Similarly, Robin Zheng argues that when we focus on the effects of racial fetish in a racially stratified world, rather than whether racial fetish is caused by negative racial stereotypes, racial fetishes can be objectionable [57]. Zheng asserts that publicly expressing racial fetishes, such as “yellow fever”—a preference for Asian women—can have a racist public meaning even if not caused by racist attitudes. She writes that “racial fetishes in a racially stratified society express racially stereotypical social meanings even if not caused by them and even if individuals with the preference do not endorse them” [57], p. 411.

I argue that, similarly, public proclamations of the causal claim can have a racist public meaning, regardless of intentions and attitudes. One racist public meaning here is that black music is dangerous. Black music subgenres have historically “haunted the penal imagination as audible signs of danger, criminality and disorder” [58], p. 30. The policing practices that are used against drill music is an enduring legacy of the pre-professional police-like forces that suppressed cultural expression during colonial times. In the US, Caribbean, Africa and Brazil, slave patrol legislation was introduced to prevent drumming, dancing and musical performances during times of slavery. This was to clamp down on activities that would encourage resistance to the violent domination, subjugation and exploitation of the enslaved [59], p. 52. This banning of black cultural expression did not end with slavery, but continued with police harassment and suppression of Afro-diasporic black music genres, such as blues, bepop, jazz and hip hop, reggae sound systems, UK garage, grime and bashment, calypso, West African Àsikò music, Brazilian Samba, Afro-Cuban music and South African Marabi music [58], p. 38. As Fatsis asserts, “to put it crudely, the slave patrols of yesteryear are the precursors of today's beat patrolling; in a historical continuum that resurrects colonial subjugation in the form of the grossly disproportionate policing against black people today” [58], p. 39.

Fatsis argues that the police's cultural and political treatment of black music entrenches the cultural position of it as noise to be eliminated rather than art to be appreciated [58]. Instead of being recognized as a valid form of creative expression, black music genres become types of incivility. Black music is consigned to the margins of refinement, propriety and respectability, and excluded from the pantheon of art. The continued critiques and policing of drill as dangerous, therefore, serve to reinforce stereotypes about black art's lack of sophistication. Black music genres/subgenres like drill predictably become maligned as sources of political danger and are policed as political disturbances that are "threatening to the sensory faculties and political institutions of 'white civil society'" [58], p. 39.

Given this context, the overpolicing of drill is not to be understood as a neutral function of the criminal justice system, but "as socio-cultural and political processes that create and convey meaning, influence cultural definitions and shape social and political reality as conductors of cultural norms and instruments of social control" [58], p. 40. One public meaning that is conveyed is that black music is dangerous noise that is not to be appreciated aesthetically. The public context that gives rise to such public meanings is the prior hyperscrutiny of black art forms. Given the hyperscrutiny of black art forms, even before rap, as contributing to perverted behavior and attitudes [60], and other racism informing the background context in which these critiques are being advanced, such condemnation of rap contributes to a culture of invalidation against black cultural expression by reinforcing the message of its dangerousness.

Another racist public meaning here is that young black men are dangerous. This serves to reinforce stereotypes about young black male criminality. The public context that gives rise to such public meanings is the prior existence of stereotypes about black criminality. When rap music is singled out and positioned as a cause of urban violence, this positions young black men as a threat to society and the "civic mainstream" [9], p. 1202. The racist public meaning of the causal claim reinforces the assumption that young black men are more likely to be criminals, and thus provides a rationale for the state to sustain radicalized forms of policing and legitimize draconian police powers [1].¹⁶ Framing rap music as a cause of violence in places like London validates the overpolicing of already marginalized communities. For instance, as a way of tackling violent youth crime, the police use techniques such as "stop and search," which have been shown to be ineffective in crime reduction, but they also use these techniques so that young black males are subjected to higher rates than others [1]. This disproportionality in the overpolicing of young black males extends to the removal of their art from streaming services and gang injunctions that restrict them in insidious ways. As we have seen, there is a justificatory burden to show that censorship (rather than other interventions) is the most effective, and this burden has not yet been met. In that context, there is also a social meaning attached to censorship approaches: that it is legitimate to surveil black youth, even when there is no evidence to suggest such surveillance is effective or fair.

One might object that even if the causal claim does have such meaning, this is a problem with our culture as, ultimately, we need to be free to investigate empirical relationships between

phenomena. They might add that if one narrowly focuses on drill rap, without focusing on the prior hyperscrutiny of black art and criminalization of black men, then it is not clear what licenses the public to infer that black music is and black men are dangerous. However, this would be akin to saying if one focuses solely on the man interrupting the woman in the important business meeting, without considering the patriarchal business culture (and patriarchy more broadly), then it is unclear what licenses an observer to infer that this interruption reinforces sexism. This response ignores what public meaning is: It is not the meaning that others can justifiably attribute to an isolated act, but it is the meaning that others can justifiably attribute to our acts given the context in which they are performed; in other words, context licenses the inferences people make about our acts and expressions. As the objection notes, the problem is with our racist culture and world that has and continues to prejudicially criminalize black cultural expression. Given our culture has denied black cultural expression the status of art, and both the makers of this art and the art itself are deemed dangerous, public proclamations of the causal claim have a racist public meaning. If the causal claim was not made in such a racist cultural and public context, but expressed in a non-racist world, then the causal claim would not carry such a racist public meaning. However, in our actual world, with its racist history, uncritically advancing the causal claim conveys a racist public meaning.

Further, it is not my view that a prejudicial public meaning should always stop us from taking action, such as investigating empirical relationships between phenomena, since it seems that whether we should care about public meaning will have to be weighed up against other considerations. For instance, in Mills' example, the man might acknowledge the unintended public meaning of his marriage, but it is hard to imagine that this should stop him from marrying someone he genuinely loves. However, in the case of the racist public meaning of advancing the causal claim, the competing considerations we must weigh up public meaning against concern whether there is evidence of a causal connection between drill and offline retaliatory violence, or whether censorship-oriented approaches to drill are justified even if we grant a causal relation. I have argued these considerations lack force. Crucially, given the arguments of Sections 2 and 3, and given the racist public meanings of uncritically advanced critiques, we should be extremely cautious about engaging in such claims and basing restrictions on them.

In conclusion, I have argued that there are various reasons against making the claim that rap music causes violence, and implementing interventionist policies from it. First, given the complicated empirical evidence and the standards required to establish a causal connection between rap music and offline retaliatory violence, there is insufficient evidence to support the causal claim. Second, the censorship-oriented approaches to tackling youth crime that law enforcement have taken are unjustified by the causal claim (even on its most charitable interpretation), and potentially damaging, because they fare poorly in comparison to public health approaches and come with a variety of financial and personal harm to those censored. Third, both the causal claim and the subsequent interventionist/censorship approaches are particularly problematic given their damaging social meanings. Given what is at stake, we need to do better,

both in establishing the causes of violence and in formulating responses to address it.

Acknowledgments

Thanks to Jules Holroyd, Lambros Fatsis, Vid Simoniti, Federico Picinali, Alex Fisher, Tommy Curry and Takeshi Morisato for their thoughtful feedback on early versions of this article. Thanks also to the anonymous reviewers for their insightful and encouraging comments.

Funding

The author has nothing to report.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Endnotes

¹ Drill music is renowned for lyrics ‘taunting rivals in vividly shocking terms, describing the harm that awaits them, and keeping a tally of “scoreboards” (members of opposition gangs who have been killed or seriously injured) as a sign of collective’s transgressive capital’ [1]. Here are some popular UK drill songs that have been removed from YouTube and other social media at the Met’s request:

Chinx (OS): Secrets not Safe (Music Video) REUPLOAD; Zone 2: No Censor (Reupload); #410 AM-Attempted 1.0 (Music Video) | Re-upload.

² In my discussion, I am interested in drill rap causing offline violence, as this is what the police seem to be concerned with—online taunts causing offline gang violence.

³ The incitement claim is an important one that is worthy of a separate investigation from the causal claim, especially given what we use to prove a causal relation (empirical evidence) is different from what we use to prove an inciting relation (attention to illocutionary acts). In this paper I focus on the causal claim. For an investigation of the incitement claim see [14].

⁴ The nature of causation constitutes one of the seminal philosophical debates. For a broad overview of the philosophy of causation see Beebe, Hitchcock and Menzies [16]. Probabilistic accounts of causation are not without critique [16], but Eaton’s framework is particularly helpful in this context because it provides a way of theorizing how violent media might cause harm and is consistent with police testimony, which expresses concern with how drill raises the risk of gang violence.

⁵ This is meant in accordance with the “ordinary use test of the word cause,” since we are concerned with legal questions (i.e., the censorship of drill music), as opposed to sophisticated accounts of causation, found, for instance, in the metaphysics literature.

⁶ Whilst I focus on the claim that drill causes retaliatory gang violence, I acknowledge that this leaves open that, even if there is no such evidence, nevertheless drill might have a causal role in a wider range of harms. I suspect that there is no evidence that it does, especially when we consider the effects of rap music on general audiences below, though it is at least possible that there is.

⁷ You might wonder why we need data about individuals given that this paper is concerned with gang and retaliatory violence. Data about individuals might provide further support for a causal link, particularly if

we compare the violent histories of or consumption of antagonising rap in violent gang members against others (see below).

⁸ It is important to note that the absence of a quantal-dose relationship does not rule out a causal relationship, since a threshold may exist where no disease/violence develops up to a certain level of exposure [6], p. 709.

⁹ An anonymous reviewer has noted we need to acknowledge that although these criteria are useful for capturing how to establish a causal connection, the disease analogy breaks down at some point [6], p. 713. For instance, whilst stopping smoking should reduce your risk of developing lung cancer according to C6, it does not apply so straightforwardly to the present case of the causal claim about drill. For instance, with regard to C6, it seems plausible that having been involved in a cultural practice that has strongly shaped your value system—like a cult—you might have a hard way ridding yourself of ways of thinking and behaviours.

¹⁰ A sentiment trajectory analysis allows researchers to examine changes in sentiment—such as the evolving emotional tone of a narrative or conversation—over time.

¹¹ External validity examines whether findings of a study can be generalized to other contexts. In other words, it asks the question: can the findings of a study be applied to life outside of the study?

¹² One might question what the actual evidence for what results from a song’s being uploaded is supposed to be. Here it is hard to see how the police’s impressions are relevant, and it seems like only the city statistician (or some sort of well-researched report over time) could know the kinds of things individuals are claiming to know. This takes us back to the epidemiology point insofar as we would expect to see a long-term pattern of spikes in violence correlating with spikes in drill songs released. Thanks to an anonymous reviewer for making this point.

¹³ The key mechanism here is the “code of the street.”

¹⁴ This served as a way to confirm that the practices observed by CBE were generalizable to other gang factions.

¹⁵ I do not want to paint a perfect picture of public health approaches. Whilst public health approaches are preferable to the hyperpunitive police-led approaches, there are critiques of how public health can easily slide into public order and how it can be too coercive albeit in the name of healthcare [47].

¹⁶ Opponents might point to the overrepresentation of young black men in crime statistics as evidence that they are more likely to be criminals. However, this is too simplistic for at least two reasons. First, it ignores that differential police presence, patrolling and profiling, combined with discrimination in the courts and correctional systems, leads to more young black men being arrested, convicted and incarcerated [61]. It is reasonable to conclude that the overpolicing of young black males and recorded offending exaggerates the extent to which black males are understood to be involved in crime [62]. Second, limitations in methodology and data employed make it difficult to examine actual levels of black men’s involvement in crime as opposed to merely recorded offending [62].

References

1. L. Fatsis, “Policing the Beats: The Criminalization of UK Drill and Grime Music,” *Sociological Review* 67, no. 6 (2019): 1300–1316.
2. T. Jalloh, “Does the Critical Scrutiny of Drill Constitute an Epistemic Injustice,” *British Journal of Aesthetics* 62, no. 4 (2022): 633–651.
3. E. Nowak, “Poetic Injustice,” *Episteme* 21, no. 3 (2023): 1–15.
4. P. D. Magnus and E. Malone, “Popular Music and Art-Interpretive Injustice,” *Inquiry* (2023): 1–16.
5. B. Hall, R. Khan, and M. Eslea, “Criminalising Black Trauma: Grime and Drill Lyrics as a Form of Ethnographic Data to Understand “Gangs” and Serious Youth Violence,” *Genealogy* 7, no. 1 (2023): 2.

6. A. Eaton, "A Sensible Antiporn Feminism," *Ethics* 117, no. 4 (2007): 674–715.
7. BBC, Crime Figures: Violent Crime Recorded by Police Rises by 19% (2019), <https://www.bbc.co.uk/news/uk-46984559>.
8. Metropolitan Police, Requests to Social Media Companies to Review or Remove Musical Content (2023), <https://www.met.police.uk/foi-ai/metropolitan-police/disclosure-2023/january-2023/requests-social-media-companies-review-remove-musical-content/>.
9. A. Lynes, C. Kelly, and E. Kelly, "THUG LIFE: Drill Music as a Periscope Into Urban Violence in the Consumer Age," *British Journal of Criminology* 60, no. 5 (2020): 1201–1219.
10. W. Pritchard, Police Are Taking Down More UK Drill and Rap Videos Than Ever—For Artists, What is the Cost? Accessed July 12, 2023 (2023), <https://djmag.com/features/police-are-taking-down-more-uk-drill-and-rap-videos-ever-artists-what-cost>.
11. Oversight Board, UK Drill Music the Oversight Board has Overturned Meta's Decision to Remove a UK Drill Music Video Clip from Instagram. Accessed July 1, 2023 (2023), <https://www.oversightboard.com/decision/IG-PT5WRTLW>.
12. A. Trendell, YouTube Deletes Half of 'Violent' Music Videos, With Police Singling Out 'Drill Music' (2018), <https://www.nme.com/news/music/youtube-deletes-half-violent-music-videos-drill-music-2327260>.
13. K. Jonathan, Why the Met Are Drilling Down on Drill Music (2023), <https://www.standard.co.uk/culture/music/met-police-drill-music-b1119500.html>.
14. T. Jalloh, "Speech Acts and Unspeakable Raps," *Journal of Applied Philosophy* 42, no. 4 (2025): 1378–1395.
15. B. Klienberga and P. McFarlane, Violent Music vs Violence and Music: Drill Rap and Violent Crime in London. arXiv.Org (2020).
16. H. Beebe, C. Hitchcock, and P. Menzies, *The Oxford Handbook of Causation* (Oxford Academic, 2009).
17. E. Williams and P. Squires, *Rethinking Knife Crime: Policing, Violence and Moral Panic* (Springer, 2021).
18. F. Stuart, *Ballad of the Bullet* (Princeton University Press, 2020).
19. F. Stuart, "Code of the Tweet: Urban Gang Violence in the Social Media Age," *Social Problems* 67, no. 2 (2020): 191–207.
20. C. E. Kubrin and R. Weitzer, "Rap Music's Violent and Misogynistic Effects: Fact or Fiction?," *Sociology of Crime Law and Deviance* 14 (2010): 121–143.
21. E. Anderson, *Code of the Street: Decency, Violence, and the Moral Life of the Inner City* (W.W. Norton & Company, 1999).
22. F. Stuart and C. L. Moore, "Gang Research in the Twenty-First Century," *Annual Review of Criminology* 5 (2022): 299–320.
23. M. Forman, *The Hood Comes First: Race, Space and Place in Rap and Hip Hop* (Wesleyan University Press, 2002).
24. D. Miranda and M. Claes, "Rap Music Genres and Deviant Behaviors in French-Canadian Adolescents," *Journal of Youth and Adolescence* 33, no. 2 (2004): 113–122.
25. C. Barongan and G. Hall, "The Influence of Misogynous Rap Music on Sexual Aggression Against Women," *Psychology of Women Quarterly* 19 (1995): 195–207.
26. J. Mahiri and E. Conner, "Black Youth Violence Has a Bad Rap," *Journal of Social Issues* 59, no. 1 (2003): 121–140.
27. R. Felson, "Mass Media Effects on Violent Behavior," *Annual Review of Sociology* 22 (1996): 102–128.
28. Z. C. Merz, J. W. Lace, T. R. Coleman, and R. M. Roth, "Challenging the Presumptive Link Between Musical Preference and Aggression," *Psychology of Music* 49, no. 6 (2021): 1515–1531.
29. K. N. Olsen, J. Terry, and W. F. Thompson, "Psychosocial Risks and Benefits of Exposure to Heavy Metal Music With Aggressive Themes: Current Theory and Evidence," *Current Psychology* 42 (2022): 21133–21150.
30. D. U. Patton, J. Lane, P. Leonard, J. Macbeth, and J. R. Smith-Lee, "Gang Violence on the Digital Street: Case Study of a South Side Chicago Gang Member's Twitter Communication," *New Media & Society* 19, no. 7 (2017): 1000–1018.
31. K. Irwin-Rodgers, J. Densley, and C. Pinkney, "Gang Violence and Social Media," in *The Routledge Handbook on Human Aggression*, ed. J. Ireland, P. Birch, and C. Ireland (Routledge, 2018), 1–21.
32. J. Scott, C. Sitton, and S. Marino, *Social Media's Impact on Crime and Retaliatory Violence* (Centre for Public Safety Initiatives, 2023), 1–20.
33. K. Irwin-Rodgers and C. Pinkney, Social Media as a Catalyst and Trigger for Youth Violence. Catch22 (2017).
34. A. Owusu-Bempah, "Prosecuting Rap: What Does the Case Law Tell us," *Journal of Popular Music* 41, no. 1 (2022): 427–455.
35. M.-M. Urbanik and K. D. Haggerty, "'#It's Dangerous': The Online World of Drug Dealers, Rappers and the Street Code," *British Journal of Criminology* 58, no. 6 (2019): 1343–1360.
36. T. McCuddy and F.-A. Esbensen, "The Role of Online Communication Among Gang and Non-Gang Youth," in *Gangs in the Era of Internet and Social Media Melde*, ed. C. Weerman and F. Weerman (Springer, 2020).
37. G. Gunatilleke, "Justifying Limitations on the Freedom of Expression," *Human Rights Review* 22 (2020): 91–108.
38. J. S. Mill, *On Liberty* (Cambridge University Press, 1859).
39. UNESCO, The Legitimate Limits to Freedom of Expression: The Three-Part Test (2021), <https://www.youtube.com/watch?v=Wg8fVtHPDag>.
40. D. Bychawska-Siniarska, *Protecting the Right to Freedom of Expression Under the European Convention on Human Rights* (Council of Europe, 2017).
41. ColumbiaLearn, MOOC FOE1x | 3.3.1 the Three Part Test: Legality | Legitimate Restrictions to Freedom of Expression (2016), <https://www.youtube.com/watch?v=2tiTaIk9fw&t=318s>.
42. ColumbiaLearn, MOOC FOE1x | 3.4.1 the Three Part Test: Valid Grounds | Legitimate Restrictions to FoE. (2016), https://www.youtube.com/watch?v=nGbLH_R56vo&t=160s.
43. ColumbiaLearn, MOOC FOE1x | 3.5.1, the Three Part Test: Necessary | Legitimate Restrictions to Freedom of Expression (2016), <https://www.youtube.com/watch?v=jywOrnEAV5Y&t=236s>.
44. Amnesty, Amnesty International UK/ Freedom of Speech (2025), <https://www.amnesty.org.uk/free-speech-freedom-expression-human-right>.
45. J. Gerards, "How to Improve the Necessity Test of the European Court of Human Rights," *International Journal of Constitutional Law* 11, no. 2 (2013): 466–490.
46. J. Brown, How Is the Government Implementing a 'Public Health Approach' to Serious Violence? (2019), <https://commonslibrary.parliament.uk/how-is-the-government-implementing-a-public-health-approach-to-serious-violence/>.
47. L. Fatsis and M. Lamb, *Policing the Pandemic: How Public Health Becomes Public Order* (Bristol University Press, 2022).
48. B.-T. Bea, This Article is More Than 6 Years Old is UK Drill Music Really Behind London's Wave of Violent Crime? (The Guardian, 2018), <https://www.theguardian.com/music/2018/apr/09/uk-drill-music-london-wave-violent-crime>.
49. A. Dunbar and C. Kubrin, "Imagining Violent Criminals: An Experiment Investigation of Music Stereotypes and Character Judgements," *Journal of Experimental Criminology* 14, no. 4 (2018): 507–528.

50. A. Dunbar, "Experimental Evaluation of Rap Music Attitudes," in *The Encyclopedia of Research Methods in Criminology and Criminal Justice*, ed. C. Barnes and D. R. Forde (Wiley, 2021), 613–621.
51. A. Dunbar, "Rap Music, Race, and Perceptions of Crime," *Sociology Compass* 13, no. 10 (2019): 1–11.
52. U. Khan, "A Guilty Pleasure: The Legal, Social Scientific and Feminist Verdict Against Rap," *Theoretical Criminology* 26, no. 2 (2022): 245–263.
53. C. E. Kubrin and E. Nielson, "Rap on Trial," *Race and Justice* 4, no. 3 (2014): 185–211.
54. C. Reyna, M. Brandt, and V. G. Tendayi, "Blame it on Hip-Hop: Anti-Rap Attitudes as a Proxy for Prejudice," *Group Processes & Intergroup Relations* 12, no. 3 (2009): 361–380.
55. A. Archer and B. Matheson, "Admiration and Education: What Should We Do With Immoral Intellectuals?," *Ethical Perspectives* 26, no. 1 (2019): 1–20.
56. C. Mills, "Do Black Men Have a Moral Duty to Marry Black Women," *Journal of Social Philosophy* 25, no. 1 (1994): 131–154.
57. R. Zheng, "Why Yellow Fever Isn't Flattering: A Case Against Racial Fetishes," *Journal of the American Philosophical Association* 2, no. 3 (2016): 400–419.
58. L. Fatsis, "Sounds Dangerous," in *Harm and Disorder in the Urban Space Social Control, Sense and Sensibility*, ed. N. Peršak and A. Di Ronco (Routledge, 2021), 30–51.
59. L. Fatsis, "From Overseer to Officer: A Brief History of British Policing Through Afro-Diasporic Music Culture," in *Southern and Post-colonial Perspectives on Policing, Security and Social Order*, ed. P. R. Cavalcanti, P. Squires, and Z. Waseem (Bristol University Press, 2023), 45–61.
60. T. Rose, *The Hip Hop Wars: What we Talk About When we Talk About Hip Hop—and Why It Matters* (Basic Books, 2008).
61. A. R. Piquero and R. W. Brame, "Assessing the Race-Crime and Ethnicity-Crime Relationship in a Sample of Serious Adolescent Delinquents," *Crime & Delinquency* 54, no. 3 (2008): 390–422.
62. C. Stott, M. Radburn, A. Kypriandes, and M. Muscat, Understanding Ethnic Disparities In Involvement in Crime—A Limited Scope Rapid Evidence Review, by Professor Clifford Stott et al. (2021), <https://www.gov.uk/government/publications/the-report-of-the-commission-on-race-and-ethnic-disparities-supporting-research/understanding-ethnic-disparities-in-involvement-in-crime-a-limited-scope-rapid-evidence-review-by-professor-clifford-stott-et-al>.