

Dirty Pictures: Performativity and the Obscene Image

Esra Soraya Padgett

epadgett@gradcenter.cuny.edu

Abstract: On December 17th, 2018, an adult content ban was enacted on Tumblr, a microblogging and online content-sharing platform. The ban enforced the removal of all so-called adult content through the use of a censorial algorithm designed to flag or remove images. This paper traces the discursive negotiations between Tumblr and its users that emerged after the ban, a process centered on the very terms of what defines obscenity and constituted by competing pragmatics for sorting obscene images. This article outlines the history of obscenity law in the U.S. and its use of both inherentist and performative approaches. It then traces the shift in Tumblr’s content ban, from a legal framework to an automated algorithmic agent for the task of discernment. The paper argues that disputes arising in the aftermath of the ban are the result of clashing sets of (meta)pragmatics—the semiotic processes of category-making, and image categorization more specifically—which, once the ken of juridical experts, are now brought to light by Tumblr’s act of censorship.

Keywords: algorithms; censorship; social media; performativity; obscenity

“Prohibition gives to what it proscribes a meaning that in itself the prohibited action never had.”

Georges Bataille, *The Tears of Eros*

Introduction

On December 17th, 2018, an adult content ban was enacted on Tumblr, a microblogging and online content-sharing platform. The ban enforced the removal of all so-called adult

content through the use of a censorial algorithm designed to flag or remove images that fit Tumblr's new Community Guidelines' definition of obscenity. In a matter of days, all "obscene" content—including posts, images, essays or entire blogs—were deleted from the site, sometimes erased entirely and without warning, other times moved to a "flagged content" (Crawford and Gillespie 2014) section of the site where users could view what had been censored from their account. The ban was met with massive outrage and came to be known by Tumblr users as "the purge." This was because the site was known as one of the only online spaces where NSFW (not safe for work) content could co-exist with other artistic images, as well as a site for hosting NSFW content that did not conform to mainstream pornographic standards—such as online communities of queer, kink, body-positive and other alternative sexual lifestyles. However, what unfolded in the days following "the purge" was a negotiation process centered on the very terms of what defines obscenity and constituted by many contradictory and competing semiotic practices. These competing semiotic practices were particularly visible in Tumblr users' refusals to define their blog content as obscene or worthy of censorship, including collective actions such as spreading petitions, organizing mass log offs through hashtags (#LOGOFFSTANDUP), the creation of memes both humorous and serious (Figure 1), as well as boycotting the site altogether. In this paper I focus on this process of negotiation as it emerged in the aftermath of "the purge." In particular, I interrogate the ways in which the obscene is a transient, and socio-historically produced semiotic category.



Figure 1. "The purge": The adult content ban is seen as just the beginning of Tumblr's demise.

This paper begins with the presumption that obscenity requires interpretation within a specific semiotic classification system, one in which many other possible categories are also produced. Thus defined, obscenity requires sorting: but what does the sorting process look like, what kinds of discursive practices does it entail? And how might the negotiation of obscenity through discursive practices—the sorting—come to shape the category of obscenity itself?

Anthropology offers a long history of inquiry into sorting processes: of dirt from order (Douglas 1984[1966]); signal from noise (Bateson 1972); of the edible (and marriageable) from the inedible (and unmarriageable) (Tambiah 1985[1969]; Silverstein 2004); or even spam from ham (Kockelman 2013). In regard to one such sorting process, Douglas (1984[1966]:36) suggests that “dirt is the by-product of a systematic ordering and classification of matter, in so far as ordering involves rejecting inappropriate elements.” Though my focus is not on *dirt* as such but on *dirty pictures*—and here I don’t mean films but obscene images—Douglas’s supposition holds: where there is obscenity, there is a system. Thus, my shift from the assumption that obscenity inheres in visual or material qualities of an image or object to an investigation of *sorting* via a particular semiotic mechanism—via what Paul Kockelman calls the “sieve” (Kockelman 2010, 2011, 2013), that is, a filter through which things, or more specifically in this case, images, are sifted into categories of obscene and non-obscene—follows from a time-tested approach within anthropology. Looking into the *sieve* does not end with questions of categoriality (is it dirt or matter?), however, but instead considers the processes through which categories emerge (Bowker and Star 1999) and the pragmatics that occasion their emergence and follow in their wake.¹

Even when the system of interpretation is clearly in view, such as within a particular legal framework, obscenity requires negotiation. This need for negotiation is in part the result of the infamously vague terms of many legal definitions of obscenity; in the United States, this is epitomized by Justice Potter Stewart’s statement in the 1964 case, *Jacobellis v. Ohio*: “I can’t define it, but I know it when I see it.” For a Supreme Court opinion, Stewart’s admission is puzzling: obscenity is *undefinable*, yet at the same time somehow *knowable*. As William Mazzarella (2013:194) asks: “What kind of knowing or understanding is imagined here?” But what kind of knowing requires a second question, who is seen as capable of this knowing? The role of sorting requires an epistemological authority: an expert (Putnam 1975; Boyer 2008; Carr 2010), agent (Kockelman 2013), or discerning subject (Reyes 2017) seen as capable of identifying an individual (person/object/image) as a kind. While for obscenity, the task of discernment was long held by judges, in “the purge,” this sorting process became automated, performed by Tumblr’s algorithm for removing images. What happens when the determination of obscenity is shifted from

juridical experts to an algorithmic agent? What falls out along the way and what remains across such a transition?

By asking this question, in this paper ultimately I am less concerned by what algorithms are (Seaver 2013, 2017) or how they work (Kockelman 2013)—though these are all important questions, of course (see Barocas, Hood and Ziewitz 2013; Gillespie 2014; Striplas 2015; Beer 2017; Kitchin 2017). Rather, my concern is with the negotiations that arise when they are seen to fail. Though online platforms frequently “pick and choose” and even surreptitiously delete content and users (Gillespie 2015, 2018), the purge was sudden and platform-wide, radically shifting the topographic landscape of the website overnight. The uniquely chaotic situation that resulted from this decision enabled broader community engagement, as users felt the shockwave of rapid and unexpected censorship across the platform. In particular, the purge catalyzed an unprecedented uptick in negotiations of obscenity, brokered in exchanges between Tumblr and its users in the weeks following December 17th. These negotiations not only redrew the line between obscene and not-obscene but confronted the pragmatics of line-drawing itself. In this paper, I argue that disputes arising in the aftermath of the purge were the result of clashing sets of (meta)pragmatics—the semiotic processes of category-making, and image categorization more specifically—which, once the ken of juridical experts, were brought to light by Tumblr’s act of censorship.

The Illicit Image vs. Images that Elicit

Historically, legal definitions of obscenity in the United States can be classified into two discursive categories: *illicit images*—which are described referentially via “inherentist” (Mazzarella 2013:195) approaches that define obscenity by what the image contains or depicts; and *images that elicit*—those categorized through performative approaches that determine an image’s obscenity by its consequentiality, that is, its tendency, or potential, to affect. Early legal standards of obscenity attempted to prevent an inherentist logic wherein specific content is deemed categorically obscene with a framework of performativity, wherein obscenity is a matter of the image-object’s performative effects relative to particular *kinds* of subjects. In order for an image to elicit something, it must elicit in someone, and thus obscenity law inscribes two subject types: the *discerning subject*—those capable of deeming something obscene—and the *susceptible subject*—those vulnerable members of a community that could be corrupted by exposure to obscenity. In his work on the Indian Censor Boards and British legal definitions of obscenity, William Mazzarella describes the Hicklin test, a British legal standard from the nineteenth century that establishes the test of obscenity as “whether the *tendency* of the matter charged as obscenity is to deprave and corrupt those whose minds are open to such immoral influences.” When obscenity is viewed as a “tendency,” the image is defined by its performative potential—as a likelihood to initiate, or proliferate obscenity among a

particular group. The Hicklin test revealed, and presumed, a second tier of sorting: how and by whom are susceptible subjects—“those whose minds are open to such immoral influences”—to be identified? This classification process was not outlined in the law, but presupposed an indexical second-order of social differentiation: where there was obscenity, there was a class of subjects vulnerable to its influence, as well as one with expertise in its identification (who were also, presumably, not vulnerable to its influence).

The Hicklin test was absorbed into early U.S. legal definitions of obscenity. When the Comstock laws (a set of federal acts to prohibit the circulation of “obscenity” via the US postal service) were deemed constitutional in 1873, it was through Judge Samuel Blatchford’s use of the Hicklin test (Wood 2007). This approach to obscenity as a tendency with specific performative effects on specific types of people continued to be used in the U.S. until 1957, when the Hicklin test was finally ruled as inappropriate by the Supreme Court in *Roth v. United States*. However, the new standard for obscenity was equally, if not more, opaque. In this case, the Court set out what would be called the Roth test, described as: “whether to the average person, applying contemporary community standards, the dominant theme of the material, taken as a whole, appeals to the prurient interest” (*Roth v. United States* 1957:489). Though the new standard of obscenity still seemed to define images by their performative effects—in this case, the forceful attraction of prurience—the Roth test also recalibrated the social personae indexed by the tests, such that both discerning and susceptible subject categories were radically expanded. First, the test named a new agent in the sorting process of obscenity, shifting from the judge’s expert knowledge to that of “the average person.” This expansion of the role of the discerning subject was mediated through the notion of the “test”—as an instrument of discernment (Reyes 2017)—that even the “average person” could use. At the same time, the susceptible subject was nearly universalized in the Roth test as less a type of person, but rather a latent tendency in us all—a “prurient interest”—that could be unwittingly appealed to by outside forces. The “average person” was now capable of being both susceptible to *and* discerning of obscenity.

The Roth test was reversed in a case just a few years later, *Jacobellis v. Ohio* (1964), when the Supreme Court was again asked to set out a clear standard for determining obscene material. Unlike previous cases which seemed to replace one “test” for another, *Jacobellis v. Ohio* did not proffer a new test, but instead refused altogether the task of setting new criteria (through which a test is conducted). As Justice Potter Stewart wrote in his opinion, “I shall not today attempt further to define the material I understand to be embraced by that shorthand description (hard-core pornography); and perhaps I could never succeed in intelligibly doing so.” Stewart’s statement framed obscenity as an object unable to be harnessed to specific criteria, making a test, at least one the average person could administer, impossible. By shifting away from defining obscene material as a set category,

Stewart instead appealed to a pragmatics through which obscenity could be identified—more specifically to the social differentiation process that decided who was capable of discernment. Though his statement, “I can’t define it, but I know it when I see it,” seems to put forth an individualist but unwavering methodology of discernment—stating that one is at least capable of *knowing obscenity when seeing obscenity*—the court’s decision also reclaimed discernment as a task for juridical experts, but only at the Supreme Court level. Since the decision reversed an obscenity charge upheld by the state, its effect was to rescind the state’s role in discerning obscenity while reclaiming it as a task for Supreme Court justices. Thus, Stewart’s infamous quote above, is really an act of “meta-discernment” (Reyes 2017:S117), through which Stewart, using the self-reflexive first-person deictic “I,” repositions himself (and others occupying the role of Supreme Court judge) as the only capable discerners of obscenity.

In the decades that followed, legal standards for obscenity seemed increasingly futile as particular legal and popular language ideologies demanded tests by which obscenity could be defined denotatively (via an explicit list of materials that would determine category membership). As the sexual revolution cascaded forward into the early 1970s, heaps of new obscenity cases followed in its wake. The Supreme Court began to feel the burden of epistemological authority. As legal scholar Ryen Rasmus (2011:259) describes it: “Eventually, the sheer volume of potentially obscene work that the Court was required to sort through led Justice William Brennan, author of the Roth opinion, to remark, ‘I’m sick and tired of seeing this goddamn shit’.” The workload was too much, and federal judges wanted to shift some of the burden back to state courts. In 1973, the Supreme Court would offer the Miller Test, which continues to be the contemporary standard today. This decision would shift again from a universal, federally-defined standard to more localized community standards—putting the task of defining obscenity back at the State level. Chief Justice Warren Burger described the Miller Test as follows: “the basic guidelines for the trier of fact must be: (a) whether the average person, applying contemporary community standards would find that the work, taken as a whole, appeals to the prurient interest, (b) whether the work depicts or describes, in a patently offensive way, sexual conduct specifically defined by the applicable state law; and (c) whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value” (*Miller v. California* 1972). The Miller test thus reprises the role of “the average person” as expert that was central to the Roth test, but this time locates them within a more localized community. The Miller test also clarifies the testing criteria for obscenity through two addendums: one which defines obscenity by content (what it “depicts”), and a second which views it by its effects as a “serious literary, artistic, political, or scientific” contribution.

In a sense, *Miller v. California* was the Supreme Court's admission of the transience of obscenity—stable definitions, at least at the national level, were simply impossible, and could not be written into law. However, this liberal definition of obscenity was by no means an end to censorship; new avenues for its legal justification had to be sought after. Though the Miller test remains the U.S. legal standard for obscenity, its notion of community standards has also been transferred to the domain of site- or platform-internal online “communities” (Boellstorf 2008). Like the legal standards that came before, these online platforms' standards presume a community that contains both discerning and susceptible subjects, roles that are constantly recalibrated from platform to platform.

But if legal standards regarding “obscenity” have not changed, what, then, occasioned Tumblr's adult content ban? To justify its new policy, on November 26, 2018 Tumblr made a statement that the Tumblr app had been removed from the Apple store due to findings of child pornography (Porter 2018). This necessitated an upgrade in their content filtration process. However, their decision to remove all adult content went far beyond the necessary “upgrade” to remove child pornography resulting in a “purge” of content that set new definitions of obscenity. This response—removal of all adult content rather than removal of *illegal* content such as child pornography—was seen by users as an overreaction, and led many to believe that Tumblr's decision was, in fact, impacted by shifts in the law—only it was not obscenity law but rather a recent anti-sex trafficking law (Martineau 2018; Padgett 2018).

Coming only a few months prior to Tumblr's announcement of the ban, most users cited two controversial bills—SESTA (Stop Enabling Sex Traffickers Act) and FOSTA (Fight Online Sex Trafficking Act)—as the real catalysts for Tumblr's decision. These bills sought to broaden the net of liability for sex trafficking by holding third-party online platforms (like Tumblr) legally responsible when evidence of trafficking (broadly defined) was found on their site. Though the bills were heavily critiqued by both sex workers' rights activists as well internet freedom and free speech advocates (including the ACLU and even the Department of Justice), the bills passed in April 2018 (Romano 2018), just six months before Tumblr's adult content ban. These laws drastically expanded and intensified the social personae to be included in the susceptible subject category—no longer was the concern immoral minds or prurient interests—but the much more somber and not-so-easily-dismissed harm done to real victims of sex trafficking. In the months that followed, a wave of anxiety rolled through the top internet companies—with many “cleaning up” their platforms in fear of being accused of sex trafficking, including shuttering parts of their site permanently (such as Craigslist's Personal ads section) (Cole 2018; Lingel 2020). Though Tumblr never admitted to being influenced by SESTA-FOSTA, the heightened concerns around both sex-trafficking and child pornography and the possibility of being held legally culpable of content-hosting raised the stakes of Tumblr's decision

significantly. These stakes are discussed further in the following sections, as I analyze the process of “purging” content along with users’ responses to the ban. In particular, I trace how, like the wavering in judicial definitions of obscenity outlined above, arguments on both sides boiled down to a dispute over the pragmatics of obscenity. This dispute evinced a tension between referential and performative approaches to obscenity, as well as conflicting views of the social relations that obscenity entails.

The Purge

Tumblr’s decision, in December 2018, to ban obscene content was outlined in their new Terms of Service guidelines (Figure 2). Though the Terms of Service were provided preemptively—as a guide for users to self-censor before the December 17th purge—the company made clear that failure to self-censor would result in the removal of all images in violation by their algorithm on that date. However, the opportunity to proleptically avoid the algorithm’s sweeping edits was not taken kindly by most users, and the new Terms of Service immediately sparked a debate over what counted as obscene, as well as the proper methodology for determining obscenity. This debate resurfaced both inherentist and performative approaches to the image. In Tumblr’s updated Terms of Service, it is clear that an inherentist logic is at work: adult content is defined denotatively as images that “show” or “depict” certain things, specifically, “real-life human genitals,” “female-presenting nipples,” or “sex acts.”

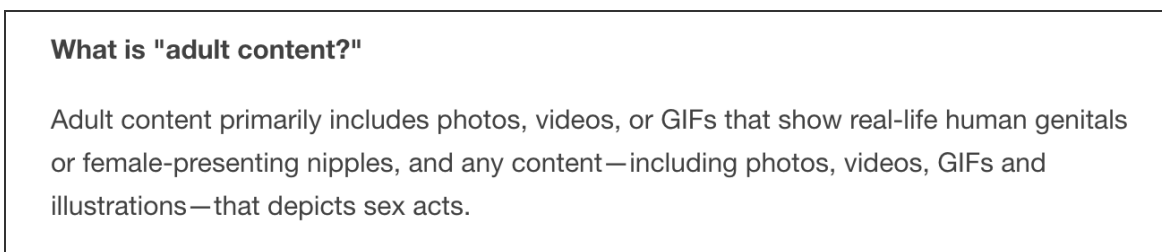



Figure 2. Virtual community standards: Tumblr’s updated Terms of Service Guidelines

Addressing these guidelines at the inherentist level, many Tumblr users rejected that this list of material content was categorically obscene. In interviews I conducted with users following the ban, an overwhelming number disparaged the usage of “female-presenting nipples” as a shibboleth of the obscene (Figure 3), some stating that the definition was misogynistic as well as transphobic—remarking on the essentialist view of gender that it perpetuated. Though appropriating the parlance of contemporary gender politics that emphasize “gender presentation” and “expression,” Tumblr’s awkward and unclear usage of inclusive language was seen as a red herring of the chaos to come. Was “female-presenting nipples” a nod to gender expression as separate from biology, or a reification of anatomy as definitive of gender identity? On the one hand, if “female-presenting nipples” meant nipples that appeared to be female regardless of the gender identity of the

person to which they were attached—thus censoring a person’s biological sex, not their gender expression—then Tumblr was being transphobic and exclusionary of gender non-conforming identities. However, if Tumblr meant nipples on bodies of those who present as female, then the algorithm, users suspected, would surely be unable to know this, and would continue to censor based solely on anatomy. Even if the machine-learning of the algorithm relied on human interaction, gender expression was highly individualized and could only be addressed on a case-by-case basis, thus making Tumblr’s inherentist approach (at least to nipples) impossible. Alongside questioning the algorithm’s ability to perform, these users were deconstructing a supposedly straightforward denotational approach to obscenity: where Tumblr had attempted to give a definitional list of explicit material—users unveiled the depth of discursive processes required for discerning the categoriality of an image, at least where gender was concerned.



 **marquis-de-la-rose-noire**
Fuck Tumblr!

Me: titty

Tumblr Staff:



Figure 3. Backlash to ban on “female-presenting nipples”

Other users’ rejections swept more broadly over the Terms of Service (ToS) guidelines. As one user, *maladyofreverie*, wrote: “As far as the ToS guidelines, I do not personally think that the human body, or even sexuality, is obscene. I feel like Tumblr is just another culture of oppression of humanity now. I do not want to be a part of a community that persecutes healthy behavior.” This user rejects the very notion that nudity or even sexuality is categorically obscene, thus disputing the very literal material constraints of the new guidelines. But this user also points us back to a performative approach to the notion of obscenity: Tumblr is not just censoring images, but “persecuting healthy behavior.” Here censorship, in its attempts to limit the negative performative effects of adult content, is seen as doing the opposite: limiting its positive performative potential.

Studies of taboo language can be helpful as a guide for understanding what it means to say that these images are performative. As Luke Fleming and Michael Lempert (2014:498) suggest regarding verbal taboos: “They are performatives in the sense that by uttering taboo expressions speakers accomplish socially recognized acts.” In taboo language, the utterance itself counts as a violation of the taboo—and this is certainly the case with taboo images and the specific sub-class of “obscene” images this paper discusses; the existence of the image on the platform of Tumblr is viewed by Tumblr as a violation of the taboo, a crossing of the boundary between what is appropriate or not (and thus must be removed). As such, the image itself is seen as constituting an act of obscenity. This is further evidenced by the presence of “avoidance registers” (Fleming and Lempert 2011), or conventions for speaking about unmentionable topics without actually mentioning them (e.g., “the F-bomb; “He-who-must-not-be-named”). In image presentation, such avoidance practices are most commonly achieved through a strategically placed black box or pixelated blur (Figure 4).



Figure 4. Avoidance registers in image presentation on Tumblr (4a) and other digital media (4b; from Reporters without Borders, “Censorship Tells the Wrong Story” campaign, 2011)

Yet the vast discrepancy in how the performative potential of images is interpreted opens up questions about the mechanics of performativity for obscene images, including the felicity conditions required for their achievement. Fleming and Lempert (2011) have stated that the performative effects of taboo linguistic registers are “rigid” and nearly “indefeasible”—that is, they are highly unlikely to “go sideways” (Nakassis 2013) or be

“unhappy” (Austin 1962)—due to the fact that verbal taboos “require few if any co(n)textual felicity conditions to accomplish performative effects” (Fleming and Lempert 2011:499). In a particularly apt comparison, Fleming and Lempert discuss the FCC ban on “obscenities” in radio broadcasting and point out that when any usage of an “obscenity” occurs, no matter who says it or the purpose of saying it—even in reported speech of someone else using the word—“obscenity happens” (Fleming and Lempert 2011:498). Thus, the supposed rigidity of taboo language is reinforced by the notion that the speaker is seen as violating a taboo regardless of whether or not they are the author of the words being said (Irvine 2011). This distinction is clarified by Erving Goffman’s (1981) notion of “production format,” in which he breaks down what is commonly thought of as the “speaker” into three separate roles: *animator*—the “sounding box” that vocalizes the utterance; *author*—who chooses the words or sentiments; and *principal*—the authority, or social identity responsible for the words uttered. In the cases discussed by Fleming and Lempert, it seems that these roles are collapsed in interpretations of taboo language—regardless of who authored the words, the animator is seen as the principal, the social identity responsible for the words, such that even quoted speech “counts” as cussing (also see Nakassis in this issue).

However, if we transfer this framework of taboo performativity to images, new questions arise: namely, if images are acts, whose image-acts (Bakewell 1998; Nakassis 2019) are they? In Tumblr’s case it appears that part of the dispute circles around the shift in production format set on by the purge. While users felt that they were the principal behind the images they post—that is, the person “committed to what the words say” (Goffman 1981:144)—Tumblr’s ban on content suggested that *it* was the principal behind all images hosted (animated) on the platform. In the decision to censor, Tumblr was taking responsibility (more specifically, the potential liability) for the image-acts of individual users.

However, it is important to note that this responsibility was not the norm for content-sharing sites and the online communities they host—nor was it the norm for Tumblr prior to the ban—who tend to separate their role as animator (content-hosting) and principal (liable for content they host). The ability to separate these roles is built into the legal codes for internet freedom—specifically section 230 of the Communications Decency Act, which states that websites are given immunity from prosecution for the content shared on their platform except for a very narrow set of exceptions, like child pornography (Tripp 2019). SESTA-FOSTA, the bill that loomed large behind the Tumblr purge, was a partial repeal of section 230 that aimed to expand the list of exceptions to include any content related to sex trafficking. However, the bill’s definition of sex trafficking was so capacious that it could be interpreted to include all sex work, including consensual forms and even talk

about consensual sex work.² The relevance of SESTA-FOSTA to the Tumblr purge was raised by many users following the announcement of the adult content ban:



Figure 5. Users in 5a and 5b cite SESTA-FOSTA as cause of Tumblr ban, and note its harms.

In the screen shots in Figure 5, users share critical opinions about SESTA-FOSTA and state the link between the laws and Tumblr. In 5a, user @tigerlizii suggests that Tumblr's decision came from "fear of being sued," while in 5b, user @paramorefan061 states that Tumblr is just one of many websites that has changed their stance on censorship since the laws' passage. While Tumblr remained silent on whether their choices were influenced by the new laws or not, the shift was clear in their approach to the production format of image-acts.

Despite these intensified circumstances and the increased awareness of SESTA-FOSTA-incited fears, many Tumblr users fiercely contested the notion of rigid performativity for images. This was accomplished in part by broad critiques of censorship as well as the specific tactics of censoring. Returning to Figure 4, the avoidance-register images shown are really examples of metapragmatic discourse about censorship: 4a points to the absurdity of Tumblr's ban on nudity, while 4b illustrates how misleading avoidance techniques can be (David Cameron is not in fact giving the middle finger in this photo, as the pixelated area suggests).

These critiques were furthered by suggestions that Tumblr was not even capable of accomplishing the kind of censorship it proposed: as user @tigerlizii writes in Figure 5a, “Tumblr should be sued” for its failure to eliminate even the most extreme violations of its own Terms of Service after the December 17th purge. Other users framed their critiques by rejecting a simple correlation between adult content (NSFW) and obscenity. Just like @maladyofreverie's refusal to agree that "the human body, or even sexuality, is obscene", @tigerlizii's (5a) distinction between the NSFW content that Tumblr is removing in fear and the actual illegal content of child pornography and sex trafficking that it has failed to remove, suggests that obscenity - as defined by Tumblr's ToS Guidelines - is not seen as happening by all. Beyond a rejection of the categorization of certain images, however, these users' resistance points toward a different pragmatics of obscenity altogether— including an alternate view of the social relations (e.g. the “community,” discerning and susceptible subjects) that underlie such pragmatics. In the following section, I examine how despite awareness of the taboos—on female breasts in public, as well as nudity and sexual activity—Tumblr users insisted on an alternate semiotic approach to obscene images, one that suggested that unlike verbal taboos, image-acts have much less certainty in their performative effects. Unlike Tumblr, these users made the claim that the sign-values of obscene images have quite a lot of room for defeasibility (Agha 2011), and projects to define images as obscene are just that—projects which can succeed or fail.

Ghost in the Machine: Interacting with Tumblr's Image Recognition Algorithm

Assessing the success or failure of Tumblr's notion of obscenity requires a closer look at how the project itself was carried out. As noted above, through an algorithm for sorting images, posts, and blogs into categories of obscene or non-obscene, Tumblr enforced its new role as principal, taking hold, by way of a non-human moderator, of the content it was now presumed responsible for. In this section, I draw on data collected through digital ethnographic work, including interviews with Tumblr users and participatory engagement with Tumblr's algorithm. I pay particular attention to what were regarded as the algorithm's failings, as well as the possibilities for intervention in the workings of the algorithm as a particular kind of sieve. I argue that Tumblr's use of the algorithm, and the censorial assemblage (of the corporation, human moderators, coders and designers) to which it belongs, enacts a situation of *forced indefeasibility*. In such a situation, an outside structure attempts to ensure (in this case via censorship) the performative outcome of a given sign.

As I spoke with users, I realized that not only were there disagreements about the new Terms of Service guidelines, but the algorithm itself was seen as stepping beyond the boundaries set by those guidelines. The algorithm, in other words, was floundering in its role as a discerning subject. This was apparent in an emergent genre of memes that

followed the ban, in which users posted images that had been incorrectly flagged by the Tumblr's censor (though the truth of the flagging is uncertain).

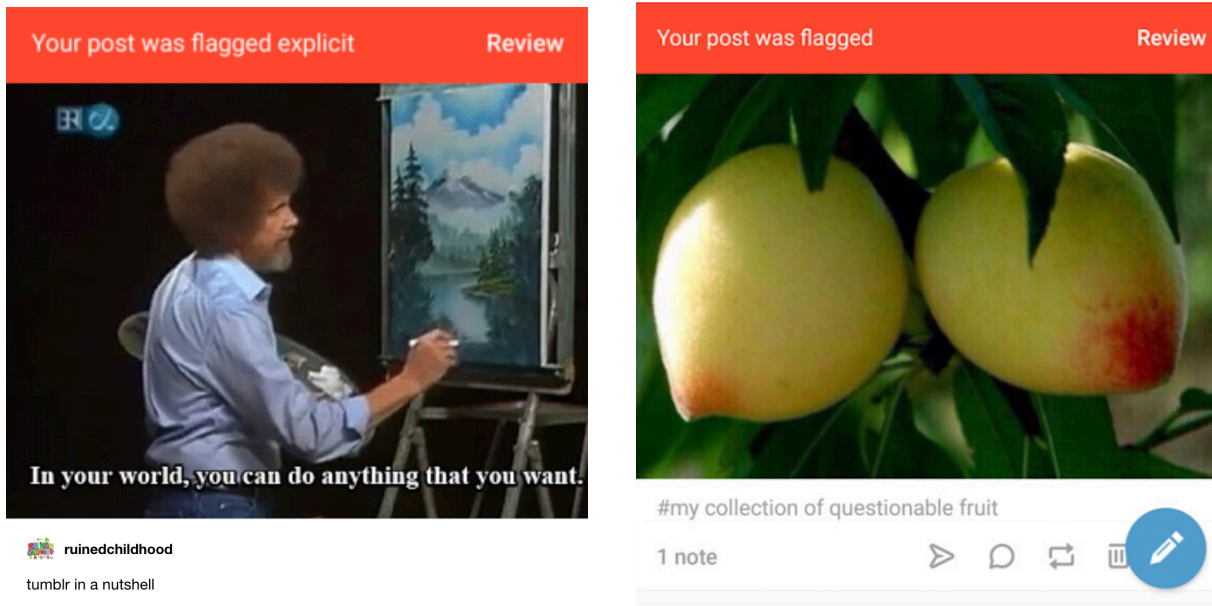


Figure 6. Post-December 17th, 2018 posts (6a and 6b) captioning “Tumblr in a nutshell”

The memes in Figure 6 serve as contestations of the adult content ban by elucidating the image recognition algorithm's failures, particularly its production of “false positives” (Captain 2019)—images that were absent of any of the Terms of Service guidelines' listed features of adult content (see Figure 2). Figure 6a pokes fun at the algorithm by showing how drastic its errors could be: a quintessentially benign image—a still from Bob Ross's painting show—has been incorrectly flagged for removal. In Figure 6b, a photo of two lemons also harbors the censors' red band, alerting the user that “your post was flagged.” But the flagged image is reposted with the user, not-original-after-all's commentary: below the photo reads “#mycollectionofquestionablefruit, while above (not featured in the figure) they write “Lemon-Presenting Nipples,” a rebuff of the Terms of Service guidelines' “female-presenting nipples.”

These memes, whether the result of actual errors or not, are meant to suggest that the algorithm is not sieving correctly. As another user, talesfromweirdland, put it:

Recently a very tame photo of Marilyn Monroe that I shared was censored, with no option to appeal it. The help desk didn't respond. I didn't see anything offensive about it, so you start wondering. There's a distinction to be made between sex, eroticism, porn, nudity, and vulgarity; but Tumblr's AI bots can't even make the distinction between Scrooge McDuck and a vagina.

Though this user does not dismiss the need for censorship altogether, in distinguishing their view of obscenity and Tumblr's, the central culprit is the approach itself: Tumblr's choice of automation in the task of censoring. Following a futile interaction with Tumblr's algorithm, this user concluded that the algorithm was not a viable discerning subject, it was not capable of "the kind of knowing or understanding" (Mazzarella 2013:194) required to sort obscene images. However, in this act of meta-discernment, user talesfromweirdland rebuked not only the algorithm for its error but the system Tumblr had put in place (to which the algorithm was a part) for its rigidity, specifically its inability to see images as context-dependent semiotic projects with uncertain outcomes. While he believed that "Tumblr's AI bots" were inept, the problem was furthered when there was no room for appeal—no response from the "help desk"—and no option other than to accept the censorship. Though this user insisted on other meanings and performative goals in posting this image, the possibilities of these alternatives were foreclosed by Tumblr's censorial assemblage—the algorithm, the user interface, the channels of communication, the presence (or lack) of human moderators. These elements worked in tandem as an infrastructure that harshly limited and overdetermined the image's performative potential as anything other than obscene, enacting a forced indefeasibility that barred other modes of interpretation going forward.

To get firsthand experience with the sorting practice of the algorithm and the extended assemblage to which it belonged, I started my own Tumblr blog where I posted images which technically exceeded the Community Guidelines of adult content. First, I posted a still from mainstream pornography, something that would obviously fit the Community Guidelines of adult content. It was immediately flagged for removal, along with the message in Figure 7.

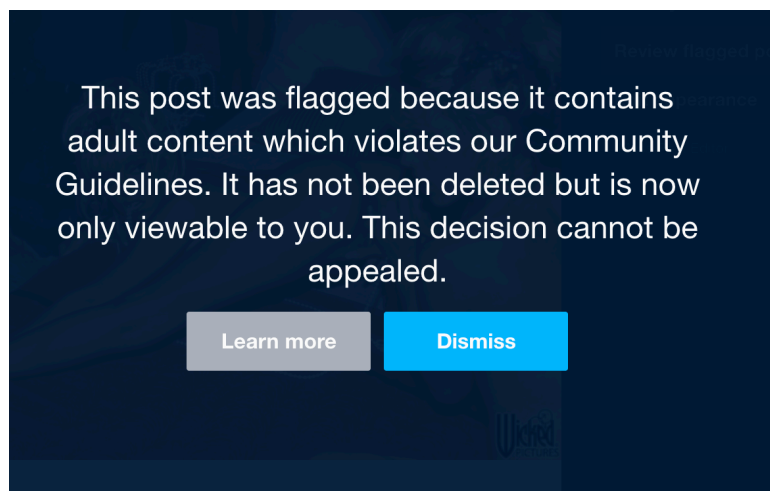


Figure 7. Content appeal on Tumblr

This response was akin to the response that user talesofweirdland had received for the Marilyn photo: both were removed without the option for appeal. In my post, the decision had been both immediate (suggesting no human intervention) and irreversible, leaving me with no way to move forward with an appeal or even inquire into the details of the flagging. This was an issue common to many of the users with whom I spoke, and one that had been particularly rampant in the days immediately following the December 17th ban.

In refusing a space for appeal, Tumblr effectively states that there is no negotiation process for certain images. In these cases, the algorithm decides that a given image is obscene, both inherently (“The image contains adult content”), as well as performatively (the image-effects are seen to be rigid, indefeasible). Like taboo language, the felicity conditions are few, or none: within the “Tumblrverse” the image-act, once censored, simply is obscene, once and for all (at least by Tumblr’s standards). But if this is a kind of rigid performativity, it is a *forced* indefeasibility: that is, rather than arising pragmatically from an image’s performative effects, the algorithm’s decision for removal is pre-emptive, recursive, and metaleptic (Butler 1997; Inoue 2006). It attempts to enforce a rigid performativity simply by naming it so. As such, its temporal motion is peculiar: forced indefeasibility both retroactively inscribes the past with new meaning while pre-emptively foreclosing the possibility of new meanings in the future.³

The Appeal

Not all instances of the algorithm’s censorial work enacted this forced indefeasibility. A second image that I posted featured a photograph from the early 1900s of a nude woman, with the watermark “Fine Art America” in the bottom right corner (Figure 8a). This image was chosen intentionally to fit into the Terms of Service guidelines’ exceptions of “nudity related to political or newsworthy speech, and nudity found in art, such as sculptures and illustrations.” This image was also flagged, but this time three minutes after posting and with the option to appeal. Though I expected to explain the reasoning behind my appeal, there was no space for doing this, just a button to “Appeal,” which I clicked. Once the image had been appealed (8b), only the message, “Your post is in content appeal. Once a decision has been made, we’ll send you an email,” appeared with no further explanation (8c). Somewhere, the message implied, deliberation was occurring: a deliberation that would result in a decision. After seven minutes had passed, the post went back on my blog and I received the email in 8d. I was given no information about how the decision was made, other than a referral back to the Community Guidelines.

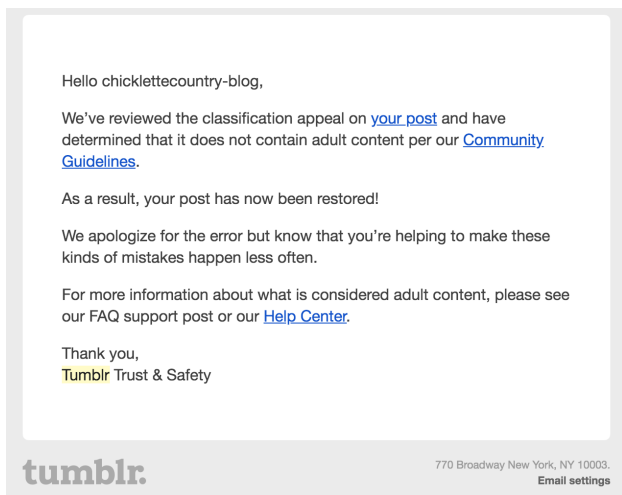
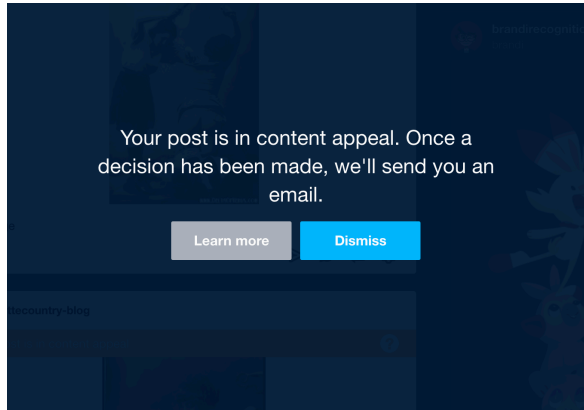
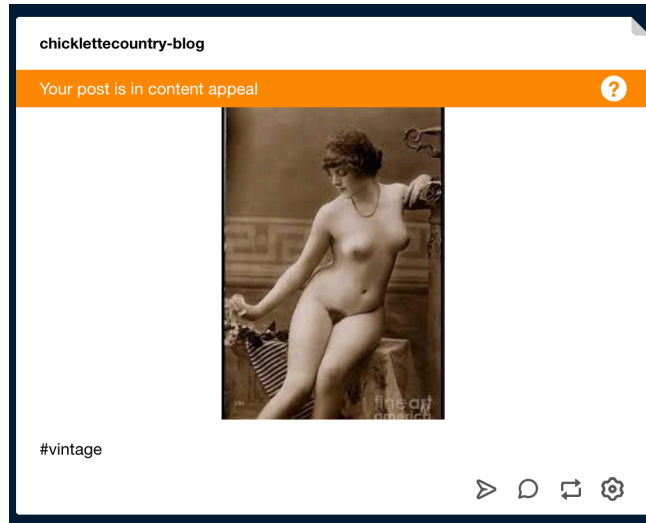


Figure 8. Content appeal on Tumblr: (8a) Original image posted; (8b) Image flagged; (8c) Tumblr's response; (8d) Tumblr's decision

While my message was written and signed by a committee (Tumblr Trust & Safety), and thus acknowledged its own impersonality to an extent, I discovered in interviews with users that this was not a consistent strategy of Tumblr at the start of the ban. This was clear in an interaction one user had in the week before “the purge,” when an image on their page was deleted (as opposed to flagged). When the user emailed the Tumblr staff email to appeal, as well as to retrieve the image, the reply “read something along the lines of ‘we’re sorry for your loss and thank you for your concern...’ and then the new guidelines, and ended with ‘thank you for your time, Ben from customer service’” (Supernove-exe). When I asked @Supernove-exe if she thought that Ben was a real person and whether she had been in contact with him again, she responded:

it read off very robotic, so I wanted to see if “ben from customer service ” was actually a real person and wrote another e-mail to the tumblr staff, in which i

plead to let the NSFW artwork stay...And sent it to the tumblr staff email, and got back the exact same thing. Word for word. “Ben” is just a robot.

This user, like many others I talked to, was actively engaged in a process of parsing human from non-human interaction. This parsing was based on a rough combination of discursive cues: the tone and style, repetition, and failure to respond accurately to, or indexically incorporate (Goodwin 2017) others’ messages in a chain of interaction. These users were consistently frustrated when messages that they believed were being sent to human staff were replied to in ways that read as non-human. One such situation was described in one of my interviews:

I wrote to the help desk to complain many of my posts were incorrectly flagged but that these posts, for some reason, lacked a “Request Review” button. I specifically wrote, as I often do: ‘Please give me a sensible reply and not a standard one.’ The reply was like, ‘You can appeal any flagged posts by clicking the Request Review button...’ That has to be a joke, hasn’t it? That’s either a bot or someone toying with you.

The lack of transparency is notable; as other scholars of algorithms have noted, obscurity or opacity are not random features but options built into an algorithmic system (Seaver 2017). Algorithms are often treated as collective products—assemblages of code, mechanical infrastructure and human design—wherein those who we might presume to be “inside” eschew responsibility by reiterating a “blackbox” (Diakopoulos 2013; Pasquale 2015; Noble 2018) ideology that treats algorithms as unknowable, even to those who create them. Thus, if the work of Tumblr’s algorithm was secret, this secrecy should be considered as a social process (Simmel 1906; Jones 2014), one that was engendered not by the algorithmic code but by both insiders and outsiders that made up the censorial assemblage. Tumblr’s secrecy was part of its tactics for enforcing censorship and was thus inseparable from their pragmatic approach to obscenity.

Though the opacity of Tumblr’s responses may have been a tactic for shirking responsibility, in the email I had received, there was something particularly striking about the line: “We apologize for the error but know that *you’re* helping to make these kinds of mistakes happen less often” (Figure 8d; emphasis added). The email was thanking me for helping the algorithm learn how to sort. Or, was it thanking me for helping the algorithm to redefine the object at the core of its task of sorting: obscenity? Was “Tumblr Trust & Safety” suggesting that I too was an expert, an agent more capable of sieving obscenity than the algorithm itself? Tumblr kept the process of learning secret, promising that they employ “humans to help train and keep our systems in check,”⁴ but showing little to no

evidence of human engagement with users. Though never coming fully into view, the note seemed to reference a system of classification that the algorithm was operating within; one which, seemingly, I and other users were helping to construct. At the same time, the severely limited interaction and the definitiveness of the appeal process' outcome both obscured the working of the algorithmic assemblage, while at the same time retaining its position as the sole authority, the only agent capable of making the decision. Like the Supreme Court in the judicial case of *Jacobellis v Ohio* decades earlier, the Tumblr algorithm was now the ultimate arbiter of obscenity, the judge. In this “case,” my involvement was really more as a defendant, putting forth an argument that would only win if approved by the algorithmic assemblage’s “expert” logics.

What Is Permitted

The appeal process did not only reveal the Tumblr algorithm’s classification of “obscene” images. It also showed that the classificatory semiosis of obscenity required its constant distinction from other non-obscene images. The exceptions—that is, Tumblr’s “What is permitted?” (Figure 9)—were critical to this system. They were its “licensed transgressions ... partly rule breaking, partly rule conserving” (Taussig 2015:167). The image I had appealed, though clearly fitting into the list of obscene content (female-presenting nipples, as well as nudity), was permissible due to its perception as *art* rather than pornography.

What is "adult content?"

Adult content primarily includes photos, videos, or GIFs that show real-life human genitals or female-presenting nipples, and any content—including photos, videos, GIFs and illustrations—that depicts sex acts.

What is permitted?

Examples of exceptions that are **permitted** are exposed female-presenting nipples in connection with breastfeeding, birth or after-birth moments, and health-related situations, such as post-mastectomy or gender confirmation surgery. Written content such as erotica, nudity related to political or newsworthy speech, and nudity found in art, such as sculptures and illustrations, are also stuff that can be freely posted on Tumblr.

Figure 9. The complete Terms of Service guidelines, including list of exceptions

Accordingly, the negotiation of obscenity was a matter of distinguishing between *kinds* of performative effects; nudity and sexual imagery are only not obscene when they are “related to political or newsworthy speech,” maternity or “health-related” or “found in art” (Figure 9); that is, when they are perceived to be serving a purpose other than their main function—as images of sexuality for sexualities’ sake, for sexual use, for reference to sex

—images are redeemable. This kind of discursive filtration is not uncommon, as Mazzarella (2013:195) writes: an image-object is *only* obscene “if its intensity cannot be referred to the soothing, moral balm of a higher social purpose” (such as art or science or political activism). Tumblr’s Terms of Service guidelines preempts this sorting of performative effects by simultaneously enunciating both the rule and its exceptions. Thus, in addition to enforcing a particular kind of indefeasibility through the algorithmic sorting process, Tumblr’s “What is permitted?” exerts its own authority to specify the conditions of defeasibility, an act Asif Agha (2017:335) has called “describing your own felicity conditions.” In his work on “money talk,” Agha discusses how early bills in the United States, always at risk of being found inauthentic, flagged those particular felicity conditions through inscriptions that explicitly stated that the bills were *not* counterfeit.

However, in the case of Tumblr’s adult content, the correspondent signaling of defeasibility —the listing of possibilities through which an obscene image’s performative effects would not be realized— was not devised by broad divisions between authentic/inauthentic or obscene/non-obscene, but rather by more specified categories. That is, the algorithm did not just sort what was obscene (or not), it sorted what was *art* (or not). Significantly, the proximity of these two semiotic categories was grafted from earlier judicial approaches to obscenity. Tumblr’s “What is permitted?” closely mimicked the third prong of the Miller Test (the obscenity law on the books today) which asks “whether the work, taken as a whole, lacks serious literary, artistic, political, or scientific value.” This implicit citation of U.S. law not only dredged up the semiotic relatedness of art and obscenity but also bolstered the notion that Tumblr’s actions were indeed in anticipation of the law.

But despite Tumblr’s absorption of a legal pragmatics to obscenity, what sparked outrage among Tumblr users was that such legal logics were enacted by an algorithmic assemblage. Even more appalling than the idea that knowledge of the obscene could be automated was that knowledge of artistic or aesthetic value could be equally automated. Many users took issue with the algorithm as a discerning subject by pointing up the algorithm’s attempts and failures to sort art. This was evidenced by the numerous posts that featured erroneously censored or flagged images of classical art as a mode of critiquing Tumblr’s censorial tactics (Figure 10).

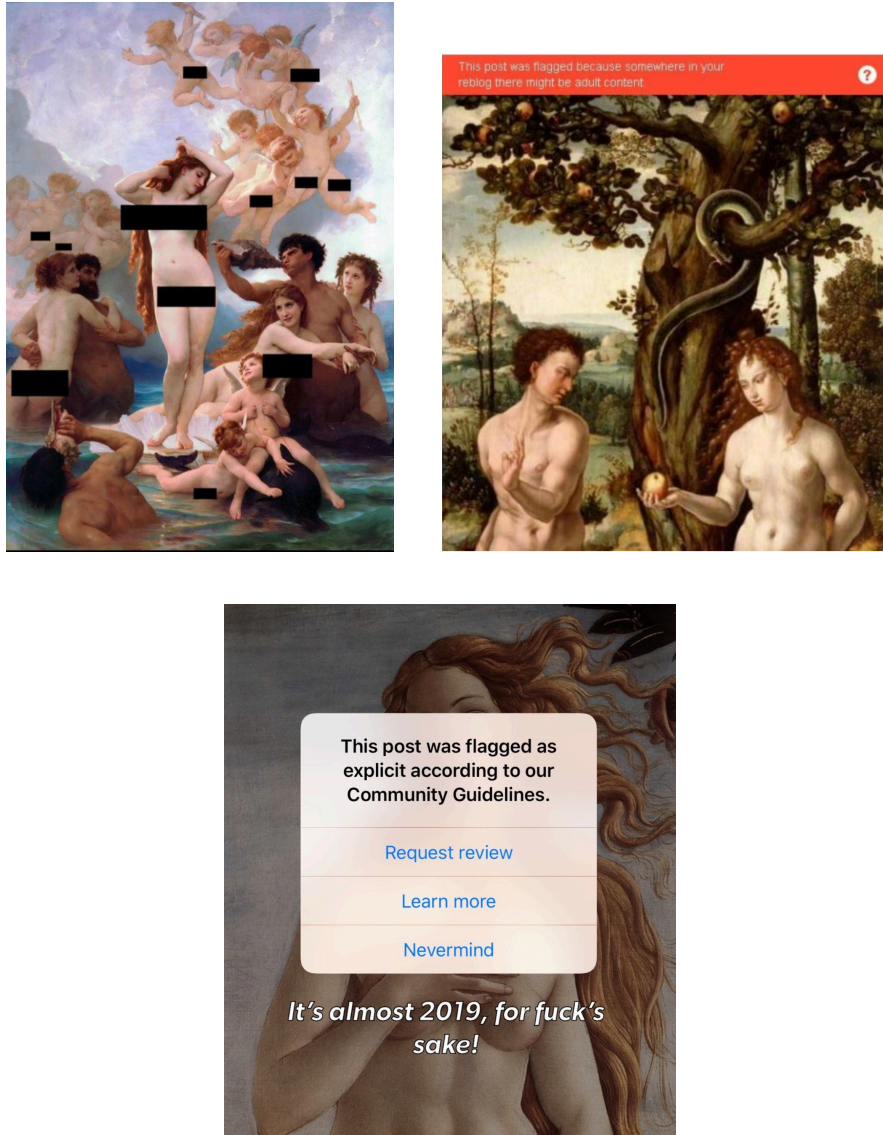


Figure 10. Licensed transgressions: memes highlight the algorithm's incorrect flagging of art.

In the memes in Figure 10, users question the authority of the algorithm by underscoring its most egregious failure: the misrecognition of art for obscenity. Within the algorithm's system of classification, art had become a semiotic correlate of obscenity; in fact, in a reversal of Douglas's description, where the obscene is a by-product of a semiotic system from which it is continuously expelled, in this case art is a secondary by-product of the categorization of the obscene: expelled and then permitted back again. In a symbiotic relation of parasite and host, art was the parasite—the exception—that which is exterior to obscenity and “yet, paradoxically of it” (Nakassis 2013:4). Tumblr's “What is Permitted” recognized this relationship but gave little to no room for the multiple claims and understandings that its users held as to what counted as exceptional.

As such, this became the vector of distinction to which so much of the dispute around the purge clung. For as we saw above, to be redeemed users had to make a case for an

image's "higher social purpose," specifically one accounted for in Tumblr's Community Guidelines. This is not to say that many of the images put through the appeal process were not already considered art by the users who posted them, but the specificity the process required in labeling these images allowed new meanings to arise. This process circles back to the very start of this paper, and to Georges Bataille's (1989:67) statement in the epigraph that "Prohibition gives to what it proscribes a meaning that in itself the prohibited action never had." This statement is three-fold: first, and more obviously, prohibition entails the meaning of *being prohibited*; second, in the process of sorting what counts as prohibited in kind, other categories are sifted through as exemptions; and finally, in the relation of the first two may emerge new categorizations and indexical meanings (such as the commentary of enraged users, and the many meme-forms featured throughout this paper). This is a common consequence of the sieve; as Kockelman (2013:36) writes: "in sieving for a feature, the substances sieved may be affected by the sieving and thereby come to take on features they did not originally have—in particular, features that allow such substances to slip through sieves." Wrung through the system of Tumblr's algorithm and an appeal process that required negotiation between users and Tumblr employees (real or automated), an image came out the other side with meanings and purposes it did not necessarily have before.

Much of the friction between Tumblr and its users, arose from the strange temporality of Tumblr's censorial approach; in particular, from the way in which Tumblr's censorship was both post-hoc and anticipatory. The algorithm made a decision that both inscribed a meaning to a past image-act, as well as foreclosed the possibility for future meanings (and performative effects) to arise. Rather than a linear temporality, this semiotic motion can be viewed through the Freudian concept of "deferred action" (*Nachträglich*, or "afterwardsness," as it is also translated and used interchangeably here). Deferred action denotes the process through which memories of the past are given new meanings in the present, such that the very material of the memory is subject to "a re-arrangement in accordance with fresh circumstances—to a re-transcription" (Laplanche and Pontalis 1973:111). Rather than considering meaning (or effects) as inhering in past memories (of objects, images, actions) "it is this revision which invests them with significance and even with efficacy or pathogenic force" (ibid.). So many images removed in "the purge" were imbued with such a pathogenic force—that of obscenity, of transgression, or even, as potential evidence of criminal acts—an efficacy that no one, neither the users nor Tumblr itself had felt previously.⁵ Yet at the same time, from the point of view of the users, the possibility of future deferred action—including re-inscriptions of an image's artistic or political value—was foreclosed, and its forced indefeasibility effectuated, by the mechanics of the censorial assemblage. Unlike a legal case that could presumably move up the chain of judicial authority, after December 17th, the semiotic motion of the obscene image-act was arrested by the impenetrability of Tumblr's censorial matrix, blocking

further negotiation that might re-define it again, including releasing it from the category of the obscene.

(Un)Licensing Transgressions

If we are to use obscenity to think towards a semiotics of the image, the temporal motion of various pragmatics around obscenity suggests that the constant forward and backward swell of meaning-making processes must always be held in view. Tumblr's decision for removal was forward-looking; they acted proleptically (in anticipation of potential legal consequences) and as such were forced to consider the future performative potentials of image-acts (to which they could be held accountable). At the same, the effects of prohibition and exception were constituted by an *afterwardsness* (the recursive, metaleptic, redefinition of past image-acts in the present) that, through the so-experienced mechanics and rigidity of a non-human agent, also halted the possibilities of further deferred action by others, at least within the confines of the Tumblrverse. This includes the licensed transgressions, those exceptions named by Tumblr in its description of its own felicity conditions as performative effects that would prevent an image from being regarded as obscene despite what the image might contain denotationally. In practice, many of these licensed transgressions were only "allowed" after a second pass at sorting via content appeal; that is, the arbitration process had its own effect of licensing or (un)licensing transgressions. Other meanings and understandings persisted through the appeal process—of the image as art, as meme, as non-obscene expression—but were often unable to penetrate the "black box" of Tumblr's censorial assemblage, leaving the *afterwardsness* of prohibition stuck in place, its temporal tide arrested.

Moreover, Tumblr users' frustration and anger often stemmed from the fact that they continued to see themselves as the principal behind the image-act of posting—it was their blog, after all—and had never surrendered that responsibility to Tumblr. Though Tumblr tried to enact a semiotic situation of forced indefeasibility even within its appeal process, users continued to resist. Understanding why this happened is a matter not of "ownership of the image" but in the question of the responsibility for the image-act (as expressed through Goffman's notion of "principal" and the participation framework wherein the image-act occurs). As we can see, the participation framework itself is multiple. While users might see the images they post as "for" the community with whom they share and connect, Tumblr feels the pressure of a wider community as well as a broader spectrum of "overhearers" (Goffman 1981). Such overhearers include both the possibility of a surreptitious audience capable of legal enforcement, as well as a mass-mediated audience—via news articles and other social media sites—on which their brand status, and business, depends.

These participation frameworks also differ in the social relations that are believed to underlie them, specifically the social personae occupying the roles of discerning and susceptible subjects. While Tumblr citationally grafts the language of obscenity law, attempting to usurp the authority of the discerning subject in the process, users do not trust or believe that Tumblr's censorial assemblage has the capacity of expertise required to occupy such a role. At the same time, if Tumblr was presumably acting in anticipation of legal consequences—as the result of SESTA-FOSTA—then this was also with a view that included potential sex trafficking victims in the role of susceptible subjects, an inclusion seen to justify their aggressive new approach to censorship. Rather than a break from judicial expertise, Tumblr's actions sought a continuity with legal standards, though in opaque terms. However, the opacity of this relationship, further obscured by the screen of the algorithmic assemblage, was what allowed Tumblr users to be at odds with this understanding, taking offense to the idea that *they* (as members of the community) were potentially susceptible subjects, while pointing up *Tumblr's* failure in actually removing illegal content.

As the production format shifts between the two views—either separating or collapsing the roles of animator and principal—the question of who is “committed” to the image-act becomes complicated by the ever-presence of polymorphic participation frameworks. As users questioned the forced indefeasibility of image-acts that Tumblr imposed, the question arose: who has the power to imbue the past (and the potential future) with new meaning? And how is it undone? It was precisely this power that the algorithmic assemblage worked to wrestle out of users' hands—taking hold of the image-acts' means of production. But that power was limited to the confines of Tumblr itself. As users found the discrepancies between Tumblr and their own pragmatics of sorting obscenity irreconcilable, a mass exodus of users from Tumblr began. In the months following the purge, Tumblr lost over a third of its web traffic—a number they have not recouped since.

Afterword/wards

Whether conducted through earlier juridical “tests” of obscenity or through an algorithmic agent, it is clear that censorship is an unfolding, and highly contested process rather than a uniform and ineluctable one. Moreover, what is revealed in looking closely at the mechanism Tumblr employed for sorting obscenity—an algorithmic assemblage—is that obscenity, at least or especially in reference to the image, is not isolatable as a stable object. In turn, as online platforms take on the responsibility of legal structures, they run up against the very same obstacles of those frameworks, namely the difficulty of defining obscenity as a single set of effects or qualities. The negotiations between users and Tumblr make clear that contestations of obscenity may begin with differing definitions of what constitutes a “dirty picture”; but they go much further, into rejections of the approach to categorization itself, with its questions of who is envisioned as a discerning subject, a

susceptible subject, as well as what the perimeter of a community is within which these roles appear. That there is no object in obscenity, nor even a single system, or sieve, for distilling the qualities of dirtiness, but instead many indefinite pragmatics of discernment, has been demonstrated by this paper. Further, the temporal motions that constitute such pragmatics are not as straightforward as they are meant to appear. Attending to the *deferred action* present in mechanisms of censorship asks that we don't search the past for the original meaning of an image-act, but ask how and why obscenity appears in a particular moment and most of all *in response to what*. Thus, *afterwardsness* dislocates obscenity as a naturalized effect of a prior image-act, instead viewing it as both an anticipatory proscription and a retroactive re-inscription—both often based on events of the present or the anticipated near-future, as with shifts in the legal code or frameworks of liability.

Perhaps this is why the purge sparked such fury. Though users suspected SESTA-FOSTA as the cause for such drastic change, Tumblr denied this explanation, instead attempting to enforce inscriptions of obscenity on images that had long been circulating on the platform. In many cases, the images censored or removed on December 17th, 2018 had been on the site for years. Not only did Tumblr withhold their reasoning, but their sweep of content required a drastic shift in power: the power *to know* what counts as obscenity and, more importantly, the power *to imbue* an image-act with new meaning. Though the algorithm supposedly relied on users for *knowing* and identifying obscenity—to “make these kinds of mistakes happen less often” (Figure 8d)—as its capacity for knowing was questioned, its power to *imbue* was rigidly enforced. Through the black box effect of the algorithmic assemblage, Tumblr's censorial regime was close to immovable—it was this impenetrability that so frustrated users. In closing the discursive pathways for negotiation, Tumblr's content ban disavowed the performativity of images through which obscenity emerges, a performativity that requires constant citation and re-citation that is “uninsured and unanticipated, persistently and interminably susceptible” (Butler and Athanasiou 2013:140) to points of contestation. The failure of the purge—at least from users' point of view—was its attempt to freeze an image in place, restricting new meanings and effects from arising. In approaching obscenity as an identifiable object, Tumblr failed to see what users saw: that obscenity is only made visible as a pragmatics, as an approach or sensibility towards an image-act, one that always contains the possibility for another's approach, another's pragmatic re-transcription. It is only through the temporal swell of this semiotic tide—in the dynamism of delineation and censorship and its dialectic return, of licensing and un-licensing transgression—that obscenity surfaces, only to withdraw, and re-surface again.

Endnotes

1. Beyond “categorization,” through which tokens are recruited to a type such as dirt or matter, *categoriality* suggests the conditions of possibility for a category’s becoming, including metapragmatic descriptions of what constitutes a category (Silverstein 2004). ↩

2. In a letter of opposition to Congress, the Freedom Network, a coalition of anti-sex trafficking advocates, made this key point: “FOSTA expands the criminalization of *consensual* commercial sex workers under the guise of addressing sex trafficking. This squanders limited federal resources and puts sex workers at risk of prosecution for the very strategies that keep them safe” (Freedom Network USA 2018). ↩

3. At the same time, taboos never have a universal social domain—their performativity is rigid only in so far as they are construed as taboo within a given space. Likewise, verbal and image taboos alike have spaces where they may get a legal or censorial response and others where they will not. The chaos of “the purge” is thus in part the result of a rapid and unexpected shift in Tumblr’s topographic landscape as a platform of images (like a radio station suddenly conforming to FCC regulations). ↩

4. From “A Better, More Positive Tumblr”: <https://staff.tumblr.com/post/180758987165/a-better-more-positive-tumblr>. ↩

5. It should be noted the Freud’s concept of deferred action was specifically invoked for memories of a sexual nature. In *Project for a Scientific Psychology* (1950[1895]), Freud gives the example of sexual events in early childhood, which have no sexual significance at the time of the occurrence but take on meaning as the individual matures, coloring events in the present day (often through displeasure). I’m not suggesting here that the images in question traverse a similar journey of maturation (or traumatization), but it is perhaps no accident that the axes of distinction between obscene/non-obscene submitted to this process of deferred action lie on similar vectors of excitation/repression (censorship), pleasure/guilt, et cetera.. ↩

Legal Cases Cited

Jacobellis v. Ohio. 1964. U.S. Supreme Court.

Miller v. California. 1972. U.S. Supreme Court.

Roth v. United States. 1957. U.S. Supreme Court.

References

Agha, Asif. 2011. Commodity Registers. *Journal of Linguistic Anthropology* 21(1):22–53.

Agha, Asif. 2017. Money Talk and Conduct from Cowries to Bitcoin. *Signs and Society* 5(2):293–355.

Austin, J. L. 1962. *How to Do Things with Words*. Oxford: Clarendon.

Bakewell, Liza. 1998. Image Acts. *American Anthropologist* 100(3):22–32.

Barocas, Solon, Sophie Hood, and Malte Ziewitz. 2013. Governing Algorithms: A Provocation Piece. (March 29, 2013). Available at SSRN: <https://ssrn.com/abstract=2245322> or <http://dx.doi.org/10.2139/ssrn.2245322>.

Bataille, Georges. 1989. *The Tears of Eros*. San Francisco: City Lights Press.

Bateson, Gregory. 1972. *Steps toward an Ecology of Mind*. New York: Ballantine.

Beer, David. 2017. The Social Power of Algorithms. *Information, Communication & Society*, 20(1):1–13.

Boellstorff, Tom. 2008. *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human*. Princeton, NJ: Princeton University Press.

Bowker, Geoffrey C. and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: The MIT Press.

Boyer, Dominic. 2008. Thinking through the Anthropology of Experts. *Anthropology in Action* 15(2):38–46.

Butler, Judith and Athena Athansiou. 2013. *Dispossession: The Performative in the Political*. Cambridge, MA: Polity Press.

Captain, Sean. 2019. Seven Weeks after NSFW Ban, Tumblr Still Bulges with Porn. *Fast Company*. <https://www.fastcompany.com/90304153/seven-weeks-after-nsfw-ban-tumblr-still-bulges-with-porn>

Carr, E. Summerson. 2010. Enactments of Expertise. *Annual Review of Anthropology* 39(1):17–32.

Cole, Samantha. 2018. Craigslist Just Nuked Its Personal Ads Section Because of a Trafficking Bill. *Vice.com*, <https://www.vice.com/en/article/wj75ab/craigslist-personal-ads-sesta-fosta>

Crawford, Kate and Tarleton Gillespie. 2014. What Is a Flag for? Social Media Reporting Tools and the Vocabulary of Complaint. *New Media & Society* 18(3):410–28.

Diakopoulos Nicholas. 2013. Algorithmic Accountability Reporting: On the Investigation of Black Boxes. *A Tow/Knight Brief*. New York: Columbia Journalism School, Tow Center for Digital Journalism.

Douglas, Mary. 1984[1966]. *Purity and Danger: An Analysis of Concepts of Pollution and Taboo*. London and New York: Routledge.

Fleming, Luke and Michael Lempert. 2011. Introduction: Beyond Bad Words. *Anthropological Quarterly* 84(1):5–13.

Fleming, Luke and Michael Lempert. 2014. Poetics and Performativity. In N. J. Enfield, P. Kockelman, and J. Sidnell, eds. *The Cambridge Handbook of Linguistic Anthropology*, pp. 485–515. Cambridge: Cambridge University Press.

Freedom Network USA. 2018. FOSTA Does Not Protect Communities at Risk of Sex Trafficking. <https://freedomnetworkusa.org/>

Freud, Sigmund. 1950. Project for a Scientific Psychology (1895). In *The Standard Edition of the Complete Psychological Works of Sigmund Freud Volume I (1886–1899): Pre-Psycho-Analytic Publications and Unpublished Drafts*, edited and translated by James Strachey, pp. 281–397. London: The Hogarth Press and the Institute of Psycho-Analysis.

Gillespie, Tarleton. 2014. The Relevance of Algorithms. In T. Gillespie, P. Boczkowski, and K. Foot, eds. *Media Technologies: Essays on Communication, Materiality, and Society*, pp. 167–94. Cambridge, MA: MIT Press.

Gillespie, Tarleton. 2015. Platforms Intervene. *Social Media + Society* 1(1). <https://doi.org/10.1177/2056305115580479>

Gillespie, Tarleton. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, CT: Yale University Press.

Goffman, Erving. 1981. *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Goodwin, Charles. 2017. *Co-Operative Action*. New York: Cambridge University Press.

Inoue, Miyako. 2006. *Vicarious Language: Gender and Linguistic Modernity in Japan*. Berkeley: University of California Press.

- Irvine, Judith T. 2011. Leaky Registers and Eight-Hundred-Pound Gorillas. *Anthropological Quarterly* 84(1):15–39.
- Jones, Graham. 2014. Secrecy. *Annual Review of Anthropology* 43:53–69.
- Kitchin, Rob. 2017. Thinking Critically about and Researching Algorithms. *Information, Communication and Society*, 20(1):14–29.
- Kockelman, Paul. 2010. Enemies, Parasites, and Noise: How to Take Up Residence in a System without Becoming a Term in It. *Journal of Linguistic Anthropology* 20(2):406–21.
- Kockelman, Paul. 2011. Biosemiosis, Technocognition, and Sociogenesis: Selection and Significance in a Multiverse of Sieving and Serendipity. *Current Anthropology* 52(5):711–39.
- Kockelman, Paul. 2013. The Anthropology of an Equation: Sieves, Spam filters, Agentive Algorithms, and Ontologies of Transformation. *HAU: Journal of Ethnographic Theory* 3(3):33–61.
- Laplanche J. and J. B. Pontalis. 1973. *The Language of Psycho-Analysis*, translated by Donald Nicholson-Smith, pp. 111–13. London: Hogarth Press and the Institute of Psycho-Analysis.
- Lingel, Jessa. 2020. *An Internet for the People: The Politics and Promise of Craigslist*. Princeton, NJ: Princeton University Press.
- Martineau, Paris. 2018. Tumblr's Porn Ban Reveals Who Controls What We See Online. *Wired.com*, <https://www.wired.com/story/tumblrs-porn-ban-reveals-controls-we-see-online/>.
- Mazzarella, William. 2013. *Censorium: Cinema and the Open Edge of Mass Publicity*. Durham, NC: Duke University Press.
- Nakassis, Constantine V. 2013. Para-s/cite, Part I. The Parasite. *Semiotic Review* 1. <https://doi.org/10.71743/gqdtqg28>.
- Nakassis, Constantine V. 2019. Poetics of Praise and Image-Texts of Cinematic Encompassment. *Journal of Linguistic Anthropology* 29(1):69–94.

- Noble, Safiya. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.
- Padgett, Esra. 2018. A Eulogy to Tumblr, One of the Last Havens for NSFW Freedom. *Playboy.com*, <https://www.playboy.com/read/a-eulogy-to-tumblr-one-of-the-last-havens-for-nsfw-freedom>.
- Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.
- Porter, Jon. 2018. Tumblr Was Removed from Apple's App Store Over Child Pornography Issues. *The Verge*, <https://www.theverge.com/2018/11/20/18104366/tumblr-ios-app-child-pornography-removed-from-app-store>.
- Putnam, Hilary. 1975. *Mind, Language and Reality*, vol. 2. London: Cambridge University Press.
- Rasmus, Ryen. 2011. The Auto-Authentication of the Page: Purely Written Speech and the Doctrine of Obscenity. *William & Mary Bill of Rights Journal* 20(1):253–85.
- Reyes, Angela. 2017. Ontology of Fake: Discerning the Philippine Elite. *Signs and Society* 5(S1):S100–27.
- Romano, Aja. 2018. A New Law Intended to Curb Sex Trafficking Threatens the Future of the Internet as We Know It. *Vox.com*, <https://www.vox.com/culture/2018/4/13/17172762/fosta-sesta-backpage-230-internet-freedom>.
- Seaver, Nick. 2013. Knowing Algorithms. Paper presented at Media in Transition 8, Cambridge, MA. Available at: <http://nickseaver.net/papers/seaverMIT8.pdf>.
- Seaver, Nick. 2017. Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems. *Big Data & Society* 4(2):1–12.
- Silverstein, Michael. 2004. "Cultural" Concepts and the Language-Culture Nexus. *Current Anthropology* 45(5):621–52.
- Simmel, George. 1906. The Sociology of Secrecy and of Secret Societies. *The American Journal of Sociology* 11(4):441–98.

Striphas, Ted. 2015. Algorithmic Culture. *European Journal of Cultural Studies* 18(4–5):395–412.

Tambiah, Stanley Jeyaraja. 1985[1969]. Animals Are Good to Think and Good to Prohibit. In *Culture, Thought, and Social Action: An Anthropological Perspective*, pp. 169–211. Cambridge, MA: Harvard University Press.

Taussig, Michael. 2015. The Obscene in Everyday Life. In *The Corn Wolf*, pp. 163–71. Chicago: University of Chicago Press.

Tripp, Heidi. 2019. All Sex Workers Deserve Protection: How FOSTA/SESTA Overlooks Consensual Sex Workers in an Attempt to Protect Sex Trafficking Victims. *Dickinson Law Review* 124(1): 219–46

Wood, Janice Ruth. 2008. *The Struggle for Free Speech in the United States, 1872–1915: Edward Bliss Foote, Edward Bond Foote, and Anti-Comstock Operations*. New York: Routledge.

© Copyright 2021 Esra Soraya Padgett

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.