

Do Not Use This Hashtag: Fat Acceptance (Mis)information and Discursive Boundary-Work as Content Moderation on Instagram

MELISSA ZIMDARS*
Merrimack College, USA

The discussions concerning the use of #fatacceptance on Instagram are important for observing both the limitations of official content moderation on social media platforms and an emerging group of users who engage in unofficial content moderation through discursive boundary-work. Discursive boundary-work takes the form of users repeatedly centering “true” or “correct” information, referencing experts and studies, and claiming the “other” is misunderstanding and misusing information or spreading (mis)information. However, these boundary-work as content moderation practices are not just about delineating between truth and falsehood; they are about users building and reinforcing norms of use, groups, relationships, and identities as well as the digital space for acceptance. Ultimately, I argue that our understandings of both content moderation and (mis)information need to be expanded and analyzed in consideration of their interplay with users’ perspectives or beliefs, collective identities, relationships with other users and groups, and larger sociocultural, political, and activist communities.

Keywords: content moderation, boundary-work, hashtag, fat acceptance, obesity epidemic, misinformation, Instagram

A study published in *The New England Journal of Medicine* on December 19, 2019 (Ward et al., 2019), quickly circulated through news media. Headlines read, “Half the population will be obese by 2030” (Park, 2019) and “Obesity in America: A Public Health Crisis” (Newman, 2019). Posts about the study appeared on Instagram with the hashtag #fatacceptance (e.g., Figure 1) either in rejection of the claims circulated by these news articles or as a way to blame fat acceptance for increasing obesity rates. The #fatacceptance hashtag is typically used to promote fat acceptance, but as is evident by this example, it is also used to argue that #fatacceptance “promotes obesity.” Some users of #fatacceptance ignore these latter posts contradicting the meaning and purpose of the fat acceptance movement while others choose to moderate how the #fatacceptance hashtag is used to spread awareness and create solidarity around fat acceptance on Instagram.

Melissa Zimdars: zimdarsm@merrimack.edu
Date submitted: 2021-12-03

Copyright © 2023 (Melissa Zimdars). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at <http://ijoc.org>.

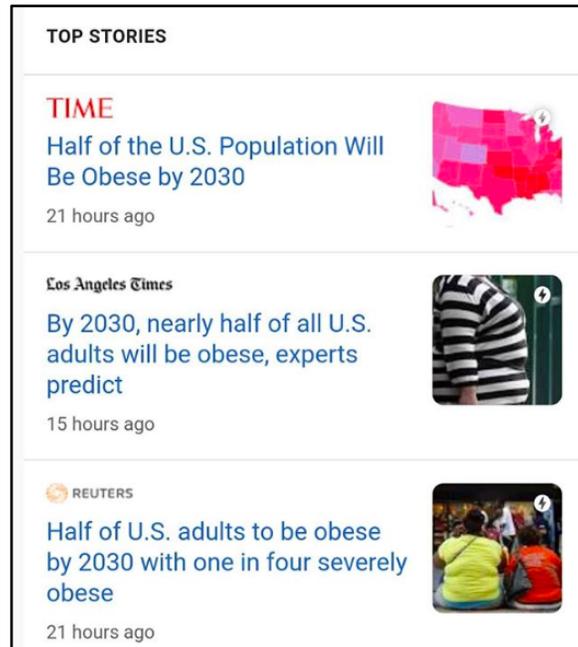


Figure 1. An image shared by an Instagram user in December 2019. The description of this Instagram image reads, "It's not an epidemic! {puzzled face emojis} anyways, I hope that we can change that, to live long and happy lives full of health! Anything I can do to help, don't be afraid to ask! #haes #fatacceptance" (personal communication, 2019).

The discussions of #fatacceptance between Instagram users, who are trying to eliminate body shame, resist fat stigma, build community, and draw attention to fat discrimination, and other users, who are trying to reach, educate, shame, antagonize, or troll proponents of fat acceptance, engender an important social media site for understanding fat embodiment and the contradictory tensions of our current sociocultural health context. Because discourses of fat acceptance challenge dominant assumptions about and research on fatness and health in the context of the obesity epidemic, this inevitably leads to debates about research on obesity and what kinds of information are "true" or "false." Moreover, the discussions about the use of #fatacceptance on Instagram are important for observing the limitations of official content moderation for addressing (mis)information, which can be defined as "false information shared by someone who believes it to be true" (Wardle, 2020, p. 71), as well as observing users engaging in unofficial content moderation through discursive boundary-work.

In the case of #fatacceptance on Instagram, discursive boundary-work takes the form of users repeatedly centering what they believe to be "true" or "correct" information, referencing experts and studies, and claiming the "other" is misunderstanding and misunderstanding information and/or spreading (mis)information. However, these boundary-work as content moderation practices are not just about the veracity of information; they are about establishing norms of hashtag use as well as building and reinforcing groups, relationships, identities, and digital spaces for fat acceptance. Ultimately, I argue that our

understandings of both content moderation and (mis)information¹ need to be considered and analyzed in consideration of their interplay with users' beliefs, identities, and relationships with other users and groups; the norms and possibilities for practice with hashtags and platforms; and larger sociocultural, political, and activist contexts and communities.

The Context of Obesity Epidemic and the Fat Acceptance Movement

Before looking at #fatacceptance in relation to (mis)information and boundary-work, I will briefly discuss the competing contexts of these Instagram discussions. Discourses of the obesity epidemic reflect dominant medical and public health research showing, according to the World Health Organization (WHO, 2021), that over 60% of people in the United States can be categorized as overweight or obese. In this research, obesity is linked to numerous health issues, such as diabetes, cardiovascular disease, and cancer (see Field et al., 2001). Across this literature, obesity is primarily framed as a matter of individual or personal responsibility (see Dressler, 2013; Font, Hernandez-Quevedo, McDonald, Ted, & Jayachandran, 2012) and as a health or medical problem. Discourses of the "obesity epidemic" are the dominant way of discussing fatness (Zimdars, 2019); thus they tend to set the terms of the debate, or undergird discussions of what information is considered to be true or false regarding the relationship between body size and health.

However, multiple areas of scholarship push back against the science underpinning obesity research (Campos, 2004; Oliver, 2006). Some examine the cultural and social constructions of fatness as a problem, and how that influences medical and health research (Boreo, 2012; Brewis, 2010; Metzl & Kirkland, 2010) or positions fat individuals as "needing" to change their "shameful" or "undesirable" bodies (Farrell, 2011; Kwan & Graves, 2013; Saguy, 2012). Others argue that discourses of the obesity epidemic engender or exacerbate fat stigma and discrimination, which is not conducive to collectively achieving positive health outcomes (Bacon & Aphramor, 2018). Ultimately, fat acceptance is an emerging alternative discourse that negotiates and challenges more established health and medical research.

In addition to academic research offering alternative views on fatness, the fat acceptance movement started in the 1960s to fight for the rights of fat individuals. Today, fat acceptance consists of multiple frameworks and strategies to achieve fat liberation and confront anti-fat phobia (Williams, 2017). A lot of fat acceptance activism takes place on digital platforms (Gurrieri & Cherrier, 2013; Meleo-Erwin, 2011), which generally serve as places of community support, "consciousness-raising," and activist organizing (Mendes, Ringrose, & Keller, 2019, p. 16). These platforms also become places where "clusters of people can coalesce, respond, and mobilize to amplify messages beyond individuals and specific communities" (Kuo, 2018, p. 495), or play a role in challenging dominant discourses (Antunovic & Linden, 2015; Clark, 2016; Thrift, 2014). While some users of #fatacceptance on Instagram post about their individual journeys of body acceptance, others use the hashtag for political and activist

¹ Throughout this article, I use "(mis)information" to capture the complexity and messiness of information circulated and debated. In the case of fat acceptance, whether information is true or false, accepted or contested, is as much about one's relationship to that information—and what discourses are articulated to that information—as the information itself.

purposes, such as countering discourses of the obesity epidemic and addressing systemic changes needed to prevent fat discrimination and stigma.

Earnest Versus Ambivalent Users

While hashtags can bring like-minded people into conversation, they can also be used by trolls or hijackers. For example, an activist #ShoutYourAbortion campaign on Twitter was eventually hijacked by anti-abortion users who circulated stigmatizing information (Kosenko, Winderman, & Pugh, 2019). Because it is difficult to tell whether users are hijacking #fatacceptance or being earnest in their intent and motivations to post, it is more accurate to consider these users as ambivalent or “simultaneously antagonistic and social, creative and disruptive, humorous and barbed. . .” (Phillips & Milner, 2017, p. 10). For this research, the intent of the users is less important than the dialogue about fatness emerging from the posts and comments. The ways the dialogue demonstrates discursive boundary-work as unofficial content moderation and the complexities of determining truth and falsehood are equally more important than the users’ intent. I cannot know the “textual origins, creators’ intent, immutable meaning” of those who are posting and commenting from analyzing their text; thus I focus on how the posts and comments “illuminate and often complicate broader cultural logics” (Phillips & Milner, 2017, p. 203). Throughout this article, I will use the term “earnest” to describe users posting in support of #fatacceptance and the term “ambivalent” to describe users posting criticism of #fatacceptance.

I have followed the #fatacceptance hashtag on Instagram since the summer of 2019 and have observed daily #fatacceptance public posts and subsequent comment-based discussions. The posts analyzed in this article are a few of the dozens of public posts and discussions that I saved because users specifically talk about “false” information, “facts,” and interpretations of fat acceptance and obesity epidemic research. I removed all identifying information from the posts and comment transcripts, and I do not analyze the profiles or past activity of Instagram users discussed here to try to maintain user privacy (Fiesler & Proferes, 2018, p.10). Instead, I analyze these images and conversations as texts and as a personal Internet archive (see Lomborg, 2013). Ultimately, this study embraces a “sociotechnical theory of media effects” to understand “cultural practice and group identity, and media, as both patterns of messages and sets of technological affordances which constrain or enable certain meanings and actions” (Marwick, 2018, p. 487). I consider the actors (Instagram users), the messages they share or engage with (Instagram posts and comments), and the affordances of Instagram, specifically how it enables a “possibility for action” (Evans, Pearce, Vitak, & Treem, 2017, p. 73). Instagram encourages user connection and discussion through hashtags that bring both similar and disparate individuals together, and their lack of official content moderation practices related to fat acceptance and the obesity epidemic keep this content in users’ feeds and creates possibilities for engagement with that content and other users.

(Mis)information and Instagram

A significant amount of research about (mis)information looks at the veracity of information, or relies on binary distinctions between truth and falsehood, fake or real (Altay, Araujo, & Mercier, 2020; Bago, Rand, & Pennycook, 2020; Fazio, 2020; Ghanem, Rosso, & Rangel, 2020; Luo, Hancock, &

Markowitz, 2020; Vosoughi, Roy, & Aral, 2018). However, contrasting ways of understanding fatness, as articulated to discourses of the obesity epidemic or fat acceptance, defy categorizations of information as true versus false. Instagram users often cite studies, facts, and specific researchers within each discursive cluster as the definitive evidence supporting one's point of view. Studies, facts, and specific researchers cited in contradiction to one's point of view are typically ignored or rejected. The contestation of multiple different truths and facts, as well as the role of discursive clusters in shaping online discussions on Instagram, can make identifying and addressing (mis)information more difficult and far messier.

Like most social media platforms, Instagram outlines its Community Guidelines regarding acceptable and unacceptable content and posting practices, such as prohibiting posts that directly violate the law or spread hateful speech about a protected group (Instagram, 2020). According to Facebook's Community Standards Transparency Report (Meta, 2021), in the second quarter of 2021, Instagram removed 26 million posts from its platform because they depicted sexual exploitation of children, terrorist propaganda, hate speech, self-harm promotion, bullying and harassment, or trafficking of drugs and firearms. While there is no prohibition on "false" content, Instagram works with third-party fact checkers to rate and label false content in an attempt to stymie its spread. However, Instagram is particularly rife with health misinformation (Vishakha & Mull, 2019) and questionable health accounts, such as a Medical Medium (with over 2 million followers) promoting "divinely-inspired" information about the healing benefits of drinking celery juice or the ways lettuce can allegedly cleanse the Epstein Barr virus from the liver. While these suspicious or too-good-to-be-true health claims flourish on the platform, Instagram now actively works to stop the spread of certain kinds of health misinformation (Bickert, 2019). Vaccine-related searches on Instagram lead to information from the World Health Organization (WHO) and U.S. Centers for Disease Control (CDC; Bickert, 2019). Similarly, Instagram also takes down or hides false content promoting "miraculous" weight-loss claims, particularly when they are thought to promote disordered eating (Farr, 2019).

When one particularly egregious and obvious bit of false information using #fatacceptance spread on the platform, many users made comments asking the original poster to provide research to back up the claims in the image. The image in Figure 2 shows a headless person squeezing their stomach with the caption, "DID YOU KNOW... If you have 5 inches of fat in your body you can stop a 9mm and prevent it to reach your vital organs? I'm not fat... I'M ARMORED" (personal communication, 2019). Instagram eventually covered the information with a warning that the image contained "False Information" (Figure 3) and linked to more information provided by a partner fact-checking organization (Figure 4). Again, it is generally easier for platforms such as Instagram to handle this kind of blatantly false information, but it is far more difficult to moderate posts when the shared information is opinion-based, contested, unsettled in research, contextual, or highly nuanced. It is also more difficult to mediate content when its interpretations are intertwined with users' perceived norms of hashtag use; perspectives or beliefs; collective identities; relationships with other users and groups; and larger sociocultural, political, and activist communities.

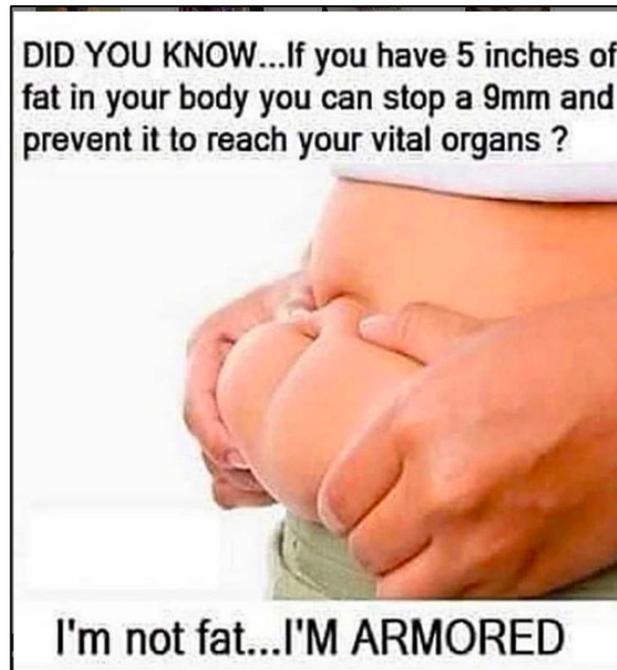


Figure 2. An image shared by an Instagram user in December 2019 (personal communication, 2019).

"Both sides" of the obesity epidemic versus fat acceptance debate view the other side as spreading misinformation. Additionally, "both sides" view the other as potentially violating Instagram's Community Guidelines. For example, some Instagram users created a Change.org petition (with 266 supporters) arguing that using #fatacceptance should trigger a content warning the same way hashtags associated with disordered eating do (Kimber, 2015). Several hashtags related to eating disorders are classified as "Self-Injury," according to Instagram. Their policy reads, "We do not allow glorification or encouragement of self-injury, including eating disorders. We do allow content that references self-injury and eating disorders if it is for the purpose of creating awareness or signposting support" (Instagram, 2020, paras. 17–18). Some ambivalent users of #fatacceptance say they believe it promotes self-injury while some earnest posters say that promoting diets and fat shame is akin to promoting self-injury. Earnest posters of #fatacceptance also argue that some ambivalent users of the hashtag are engaging in hate speech and even violence against fat individuals,² which would go against Instagram's Community Guidelines on "Hate Speech, Bullying, and Abuse."³

² Although weight is not considered a federally protected class, the state of Michigan and several cities throughout the United States ban weight discrimination.

³ Many fat acceptance activists push back on framings of fatness in relation to disease, but "obesity" is categorized as a disease by the American Medical Association. Fat acceptance activists and advocates have a mixed view on categorizing fatness in relation to disability. For some fat individuals, fatness is a disability, but for others it is not.

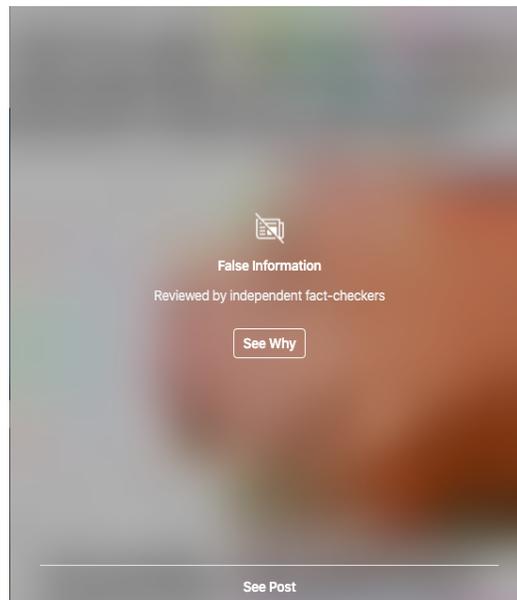


Figure 3. This image covered Figure 2 shortly after it was posted in December 2019 (personal communication, 2019).

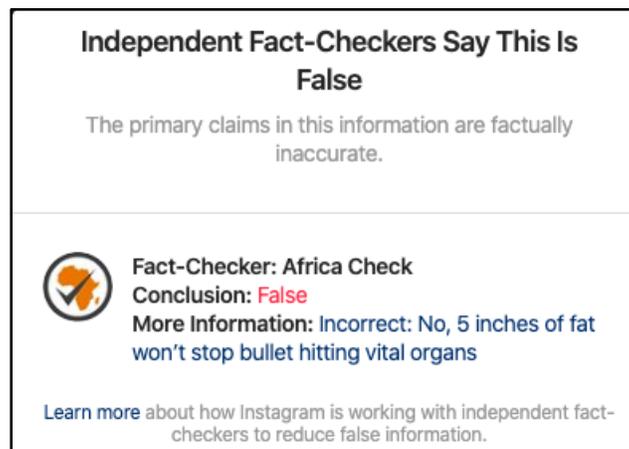


Figure 4. This image appeared when users clicked "See Why" on Figure 3 in December 2019 (personal communication, 2019).

These conceptualizations of harm and self-injury are contested and framed according to one's views on obesity or fat acceptance. These differing orientations to #fatacceptance reflect contradictory discourses characterizing health and medical information regarding the size of our bodies, which is distinct from anti-vaccination rhetoric or the notion that adipose tissue can stop a bullet because both are obviously false claims.

The difficulty of moderating something like the #fatacceptance hashtag shows the need to think about (mis)information from a "frame checking" (Cloud, 2018) perspective rather than a "fact-checking" perspective. Most of the information posted to #fatacceptance is not easily "proved" to be either "true" or "false" because the information is complicated, contradictory, and requiring of nuance and context. Instead of focusing on specific, incorrect bits of information, we need to think of (mis)information in terms of how it engages with the complexities of truth and falsehood, how "knowledge functions in service of power," and "who gets to shape facts for public sense making" (Cloud, 2018, pp. 52–74).

Unofficial Content Moderation and Discursive Boundary-Work

Social media platforms often rely on the same "solution" for (mis)information: content moderation (Gillespie, 2020). Gillespie et al. (2020) define content moderation as

the detection of, assessment of, and interventions taken on content or behavior deemed unacceptable by platforms or other information intermediaries, including the rules they impose, the human labor and technologies required, and the institutional mechanism of adjudication, enforcement, and appeal that support it. (p. 2).

According to Gillespie et al. (2020), moderation can function in many different contexts and take many different forms (p. 3). It can be done by official volunteers, such as on Reddit, by individuals employed by commercial content moderation companies, such as with Facebook, and through algorithmic automation and filtering of keywords. Content moderators, whether paid or volunteering in an official capacity, act as "digital gatekeepers" and "must balance competing demands, such as free expression, profit motives, and brand protection when deciding on the acceptability of content" (Roberts, 2016, p. 148).

Content moderation is often criticized for being unclear, subjective, and discriminatory (Nurik, 2019). Moderation is also criticized for lacking transparency, being inconsistent in the application of rules or standards, erroneous in automatic filtering processes, and ineffective at dealing with certain kinds of information (Roberts, 2016). For example, digital platforms like Tumblr, Pinterest, and Instagram attempt to block pro-eating disorder (pro-ED) content from being circulated, but users simply avoid using hashtags that are likely to get blocked or filtered and instead use other ways to express themselves (Gerrard, 2018). Gillespie (2020) argues that content moderation cannot solve every type of digital platform problem because it is sometimes systemic and can "reflect the very nature of the platform itself" (p. 330). According to Gillespie (2020), content moderation generally works for two particular issues, "deliberate and obvious violations of the platform's guidelines" (like the claim that adipose tissue can stop a bullet) and in reviewing contested violations, but not for a whole host of other issues (pp. 330–331).

In the absence of paid moderators for flagged content that obviously violates platform standards and guidelines, or official volunteer moderators who run Reddit subreddits and Facebook pages and groups, users connected through hashtags work to moderate content outside of official channels through discursive boundary-work. In the case of #fatacceptance, one group of users discursively moderates

content based on the earnest use of the hashtag in accordance with the fat acceptance movement. The other type of user discursively moderates fat acceptance content, the activist movement of fat acceptance, and the emerging alternative discourse of accepting fat embodiment while we are in an obesity epidemic. Many users across both groups appear to “believe” that the content they are circulating is factual, and both groups of users contest the factuality of the other users advancing a different understanding of fatness. Ultimately, I argue that this is another example of content moderation “by other means” (Gillespie et al., 2020, p. 6).

Boundary-work is a concept developed by Gieryn (1983), who defines it as the “attribution of selected characteristics to the institution of science (i.e., to its practitioners, methods, stock of knowledge, values and work organization) for purposes of constructing a social boundary that distinguishes some intellectual activities as ‘non-science’” (p. 782). Boundary-work is increasingly discussed in sociology, philosophy, rhetoric, and science communication (Scott, 2016, p. 61), and it is used more broadly to understand issues of knowledge formation and validity (what counts or does not count), expertise or authority, and norm-setting. For example, in looking at vaccines and claims about their links to autism, Scott (2016) shows that “rhetors on either side of the controversy disagree about what kinds of evidence should be admissible into the discussion around vaccines and autism—in other words, they dispute the boundary between evidence and nonevidence” (p. 63). Boundary-work is also used in journalism studies to discuss the legitimation of journalism and issues of credibility, objectivity, and fact checking (Graves, 2016). According to Carlson and Lewis (2019), boundary-work can take the form of “credibility contests” and tensions regarding “who gets to say what is true” between different groups (pp. 124–125).

In *Cultural Boundaries of Science*, Gieryn (1999) explores how constructions and boundaries of science move from research labs and academic journals to boardrooms, living rooms, newspapers, talk shows, and “cyberspace.” He argues that with public contestations of science or scientific ideas, which are “lay-of-the-culture representations of science,” local contingencies put forward “compelling narrations for why my science (but not theirs) is bona fide” (p. 4). To Gieryn (1999), the “domain of ‘doing science’” includes ways it is discussed or debated in mass media (p. 187), and I argue that the same is true of debates about health and fatness on *social media*. On Instagram, #fatacceptance is an example of debate over the boundaries and validities of different kinds of science in the form of research on health, obesity, and fat acceptance. By creating and repeatedly rearticulating rules around the hashtag’s use (and other posting and commenting behaviors) and attempting to moderate #fatacceptance content by *unofficial* means, users set boundaries around knowledge.

Matias (2019) uses boundary-work in analyzing how online “volunteer moderators define and redefine what it means to be a moderator in conversation with platform operators, their communities, and other moderators” (p. 2). Unlike Matias, I am not using the concept of boundary-work to determine what it means to be an *official* moderator on a social media platform. Rather, I am using boundary-work outside of official content moderation practices and policies to explore how users, or “information intermediaries” (Gillespie et al., 2020), *unofficially* moderate content in the absence of formal policies and practices. I argue that boundary-work—in the form of repeating claims, figures, and facts or trying to set standards for using a particular hashtag, among other things—is a way for earnest users of #fatacceptance to attempt

legitimizing and protecting certain types of hashtag use. It is also how users attempt to construct authority about fat acceptance or the obesity epidemic, who is deemed to be spreading problematic (mis)information, and who is “allowed” to use a hashtag.

Because Instagram does not allow earnest users to claim a particular group space or page for #fatacceptance, each post and subsequent comments can become a site of struggle for “space” or a place to share and connect #fatacceptance information, often requiring the repeated expenditure of time and energy in countering ambivalent users to meaningfully engage with other earnest users, create sense of solidarity or community among earnest users, and assert autonomy over one’s body and authority over knowing what is best, healthiest for one’s body. Gieryn (1983) says that the acquisition or expansion of symbolic or material resources can but does not need to be a part of boundary-work. Although the ability to post and engage with like-minded users or reach receptive users about #fatacceptance may not appear as a conflict over resources, the spreading awareness and organizing around fat acceptance more broadly is about fighting the discrimination against fat individuals in education, careers, healthcare, and other fundamental areas to living that can *literally* threaten the livelihoods and even lives of fat individuals (see Bacon & Aphramor, 2018; Boreo, 2012; Brewis, 2010; Campos, 2004; Farrell, 2011; Kwan & Graves, 2013; Oliver, 2006; Saguy, 2012). To achieve these goals and change these circumstances of life, fat acceptance activists, through scholarly research and the use of #fatacceptance on Instagram, among many other things, aim at expanding their ideas and challenging the authority or expertise claimed by those who reify discourses of the obesity epidemic. Viewing (unofficial) content moderation as discursive boundary-work can help us center thinking through the “ripple effects that moderation can have on the social fabric of communities...” (Gillespie et al., 2020, p. 16) and other aspects of our lives.

#Fatacceptance and Discursive Boundary-Work as Unofficial Content Moderation

Ambivalent users typically try to convince earnest users of the #fatacceptance hashtag of what they consider to be common sense: Fat acceptance is illogical or unreasonable; it promotes unhealthiness; and the movement is based on faulty information and denying scientific evidence showing that obesity is a major medical and public health problem. Ambivalent users of hashtag also accuse earnest users of #fatacceptance of circulating misinformation about health and fat embodiment. Sometimes these users circulate clearly false information like the example of Figure 2, but more often the information posted, typically in snack-size memetic forms with images and short captions, circulate accurate yet oversimplified, decontextualized, or in some way limited information—the very kind of communication that Instagram tends to enable as an image-based platform. Earnest users of #fatacceptance who respond to ambivalent posts claim in differing ways that the users do not understand the fat acceptance movement and its actual claims and political goals, alleging that they are the ones spreading misinformation and causing harm.

Earnest users who post with #fatacceptance feature predominantly women taking full-bodied selfies (usually to show that fat bodies do not need to be hidden from public view) as well as text-based, meme-like images challenging diet culture, promoting Health at Every Size and stating that all bodies deserve respect; infographics critiquing aspects of obesity epidemic research; or body positive slogans and sayings. Ambivalent

users of #fatacceptance often use the hashtag in conjunction with others that seem to convey their actual points of view, such as #fatacceptanceisbullshit, #fatacceptancekills, and #fatacceptanceisdeathacceptance. There is also another group of users who post weight-loss pictures, typically "before" and "after" photos, and who use the #fatacceptance hashtag along with other contradictory hashtags like #weightloss, #myweightlossjourney, and #loseweight.

Like other social media platforms, Instagram does not intervene in the kinds of content and hashtag disputes occurring over #fatacceptance. This lack of intervention is similar to Reddit's hands-off approach that positions itself as an "impartial" or "neutral" platform, which often allows toxicity to flourish on the platform (Massanari, 2015, p. 339). The hands-off approach of platforms can also necessitate for users to repeatedly try moderate content themselves. In this sense, the boundary-work surrounding the #fatacceptance hashtag also doubles as and functions similarly to our understanding of identity work on social media platforms, which is "often manifested through the repetitive, active, and explicit drawing of personal characteristics and collective boundaries" (boyd, 2014, as cited in Gal, 2019, p. 731). Identity work and boundary-work require this repetitive or constant maintenance by users because of continuous content posting from users (requiring continuous boundary-work as moderation) and from context collisions (Davis & Jurgenson, 2014) wherein earnest users imagine themselves speaking to one audience (other earnest users) while (unintentionally) speaking to multiple audiences (earnest and ambivalent users) via #fatacceptance.

Because discursive boundary-work as unofficial content moderation is fragmented across numerous posts using the same hashtag, each post seems to begin the debate and discussion anew, even if the same users regularly debate and discuss the #fatacceptance hashtag, and users across posts employ the same strategies again and again. In this way, discursive boundary-work as content moderation becomes a bit of a Sisyphean task wherein earnest users are unlikely to achieve moderation outcomes whether through their own repeated boundary-work or through using Instagram's official channels for reporting ambivalent users' content and behavior. This creates a situation in which official content moderation policies and practices are ineffective for addressing #fatacceptance misinformation because it does not fall under the purview of those policies and there is thus platform inaction. Since actual platform change and impactful content moderation are unlikely, these repeated user debates and repeated boundary erections around certain kinds of knowledge and behavior become important even if they are also unlikely to significantly affect the broader flows or visibility of fat acceptance misinformation on Instagram. These *acts* of discursive boundary-work are themselves impactful for earnest users as they continue the work of legitimizing and centering alternative discourses of health and the body while building and reinforcing relationships, activist communities, and digital spaces for acceptance.

Both earnest and ambivalent users of #fatacceptance use similar strategies to discursively moderate fat acceptance-related content, with one group trying to shut down or argue against the hashtag (ambivalent users) and the other trying to defend and protect its use (earnest users). Both groups moderate content by engaging in discursive boundary-work in the following ways: centering "true" or "correct" information (e.g. facts, statistics); referencing experts and studies (e.g. information circulated by researchers, medical doctors, and governmental and nongovernmental organizations); claiming the "other"

is misunderstanding or misrepresenting information; and arguing some form of “if you don’t like it” (e.g., the general content associated with the hashtag, the particular post, etc.), then “leave.”



Figure 5. An image shared by an Instagram user in August 2019 (personal communication, 2019).

For example, Figure 5 is a black and white image featuring a fat individual sitting on the beach overlaid with the words “fat acceptance is assisted suicide” (personal communication, 2019). The description accompanying Figure 5 reads, in part:

Original Poster 1: This is likely to be controversial, but I don’t really care. I think it needs to be said: The Fat Acceptance and Body Positivity movements have been trying to sell the lie that you can be healthy at any size for years now, while ignoring the fact that #obesity kills more people every year than tobacco smoking, terrorism, car crashes and starvation all over the world. . . #obesitykills #fatacceptanceisbullshit #thesouthronbear #fitness #health #fatacceptanceisassistedsuicide #obesity #healthyatany size #fatacceptance #bodypositive #bodypositivity #workforchange #bethechange you want to see #preventable death. (personal communication, 2019)

Similarly to others who spread information contradicting or undermining the fat acceptance movement, Original Poster 1 tries to persuade earnest users of the #fatacceptance hashtag that the concept itself is a lie. Although they do not necessarily center information they instead believe to be “true” or “correct,” they do suggest that earnest posters of #fatacceptance are posting “untrue” and “incorrect” information and are in denial about the relationship between obesity and health. Commenter 1 replies by arguing that the Original Poster is the one misinformed about fat acceptance by saying, “seems you’ve missed the point of Fat Acceptance and Body Positivity. It is about the fact no person should be devalued, looked down upon, or in any way shamed or humiliated because of their size. Period” (personal communication, 2019). To which Original Poster 1 responds, “I never said they should

be looked down upon, shamed or humiliated. I didn't miss the point, I just don't agree with the message it's sending by proxy. Obesity is unhealthy. Period" (personal communication, 2019). The discussion and discursive boundary-work continue when another commenter, presumably another earnest user of #fatacceptance, also assumes that Original Poster 1 is misinformed about #fatacceptance. Commenter 2 offers "correct" information and recommends that the Original Poster read "doctor-authored literature" (personal communication, 2019):

Commenter 2: Understanding of HAES is not that anyone *is* healthy at any size, but that you cannot judge a person's healthfulness by their size, that healthy behaviors at any size are better predictors of health outcomes than simple measurements like BMI, and that a person of any size can engage in healthy behaviors and improve their actual medical health indicators, and that has better health outcomes than the singular goal of losing weight at any cost, which can be incredibly harmful. . . Body positivity and fat acceptance are movements to undermine the social shame and stigma from existing as a fat person. Fatness should NOT determine our worth, or whether or not we deserve basic respect. . . I'll give that there are spaces where these messages are getting fractured, but read the original, doctor-authored literature before you attack both movements for what they aren't. None of it has to do with telling fat people that they already are healthy, it has to do with telling other people it's not okay to treat fat people poorly or assume anything about their lifestyle (personal communication, 2019).

Other commenters join the discussion, replying to Original Poster 1: "unless you provide actual scientific sources with real evidence, you're just talking without any validity"; and "unless you're a medical professional, who has seen a person's complete medical history, then how do you know shit just based on someone's appearance? It's simply ignorant judgment." Another commenter says, "Fat people are constantly devalued and dehumanized. That is what #bodypositive hashtags are meant for." Typical discursive boundary-work strategies are evident in this post and the subsequent discussion: centering "correct" information, accusations of being misinformed, and claims that the hashtag is not being "properly" used. As I mentioned before, these strategies are as much or more about how users construct sociocultural boundaries, relationships, and their own identities as they are about actually asserting something as "true" while rejecting something else as "false" or convincing someone else about what is "true" or "false." Discursive boundary-work as moderating (mis)information—whether it is information or misinformation depends on a user's views of fat acceptance or the obesity epidemic—is about maintaining and even bolstering one's perspective rather than defending or rejecting the veracity of different claims.

Many ambivalent #fatacceptance posts feature text-based images, such as Figure 6 attempting to refute a study's conclusion that "95% of diets fail" (personal communication, 2019). The posts and comments with Figure 6 also demonstrate several of the implicit strategies of moderating #fatacceptance content via discursive boundary-work. Although the discussion around Figure 6 again references vague notions of truth and falsehood, other times the debate involves more specific information, facts, or references.

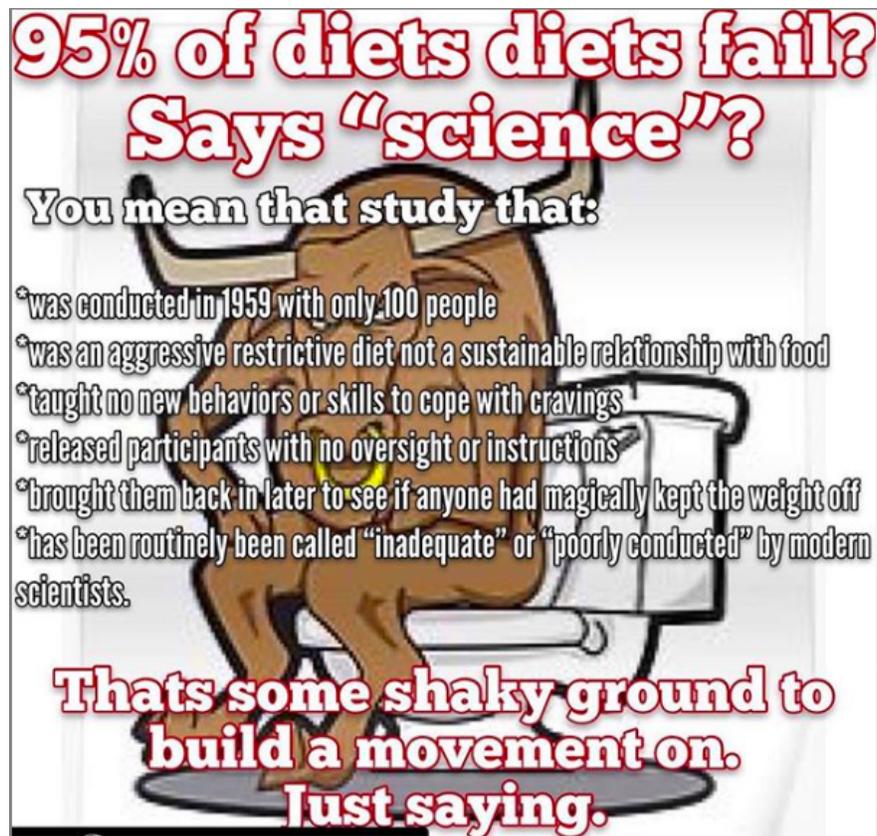


Figure 6. An image shared by an Instagram user in July of 2019. The Image reads, "Every time I see a post from someone in the #fatacceptance or #haes group they always quote "studies show 95% of diets fail" and yet can never back up those studies. This. Is. Why. Don't quote a study from 1959 that is now wholly irrelevant and inadequate. Stop twisting science to fit your standards and agenda. #effyourlogicstandards #purebullfit #everydamndayfitness #stopmakingexcuses #ameliamaddnes #loveyourbodyenoughtotakecareofit #fallogic" (personal communication, 2019).

Figure 6 inspired a significant amount of discussion between the ambivalent Original Poster 2 and several earnest commenters, as follows (personal communication, 2019):

Commenter 1: Actually this is not why, this is false information. Also if you don't wish to see posts from fat acceptance maybe don't follow or use the hashtag.

Original Poster 2: I'm not going to sit by while people spread mass misinformation and contribute to a problem that literally kills people and effects the entire population. Look again my repost is from #purebullfit. Definitely not a FA person. Also how is this post false?

Commenter 1: No, you're going to spread false information yourself! Yay you!!

Original Poster 2: It's not false. Just because you refuse to accept a fact does not invalidate that fact. The 95% failure rate comes from the study conducted by Dr Albert Stunkard Mavis McLaren-Hume. Look it up.

Commenter 1: I sent facts from 2017 to [other user] who posted that shit. I don't owe it to every troll to do their research.

Original Poster 2: I've done the research. And yet still don't see how a study that was done inadequately to the point that the writer of the study Dr Albert Stunkard himself in 1999 admitted it was wrong is some how a fact? Also you do realize that once you have to resort to name calling you've lost your argument right? It's very easy, if you are going to respond to a person telling them are wrong you need to back it up. If you can't do that like an adult without name calling then obviously your argument doesn't hold water. Good luck with that though. {smiley face emoji}.

Commenter 2: Posting this content on this take is abuse. Blocked and reported.

Original Poster 2: If you didn't block me yet, just wondering how exactly is posting facts abuse?

In this exchange, Original Poster 2 relies on similar claims of knowledge authority as other ambivalent posters, but the commenters more explicitly work to construct a boundary around and norms of use for #fatacceptance by both asking Original Poster 2 to stop following the hashtag and by blocking and reporting the user through official Instagram channels. Blocking users from seeing and engaging with future content is, of course, an example of users going through fan official moderation tool. However, it does not prevent that user from engaging with other #fatacceptance content or from continuing to circulate content deemed by #fatacceptance users as false but nevertheless allowed on Instagram per its Community Guidelines. Additionally, this exchange centers the discussion around a specific researcher, "Dr. Albert Stunkard," with each side citing aspects of study to support the opposite point of view. Users work to center alleged "true" information as a correction to "false" information and claim the "other side" is misunderstanding or spreading misinformation, exemplifying the "credibility contests" (Carlson & Lewis, 2019, p. 124) that occur between competing groups engaging in boundary-work as unofficial content moderation.

With this example, discursive boundary-work occurs on two levels: the question of who can or should use the #fatacceptance hashtag and the idea of fat acceptance in the context of the obesity epidemic. Unlike other examples of ambivalent posts, this one results in less of a dialogue between users and more of a sustained defense by earnest users of both the hashtag and fat acceptance itself, as well as continuous attempts to reach, educate, troll, or shame earnest users by ambivalent users. As is also evident here, the actions of both ambivalent and earnest users of #fatacceptance generally do not result in finding common ground or any type of factual consensus between users because the goal

of discursive boundary-work as unofficial content moderation and users' relationships to different discursive clusters of (mis)information are about reinforcing one's perspective, identity, group, or digital space in relation to fat acceptance. It is possible that discursive boundary-work as content moderation may encourage earnest users to keep posting to the hashtag because they know others are moderating and defending its use, and this may further entrench users' existing beliefs and create a sense of solidarity and community among users. Yet more importantly, discursive boundary-work functions as an unofficial way for users to moderate content in the absence official moderation practices or policies while users engage with (mis)information perceived to be problematic.

Conclusion

This article offers a contextualized discursive analysis of #fatacceptance Instagram posts as popular culture texts and as people's public social media practices in relation to unofficial content moderation and (mis)information. By analyzing #fatacceptance on Instagram, I examine individual actors and "clusters in overlapping and distinct public spheres, how they are connected, and how these connections enable discourse to circulate within and outside enclaves and counterpublics" (Kuo, 2018, p. 500). By exploring these competing discourses articulated by both earnest and ambivalent users of #fatacceptance, I offer discursive boundary-work as a way to view unofficial content moderation in addition to or in the absence of official platform policies, paid administrators or moderators, or volunteer moderators in some kind of official capacity. However, it will be important to further analyze whether and how discursive boundary-work can apply to other health-related or non-health-related contexts across social media platforms.

Instagram is not delegating policy and governance power to moderators in the case of health (mis)information related to #fatacceptance. It is thus up to individual users to moderate and try to construct boundaries in a digital space used by both activists and proponents of fat acceptances as well as ambivalent users. This is especially consequential when the hashtag is connected to fighting against stigma and discrimination, and when ambivalent hashtag use can cause harm or to some users or even rise to a level of abusive behavior. We need to think of content moderation not just in terms of official digital platform policies and the workers or volunteers whose task is to assess the flagging of content and enforce a particular set of community standards but also as a kind of discursive boundary-work that happens when speech occurs outside of what is outlined in a platform's community standards, outside of discrete boundaries between truth and falsehood, and outside of a system of groups, pages, designated hashtags, or official moderation roles.

Finally, official forms of content moderation are an ineffective tool for platforms to manage many kinds of (mis)information, especially when (mis)information is opinion-based, contested, unsettled in research, contextual, or highly nuanced. It is also an ineffective tool because moderating (mis)information, as is evident in this analysis, is less about determining truth or falsehood—and how true or false information circulates—and more about understanding why and how different groups and users of #fatacceptance understand (mis)information and use (mis)information to unofficially moderate digital spaces and reify norms of use, groups, relationships, and identities.

References

- Altay, S., Araujo, E. de, & Mercier, H. (2020). "If this account is true, it is most enormously wonderful": Interestingness-if-true and the sharing of true and false news. *Digital Journalism*, 10(3), 373–394. doi:10.1080/21670811.2021.1941163
- Antunovic, D., & Linden, A. D. (2015). Disrupting dominant discourses: #HERESPROOF of interest in women's sports. *Feminist Media Studies*, 15(1), 157–159. doi:10.1080/14680777.2015.987426
- Bacon, L., & Aphramor, L. (2018). *Body respect: What conventional health books get wrong, leave out, and just plain fail to understand about weight*. Dallas, TX: BenBella Books.
- Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, 149(8), 1608–1613. doi:10.1037/xge0000729
- Bickert, M. (2019, March 7). *Combatting vaccine misinformation*. Meta. Retrieved from <https://about.fb.com/news/2019/03/combating-vaccine-misinformation/>
- Boreo, N. (2012). *Killer fat: Media, medicine, and morals in the American "obesity epidemic."* New Brunswick, NJ: Rutgers University Press.
- boyd, d. (2014). *It's complicated: The social lives of networked teens*. New Haven, CT: Yale University Press.
- Brewis, A. A. (2010). *Obesity: Cultural and biocultural perspectives*. New Brunswick, NJ: Rutgers University Press.
- Campos, P. (2004). *The obesity myth: Why America's obsession with weight is hazardous to your health*. New York, NY: Penguin.
- Carlson, M., & Lewis, S. C. (2019). Boundary work. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The handbook of journalism studies* (pp. 123–135). New York, NY: Routledge.
- Clark, R. (2016). "Hope in a hashtag": The discursive activism of #WhyIStayed. *Feminist Media Studies*, 16(5), 788–804. doi:10.1080/14680777.2016.1138235
- Cloud, D. (2018). *Reality bites: Rhetoric and the circulation of truth claims in U.S. political culture*. Athens, OH: Ohio University Press.
- Davis, J. L., & Jurgenson, N. (2014). Context collapse: Theorizing context collusions and collisions. *Information, Communication, & Society*, 17(4), 476–485. doi:10.1080/1369118X.2014.888458

- Dressler, H. (2013). Food choice, eating behavior, and food liking differs between lean/normal and overweight/obese low-income women. *Appetite*, 65(1), 145–152. doi:10.1016/j.appet.2013.01.013
- Evans, S. K., Pearce, K. E., Vitak, J., & Treem, J. W. (2017). Explicating affordances: A conceptual framework for understanding affordances in communication research. *Journal of Computer Mediated Communication*, 22(1), 35–52. doi:10.1111/jcc4.12180
- Farr, C. (2019, September 18). *Instagram cracks down on miraculous weight loss post*. CNBC. Retrieved from <https://www.cnbc.com/2019/09/18/instagram-cracks-down-on-miraculous-weight-loss-posts.html>.
- Farrell, A. E. (2011). *Fat shame: Stigma and the fat boy in American culture*. New York, NY: New York University Press.
- Fazio, L. K. (2020). Pausing to consider why a headline is true or false can help reduce the sharing of false news. *The Harvard Kennedy School (HKS) Misinformation Review*, 1(2). doi:10.37016/mr-2020-009
- Field, A. E., Coakley, E. H., Must, A., Spadano, J. L., Laird, N., Dietz, W. H. . . . Colditz, G. A. (2001). Impact of overweight on the risk of developing common chronic disease during a 10-year period. *Archives of Internal Medicine*, 161(13), 1581–1586. doi:10.1001/archinte.161.13.1581
- Fiesler, C., & Proferes, N. (2018). "Participant" perceptions of Twitter research ethics. *Social Media + Society*, 4(1), 1–14. doi:10.1177/2056305118763366
- Font, J. C., Hernandez-Quevedo, C., McDonald, J. T., Ted, J., & Jayachandran, V. N. (2012). Understanding healthy lifestyles: The role of choice. *Applied Economic Perspectives and Policy*, 35(1), 1–6. doi:10.1093/aep/pps051
- Gal, N. (2019). Ironic humor on social media as participatory boundary-work. *New Media & Society*, 21(3), 729–749. doi:10.1177/1461444818805719
- Gerrard, Y. (2018). Beyond the hashtag: Circumventing content moderation on social media. *New Media & Society*, 20(12), 4492–4511. doi:10.1177/1461444818776611
- Ghanem, B., Rosso, P., & Rangel, F. (2020). An emotional analysis of false information in social media and news articles. *ACM Transactions on Internet Technology*, 20(2), 1–18. doi:10.1145/3381750
- Gieryn, T. F. (1983) Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American Sociological Review*, 48(6), 781–795. doi:10.2307/2095325

- Gieryn, T. F. (1999) *The cultural boundaries of science*. Chicago, IL: University of Chicago Press.
- Gillespie, T. (2020). Platforms throw content moderation at every problem. In M. Zimdars & K. McLeod (Eds.), *Fake news: Understanding media and misinformation in the digital age* (pp. 329–340). Cambridge, MA: MIT Press.
- Gillespie, T., Aufderheide, P., Carmi, E., Gerrard, Y., Gorwa, R., Matamoros-Fernandez, A, . . . Myers West, S. (2020). Expanding the debate about content moderation: Scholarly research agendas for the coming policy debates. *Internet Policy Review*, 9(4). doi:10.14763/2020.4.1512
- Graves, L. (2016). Boundaries not drawn: Mapping the institutional roots of the global fact-checking movement. *Journalism Studies*, 19(5), 613–631. doi:10.1080/1461670X.2016.1196602
- Gurrieri, L., & Cherrier, H. (2013). Queering beauty: Fatshionistas in the fatosphere. *Qualitative Market Research*, 16(3), 276–295. doi:10.1108/13522751311326107
- Instagram. (2020). *Community guidelines*. Retrieved from https://help.instagram.com/477434105621119?helpref=page_content
- Kimber, A. (2015). *Add a #sensitive hashtag to #obesity*. Change.org. Retrieved from <https://www.change.org/p/instagram-add-a-sensitive-hashtag-to-obesity>
- Kosenko, K., Winderman, E., & Pugh, A. (2019). The hijacked hashtag: The constitutive features of abortion stigma in the #ShoutYouAbortion campaign. *International Journal of Communication*, 13, 1–21.
- Kuo, R. (2018). Racial justice activist hashtags: Counterpublics and discourse circulation. *New Media & Society*, 20(2), 495–514. doi:10.1177%2F1461444816663485
- Kwan, S., & Graves, J. (2013). *Framing fat: Competing constructions in contemporary culture*. New Brunswick, NJ: Rutgers University Press.
- Lomborg, S. (2013). Personal internet archives and ethics. *Research Ethics*, 9(1), 20–31. doi:10.1177/1747016112459450
- Luo, M., Hancock, J. T., & Markowitz, D. M. (2020). Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues. *Communication Research*, 49(2), 171–195. doi:10.1177%2F0093650220921321
- Marwick, A. (2018). Why do people share fake news?: A sociotechnical model of media effects. *Georgetown Law Technology Review*, 2(2), 474–512.

- Massanari, A. (2015). #Gamergate and the fapping: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society*, 19(3), 329–346.
doi:10.1177/1461444815608807
- Matias, N. J. (2019). The civic labor of volunteer moderators online. *Social Media + Society*, 5(2), 1–12.
doi:10.1177%2F2056305119836778
- Meleo-Erwin, Z. C. (2011). "A beautiful show of strength": Weight loss and the fat activist self. *Health*, 15(2), 188–205. doi:10.1177/1363459310361601
- Mendes, K., Ringrose, J., & Keller, J. (2019). *Digital feminist activism: Girls and women fight back against rape culture*. Oxford, UK: Oxford University Press.
- Meta. (2021). *Community standards enforcement report*. Retrieved from <https://transparency.fb.com/data/community-standards-enforcement/>
- Metzl, J., & Kirkland, A. (2010). *Against health: How health became the new morality*. New York, NY: New York University Press.
- Newman, K. (2019, September 19). Obesity in America: A public health crisis. *US News and World Report*. Retrieved from <https://www.usnews.com/news/healthiest-communities/articles/2019-09-19/obesity-in-america-a-guide-to-the-public-health-crisis-health-crisis>
- Nurik, C. (2019). "Mean are scum": Self-regulation, hate speech, and gender-based censorship on Facebook. *International Journal of Communication*, 13, 2878–2898.
- Oliver, J. E. (2006). *Fat politics*. Oxford, UK: Oxford University Press.
- Park, A. (2019, December 18). Half of the U.S. population will be obese by 2030. *Time*. Retrieved from <https://time.com/5751551/us-obesity-by-state/>
- Phillips, W., & Milner, R. M. (2017). *The ambivalent internet: Mischief, oddity, and antagonism online*. Cambridge, UK: Polity.
- Roberts, S. T. (2016). Commercial content moderation: Digital laborers' dirty work. *Media Studies Publications*, 12. Retrieved from <https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1012&context=commpub>
- Saguy, A. (2012). *What's wrong with fat?* Oxford, UK: Oxford University Press.
- Scott, J. B. (2016). Boundary-work and the construction of scientific authority in the vaccines-autism controversy. *Journal of Technical Writing and Communication*, 46(1), 59–82.
doi:10.1177%2F0047281615600638

- Thrift, S. C. (2014). #YesAllWomen as feminist meme event. *Feminist Media Studies*, 14(6), 1090–1092. doi:10.1080/14680777.2014.975421
- Vishakha, D., & Mull, A. (2019, January 10). Health misinformation is rampant on Instagram. *The Atlantic*. Retrieved from <https://www.theatlantic.com/video/index/579880/instagram-wellness/>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. doi:10.1126/science.aap9559
- Ward, Z. J., Bleich, S. N., Cradock, A. L., Barrett, J. L., Giles, C. M., Flax, C., . . . Gortmaker, S. L. (2019). Projected U.S. state-level prevalence of adult obesity and severe obesity. *New England Journal of Medicine*, 381(25), 2440–2450. doi:10.1056/NEJMsa1909301
- Wardle, C. (2020). Journalism and the new information ecosystem: Responsibilities and challenges. In M. Zimdars & K. McLeod (Eds.), *Fake news: Understanding media and misinformation in the digital age* (pp. 71–85). Cambridge, MA: MIT Press.
- Williams, A. A. (2017). Fat people of color: Emergent intersectional discourse online. *Social Sciences*, 6(1), 15. doi:10.3390/socsci6010015
- World Health Organization. (2021, June 9). *Obesity and overweight*. Retrieved from www.who.int/mediacentre/factsheets/fs311/en/
- Zimdars, M. (2019). *Watching our weights: The contradictions of televising fatness in the "obesity epidemic."* New Brunswick, NJ: Rutgers University Press.