

The McMaster Journal of Communication  
Vol 12 No 2  
2020

Dealing with negative  
social media comments:  
The giant bullhorn that punctures  
holes in organizations' walls,  
like an arrow to the chest

Antonio Rino

Mount Saint Vincent University

Copyright ©  
2020 by the author.

The McMaster Journal of Communication is produced by The Berkeley Electronic Press  
(bepress).

<https://journals.mcmaster.ca/mjc>

# Dealing with negative social media comments: The giant bullhorn that punctures holes in organizations' walls, like an arrow to the chest

Antonio Rino<sup>1</sup>  
Mount Saint Vincent University

## Abstract

People in organizations that use social media for business tend to focus on a negative social media comment, even in a sea of positivity. If there is one negative comment and 99 that are positive, it can feel as though the negative comment gets at least 99 percent of the attention. This research explores strategic ways to avoid social media negativity and brings together insights from academic and industry sources to understand online negativity, including why a company would even consider having a social media presence.

*Keywords:* social media, community manager, online negativity, negativity bias, negative comments, online emotions, user categorization

---

<sup>1</sup> Currently enrolled graduate student of the Master of Arts (Communications) program at Mount Saint Vincent University. Email: [tony@pastayoyo.com](mailto:tony@pastayoyo.com)

The McMaster Journal of Communication  
2020  
Vol 12, No 2, ISSN 1710-257X

# Dealing with negative social media comments: The giant bullhorn that punctures holes in organizations' walls, like an arrow to the chest

Antonio Rino  
Mount Saint Vincent University

**A** negative comment on a corporate social media post can pierce like an arrow to the chest and puncture holes in an organization's walls. A single negative voice in a sea of positive feedback can feel as though it is blaring from a giant bullhorn, striking fear into corporate community managers that an avalanche of negativity will overtake positivity like a contagious bandwagon. Why would a corporation consider telling its story in the online battlefield of social media and risk exposing its reputation to a cesspool of negativity? Through research, industry advice and best practices – including from the researchers and experts who use the foregoing colourful idioms and metaphors to describe negative online comments – I will explore answers to why negativity is an online barrier. To answer my main question of why an organization would consider engaging on social media in the face of prolific negativity and hate speech, I will review the evolution of online emotions and the rise of negativity on social media. I will define negative online comments in the corporate context using research on trolls, cyberbullying and online personal attacks. Using the psychology of Pareto's 80/20 rule and negativity bias, I will provide quantitative and qualitative perspectives on negativity to show why companies pay greater attention to negative comments than positive ones, and how analysis of negativity can help a company develop emotional intelligence. I will also present examples from research and industry to understand and combat negativity and review research on user comments that classifies users to better understand their motivations. Using research on tone and voice in online conversation, I will share cautionary case studies that demonstrate how companies that are not self-aware can incite negative comments. Finally, I will review research

on platform content moderation techniques to understand how social media platforms like Facebook manage negativity and I will suggest similar solutions for corporations, including not only the online community's ability but also our collective responsibility to moderate and overcome the online positivity deficit.

### **Online emotions and the giant bullhorn of negativity**

In its infancy, the internet was an exciting way for people to connect with other individuals locally and around the world. Companies were also excited and developed corporate websites as digital brochures to connect with people by promoting their products and services. The early days of the web were like a digital *Mad Men* era where companies retained control over static digital messages to their consumers (David, 2015). With the rise of conversations on social media platforms, companies no longer have control over one-way digital message delivery with their customers. People are talking freely about companies online using social media platforms like Facebook. The internet is no longer seen as exclusively positive. Unfortunately, online discussion has become an avenue for negativity and abuse (Wulczyn et al., 2017, p. 1). Social media is an environment that has become loaded with emotions, and those emotions are the driving force for posting content (Jalonen, 2014, p. 2). "Metaphorically, social media punctures holes into organisations' walls, making them transparent in an unforeseen way" (Jalonen, 2014, p. 7). Customers are using social media to make emotionally charged complaints about companies, whether rational or irrational. Emotional social media comments and online negativity are changing the way companies converse with their customers.

The positive excitement during the early days of the internet is what Siva Vaidhyanathan calls a myth. In *Antisocial media: How Facebook disconnects us and undermines democracy*, Vaidhyanathan (2018) explained that the internet was going to be like a digital saviour that would make things better for us, partly because it would help us make rational decisions. "Conversation, deliberation, argumentation, information, mutual recognition, and communication would generate an ideal setting for making good decisions" (Vaidhyanathan, 2018, p. 207). Instead, people's negative online conversations about peripheral issues and concerns can be amplified disproportionately against a positive majority of discussion like a giant bullhorn, driving an unrealistic public view (David, 2015). The rise of negativity has become so strong, it has become a business opportunity for reputation management companies like Ontarioseo.ca that caution "your online reputation can be tarnished instantly by people posting negative feedback about your company, whether it's true or not" (Ontario SEO Targeted Digital Marketing, 2017, para. 3). The perception is that the positive majority is not as loud as the negative minority and "people are much more likely to recognise and be influenced by negative information shared in social media" (Jalonen, 2014, p. 2). Just one negative comment can have the power to create a chain-reaction avalanche of negative emotion (Tadic et al., 2013). And just one negative comment, particularly if it is a company's first negative online experience, can hit a business "like an arrow to the chest" and evoke an emotional reaction within a community manager like "fighting the commenter or trying to get the comment deleted" (Sweeney, 2018a, para. 3). Negative online emotions are a powerful force and can deter organizations from participating online, but my review of research here will provide perspectives and solutions to help organizations overcome the barrier of negativity.

### **The stopping power of a negative comment**

As stated by Britton (2017), a negative comment can take the form of a hurtful personal attack, name-calling, foul language, hate speech, a confrontational rant or otherwise inappropriate content. The source of a negative comment can range from a disgruntled customer to a trouble-making online user known as a troll. Joanne Sweeney, a recognized online reputation management expert defines a troll as someone who creates “a greater emotion than is appropriate to the situation in hopes that the target will embarrass themselves and be goaded into overreacting” (Sweeney, 2018a, para. 9). Unlike a disgruntled customer with a legitimate complaint, a troll’s intent is to deliberately damage a company (Jalonen, 2014). Negative comments can also be posted by cyberbullies who have previous knowledge or experience with the company. A cyberbully poses a strong threat to a company because “their negative comments, outright lies, or offensive ranting may be viewed by everyone else who follows your company and can go viral” (Britton, 2017, para. 4).

Negativity is a powerful force that can deter a company from participating in online conversations. “This does not necessarily mean that negative posts and comments outweigh positive ones in numbers, but that negative emotion is more influential” (Jalonen, 2014, p. 3). The harshness of a negative comment can damage confidence and corrode self-esteem, but it should not have stopping power. “It can make you second guess yourself as to what you talk about and what you write about ... [Online abuse] is not free speech, it’s actually limiting everyone else’s free speech” (Tenove et al., 2018, para. 15).

### **Perspectives on online negativity: Pareto’s 80/20 law and negativity bias**

I suggest two behavioural theories to offer a helpful perspective on the disproportionate power of negativity: Pareto’s 80/20 law and negativity bias. “The harshest trolls aren’t necessarily more numerous — they’re just louder than everyone else” (Chen, 2017, para. 4). This phenomenon can be interpreted using Pareto’s 80/20 law: “80% of everything is caused/explained by 20% of the actor/units involved in the process” (Matei & Bruno, 2015, p. 178). Additionally, even when the number of positive comments far outweigh the negative, a psychological theory known as the negativity bias shows that negativity has a stronger influence. “In psychology, negativity bias refers to a phenomenon in which humans have a greater recall of unpleasant memories compared to positive memories” (Jalonen, 2014, p. 2). Pareto’s 80/20 law and negativity bias each offer important insights that provide perspectives to help companies understand negativity and overcome the fear of participating online.

Pareto’s 80/20 law and negativity bias have nuanced perspectives on negative social media comments. Pareto’s 80/20 law is more of a quantitative measure, as the name suggests. Pareto’s law uses percentages to explain how a few comments can have disproportionate influence. Negativity bias, on the other hand, is more of a qualitative measure. Negativity bias asserts that negative comments have more influence than positive comments.

Vilfredo Pareto (1848-1923) was an Italian economist and philosopher who noticed that 20 percent of Italians owned 80 percent of the land and wealth. He conceived his theory of disproportion, where a small percentage of a total can control the majority, known as the Pareto

principle or Pareto's 80/20 law (the Pareto principle, n.d.). In economic theory, "Pareto's 80/20 rule explained that 80% of total benefits can be achieved with 20% of total effort" (Banduka et al., 2016, p. 488). The numbers in the theory don't need to be specifically 80/20 and they don't need to add up to 100. Pareto's focus was to identify the disparity. The Pareto principle is also used to explain human behaviour. For example, "80 percent of results or outputs flow from 20 percent of causes, and sometimes from a much smaller proportion of powerful forces" (Koch, 1999, p. 2).<sup>2</sup>

In a study called "Ex Machina: Personal Attacks Seen at Scale," Wulczyn, Thain, and Dixon (2017) found that a very small number of "highly toxic" Wikipedia contributors were responsible for a large percentage of attacks (p. 6). Wulczyn et al. analyzed attacks on Wikipedia contributors in 2015. Their analysis shows that while 80 percent of attacks came from 9,000 light to moderately active users, 9 percent of attacks came from just 34 highly active users. Wulczyn et al. also observed that "significant progress could be made by moderating a relatively small number of frequent attackers" (Wulczyn et al., 2017, p. 6). Wulczyn et al.'s findings illustrate Pareto's 80/20 rule, showing that a small minority of highly vocal users can create a disproportionate amount of negativity.

Negativity bias is a term coined by psychologists Paul Rozin and Edward B. Royzman in their 2001 research titled "Negativity Bias, Negativity Dominance, and Contagion." Rozin and Royzman (2001) hypothesized that there is a general bias in animals and humans to give greater weight to negative entities and that "negative entities are more contagious than positive entities" (p. 1). In negativity bias, "negative entities are stronger than the equivalent positive entities, and the negativity of negative events grows at a steeper rate than the positivity of positive events" (Rozin & Royzman, 2001, p. 1).

Franklin Waddell and S. Shyman Sundar illustrated negativity bias in their 2017 study "#thisshowsucks! The Overpowering Influence of Negative Social Media Comments on Television Viewers." Waddell and Sundar asked: "Why do television viewers assume that the comments of a few reflect the sentiment of the majority?" (Waddell & Sundar, 2017, p. 393). In their analysis of how viewers were affected by social media comments made during the airing of television programming, Waddell and Sundar (2017) found that positive social media comments did not influence an individual's level of enjoyment. The results of the study found that "negative social media comments appear to indirectly affect individuals' own level of program enjoyment based on the assumed sentiment of other program viewers" (Waddell & Sundar, 2017, p. 404). Waddell and Sundar (2017) also noted a "bandwagon affect" (p. 304), where users who read negative comments were more likely to shift their attitudes negatively "leading viewers both to perceive that others have evaluated the program negatively and to lower their own evaluation of the program" (p. 406).

---

<sup>2</sup> There is another 80/20 rule of social media marketing used to create social media content which states that "80% of your social media posts should inform, educate, and entertain your audience, while only 20% should directly promote your business" (Hall, 2018, para. 1). While not the same as the Pareto principle, the 80/20 marketing rule illustrates the significance of the smaller proportion.

Companies who participate in social media conversations should consider the quantity and quality characteristics of the Pareto principle and negativity bias to help frame the tone of their online feedback. “A few loud and cantankerous voices should not drown out reality” (Patel, n.d.). While the Pareto principle can help companies understand that a few negative comments often do not represent the majority, companies should recognize negativity bias and the influential power of a few negative comments. Recognizing the influence of negativity, companies can understand that not every commenter with a complaint is a cyberbully. “The vast majority of people commenting on your Facebook page or tweeting at your company likely have legitimate concerns or complaints” (Britton, 2017, para. 8). With a deeper understanding of the quantity and quality of emotionally charged social media comments, companies can gain more insights into online negativity and begin to develop online emotional intelligence.

### **Categorizing user negativity through emotional intelligence**

Since the early days of social media, researchers, organizations and social media industry experts have analyzed user comments and developed response strategies that classify users in an effort to understand motivations and process whether or not a response is required. Sweeney (2018a) posited there are a lot of reasons why a stakeholder may post a negative comment. Some customers have had a bad experience and want everyone on social media to know, others are simply seeking customer support, “and more than a few out there just looking to see if they can spark a negative reaction by being mean” (Sweeney, 2018a, para. 6). To combat negativity, an organization needs a specialized and dedicated resource beyond a traditional customer service team, sometimes known as a social media community manager. A dedicated online community manager is needed not only to monitor and respond to online comments, but also to understand the reason behind them (Sweeney, 2018b). I present four sources spanning the past 11 years to illustrate the dynamics of social media emotions, the importance of developing emotional intelligence to understand user comments and how motivations behind social media comments have changed.

In 2008, when conversations on blogs and social media were developing a strong online voice, some of the first companies and experts to step into social media conversations turned to a comment classification process called the “Air Force Blog Assessment” chart, developed by the United States Air Force Public Affairs Agency, Emerging Technology Division (see Fig. 1).

Silicon Valley’s Jeremiah Owyang introduced the “Air Force Blog Assessment” chart to his business audience in his web strategy blog, saying “There’s a lot to be learned from this military diagram that can be applied to corporations or even your personal blogging efforts” (Owyang, 2008, para. 3). Another web strategist, David Meerman Scott used a more colourful assessment, calling social media a “battlefield” and stated the chart is needed because “many corporations are scared witless about social media” (Scott, 2008, para. 4).

The chart is a decision tree that helps assess, evaluate and respond to comments systematically. If a comment is positive and there is concurrence, the chart steps through a simple process that might lead to a positive response. If the comment is negative, the chart bases the possibility of responding on four user categorizations: Trolls, Rager, Misguided and Unhappy Customer. Underlying the process are five strategic considerations for responses: Transparency,

Sourcing, Timeliness, Tone and Influence. As one of a small group of early corporate social media community managers working in the oil and gas industry in 2008, I used the chart regularly – almost to the point of worship – to help me assess and deal with online comments. In the early days of understanding social media comments, the classification on the “no” side of the chart was especially helpful in processing responses.

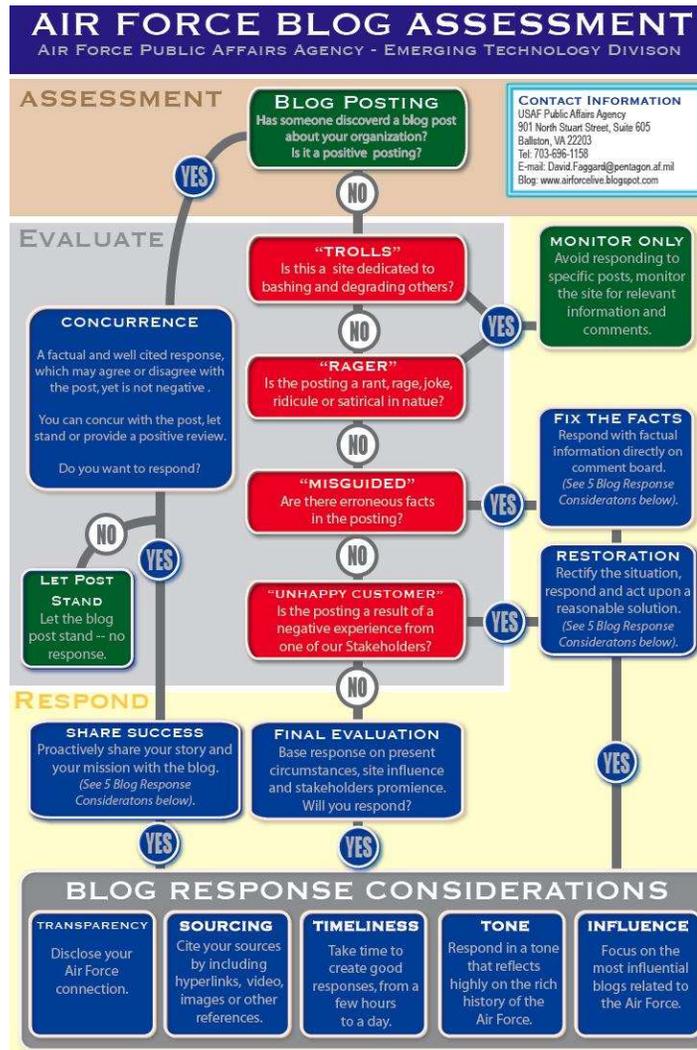


Figure 1. Air Force Blog Assessment. From Air Force Public Affairs Agency - Emerging Technologies Division, 2008, <https://freshspot.typepad.com/.a/6a00d83451f23a69e20105365f0d62970b-popup>.

While the “Air Force Blog Assessment” chart is still a good elementary resource for corporations and community managers, responding to comments on social media has evolved since 2008. As already noted from Vaidhyathan, Sweeney, Waddell & Sundar, Wulczyn et al. and others, emotions on social media, including increased negativity, have become a powerful motivating force for online comments and require deeper and more intelligent consideration to assess their origin and motivation. Consequently, the decision tree needs more branches.

“Motivation for content generation in social media” is a 2015 research article by Janne Matikainen from the University of Helsinki, Finland. Matikainen surveyed more than 1,000 Finnish web users and found that people are motivated to post comments in online communities in three ways: 1) they want to be involved in the development of the internet; 2) they value sharing information about their lives through self-expression; and, 3) they want to belong to an online community (Matikainen, 2015). While Matikainen did not ask specific questions about emotions, positivity or negativity, one respondent mentioned “nasty” online communities illustrating why some organizations do not want to participate in online social media discussions. “There are of course also communities that just feel repelling somehow. Nasty comments can make you feel that way, and a lot of people have stopped blogging because the comments have been so out of line” (Matikainen, 2015, p. 50). The three classifications in Matikainen’s study are helpful in understanding what motivates users to post online comments. But the study only touches the surface of understanding emotionally motivated content.

In his 2014 conference paper titled “Negative emotions in social media as a managerial challenge,” Harri Jalonen posited that online emotional intelligence is an important characteristic required of a company participating in social media discussions (Jalonen, 2014). Jalonen (2014) said an organization develops online emotional intelligence through “the ability to recognise the emotions shared and diffused in social media, and to understand their meaning for the business and to behave on the basis of that understanding” (p. 7). Jalonen (2014) categorized six themes “which challenge companies’ ability to act in an emotionally intelligent way in the social media age” (p. 4), including: a) demographic considerations where age, gender and education play a role in expressing negativity emotions; b) cultural factors and how social media is used in a particular society; c) the object of expressing negative personal experiences and sharing negative emotions on social media rant-sites; d) how different social media platforms induce negativity differently; and, e) the dynamics of sharing negative emotion, particularly when the avalanche effect influences users to share viral negative posts. In his sixth theme on user motivation, Jalonen (2014) identified four reasons people express negative emotions on social media: 1) To draw attention to their dissatisfaction in order to get a solution or receive compensation; 2) for altruistic reasons to help others by disclosing their negative experiences in order to prevent others from suffering a similar incident; 3) to help companies improve by talking about their products to assure that the issue is “structurally solved;” and, 4) to express out of the need for publicity and attention, usually as a troll (p. 6).

### **Organizational introspection and tone and voice in online conversations**

In 2018, illustrating that social media activity has grown and corporate engagement has advanced in sophistication, Barcelos, Dantas and Sénécal’s study took an introspective view on social media conversations. In “Watch Your Tone: How a Brand’s Tone of Voice on Social Media Influences Consumer Responses,” Barcelos et al. (2018) recognized that an organization’s online corporate voice has the ability to influence positive or negative emotions and suggested using “human voice manipulation” when interacting with customers (p. 71). Barcelos et al. (2018) used case studies to determine if a personal or corporate voice is more appropriate based on whether an organization’s goods or services are seen as hedonic or utilitarian. A hedonic organization has products or services with emotional benefits from purchasing or experience, like a pizzeria, wine shop, book store or hotel (Barcelos et al., 2018). “It is a good idea for a

brand to use a human voice in its interactions with customers on social media if its products or services are primarily hedonic” (Barcelos et al., 2018, p. 73). A utilitarian organization focuses on solving rational and objective needs like medical services, financial consulting or business-to-business services (Barcelos et al., 2018). Barcelos et al. (2018) said that for a utilitarian organization, “it is better to adopt a corporate voice and maintain a certain distance when dealing with customers” (p. 73). The results of Barcelos et al.’s study show that businesses need to understand whether they are hedonic or utilitarian because using the wrong voice can incite negative social comments from their customers. By extension, I suggest Barcelos et al.’s findings apply to non-retail organizations that interact with an audience of stakeholders rather than customers.

Barcelos et al.’s findings illustrate a problem with businesses categorizing themselves as hedonic or utilitarian. Social media conversations are often in human voice and “brands are increasingly employing an informal style in their social media communications” (Barcelos et al., 2018, p. 60). Utilitarian organizations, therefore, need to be careful of not only their tone of voice, but also the kind of content that they post. For example, Alberta-based utility company ATCOenergy posted a Facebook message with an energy conservation message on February 2, 2019, referencing Groundhog Day (see Fig. 2). The post shows a photo of a groundhog with its front paw raised and the message: “What did the groundhog say this morning? He said, ‘Stop wasting energy. Turn down your thermostat at night and save on your electricity bill.’ Smart groundhog” (para. 1).

ATCOenergy, a utility company which could be categorized as offering a utilitarian service, used a human voice and informal photo in the post to convey a rational message about saving energy.



Figure 2 ATCOenergy. (2019, February 2). What did the groundhog say this morning? He said, "Stop wasting energy. Turn down your thermostat at night and save on your electricity bill." Smart groundhog. <http://ow.ly/TpHl30nvimb> [Facebook Status Update]. Retrieved from: <https://www.facebook.com/ATCOenergy/photos/a.1554877084834418/2213146849007435>

The Facebook post incited several negative messages including from user Jojo Bicks who replied with two angry emojis: "Pffft. Doesn't matter how little energy we use, still get gouged with all of the 'extra' charges" (ATCOenergy, 2019, para. 2). The negative comment sparked another negative comment and a subsequent negative reply from user Jojo Bicks, before ATCOenergy stepped into the conversation with a response. In this example, the text in ATCOenergy's Facebook post uses a hedonic voice with an informal example of a groundhog instead of a utilitarian voice with a rational and objective tone, resulting in a negative emotional response as predicted by Barcelos et al.'s findings. Interestingly, ATCOenergy's response to the negative comment did not continue with the hedonic tone or the Groundhog Day theme of the original post and switched to a more rational and objective corporate response: "Hi Jojo. We know the other fees can be a large part of your bill. Unfortunately, we don't have any control over them, since they come from other sources. That's why we try to save you some money where we can" (ATCOenergy, 2019, para. 5). Given the utilitarian nature of ATCOenergy's business as a utility, it can be argued that the informal metaphor, which capitalizes on the fact that February 2 is recognized as Groundhog Day, is unnecessary for a non-hedonic business and

illustrates the danger of brands using an increasingly informal style on social media, as Barcelos et al. warn (Barcelos et al., 2018).

Matikainen, Jalonen and Barcelos et al.'s findings, including Jalonen's (2014) invocation of online emotional intelligence, offer a deep perspective into motivations behind online conversations, emotion and negativity. Although only three studies were presented in the field of understanding online behaviour, the findings suggest the need to add several branches to the "Air Force Blog Assessment" decision tree, particularly on the "no" side. Adding to Trolls, Rager, Misguided and Unhappy Customer, assessing comments could include understanding users' desires to be involved, self-expression and belonging to community. Motivating factors behind the "No" categorization of the Air Force chart that require emotional intelligence include a deeper understanding of user demographics, cultural factors, negative self-expression, and whether the content being shared is trending or viral. The social media network where the comment was published (i.e. Facebook, LinkedIn, Twitter, etc.) is also an important consideration in understanding where emotions originated because "different social media sites differ in their ability to induce expressing and sharing negative emotions" (Jalonen, 2014, p. 5).

Another branch in the social media assessment tree could be based on Jalonen's (2014) four reasons why people express negativity towards organizations: to receive compensation, help others, solve a problem or seek attention. And another branch is needed on the Respond phase of the chart to accommodate Barcelos et al.'s (2018) findings that organizations need to perform an introspective hedonic or utilitarian assessment to determine if a personal or corporate voice is more appropriate in order to avoid inciting negative emotions.

Barcelos et al.'s (2018) research on developing an online voice also illustrates the need organizations have for a nuanced understanding of how people perceive their product or service online and the need to continuously monitor and respond to social media conversations with emotional intelligence. In summary, while the "Air Force Blog Assessment" chart is a good first step in understanding how to handle and process social media comments, the three studies presented add more intricate response considerations to the decision tree, requiring deeper understanding, dedicated resources and emotional intelligence. These additional considerations require additional overhead for organizations and may deter them from participating in social media.

Jalonen's thought-provoking 2014 research is an indication of how negativity in online commenting has advanced since the 2008 "Air Force Blog Assessment" chart. But the results from one of the studies Jalonen cited are now outdated, illustrating that the negativity in social media comments is changing and dynamic. In his theme, "Platforms of expressing and sharing negative emotions," Jalonen (2014) cited Leung who, in a 2013 study, found that "Facebook is not a channel of choice for venting negative feelings" and "not a place for voicing conflicting views. When people want to display negative emotions and voice conflicting views, they prefer blogs and forums" (p. 7). Any Facebook user or corporate social media community manager today might laugh in disbelief at the notion that Facebook is not a place for venting negative feelings or voicing conflicting views because Facebook comments on controversial or partisan public posts are often filled with strong negative emotions and even hate speech.

### **Moderating the cesspool of hate speech**

Negative comments, negativity and hate speech are so prolific, they have become an embedded characteristic of social media, and companies and organizations are not immune to the negativity. In their research article titled “A Web of Hate: Tackling Hateful Speech in Online Social Spaces,” Saleem, Dillon, Benesch and Ruths discussed the effects of online negative content and hate speech on users. Negativity can be so overwhelming that it can frighten, intimidate, or silence users (Saleem et al., 2017), and can deter corporations from participating online. Hate speech and negative comments are a difficult challenge to overcome for corporations who do not want their brand to be associated with negativity. Two content moderation solutions to help mitigate the proliferation of negativity are keyword filtering and human decision-making.

Automated tools that filter and block vulgar words are a limited and basic form of content moderation. Facebook, for example, has a profanity filter built into the settings of its company page preferences. Facebook describes the tool this way: “You can choose whether to block profanity from your Page, and to what degree. Facebook determines what to block by using the most commonly reported words and phrases marked offensive by the community” (Facebook, n.d., para. 1). Saleem et al. said that keyword-based filtering is effective but not perfect. Filtering keywords for commonly-known slurs, vulgarity and offensive terms will catch some hateful speech, but “it is common to express hate in less explicit terms, without resorting to standard slurs and other offensive terms” (Saleem et al., 2017, p. 4), including nuances and misspellings. Saleem et al.’s research concluded that automated keyword-based filtering, like Facebook’s profanity filter, is not enough to understand and censor hateful speech. “In all cases considered, a classifier trained on community-based data outperforms a keyword-based classifier” (Saleem et al., 2017, p.7-8). However, training human classifiers has its own challenges.

In a 2018 book, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*, author Tarleton Gillespie investigated how social media platforms regulate what people post online. Chapter 8 begins with a striking comment by the chairman and CEO of Facebook, Mark Zuckerberg, whose words acknowledge the tremendous volume of inappropriate content and the huge amount of work required to govern Facebook’s content:

Facebook is not just technology or media, but a community of people. That means we need Community Standards that reflect our collective values for what should and should not be allowed. In the last year, the complexity of the issues we’ve seen has outstripped our existing processes for governing the community. (Zuckerberg, 2017, as cited in Gillespie, 2018, p. 197)

Facebook employs large numbers of content moderators as human decision-makers to moderate hate speech and negative content across the platform that violate its own standards (Gillespie, 2018). Gillespie (2018) said Facebook and other social platforms are obligated to moderate negative comments. He said, for example, that anyone can create a website and allow users to comment openly but without moderation, that website would become a cesspool of negative comments. Gillespie (2018) said that because negative content is so prolific, content

moderation is not a casual option, but “in fact an essential, constant, and definitional part of what platforms do. I mean this literally: moderation is the essence of platforms, it is the commodity they offer” (p. 207). Constant moderation reflects the ever-changing meanings in hate speech and the limitations of profanity filters underline the need for human decision-makers. In “The New Governors: The People, Rules, and Processes Governing Online Speech,” Kate Klonick (2018) said “Content moderators act in a capacity very similar to that of a judge” (p. 1599). Klonick analyzed content moderation on platforms like Facebook, Twitter and YouTube and how the platforms use human decision-making to curate content posted by their users. Although her research acknowledged that platforms like Facebook have not made their content moderation guidelines public, Klonick (2018) said “moderators are trained to exercise professional judgment concerning the application of a platform’s internal rules” (p. 1642). However, Saleem et al. recognized that the rules themselves can be difficult to enforce because the dynamic changes and nuances of hate speech can lead to confusion, making content difficult to judge. “Each person seems to have an intuition for what hate speech is, but rarely are two people’s understandings the same” (Saleem et al., 2017, p. 2). Between hate speech’s dynamics and human interpretation, moderating negativity is a tremendous challenge.

A Facebook post from SNC-Lavalin illustrates ongoing challenges companies have moderating negative comments on Facebook (see Fig. 3). In early 2019, SNC-Lavalin was facing criminal charges in connection with payments it made to the Libyan government (Reynolds, 2019). During the same time period, it was alleged that staff from Canada’s Prime Minister’s office attempted to pressure Canada’s attorney general to intervene in the SNC-Lavalin prosecution (Gollom, 2019). The SNC-Lavalin affair, as it has become known, is a complicated matter receiving intense media coverage and placing the company under public scrutiny.

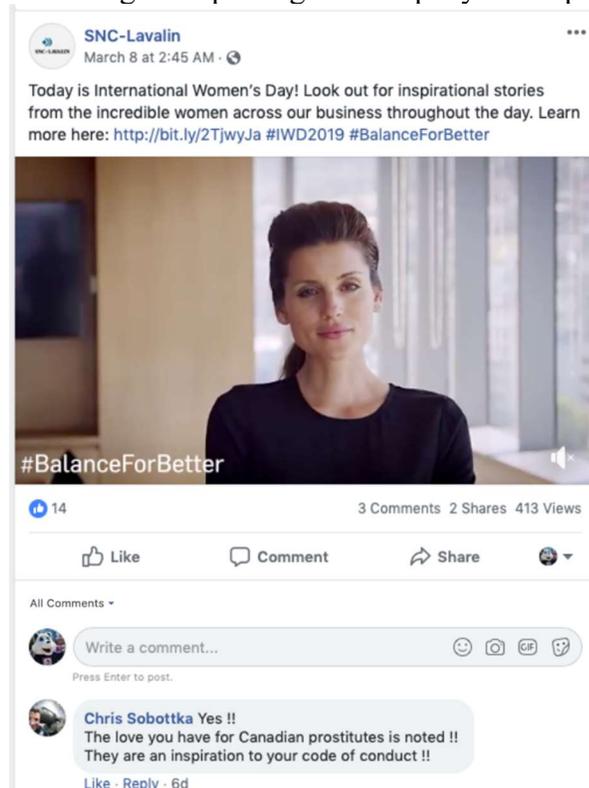


Figure 3. SNC-Lavalin (2019, March 8). Today is International Women’s Day! Look out for inspirational stories from the incredible women across our business throughout the day. Learn more here: [#IWD2019 #BalanceForBetter](http://bit.ly/2TjwyJa). [Facebook Status Update]. Retrieved from: <https://www.facebook.com/snclavalin/videos/359880481524932/>

On March 8, 2019, to celebrate International Women’s Day, SNC-Lavalin posted a video on its company Facebook page spotlighting women who work within its organization. The video post was accompanied with the following text: “Today is International Women’s Day! Look out for inspirational stories from the incredible women across our business throughout the day. Learn more here: [#IWD2019 #BalanceForBetter](http://bit.ly/2TjwyJa)” (para. 1). In examining the post in Figure 3 – and based on my experience as a corporate social media community manager – it is evident that the accumulated statistics below the video say “3 comments” and that, even with the “All Comments” option selected, only one comment is visible. While I cannot find an official Facebook reference as to why this occurs, in my experience moderating comments on Facebook business pages, this discrepancy is usually the result of comments being removed by a moderator or hidden with Facebook’s automated profanity filter. In this case, two comments are unaccounted for and were likely hidden or removed, yet Facebook still inexplicably acknowledges their existence in the total number of comments. Some online discussion threads also suggest that comment counts include hidden or moderated comments (Marigold, 2013). This presents a challenge for corporations and community managers because savvy users will know that content is being moderated.

The post also incited a negative comment from Facebook user Chris Sobottka: “Yes !! The love you have for Canadian prostitutes is noted !! They are an inspiration to your code of conduct !! (sic)” (SNC-Lavalin, 2019, para. 2). Sobottka’s comment, which references further allegations in the SNC-Lavalin affair, is negative because it uses sarcasm to highlight alleged corruption and could also be interpreted as equating prostitution with the SNC-Lavalin employees who are featured in the video. Yet, SNC-Lavalin has chosen not to remove the comment. We can conclude from the discrepancy in the comment count that the post is being moderated, and looking back at Sweeney, Chen and Jalonen’s categorizations of trolls – considering the “Air Force Blog Assessment” chart to assist in dealing with comments – it is surprising that Sobottka’s comment has not been removed. It may be that SNC-Lavalin’s community manager is too overwhelmed with negative comments to notice this one.

Further, given Jalonen’s (2014) invocation of corporate emotional intelligence and Barcelos et al.’s (2018) research on the need for corporations to take an introspective view of their relationship with customers and stakeholders, it’s surprising that SNC-Lavalin – who is under intense media scrutiny for an alleged scandal and who would likely be categorized as utilitarian based on Barcelos et al.’s (2018) research – would invite negativity by publishing a hedonic feel-good story like their International Women’s Day Facebook post for public consumption. Given the circumstances, the video is better suited for internal communications within SNC-Lavalin. The SNC-Lavalin Facebook post is an excellent case study of how a Facebook post can incite negativity because it serves as a cautionary example of what not to do based on the research I have explored. Taking a closer look at SNC-Lavalin’s Facebook page, it does not have rules or community standards about the expectations of visitors who comment on the page and how abuse or negativity will not be tolerated.

As platforms have adapted to deal with negativity and hate speech by adopting community standards and moderating content based on collective values, corporations and organizations with a social media presence must also adapt with their own community standards and moderation practices. However, corporations “must recognize that moderation is hard work” (Gillespie, 2018, p. 198). Corporations need to consider their governance ethically and with emotional intelligence (Jalonen, 2014), including an introspective view of their business and an understanding of their customer/stakeholder relations (Barcelos et al., 2018). Corporations and organizations can overcome the negativity bias of participating on social media by reviewing the techniques that platforms use to moderate content and using content moderation within their response strategies.

A benefit to content moderation is that it could stop the avalanche of negative comments. Wulczyn et al. (2017) pointed out that “personal attacks cluster in time – perhaps because one personal attacks [sic] triggers another” (p. 7) and that “early intervention by a moderator could have a disproportionately beneficial impact” (p. 7). Moreover, Gillespie (2018) asserted that platforms and users, including organizations with a corporate presence on social media, should take on the greater responsibility in moderating negativity and hate speech, “attending to these unresolvable tensions, acknowledging and staying with them – not just trying to sweep them away” (p. 212-213).

### **Conclusion**

I have reviewed scholarly research and industry best practice on online negativity to explore why an organization would consider engaging on social media in the face of prolific negativity and hate speech, and to provide perspectives and insights that will help organizations overcome the online barrier of negativity. Through Jalonen’s (2014) research on negative emotions in social media and Wulczyn et al.’s (2017) analysis of negative comments on Wikipedia, I demonstrated the evolution of online emotions and the influencing power of even one negative comment. I defined negative comments in the corporate context, using analyses of trolls and cyberbullies by Sweeney, Britton and Jalonen.

To provide a quantitative perspective on the influence of negative comments, I reviewed research by Matei & Bruno (2015) on Pareto’s 80/20 rule, supported by Wulczyn et al.’s (2017) case study on how a few Wikipedia users generated many negative comments. For a qualitative perspective on the influencing power of negativity, I used Rozin and Royzman’s (2001) theory on negativity bias, supported by Waddell and Sundar’s (2017) case study on how negative comments have the power to influence television viewers. I shared techniques and research on how social media community managers categorize user comments to help understand negativity, including “The Air Force Blog Assessment Chart,” and Maitekainen’s (2015) classifications of what motivates a user to post comments in online communities. I introduced Jalonen’s (2014) assertion that emotional intelligence is an important characteristic for an organization on social media as well as Jalonen’s (2014) categorizations of what challenges a company to act in an emotionally intelligent way.

To help companies become more self-aware, I presented an ATCOenergy case study and reviewed research by Barcelos et al. (2018) showing that if a business can understand whether it

is hedonic or utilitarian, it can manipulate its online voice and reduce the chance of receiving negative comments. To help organizations understand the use of content moderation, I reviewed automated profanity filters in Saleem et al.'s (2017) research on tackling hateful speech online. Using Gillespie's (2018) perspective on prolific online negativity and Klonick's (2018) analysis of how major social media platforms use human content moderators, I showed that despite the challenge of content moderation, it is not an option for organizations. Reviewing my research and using an SNC-Lavalin Facebook post, I presented a cautionary example of what companies should not do online.

Despite reviewing an intriguing list of scholarly and industry sources, my research is an overview of a snapshot in time. Negativity on social media has changed (particularly as described by Leung's observations in 2013 about Facebook not being a platform for negativity, as cited by Jalonen, 2014). As social media continues to evolve – perhaps with new platform governance, government regulations, a new understanding of motivations, a new generation of users or new platforms – negative online comments will evolve. Future research in the emerging field of understanding online users and their motivations for posting negative comments would be helpful in profiling and categorizing users for a deeper understanding of the source of online negativity. Also, although I did not review corporate security in my research, the existence of automated profanity filters tweaked my interest in questioning whether comments that are automatically hidden and possibly ignored by a community manager might contain valuable, unseen information for the development of a company's emotional intelligence, as well as information on potential security threats to the organization or members of its staff. Further research on the effects of automated content moderation and a user's strategic use of negativity would be another interesting way of understanding why online negativity exists.

The "Air Force Blog Assessment" chart (See Fig. 1), developed in 2008 is still a valuable process for organizations to tactically understand and process online comments at a basic level. Based on the research I have presented in this paper, and considering the rise of emotionally charged online negativity since 2008, I suggest more questions and considerations for an organization when responding to online comments, including:

- Is your organization paying more attention to one negative comment when the vast majority are positive, experiencing the effects of Pareto's 80/20 rule? (Wulczyn et al., 2017; Matei & Bruno, 2015; David, 2015)
- Is your organization being influenced by the negative content of one comment when the overall quality of feedback is positive, experiencing the effects of negativity bias? (Rozin & Royzman, 2001; Waddell & Sundar, 2017; Britton, 2017)
- Has your organization developed emotional intelligence? (Jalonen, 2014)
- What user characteristics are motivating the customer/stakeholder? (Matikainen, 2015; Sweeney, 2018a)
- Where is the user from and what are the user's demographics? Are there local concerns or circumstances related to their demographics? What social media network was the comment published on? (Jalonen, 2014)

- Is your organization hedonic or utilitarian? What voice is best suited to your organization's social media content? (Barcelos et al., 2018; ATCOenergy case study, 2019)
- Is a profanity filter enough to moderate negative comments on your social media company page or do you require human decision-making? (Saleem et al., 2017)
- Do you require a trained and dedicated resource for content moderation? (Gillespie, 2018; Klonick, 2018)
- Is your company currently the focus of negative media attention? Is this the best time for your organization to create content for social media? (SNC-Lavalin case study, 2019)

With the rise of prolific online negativity, a negative comment from a customer or stakeholder can feel amplified like a giant bullhorn. Given the stopping power of negative comments, it is easy to understand why an organization would choose not to risk tarnishing its image by engaging on social media. Ultimately, it is an organization's choice to participate in social media and, given the barrier of online negativity, the choice not to participate could be a valid strategic decision. But organizations that do choose to participate on social media must prepare to face online negativity, take on the responsibility of managing negative comments and not sweep negativity away. An organization that is well-prepared to engage on social media, applying learnings from research and case studies, including what I have presented in this paper, will find ways to silence the bullhorn of negativity before it can sound.

## References

- Air Force Public Affairs Agency – Emerging Technologies Division. (2008). *Air Force Blog Assessment*. [Image]. Retrieved from <https://freshspot.typepad.com/.a/6a00d83451f23a69e20105365f0d62970b-popup>
- ATCOenergy. (2019, February 2). What did the groundhog say this morning? He said, "Stop wasting energy. Turn down your thermostat at night and save on your electricity bill." Smart groundhog. In *Facebook*[Status Update]. Retrieved from <https://www.facebook.com/ATCOenergy/photos/a.1554877084834418/2213146849007435/?type=3&theater>
- Banduka, N., Mačuzić, I., Stojkić, Ž., Bošnjak, I., & Peronja, I. (2016). Using 80/20 principle to improve decision making at PFMEA. *Annals of DAAAM & Proceedings*, 27, 487–492. <https://doi-org.ezproxy.msvu.ca/10.2507/27th.daaam.proceedings.073>
- Barcelos, R. H., Dantas, D. C., & Sénécal, S. (2018). Watch your tone: How a brand's tone of voice on social media influences consumer responses. *Journal of Interactive Marketing (Elsevier)*, 41, 60–80. <https://doi-org.ezproxy.msvu.ca/10.1016/j.intmar.2017.10.001>
- Britton, C. (2017). How business leaders should respond to cyberbullying [Web log post]. Retrieved from <https://www.rockdovesolutions.com/blog/how-business-leaders-should-respond-to-cyberbullying>
- Chen, D. (2017). The 80-20 rule of trolls. *The Life and Times of David Chen*. Retrieved from <https://davechen.net/2017/02/the-80-20-rule-of-trolls/>
- David, E.D. (2015). Cyberbullying's new target: Big companies. *CNBC*. Retrieved from <https://www.cNBC.com/2015/03/27/cyberbullyings-new-target-big-companies.html>
- Facebook. (n.d.). *Profanity Filter Information*. Retrieved from <https://www.facebook.com>
- Gillespie, T. (2018). What platforms are, and what they should be. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media* (pp. 197-214). New Haven, CT: Yale University Press.
- Gollom, M. (2019, February 13). What you need to know about the SNC-Lavalin affair. CBC explains. *CBC News*. Retrieved from <https://www.cbc.ca/news/politics/trudeau-wilson-raybould-attorney-general-snc-lavalin-1.5014271>
- Jalonen, H. (2014). Negative emotions in social media as a managerial challenge. *Proceedings of the European Conference on Management, Leadership & Governance*, 128–135. Retrieved from <http://ezproxy.msvu.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=99239842&site=ehost-live&scope=site>

- Marigold. (2013, June 23). Re: Hidden comments on Facebook posts [Online discussion post]. Retrieved from <https://webapps.stackexchange.com/questions/45393/hidden-comments-on-facebook-posts>
- Hall, T. (2018). The new 80/20 rule of social media marketing: Creating your own ratio for social media that accelerates your brand strategy. *Inc.* Retrieved from <https://www.inc.com/tanya-hall/the-new-80/20-rule-of-social-media.html>
- Klonick, K. (2018). The new governors: The people, rules, and processes governing online speech. *Harvard Law Review*, 131(6), 1599–1670. Retrieved from <http://ezproxy.msvu.ca/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=129083311&site=ehost-live&scope=site>
- Koch, R. (1999). The 80/20 principle: the secret to achieving more with less. *Crown Business*, ISBN:0385491743, <http://leadershipcoachingblog.com/wp-content/uploads/2012/03/the-80-20-principle-to-achieve-more-with-less-effort1.pdf>
- Leung, L. (2013). Generational differences in content generation in social media: The roles of the gratifications sought and of narcissism. *Computers in Human Behavior*, 29(3), 997–1006.
- Matei, S. A., & Bruno, R. J. (2015). Pareto’s 80/20 law and social differentiation: A social entropy perspective. *Public Relations Review*, 41(2), 178–186. <https://doi-org.ezproxy.msvu.ca/10.1016/j.pubrev.2014.11.006>
- Matikainen, J. (2015). Motivations for content generation in social media. *Participations Journal of Audience & Reception Studies*, 12(1), 41-58. <http://www.participations.org/Volume%2012/Issue%201/4.pdf>
- Ontario SEO Targeted Digital Marketing. (2017). *Social Media Monitoring*. Retrieved from <https://www.ontarioseo.ca/social-media-monitoring>
- Owyang, J. (2008). Diagram: How the Air Force response to blogs. *Jeremiah Owyang*. Retrieved from <http://www.web-strategist.com/blog/2008/12/31/diagram-how-the-air-force-response-to-blogs/>
- Patel, N. (2019). Should You Delete Your Business Facebook Page? *NEILPATEL*. Retrieved from <https://neilpatel.com/blog/should-you-delete-facebook/>
- Reynolds C. (2019, February 12) SNC-Lavalin could face criminal charges over Montreal bridge contract, court documents reveal. *National Post*. Retrieved from <https://nationalpost.com/news/canada/snc-lavalin-faces-criminal-probe-over-montreal-bridge-contract-documents-reveal>

- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296–320. [https://doi.org/10.1207/S15327957PSPR0504\\_2](https://doi.org/10.1207/S15327957PSPR0504_2)
- Saleem, H., Dillion K., & Benesch, S., & Ruths, D. (2017). A web of hate: Tackling hateful speech in online social spaces. Retrieved from [https://www.researchgate.net/publication/320163517\\_A\\_Web\\_of\\_Hate\\_Tackling\\_Hateful\\_Speech\\_in\\_Online\\_Social\\_Spaces](https://www.researchgate.net/publication/320163517_A_Web_of_Hate_Tackling_Hateful_Speech_in_Online_Social_Spaces)
- Scott, D. M. (2008). The US Air Force: Armed with social media. *David Meerman Scott*. Retrieved from <https://www.davidmeermanscott.com/blog/2008/12/the-us-air-force-armed-with-social-media.html>
- SNC-Lavalin. (2019, March 8). Today is International Women’s Day! Look out for inspirational stories from the incredible women across our business throughout the day. Learn more here: [#IWD2019 #BalanceForBetter](http://bit.ly/2TjwyJa). In *Facebook* [Status Update]. Retrieved from <https://www.facebook.com/snclavalin/videos/359880481524932/>
- Sweeney, J. (2018a, May 30). How to deal with negative comments in your social media campaign. *Digital Training Institute*. Retrieved from <https://digitaltraininginstitute.ie/how-to-deal-with-negative-comments/>
- Sweeney, J. (2018b, August 1). The perfect social media community manager. *Digital Training Institute*. Retrieved from <https://digitaltraininginstitute.ie/perfect-social-media-community-manager/>
- Tadić, B., Gligorijević, V., Mitrović, M., & Šuvakov, M. (2013). Co-evolutionary mechanisms of emotional bursts in online social dynamics and networks. *Entropy*, 15(12), 5084–5120.
- Tenove, C., Tworek, H.J.S., & McKelvy, F., (2018). Poisoning democracy: How Canada can address harmful speech online. *Public Policy Forum*. Retrieved from <https://ppforum.ca/publications/poisoning-democracy-what-can-be-done-about-harmful-speech-online/>
- The Pareto Principle. (n.d.). *California State University Northridge*. Retrieved from <http://www.csun.edu/~jmotil/Pareto.pdf>
- Vaidhyanathan, S. (2018). Conclusion: The nonsense machine. *Antisocial media: How Facebook disconnects us and undermines democracy* (pp.196-221). Oxford, UK: Oxford University Press.
- Waddell, T. F., & Sundar, S. S. (2017). pass:[#]thisshowsucks! The overpowering influence of negative social media comments on television viewers. *Journal of Broadcasting & Electronic Media*, 61(2), 393–409. <https://doi-org.ezproxy.msvu.ca/10.1080/08838151.2017.1309414>

Wulczyn, E., Thain, N., & Dixon, L. (2017). Ex Machina: Personal attacks seen at scale.  
*Proceeding WWW '17 Proceedings of the 26th International Conference on World Wide  
Web* (pp. 1391-1399). 10.1145/3038912.3052591