

# **Trending Now: The Role of Defamation Law in Remediating Harm from Social Media Backlash**

## *Abstract*

*Defamatory comments on social media have become commonplace. When the online community is outraged by some event, social media users often flood the Internet with hateful and false comments about the alleged perpetrator, feeling empowered by their numbers and anonymity. This wave of false and harmful information has caused many individuals to lose their jobs and suffer severe emotional trauma.*

*This Comment explores whether the target of social media backlash can bring a successful defamation claim against the users who have destroyed their reputations on and offline. Notably, one of the biggest hurdles these plaintiffs will face is the public figure doctrine, which requires public figures to prove by clear and convincing evidence that the defamatory statements were made with actual malice, a much higher standard than the usual negligence standard. This Comment concludes that courts should hold that the average social media user does not meet the public figure test and therefore only needs to prove the defamatory social media posts were made negligently.*

TABLE OF CONTENTS

I. INTRODUCTION .....	431
II. COMMUNICATION ON SOCIAL MEDIA SITES .....	434
A. <i>The Prevalence of Online Abuse and Harassment</i> .....	437
B. <i>A New Defamatory Phenomenon: Social Media Backlash</i> .....	438
III. DEFAMATION CLAIMS TODAY .....	441
A. <i>The Development of Defamation Law</i> .....	441
B. <i>Public Figures Versus Private Individuals</i> .....	445
C. <i>Lower Courts Struggle to Apply the Public-Figure Test</i> .....	449
1. Access to Media .....	449
2. Public Controversies .....	450
3. Voluntary Involvement in Public Controversies .....	451
IV. OTHER ISSUES PLAINTIFFS WILL ENCOUNTER .....	452
A. <i>Twitter and Facebook Will Not Be Held Liable</i> .....	452
B. <i>Anonymous Social Media Speakers</i> .....	453
C. <i>Opinion or Actionable False Statement?</i> .....	454
V. HOW THE PUBLIC-FIGURE TEST SHOULD BE EMPLOYED IN SOCIAL MEDIA BACKLASH SCENARIOS .....	459
A. <i>Targets of Social Media Backlash Do Not Have the         Requisite Access to Media to Counteract False Statements</i> ....	461
B. <i>Mob Shaming of a Social Media User Should Not Be         Considered a Matter of Public Concern</i> .....	467
C. <i>A Single, Isolated Social Media Post Is Not Enough to         Become Voluntarily Involved in a Controversy</i> .....	470
VI. CONCLUSION .....	474

*The snowflake never needs to feel responsible for the avalanche.*<sup>1</sup>

## I. INTRODUCTION

#HasJustineLandedYet: these twenty characters would dominate social media for hours but would scar one woman’s reputation and livelihood for years to come.<sup>2</sup> On December 20, 2013, Justine Sacco was in Heathrow Airport about to board a plane to Africa.<sup>3</sup> She had been traveling over the holidays, shooting off tweets to keep in touch with her followers.<sup>4</sup> Justine attempted a joke: “Going to Africa. Hope I don’t get AIDS. Just kidding. I’m white!”<sup>5</sup> A half hour later, she boarded her eleven-hour flight, disappointed that she had received no replies from her 170 followers.<sup>6</sup> Between takeoff and touchdown, the ill-advised tweet would become *the* trending topic on Twitter and not for reasons she would hope: one hundred thousand tweets condemned Justine as an insensitive racist who was too incompetent to work in public relations.<sup>7</sup> Social media users demanded her job, and they got it—Justine’s employer fired her before she could even check her Twitter feed.<sup>8</sup>

Justine’s story is not a unique one—it is just one glimpse of the extent to which digital public shaming can ruin a person’s reputation, both on and offline.<sup>9</sup> Amongst the many instances of emotional abuse on Twitter exists the social media backlash scenario: when an individual posts a reprehensible

---

1. JON RONSON, SO YOU’VE BEEN PUBLICLY SHAMED 56 (2015).

2. *See id.* at 67–79.

3. *See id.* at 67–68.

4. *See id.* at 68.

5. *Id.*

6. *See id.*

7. *See id.* at 69–70. Some of the tweets include: “How did @JustineSacco get a PR job?! Her level of racist ignorance belongs on Fox News. #AIDS can affect anyone!”, “No words for that horribly disgusting, racist as fuck tweet from Justine Sacco. I am beyond horrified”, “I’m an IAC employee and I don’t want @JustineSacco doing any communications on our behalf ever again. Ever”, “Everyone go report this cunt @JustineSacco”, “This is an outrageous, offensive comment. Employee in question currently unreachable on an intl flight.” *Id.*

8. *See id.*

9. *See infra* notes 69–73 and accompanying text.

tweet and social media users backlash with hateful criticisms, threats, and false statements.<sup>10</sup> These online defamers create digital mobs of condemnation where no single user feels responsible for the ensuing real-life consequences of their actions.<sup>11</sup> And if they realize their collective power, they may believe the individual deserved to have her reputation damaged for her perceived moral transgressions.<sup>12</sup>

This digital phenomenon has threatened reputations to unprecedented magnitudes.<sup>13</sup> Social media users are free to instantaneously praise or condemn others on a whim, and social media platforms encourage just that through the now-ubiquitous “hashtag,” which facilitates widespread, real-time commentary on virtually any topic.<sup>14</sup> Further, social media backlash uniquely removes the responsibility one particular poster of defamatory content might feel because typically there are thousands of users tweeting abusive and defamatory comments at one time.<sup>15</sup> Defamatory tweets have caused real consequences in real people’s lives.<sup>16</sup> But unapologetic users have felt empowered by their numbers and anonymity, forgetting that the rule of law applies even to online conduct.<sup>17</sup> To ensure that mob justice does not rule over legal justice, courts must carefully apply defamation law to social media backlash.<sup>18</sup> Perhaps it is time for the snowflake to start being held responsible for the avalanche.

Throughout history, defamation law has provided a means for recovery when a person’s reputation has been wrongfully disparaged.<sup>19</sup> Congress and the Supreme Court have greatly valued online speech, but at what point should the First Amendment no longer protect this speech?<sup>20</sup> When does speech contribute such a low level of value to the marketplace that the court system should punish the speaker and repair the plaintiff’s shambled

---

10. *See infra* Part II.

11. *See infra* Part II.

12. *See infra* Part II.

13. *See infra* Part II.

14. *See infra* Part II.

15. *See infra* Part II.

16. *See infra* Part V.

17. *See infra* Part V.

18. *See infra* Part V.

19. *See infra* Section III.A.

20. *See infra* Part III.

reputation?<sup>21</sup>

Victims of defamatory statements on social media have increasingly begun suing for defamation, and courts faced with cases involving harmful comments on social media have been applying traditional defamation principles.<sup>22</sup> These well-established principles focus on balancing people's First Amendment free speech rights and the state's interest in protecting people's reputations.<sup>23</sup> The First Amendment was first applied to defamatory speech in *New York Times Co. v. Sullivan*,<sup>24</sup> which established protections for individuals who speak out against public officials by requiring the officials to prove that the allegedly defamatory statements were made with "actual malice."<sup>25</sup> In *Curtis Publishing Co. v. Butts*,<sup>26</sup> the actual malice requirement was extended to public figures, defined as individuals who are "intimately involved in the resolution of important public questions" or "shape events in areas of concern to society."<sup>27</sup>

If the target of social media backlash, such as Justine Sacco, were to sue for defamation, the court would have to determine whether that average social media user should be considered a public figure.<sup>28</sup> If found to be a public figure, the plaintiff would be required to prove that the defendant made the statements with actual malice, meaning "with knowledge that it was false or with reckless disregard of whether it was false or not."<sup>29</sup> The actual malice standard is difficult to meet and, in the social media backlash context, would likely mean that the social media user would not recover

---

21. *See infra* Part III. The First Amendment allows "restrictions upon the content of speech in a few limited areas, which are 'of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality.'" *Virginia v. Black*, 538 U.S. 343, 358–59 (2003) (quoting *Chaplinsky v. New Hampshire*, 315 U.S. 568, 572 (1942)).

22. *See infra* Part V.

23. *See infra* Part III.

24. 376 U.S. 254 (1964).

25. *Id.* at 279–80.

26. 388 U.S. 130 (1967).

27. *See id.* at 342 ("Those who, by reason of the notoriety of their achievements or the vigor and success with which they seek the public's attention, are properly classed as public figures . . . may recover for injury to reputation . . .").

28. *See infra* Part V.

29. *Sullivan*, 376 U.S. at 280.

against the abusive posters.<sup>30</sup> This Comment will argue that average social media users should not be considered public figures because they do not meet the three characteristics the Supreme Court has emphasized—access to media, existence of a public controversy, and voluntary involvement in a public controversy—and thus, they do not need to prove actual malice.<sup>31</sup>

Part II of this Comment will survey instances of social media backlash and explain how defamatory communications on social media sites are uniquely harmful to an individual's reputation.<sup>32</sup> Part III will discuss the development of defamation law, including the rationale and application of the public-figure test.<sup>33</sup> Part IV will briefly outline other limitations that targets of social media defamation face.<sup>34</sup> Part V will discuss why the average social media user should not be deemed a public figure.<sup>35</sup> Part VI concludes that there is a need to carefully apply defamation law to social media backlash scenarios.<sup>36</sup>

## II. COMMUNICATION ON SOCIAL MEDIA SITES

When compared to other Internet sites, social media is uniquely structured to encourage a large number of individuals to instantaneously discuss matters of little or much social importance.<sup>37</sup> Since their creation, social media sites have steadily increased in popularity, and as a result, users have become so connected to their online selves that their social media profiles have become extensions of their physical lives.<sup>38</sup> And as more and

---

30. See *infra* notes 103–07 and accompanying text.

31. See *infra* Part V.

32. See *infra* Part II.

33. See *infra* Part III.

34. See *infra* Part IV.

35. See *infra* Part V.

36. See *infra* Part VI.

37. See *infra* notes 45–48 and accompanying text.

38. See Matthew Lafferan, *Do Facebook and Twitter Make You a Public Figure?: How to Apply the Gertz Public Figure Doctrine to Social Media*, 29 SANTA CLARA COMPUTER & HIGH TECH. L.J. 199, 209–10 (2013). One anthropologist even argues that our social media presence has become like a second self:

Whether you like it or not, you're starting to show up online, and people are interacting with your second self when you're not there. And so you have to be careful about leaving your front lawn open, which is basically your Facebook wall, so that people don't write on it in the middle of the night—because it's very much the equivalent. And

more individuals peruse the Internet to evaluate their interests, characteristics, and values, their digital footprint has come to define them in the physical world.<sup>39</sup>

As of June 30, 2016, there were 313 million monthly active users on Twitter.<sup>40</sup> Twitter's mission is "[t]o give everyone the power to create and share ideas and information instantly, *without barriers*."<sup>41</sup> Users can tweet messages with up to 140 characters of text and follow other users.<sup>42</sup> The

suddenly we have to start to maintain our second self. You have to present yourself in digital life in a similar way that you would in your analog life. So, in the same way that you wake up, take a shower and get dressed, you have to learn to do that for your digital self.

Amber Case, *We Are All Cyborgs Now*, TED (Jan. 2011), [http://www.ted.com/talks/amber\\_case\\_we\\_are\\_all\\_cyborgs\\_now/transcript?language=en](http://www.ted.com/talks/amber_case_we_are_all_cyborgs_now/transcript?language=en).

39. See RONSON, *supra* note 1, at 2.

40. See *Company*, TWITTER, <https://about.twitter.com/company> (last updated June 30, 2016). In comparison, Facebook had 1.79 billion monthly active users as of September 30, 2016. See *Stats*, FACEBOOK NEWSROOM, <http://newsroom.fb.com/company-info/> (last visited Nov. 7, 2016). As of 2016, LinkedIn had 467 million registered members. See *About Us*, LINKEDIN <https://press.linkedin.com/about-linkedin> (last visited Nov. 7, 2016). This Comment will focus mainly on Twitter; however, the analysis is applicable to any of the three main social media sites: Twitter, Facebook, and LinkedIn. See Lafferman, *supra* note 38, at 209.

Instagram, a social media application centered around the sharing of photos, has also become extremely popular, now boasting over 500 million monthly active users. See *Press News*, INSTAGRAM <https://www.instagram.com/press/?hl=en> (last visited Nov. 7, 2016); see also Arjun Kharpal, *Facebook's Instagram Hits 400M Users, Beats Twitter*, CNBC (Sept. 23, 2015, 5:58 AM), <http://www.cnbc.com/2015/09/23/instagram-hits-400-million-users-beating-twitter.html>. But, because defamation necessarily involves false *statements*, it is unlikely that a plaintiff will have a successful defamation suit regarding photos posted on Instagram. See, e.g., *Binion v. O'Neal*, No. 14-13454, 2015 WL 3544518 (E.D. Mich. Apr. 2, 2015). In *Binion*, the plaintiff, a man who suffers from a rare genetic condition, sued for defamation after another Instagram user reposted his photo. See *id.* at \*2. Because the plaintiff did not allege a single "statement" and the argument that the photo itself ridiculed the plaintiff was meritless, his complaint was dismissed. *Id.* at \*5. The court noted that "[a]t most, Burke posted a true and accurate photograph of Plaintiff—the same photograph that Plaintiff posted on his public Instagram account." *Id.* But see *Rapper The Game Must Pay Nanny \$200,627 over Instagram Libel*, L.A. DAILY NEWS (June 29, 2015), <http://www.dailynews.com/general-news/20150629/rapper-the-game-must-pay-nanny-200627-over-instagram-libel-faces-arraignment-in-separate-case> (finding a successful libel claim for *statements* posted on Instagram).

41. *Company*, *supra* note 40 (emphasis added). Facebook's mission is "to give people the power to share and make the world more open and connected. People use Facebook to stay connected with friends and family, to discover what's going on in the world, and to share and express what matters to them." *Our Mission*, FACEBOOK NEWSROOM, <http://newsroom.fb.com/company-info/> (last visited Nov. 17, 2016).

42. See *Posting a Tweet*, TWITTER HELP CTR., <https://support.twitter.com/articles/15367> (last

average user has 208 followers.<sup>43</sup> If a user's page is public, anyone may read a user's tweets without having a personal Twitter account.<sup>44</sup> A popular function of Twitter involves groups of users collectively commenting on a single topic through the use of a hashtag—created by using the “#” symbol—categorizing tweets into easily searchable groups.<sup>45</sup> Clicking on a hashtag takes the user to all the tweets that are in the same category, thereby increasing the tweet's publicity.<sup>46</sup> If the hashtag is popular enough,<sup>47</sup> it can become a trending topic on Twitter.<sup>48</sup> Users can also retweet or reply to another user's tweet<sup>49</sup> or mention other users, further increasing instantaneous social discussions.<sup>50</sup>

---

visited Nov. 7, 2016) (“A Tweet may contain photos, videos, links and up to 140 characters of text.”); *Getting Started with Twitter*, TWITTER HELP CTR., <https://support.twitter.com/articles/215585> (last visited Nov. 7, 2016) (“Follow accounts to see all their Tweets.”).

43. Craig Smith, *Twitter Statistics and Facts*, DMR, <http://expandedramblings.com/index.php/march-2013-by-the-numbers-a-few-amazing-twitter-stats/> (last updated Nov. 2, 2016).

44. See Adeline A. Allen, *Twibel Retweeted: Twitter Libel and the Single Publication Rule*, 15 J. HIGH TECH. L. 63, 75–76 (2014).

45. *Id.* at 76; see also *Using Hashtags on Twitter*, TWITTER HELP CTR., <https://support.twitter.com/articles/49309?lang=en> (last visited Nov. 7, 2016).

46. Allen, *supra* note 44, at 76.

47. There is no definitive way to know if a popular hashtag will turn into a trending topic on Twitter. See *FAQs About Trends on Twitter*, TWITTER HELP CTR., <https://support.twitter.com/articles/101125?lang=en> (last visited Nov. 7, 2016).

48. See, e.g., Patrick H. Hunt, *Tortious Tweets: A Practical Guide to Applying Traditional Defamation Law to Twibel Claims*, 73 LA. L. REV. 559, 580 (2013) (“By using hashtags to mark a tweet's relevance to a particular category, the user potentially increases his tweet's publicity.”). A trending topic is one that is “popular now . . . to help [users] discover the hottest emerging topics of discussion on Twitter that matter most.” *FAQs About Trends on Twitter*, *supra* note 47.

49. See *FAQs About Retweets*, TWITTER HELP CTR., <https://support.twitter.com/articles/77606?lang=en> (last visited Nov. 7, 2016). “A Retweet is a re-posting of [someone else's] Tweet. Twitter's Retweet feature helps you and others quickly share that Tweet with all of your followers.” *Id.*

50. See *What Are Replies and Mentions?*, TWITTER HELP CTR., <https://support.twitter.com/articles/14023?lang=en> (last visited Nov. 7, 2016). Replies are direct responses to other tweets, while mentions are tweets that include another user's name in the text. See *id.*

A. *The Prevalence of Online Abuse and Harassment*

As Twitter has become more popular, Twibel<sup>51</sup> has been on the rise.<sup>52</sup> Social media makes name-calling easy,<sup>53</sup> and many users rant without thinking of the consequences of their words.<sup>54</sup> Twitter has been criticized for allowing these abusive, harassing, and violent tweets to abound on its site.<sup>55</sup> On December 29, 2015, Twitter announced that it would take down “abusive behavior and hateful conduct”<sup>56</sup> after women and minorities complained of its platform having an ineffective abuse policy.<sup>57</sup> Dick

51. See Hunt, *supra* note 48, at 559 n.1. Twibel is the term used to describe libelous posts on Twitter. See *id.*

52. See *id.* at 588.

53. See, e.g., Kraig J. Marton et al., *Protecting One’s Reputation—How to Clear a Name in a World Where Name Calling Is So Easy*, 4 PHOENIX L. REV. 53, 56 (2010).

54. See, e.g., Eric Posner, *A Terrible Shame: Enforcing Moral Norms Without the Law Is No Way to Create a Virtuous Society.*, SLATE (Apr. 9 2015, 11:14 AM), [http://www.slate.com/articles/news\\_and\\_politics/view\\_from\\_chicago/2015/04/internet\\_shaming\\_the\\_legal\\_history\\_of\\_shame\\_and\\_its\\_costs\\_and\\_benefits.2.html](http://www.slate.com/articles/news_and_politics/view_from_chicago/2015/04/internet_shaming_the_legal_history_of_shame_and_its_costs_and_benefits.2.html).

The major effect of social media is that it enables people to broadcast an opinion—or, more accurately, a gut reaction—to the whole world, instantly, without pausing to give it any thought. This, combined with pervasive anonymity and traditional animosity to anyone who acts or thinks unconventionally, has awoken atavistic instincts that are multiplied a hundredfold through herd mentality.

*Id.* However, the law may hold ranting Twitter users more accountable than they might think. See *In re Does 1–10*, 242 S.W.3d 805, 820 (Tex. App. 2007) (“[A]nonymous (electronic) speakers may not freely defame individuals without facing civil responsibility for their acts.”).

55. See Sara Ashley O’Brien, *Twitter Crackdown on Hate Speech Backfires*, CNN (Jan. 10, 2016), <http://money.cnn.com/2016/01/10/technology/twitter-hate-speech-crackdown-milo-yiannopoulos/>; see also Johnny Holschuh, *#civilrightscybertorts: Utilizing Torts to Combat Hate Speech in Online Social Media*, 82 U. CIN. L. REV. 953, 953–54 (2014) (“The rise of Twitter and Facebook provide new and unique forums through which individuals can target other individuals and spread hate, fear, and intimidation.”).

56. O’Brien, *supra* note 55.

57. See Yoree Koh, *Twitter, in Punishing a Controversial User, Stokes Freedom of Speech Debate*, WSJ (Jan. 11, 2016, 6:20 PM), <http://blogs.wsj.com/digits/2016/01/11/twitter-in-punishing-a-controversial-user-stokes-freedom-of-speech-debate/>. Twitter’s updated abuse and harassment policy states:

[Users] may not incite or engage in the targeted abuse or harassment of others. Some of the factors that we may consider when evaluating abusive behavior include: if a primary purpose of the reported account is to harass or send abusive messages to others; if the reported behavior is one-sided or includes threats; if the reported user is inciting others to harass another user; and if the reported user is sending harassing messages to an account from multiple accounts.

Costolo, Twitter's former CEO, even took personal responsibility for the rampant problems: "I'm frankly ashamed of how poorly we've dealt with this issue during my tenure as CEO. It's absurd. There's no excuse for it. I take full responsibility for not being more aggressive on this front. It's nobody else's fault but mine, and it's embarrassing."<sup>58</sup> Twitter has since updated its abuse policies, but it is currently unclear how proactively Twitter will seek out and remove abusive speech.<sup>59</sup>

### B. *A New Defamatory Phenomenon: Social Media Backlash*

Abusive and hateful comments take on additional power when they are part of a large group.<sup>60</sup> One commentator has described social media backlashes as the "digital mob."<sup>61</sup> Merriam-Webster defines a mob as a "large or disorderly crowd; especially: one bent on riotous or destructive action."<sup>62</sup> The term is fitting: it has become a regular occurrence for social media users to jointly and aggressively punish those who, in one way or another, exhibited bad judgment in what they posted online.<sup>63</sup> Many of the digital mob's hateful comments would never be said to these people's

*The Twitter Rules*, TWITTER HELP CTR., <https://support.twitter.com/articles/18311> (last visited Nov. 7, 2016). As a point of comparison, Facebook's bullying and harassment policy states:

We don't tolerate bullying or harassment. We allow you to speak freely on matters and people of public interest, but remove content that appears to purposefully target private individuals with the intention of degrading or shaming them. This content includes, but is not limited to: [p]ages that identify and shame private individuals, [i]mages altered to degrade private individuals, [p]hotos or videos of physical bullying posted to shame the victim, [s]haring personal information to blackmail or harass people, and [r]epeatedly targeting other people with unwanted friend requests or messages.

*Community Standards*, FACEBOOK, <https://www.facebook.com/communitystandards> (last visited Nov. 7, 2016) (follow "Bullying and Harassment" hyperlink).

58. Nitasha Tiku & Casey Newton, *Twitter CEO: 'We Suck at Dealing with Abuse'*, VERGE (Feb. 4, 2015, 9:25 PM), <http://www.theverge.com/2015/2/4/7982099/twitter-ceo-sent-memo-taking-personal-responsibility-for-the>.

59. *See id.*

60. *See, e.g.*, Conor Friedersdorf, *A Social-Media Mistake Is No Reason to Be Fired*, ATLANTIC (Feb. 18, 2015), <http://www.theatlantic.com/politics/archive/2015/02/being-shamed-on-social-media-should-not-get-people-fired/385597/>.

61. *Id.*

62. *Mob*, MERRIAM-WEBSTER, <http://www.merriam-webster.com/dictionary/mob> (last visited Nov. 7, 2016).

63. *See* Friedersdorf, *supra* note 60.

faces,<sup>64</sup> yet users feel free to broadcast their negative opinions of them on social media platforms at virtually no cost.<sup>65</sup> Even with our efforts to preserve how we are viewed online,<sup>66</sup> each of us is vulnerable to the great power of social media backlash. Feeling free to abuse and defame others online, social media users engage in public shamings every day.<sup>67</sup> Well-known victims<sup>68</sup> of shamings include Justine Sacco,<sup>69</sup> Walter Palmer,<sup>70</sup> Kendall Jones,<sup>71</sup> Jonah Lehrer,<sup>72</sup> and Lindsay Stone.<sup>73</sup> These individuals and others not mentioned have lost jobs and suffered emotional harm as a result of their shamings.<sup>74</sup> A social media user's ill-considered response to another

64. See, e.g., *supra* note 7; Erik P. Lewis, *Unmasking "Anon12345": Applying an Appropriate Standard When Private Citizens Seek the Identity of Anonymous Internet Defamation Defendants*, 2009 U. ILL. L. REV. 947, 953 (2009) (arguing that when users post anonymously, they "might use language in ways that they would never choose to speak in public, or even privately with close friends"). This is especially true because many users only communicate through social media and do not actually know each other in the real world. See *id.*

65. See Marton et al., *supra* note 53, at 60.

66. See, e.g., *id.* at 59–60. Despite our efforts to only make politically correct comments on Facebook or post attractive or conservative photos on Instagram, even the most innocuous post can reveal information we might try to hide from the Internet. See Jennifer Goldbeck, *Smart People Prefer Curly Fries*, SLATE (Oct. 7, 2014, 7:48 AM), [http://www.slate.com/articles/technology/future\\_tense/2014/10/youarewhatyoulike\\_find\\_out\\_what\\_algorithms\\_can\\_tell\\_about\\_you\\_based\\_on\\_your.html](http://www.slate.com/articles/technology/future_tense/2014/10/youarewhatyoulike_find_out_what_algorithms_can_tell_about_you_based_on_your.html). New computer algorithms can find patterns in massive amounts of social media activity and predict a certain user's set of traits, such as his or her gender, religion, and sexual orientation. See *id.* For some, these are traits that close family and friends might not even know. See *id.*

67. See *infra* notes 69–73 and accompanying text.

68. Significantly, none of these individuals, except for Jonah Lehrer, were well-known before their public shaming. See *infra* notes 70–73.

69. See *supra* notes 2–8 and accompanying text.

70. Todd Leopold, *The Price of Public Shaming in the Internet Age*, CNN (Apr. 16, 2015), <http://www.cnn.com/2015/04/16/living/feat-public-shaming-ronson/>. Walter Palmer was vilified online after the online community discovered that he hunted and killed Cecil the lion in Zimbabwe. *Id.* "Twitter users called for him to be shot and skinned after the famous lion was found dead after an alleged 40-hour hunt." *Id.*

71. *Id.* Kendall Jones sparked online outrage after she posted photos of her African safari kills online. *Id.*

72. *Id.* Jonah Lehrer, a successful journalist, was accused of plagiarizing passages of a book, and when he attempted to publicly apologize, a live Twitter feed brutally ripped him apart in his plain sight. *Id.*

73. *Id.* After Lindsey Stone posted a photo on Facebook of herself flipping off an Arlington National Cemetery tombstone, her ill-advised attempt at humor caused her to lose her job working with autistic children. *Id.*

74. See *supra* notes 69–73; see, e.g., Todd Leopold, *The Price of Public Shaming in the Internet Age*, CNN (Apr. 16, 2015, 12:37 PM), <http://www.cnn.com/2015/04/16/living/feat-public-shaming->

user's online activity can threaten a person's reputation more than a verbally spoken statement would.<sup>75</sup> Through the use of hashtags, defamatory comments easily spread through Twitter and encourage other users to take part in the public shaming.<sup>76</sup> If thousands of individuals use similar hashtags to shame another, the event can become a top trending topic on Twitter, which further widens the community of users that have read the false and defamatory comments.<sup>77</sup> Additionally, these defamatory comments are stored indefinitely online so that the targeted user will never be able to fully regain a positive reputation.<sup>78</sup> Even individuals that do not have a Twitter account can easily access these comments through the use of a search engine, such as Google.<sup>79</sup>

---

ronson/ (mentioning other victims of social media backlash).

75. See Posner, *supra* note 54 (“[T]hese ill-considered reactions are stored indefinitely, while being immediately accessible to anyone, thanks to the efficiency of search engines.”); see *infra* notes 76–79 and accompanying text.

76. See *supra* notes 45–48 and accompanying text.

77. See *supra* notes 45–48 and accompanying text.

78. See Posner, *supra* note 54. Even if a site takes down a harassing comment, there are online platforms that document social media takedowns. See *id.*; see, e.g., ONLINECENSORSHIP.ORG, <https://onlinecensorship.org/about/what-we-do> (last visited Sept. 21, 2016) (“Has your post . . . been blocked or deleted? Onlinecensorship.org gives social media users a place to report takedown notices and catalog content that has been removed.”). Because reputation-harming comments are typically online indefinitely unless the site takes them down, some have resorted to “search engine optimization,” where the defamed person attempts to increase the “good information” about herself so that Internet users see that information first and must continue using the search engine to find the defamatory information. Marton et al., *supra* note 53, at 81. However, using search engine optimization is controversial and may put an individual at risk of more damage to his or her reputation. See *id.* at 82.

Jon Ronson, a journalist who wrote the bestselling book *So You've Been Publicly Shamed*, offered one of his interviewees, Lindsay Stone, one of these online reputation-management services called Reputation.com as an incentive to talk to him. See RONSON, *supra* note 1, at 218–19, 224. Michael Fertik, the owner of the company, thought of the concept of online reputation management while clerking for the U. S. Court of Appeals for the Sixth Circuit. See *id.* at 219. His services typically cost hundreds of thousands of dollars. *Id.* at 226.

His plan was to create Lindsey Stone Tumblrs and LinkedIn pages and WordPress blogs and Instagram accounts and YouTube accounts to overwhelm that terrible photograph, wash it away in a tidal wave of positivity, away to a place on Google where normal people don't look—a place like page two of the search results. According to Google's own research into our “eye movements,” 53 percent of us don't go beyond the first two search results, and 89 percent don't look down past the first page. What the first page looks like . . . determines what people think of you.

*Id.* at 264–65.

79. See Posner, *supra* note 54; see also Marton et al., *supra* note 53, at 81 (“Today, many people

Most commentators have identified this phenomenon as a form of online bullying<sup>80</sup> or view it as a way the public has taken the law into its own hands.<sup>81</sup> But to what extent does online bullying have remedy in the law?<sup>82</sup> One potential avenue for recovery is the law of defamation, which traditionally has been invoked to remedy harm to a person's reputation.<sup>83</sup>

### III. DEFAMATION CLAIMS TODAY

#### A. *The Development of Defamation Law*

Most scholars agree that one of tort law's primary goals is to remedy harmed individuals.<sup>84</sup> When an individual's reputation has been harmed by the comments of others, relief is primarily achieved through defamation law.<sup>85</sup> Because the victims of social media shaming have undoubtedly

---

find defamatory material online by using search engines, like Bing, Google, or Yahoo.”).

80. See, e.g., Lindsay Hoffman, *Public Shaming on the Internet*, HUFFINGTON POST (May 11, 2015), [http://www.huffingtonpost.com/lindsay-hoffman/public-shaming-on-the-internet\\_b\\_6849036.html](http://www.huffingtonpost.com/lindsay-hoffman/public-shaming-on-the-internet_b_6849036.html).

81. See Friedersdorf, *supra* note 60.

82. See Holschub, *supra* note 55. One commentator has analyzed how and when social media speech qualifies as hate speech and how tort claims might provide a mechanism to counteract these attacks. *Id.*

Another potential avenue for victims of social media bullying is the “true threats” doctrine. See *Elonis v. United States*, 135 S. Ct. 2001, 2007–08 (2015). However, a recent Supreme Court opinion made it more difficult to prove that social media comments amount to a true threat. See *id.* at 2012. In *Elonis*, a husband posted threatening comments on Facebook, including violent language and graphic images, after his wife left him. See *id.* at 2004–05. Some posts included explicit statements that he was simply exercising his First Amendment rights. *Id.* at 2005. The Supreme Court rejected that *Elonis*'s statements should be viewed from a reasonable person's perspective—that a reasonable person would regard the posts as a true threat. See *id.* at 2011. Rather, a defendant has the required mental state to be convicted of a true threat if he or she “transmits a communication for the purpose of issuing a threat, or with knowledge that the communication will be viewed as a threat.” *Id.* at 2012. The Court decided this case solely on statutory grounds—whether the defendant violated 18 U.S.C. § 875(c) (1994), which makes true threats a federal crime—and did not reach the First Amendment issues. See *id.*

83. See *infra* Section III.A.

84. See, e.g., Lafferman, *supra* note 38, at 224–25.

85. See RESTATEMENT (SECOND) OF TORTS § 559 (AM. LAW INST. 1977). This Comment focuses exclusively on defamation, but victims of social media backlash might also claim intentional infliction of emotional distress (IIED), which provides for recovery when “[a]n actor who by extreme and outrageous conduct intentionally or recklessly causes severe emotional harm to another

suffered damage to their reputations, determining whether defamation law can serve as a basis for recovery is important.<sup>86</sup>

A defamatory communication can be defined as a statement that “tends so to harm the reputation of another as to lower him in the estimation of the community or to deter third persons from associating or dealing with him.”<sup>87</sup> Once Justine Sacco was labeled a white, privileged racist, the online community undoubtedly thought less of her, some actively hating her; there is no doubt that the hateful comments made about her deterred others from associating with her.<sup>88</sup> To successfully recover under a defamation claim, a plaintiff must typically prove the following elements: (1) the defendant published a defamatory statement; (2) that was false; (3) that was of and concerning the plaintiff; (4) the defendant had some degree of fault; and (5) the statement damaged the plaintiff.<sup>89</sup> Slander is oral defamation, while libel

---

is subject to liability for that emotional harm and, if the emotional harm causes bodily harm, also for the bodily harm.” See RESTATEMENT (THIRD) OF TORTS § 46 (AM. LAW INST. 2012). The largest hurdle for recovering under an IIED claim would be whether an abusive tweet would qualify as “extreme and outrageous conduct.” See Holschuh, *supra* note 55, at 963–64.

The victim of social media backlash might also claim tortious interference with contract because many social media users demand that employers fire the employee for his or her conduct. See RESTATEMENT (SECOND) OF TORTS § 766 (AM. LAW INST. 1979) (“One who intentionally and improperly interferes with the performance of a contract (except a contract to marry) between another and a third person by inducing or otherwise causing the third person not to perform the contract, is subject to liability to the other for the pecuniary loss resulting to the other from the failure of the third person to perform the contract.”).

86. See RONSON, *supra* note 1. For example, Justine Sacco lost the career of her dreams and suffered emotional distress after the social media backlash on Twitter. *Id.* at 79. Similarly, Lindsay Stone lost her job, “fell into a depression, became an insomniac, and barely left home for a year.” *Id.* at 210; see also Holschuh, *supra* note 55, at 953–54 (“The damage caused by social media hate speech is not trivial and can cause serious physical, emotional, and economic harm to victims.”).

87. RESTATEMENT (SECOND) OF TORTS § 559 (AM. LAW INST. 1977). Another court defines defamation as “the making of a false statement which tends to expose the plaintiff to public contempt, ridicule, aversion or disgrace, or induce an evil opinion of him in the minds of right-thinking persons, and to deprive him of their friendly intercourse in society.” *Franklin v. Daily Holdings, Inc.*, 21 N.Y.S.3d 6, 10 (App. Div. 2015).

88. See *supra* notes 2–8 and accompanying text.

89. RESTATEMENT (SECOND) OF TORTS § 558 (AM. LAW INST. 1977). These elements, however, vary from state to state because defamation is a state common law issue. Hunt, *supra* note 48, at 564–65. “Communications are often defamatory because they tend to expose another to hatred, ridicule or contempt. A defamatory communication may tend to disparage another by reflecting unfavorably upon his personal morality or integrity.” *Id.* § 559 cmt. b. The plaintiff has the burden of proving that the statement is false. See *id.* § 581A cmt. b.

is printed, so libel often arises in the social media context.<sup>90</sup> Social media libel has been termed “Twibel” because of its unique nature and how frequently it occurs.<sup>91</sup>

A discussion of how defamation law has developed will assist in determining how to apply the existing framework to the social media context.<sup>92</sup> Defamation law has had two distinct phases.<sup>93</sup> Initially, under the common law, expressing an “opinion could be defamatory if the expression was sufficiently derogatory of another as to cause harm to his reputation, so as to lower him in the estimation of the community or to deter third persons from associating or dealing with him.”<sup>94</sup> The traditional common law scheme required the speaker to establish the statement’s truth as an affirmative defense.<sup>95</sup> The justification for this strict liability regime was to limit the traffic of information to the truth when dealing with the reputations of others.<sup>96</sup> Under this early common law, recovery for social media shamings would be easily achievable.<sup>97</sup> Justine could have effortlessly established that statements condemning her as a racist and incompetent at her job so harmed her reputation that the online community thought less of her, and she would have won her case.<sup>98</sup>

90. See RESTATEMENT (SECOND) OF TORTS § 568 (AM. LAW INST. 1977) (distinguishing between libel and slander as forms of defamatory communications). One commentator argues that online defamation should be analyzed under slander rather than libel. See Glenn H. Reynolds, *Libel in the Blogosphere: Some Preliminary Thoughts*, 84 WASH. U. L. REV. 1157, 1165 (2006).

91. Ellyn M. Angelotti, *Twibel Law: What Defamation and Its Remedies Look Like in the Age of Twitter*, 13 J. HIGH TECH. L. 430, 432 (2013).

92. See *infra* notes 218–21 and accompanying text.

93. See *infra* notes 94–104 and accompanying text.

94. RESTATEMENT (SECOND) OF TORTS § 566 cmt. a (AM. LAW INST. 1977).

95. See John J. Watkins & Charles W. Schwartz, *Gertz and the Common Law of Defamation: Of Fault, Nonmedia Defendants, and Conditional Privileges*, 15 TEX. TECH L. REV. 823, 825 (1984).

96. See *id.*

97. See *supra* notes 94–96 and accompanying text.

98. See *supra* note notes 2–9 and accompanying text. Even earlier instances of public shaming can be seen in public punishment, where criminals were put in the stocks or struck at the whipping post. RONSON, *supra* note 1, at 51. In the early nineteenth century, all but one state—Delaware—abolished public punishments. *Id.* at 54–55. In an effort to persuade Delaware to finally destroy the whipping post, one *New York Times* article stated:

If it had previously existed in [the convicted person’s] bosom a spark of self-respect this exposure to public shame utterly extinguishes it. Without the hope that springs eternal in the human breast, without some desire to reform and become a good citizen, and the feeling that such a thing is possible, no criminal can ever return to honorable courses.

The defamation law landscape drastically changed in 1964 when a landmark Supreme Court decision established that the First Amendment protects defamatory speech to a certain extent, thereby placing greater emphasis on free speech than injured reputations.<sup>99</sup> In *New York Times Co. v. Sullivan*,<sup>100</sup> an elected Commissioner, who was in charge of supervising a police department, brought a libel action against four Alabama clergymen and the New York Times for an advertisement that contained inaccurate statements about police acts toward African Americans that the Commissioner claimed harmed his reputation.<sup>101</sup> The Court ruled in favor of the New York Times and held that the lower court's decision was "constitutionally deficient for failure to provide the safeguards for freedom of speech and of the press that are required by the First and Fourteenth Amendments in a libel action brought by a public official against critics of his official conduct."<sup>102</sup> In essence, the Court established greater protections for publishers of information by distinguishing between private individuals and public officials, requiring public officials to prove that the statement was made with actual malice, a higher standard.<sup>103</sup> In doing so, the *Sullivan* Court sought to encourage discourse in the marketplace by allowing

---

The boy of eighteen who is whipped at New Castle [a Delaware whipping post] for larceny is in nine cases out of ten ruined. With his self-respect destroyed and the taunt and sneer of public disgrace branded upon his forehead, he feels himself lost and abandoned by his fellows.

*Id.* at 55 (quoting Robert G. Caldwell, Editorial, *Red Hannah: Delaware's Whipping Post*, N.Y. TIMES, 1867).

99. See Joe Trevino, *From Tweets to Twibel: Why the Current Defamation Law Does Not Provide for Jay Cutler's Feelings*, 19 SPORTS L.J. 49, 52 (2012) (citation omitted) ("In *Sullivan*, the Court 'revolutionized the modern law of libel by declaring for the first time that state libel laws were subject to First Amendment restraints."); see also Elynn M. Angelotti, *Twibel Law: What Defamation and Its Remedies Look Like in the Age of Twitter*, 13 J. HIGH TECH. L. 430, 432 (2013) ("Throughout its history, defamation law has aimed to strike a balance between free speech and protecting the reputations of others.")

100. 376 U.S. 254 (1964).

101. See *id.* at 256–59.

102. *Id.* at 264.

103. See *id.* at 279–80. The majority specifically stated:

The constitutional guarantees require, we think, a federal rule that prohibits a public official from recovering damages for a defamatory falsehood relating to his official conduct unless he proves that the statement was made with "actual malice"—that is, with knowledge that it was false or with reckless disregard of whether it was false or not.

*Id.*

individuals to verbalize or publish comments about public officials without full knowledge of the information's falsity.<sup>104</sup>

So, after the *Sullivan* decision, public officials must demonstrate that the publisher of the allegedly defamatory content acted with "actual malice."<sup>105</sup> Actual malice requires that the defendant made the statement "with knowledge that it was false or with reckless disregard of whether it was false or not."<sup>106</sup> This requirement makes it much more difficult for public officials to succeed in defamation actions, thereby protecting more speech under the First Amendment.<sup>107</sup> On the other hand, private individuals only have to prove that the speaker published the statement negligently—a lower burden, which typically means that the individual failed to exercise the "standard of care that a reasonably prudent person would have exercised in a similar situation."<sup>108</sup>

### B. Public Figures Versus Private Individuals

Three years after *Sullivan*, the Court extended the requirement of proving actual malice from public officials to public figures in *Curtis Publishing Co. v. Butts*.<sup>109</sup> In explaining which plaintiffs are required to meet the higher standard of actual malice, Chief Justice Warren noted that a public figure is one who is "intimately involved in the resolution of

104. See Ann E. O'Connor, Note, *Access to Media All A-Twitter: Revisiting Gertz and the Access to Media Test in the Age of Social Networking*, 63 FED. COMM. L.J. 507, 512 (2011). Stated another way, the Court was worried that if officials only had to prove that a defendant acted negligently when making the allegedly defamatory statement, the law would have a chilling effect on individuals who criticized public officials. See Lafferman, *supra* note 38, at 212.

105. *Sullivan*, 376 U.S. at 279–80; see also *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 347 (1974).

106. *Sullivan*, 376 U.S. at 280; see Trevino, *supra* note 99, at 53. Actual malice concerns the speaker's knowledge of the truth, rather than the attitude toward the defendant, so "a defendant's 'free-floating' feeling of ill will toward a plaintiff is generally irrelevant to proving actual malice." *Cummins v. Bat World Sanctuary*, No. 02-12-00285-CV, 2015 WL 1641144, at \*5 (Tex. App. Apr. 9, 2015), *reh'g overruled* (Apr. 30, 2015), *review denied* (Aug. 28, 2015).

107. See *Sullivan*, 376 U.S. at 280; see Trevino, *supra* note 99, at 53.

108. *Negligent*, BLACK'S LAW DICTIONARY (9th ed. 2009).

109. 388 U.S. 130, 155 (1967). Chief Justice Warren believed that public officials and public figures should be held to the same requirements in defamation cases because their "views and actions with respect to public issues and events are often of as much concern to the citizen as the attitudes and behavior of 'public officials' with respect to the same issues and events." See *id.* at 162 (Warren, C.J., concurring).

important public questions or, by reason of their fame, shape[s] events in areas of concern to society at large.”<sup>110</sup> In *Gertz v. Welch*,<sup>111</sup> the Court further clarified the public-figure doctrine.<sup>112</sup> The issue was whether a newspaper or broadcaster that published a false defamatory statement about a private individual could claim constitutional privilege against liability.<sup>113</sup> Acknowledging that there is a legitimate state interest in protecting individuals’ reputations, the Court noted that in defamation actions, courts must balance the needs of the press and the individual on a case-by-case basis.<sup>114</sup> Important to the development of the public-figure test, the Court recognized at least two, maybe three,<sup>115</sup> different types of public figures: general-purpose, limited-purpose, and involuntary.<sup>116</sup> General-purpose

110. *Id.* at 164.

Viewed in this context, then, it is plain that although they are not subject to the restraints of the political process, “public figures,” like “public officials,” often play an influential role in ordering society. And surely as a class these “public figures” have as ready access as “public officials” to mass media of communication, both to influence policy and to counter criticism of their views and activities. Our citizenry has a legitimate and substantial interest in the conduct of such persons, and freedom of the press to engage in uninhibited debate about their involvement in public issues and events is as crucial as it is in the case of “public officials.” The fact that they are not amenable to the restraints of the political process only underscores the legitimate and substantial nature of the interest, since it means that public opinion may be the only instrument by which society can attempt to influence their conduct.

*Id.*

111. 418 U.S. 323 (1974).

112. *See id.* The *Gertz* Court also overruled an intervening decision, *Rosenbloom v. Metromedia, Inc.*, 403 U.S. 29 (1971), which had extended the actual malice requirement to private persons who commented on events of public interest. *Gertz*, 418 U.S. at 368–69. The plurality in *Rosenbloom* protected First Amendment principles to an even more drastic extent than in *Sullivan*: “We honor the commitment to robust debate on public issues, which is embodied in the First Amendment, by extending constitutional protection to all discussion and communication involving matters of public or general concern, without regard to whether the persons involved are famous or anonymous.” *Rosenbloom*, 403 U.S. at 43–44. Some commentators argue that the *Rosenbloom* standard provides a superior framework for defamation in the digital age. *See* David Lat & Zach Shemtob, *Public Figurehood in the Digital Age*, 9 J. TELECOMM. & HIGH TECH. L. 403 (2011).

113. *See Gertz*, 418 U.S. at 332.

114. *See id.* at 343.

115. There is debate as to whether the involuntary public figure is actually a practically workable category, because many courts have declined to use it. *See, e.g., Wells v. Liddy*, 186 F.3d 505, 538 (4th Cir. 1999) (“So rarely have courts determined that an individual was an involuntary public figure that commentators have questioned the continuing existence of that category.”); Lafferman, *supra* note 38, at 213.

116. *See Gertz*, 418 U.S. at 340–45.

public figures are individuals who assume roles of “especial prominence” in society.<sup>117</sup> More commonly, limited-purpose public figures “thrust themselves to the forefront of particular public controversies in order to influence the resolution of the issues involved,” inviting “attention and comment.”<sup>118</sup> Lastly, individuals can become involuntary public figures “through no purposeful action of [their] own,” but that is “exceedingly rare.”<sup>119</sup>

Later Supreme Court cases helped define the outer limits of the public-figure doctrine.<sup>120</sup> In *Time, Inc. v. Firestone*,<sup>121</sup> the Court declined to extend public-figure status to an individual merely based on local notoriety.<sup>122</sup> The Court also refused to define public controversies as “all controversies of

117. *Id.* at 345. For example, courts have found that Johnny Carson, Carol Burnett, and William Buckley, Jr. are general-purpose public figures. See Ashley Messenger & Kevin Delaney, *In the Future, Will We All Be Limited-Purpose Public Figures?*, 30 COMM. LAW. 1, 4 (2014). The D.C. Circuit set forth a test to help determine whether an individual is a general-purpose public figure:

The judge can examine statistical surveys, if presented, that concern the plaintiff’s name recognition. Previous coverage of the plaintiff in the press also is relevant. The judge can check whether others in fact alter or reevaluate their conduct or ideas in light of the plaintiff’s actions. He also can see if the plaintiff has shunned the attention that the public has given him and determine if those efforts have been successful. . . . No one parameter is dispositive; the decision still involves an element of judgment. Nevertheless, the weighing of these and other relevant factors can lead to a more accurate and a more predictable assessment of a person’s overall fame and notoriety in the community.

*Waldbaum v. Fairchild Publications, Inc.*, 627 F.2d 1287, 1295 (D.C. Cir. 1980).

118. *Gertz*, 418 U.S. at 345. Put another way, the court considers “whether the plaintiff voluntarily assumed a prominent role in a public controversy and the attendant risk of enhanced public scrutiny that accompanies it.” Lyriisa Barnett Lidsky, *Silencing John Doe: Defamation and Discourse in Cyberspace*, 49 DUKE L.J. 855, 909 (2000) (citing *Gertz*, 418 U.S. at 344–45). The Supreme Court has alluded that speech not involving public controversies has a reduced constitutional value. See *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 761 (1985). As discussed in Part V, the limited-purpose public figure is most relevant to social media defamation. See *infra* Part V.

119. *Gertz*, 418 U.S. at 345; see also Lafferman, *supra* note 38, at 220–22. In the social media context, one commentator has stated that “[t]hese private figures may involuntarily become public figures by participating in an event that is covered by the media.” Joshua Azriel & Charles Mayo, *Fifty Years After New York Times Co. v. Sullivan and Forty Years After Gertz v. Welch: How These Twentieth Century Supreme Court Rulings Impact Twenty-First Century Online Social Media Libel Claims*, 20 B.U. J. SCI. & TECH. L. 191, 203 (2014).

120. See *infra* notes 122–25 and accompanying text.

121. 424 U.S. 448 (1976).

122. See *id.* at 453.

interest to the public” because it was concerned that applying the public-figure doctrine so broadly would capture too large a class of people.<sup>123</sup> Further, in *Hutchinson v. Proxmire*,<sup>124</sup> the Supreme Court stressed that an individual should only be deemed a public figure if the court can identify that the plaintiff voluntarily participated in public activities.<sup>125</sup>

The Court has announced that public figures should be required to meet a higher standard for public policy reasons.<sup>126</sup> Because public officials and public figures inherently have better access to channels of effective communication than private individuals, they can rebut false statements made about them more easily and are therefore less vulnerable to false communications.<sup>127</sup> In contrast, private individuals “are not only more vulnerable to injury than public officials and public figures; they are also more deserving of recovery.”<sup>128</sup> Additionally, public figures have assumed the risk of injury by voluntarily assuming roles of fame or influence in society.<sup>129</sup> The Court also stressed the potential chilling effect on free speech if an individual who criticized a public official was worried about

123. *Id.* at 454; *see also* *Wolston v. Reader’s Digest Ass’n*, 443 U.S. 157, 167 (1979) (“A private individual is not automatically transformed into a public figure just by becoming involved in or associated with a matter that attracts public attention.”).

124. 443 U.S. 111 (1979) (rejecting the proposition that an adjunct professor was a public official).

125. *See* *Lafferman*, *supra* note 38, at 215.

126. *See infra* notes 127–31 and accompanying text.

127. *See* *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 344 (1974).

The first remedy of any victim of defamation is self-help—using available opportunities to contradict the lie or correct the error and thereby to minimize its adverse impact on reputation. Public officials and public figures usually enjoy significantly greater access to the channels of effective communication and hence have a more realistic opportunity to counteract false statements [than] private individuals normally enjoy. Private individuals are therefore more vulnerable to injury, and the state interest in protecting them is correspondingly greater.

*Id.*

128. *Id.* at 345.

129. *See id.*; *see also* *Lafferman*, *supra* note 38, at 213. The Court recognized that public officials or public figures “must accept certain necessary consequences of [their] involvement in public affairs” and risk “closer public scrutiny than might otherwise be the case.” *Gertz*, 418 U.S. at 344; *see also* *Wolston v. Reader’s Digest Ass’n*, 443 U.S. 157, 164 (1979) (citing *Gertz*, 418 U.S. at 345) (“[P]ublic figures are less deserving of protection than private persons because public figures, like public officials, have ‘voluntarily exposed themselves to increased risk of injury from defamatory falsehood concerning them.’”).

liability.<sup>130</sup> Because bringing a defamation claim as a public figure requires proof that the statement was made with actual malice, the determination of whether a harmed individual is a public or private figure essentially controls the claim's likelihood of success.<sup>131</sup>

### C. Lower Courts Struggle to Apply the Public-Figure Test

While the Supreme Court has provided some guidance, lower courts have struggled to apply the public-figure test and have had inconsistent results.<sup>132</sup> For the limited-purpose public-figure test in particular, lower courts have developed a variety of different tests.<sup>133</sup>

#### 1. Access to Media

One inconsistency relevant to social media defamation is how lower courts have applied the access-to-media element of the public-figure test.<sup>134</sup> In analyzing the plaintiff's access to media, the Supreme Court "seems to be weighing the plaintiff's ability to command media attention in order to redress claims leveled against him."<sup>135</sup> The Court was effectively worried about private individuals who would not have the ability to ward off negative comments about themselves.<sup>136</sup> Additionally, in *Hutchinson v.*

130. See *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 269 (1964) (Goldberg, J., concurring); see also O'Connor, *supra* note 104, at 512 ("The aim was to allow discourse concerning public officials as it advanced the introduction of important ideas into the marketplace."). The Court in *Sullivan* emphasized the "profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open, and that it may well include vehement, caustic, and sometimes unpleasantly sharp attacks on government and public officials." *Sullivan*, 376 U.S. at 270.

131. See *supra* notes 105–08 and accompanying text.

132. See O'Connor, *supra* note 104, at 514–15. For example, courts have struggled with determining what constitutes a public controversy. See *Messenger & Delaney*, *supra* note 117, at 4. Whether an individual is a public figure is a question of law for the judge. See *Time, Inc. v. Firestone*, 424 U.S. 448, 463 (1976).

133. See Lafferman, *supra* note 38, at 216–18.

134. See, e.g., Erik Walker, Comment, *Defamation Law: Public Figures—Who Are They?*, 45 BAYLOR L. REV. 955, 975–76 (1993).

135. O'Connor, *supra* note 104, at 520; see *supra* note 127 and accompanying text.

136. See O'Connor, *supra* note 104, at 520.

*Proxmire*,<sup>137</sup> the Court held that an individual was not a public figure because he did not have *regular and continuing access* to the media.<sup>138</sup> While the Supreme Court has stressed that the limited-purpose public figure has unique access to the media, courts, in developing their own tests, have neglected that consideration and have often only focused on whether the individual thrust himself into a public controversy.<sup>139</sup> For example, the *Waldbaum* Court set forth a limited-purpose public-figure test that completely fails to consider a plaintiff's access to media.<sup>140</sup>

## 2. Public Controversies

There has also been difficulty in determining when a public controversy exists.<sup>141</sup> The First Amendment protects speech when the speech pertains to matters of public concern or controversy.<sup>142</sup> And a public controversy is more than merely “a matter that attracts public attention.”<sup>143</sup> The Supreme

137. 443 U.S. 111 (1979).

138. *See id.* at 136. The Court failed to describe what “regular and continuing” means. *See Walker, supra* note 134, at 976. This Author will analyze the average social media user’s access to media by evaluating both whether the user has the ability to counteract false statements and whether the user has regular and continuing access to the media. *See infra* Section V.A.

139. O’Connor, *supra* note 104, at 520; *see also Walker, supra* note 134, at 976. So, practically speaking, these courts will determine that an individual is a public figure without regard to whether that plaintiff has the ability to respond to the defamatory comments made about him. *Id.*

140. *Waldbaum v. Fairchild Publ’ns, Inc.*, 627 F.2d 1287, 1296–98 (D.C. Cir. 1980). The first step of this limited-purpose public-figure test is to isolate the public controversy. *Id.* at 1296. Next, the court must analyze the plaintiff’s role in the controversy. *Id.* at 1297. “Finally, the alleged defamation must have been germane to the plaintiff’s participation in the controversy.” *Id.* at 1298; *see also Parsi v. Daiouleslam*, 595 F. Supp. 2d 99, 105 (D.D.C. 2009); *Framsted v. Mun. Ambulance Serv.*, 347 F. Supp. 2d 638, 662 (W.D. Wis. 2004). *But see Contemporary Mission, Inc. v. N.Y. Times Co.*, 842 F.2d 612, 617 (2d Cir. 1988) (requiring that the plaintiff “(1) successfully invited public attention to his views in an effort to influence others prior to the incident that is the subject of litigation; (2) voluntarily injected himself into a public controversy related to the subject of the litigation; (3) assumed a position of prominence in the public controversy; and (4) maintained regular and continuing access to the media”).

141. *Walker, supra* note 134, at 969 (stating that “[l]ower courts have applied divergent standards in determining whether public controversies are present in defamation cases”).

142. *See N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 269 (1964) (“The general proposition that freedom of expression upon public questions is secured by the First Amendment has long been settled by our decisions.”).

143. *Wolston v. Reader’s Digest Ass’n*, 443 U.S. 157, 167 (1979).

Court clarified in *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*<sup>144</sup> that courts can determine when speech addresses matters of public concern by looking at the “content, form, and context” of the speech’s publication.<sup>145</sup> Such a determination is made on a case-by-case basis.<sup>146</sup> Unsurprisingly, the “content, form, and context” test has provided courts with very little guidance for finding public controversies.<sup>147</sup> Some courts have read *Gertz* to mean that it is not their role to evaluate whether issues are matters of public interest and have found that a public controversy exists “merely if the events in question have generated widespread public interest.”<sup>148</sup> Other courts have found newsworthiness insufficient to determine a matter of public concern, and instead usually have followed the *Waldbaum* Court, which determined that a public controversy is “a specific public dispute that has foreseeable and substantial ramifications for persons beyond its immediate participants.”<sup>149</sup>

### 3. Voluntary Involvement in Public Controversies

As outlined in *Gertz*, individuals can become limited-purpose public figures when they voluntarily thrust themselves into a public controversy.<sup>150</sup> Some courts liberally apply the voluntariness principle and hold that “an individual voluntarily rises to the forefront of a controversy by voluntarily engaging in a pattern of activity destined to invite attention and

144. 472 U.S. 749 (1985).

145. *Id.* at 761 (citing *Connick v. Myers*, 461 U.S. 138, 147–48 (1983)); *see also* *Flamm v. Am. Ass’n of Univ. Women*, 201 F.3d 144, 150 (2d Cir. 2000); *Messenger & Delaney*, *supra* note 117, at 4. In *Dun & Bradstreet, Inc.*, the Court held that the speech, consisting of false credit reports, was not a matter of public concern because it only concerned the interest of the speaker and the specific business audience. 472 U.S. at 762.

146. *Dun & Bradstreet, Inc.*, 472 U.S. at 761.

147. *See infra* notes 148–49 and accompanying text.

148. Walker, *supra* note 134, at 969. In *Gertz v. Robert Welch, Inc.*, the Court doubted judges’ ability to determine which issues are of public concern. 418 U.S. 323, 346 (1974).

149. *Waldbaum v. Fairchild Publications, Inc.*, 627 F.2d 1287, 1292 (D.C. Cir. 1980). In the Sixth Circuit, the dispute must affect the general public “in an appreciable way” and must have “received public attention because its ramifications will be felt by persons who are not direct participants.” *Thomas M. Cooley Law Sch. v. Kurzon Strauss, LLP*, 759 F.3d 522, 529–30 (6th Cir. 2014). Still other courts have taken the position that there must be public disagreement over an issue for it to be an issue of public concern. Walker, *supra* note 134, at 970.

150. *See supra* note 129 and accompanying text.

comment.”<sup>151</sup> That the individual might not desire publicity is of no concern to these courts.<sup>152</sup> Other courts require that individuals actively seek out public attention and hold that individuals will be deemed public figures if they “take affirmative steps to attract public attention, such as using the media or a similar device, in order to influence the controversy’s outcome.”<sup>153</sup>

#### IV. OTHER ISSUES PLAINTIFFS WILL ENCOUNTER

##### A. *Twitter and Facebook Will Not Be Held Liable*

Properly applying the current defamation framework to social media platforms is of even more importance when considering the limitations of other forms of recovery.<sup>154</sup> Section 230 of the Communications Decency Act creates federal immunity for all Internet service providers for defamatory statements users post on its platforms.<sup>155</sup> Therefore, Facebook and Twitter are not held liable for the defamatory comments social media users make on their platforms.<sup>156</sup> Instead, the target of online defamatory statements is forced to go after the private individual who made the statement.<sup>157</sup> Section 230 may also extend federal immunity to reposters of defamatory information online.<sup>158</sup> In essence, if an individual posted a

151. Walker, *supra* note 134, at 972.

152. *Id.* (“Whether the individual wants to be a public figure or desires publicity is irrelevant.”).

153. *Id.* at 973.

154. *See infra* notes 155–59 and accompanying text.

155. *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 330 (4th Cir. 1997) (“[Communications Decency Act (CDA)] creates a federal immunity to any cause of action that would make service providers liable for information originating with a third-party user of the service.”); Lafferman, *supra* note 38, at 206. Congress sought to protect Internet service providers from traditional publisher liability in order to encourage Internet growth. *See Vazquez v. Buhl*, 90 A.3d 331, 335 (Conn. App. Ct. 2014). The relevant portion of the statute states: “No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.” 47 U.S.C. § 230(c)(1) (2012). An information content provider is defined as “any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service.” *Id.* § 230(f)(3).

156. *See supra* note 155 and accompanying text.

157. Lafferman, *supra* note 38, at 206.

158. Matthew E. Kelley & Steven D. Zansberg, *A Little Birdie Told Me, “You’re A Crook”*: *Libel*

defamatory tweet online, and a different user simply retweeted that statement without adding any new information to the statement, the user would likely have federal immunity, thus further limiting plaintiffs' avenues of recovery.<sup>159</sup>

### B. Anonymous Social Media Speakers

To have a successful lawsuit, a plaintiff must name the correct defendants.<sup>160</sup> However, for an Internet defamation case, a speaker on social media may not publicly disclose his real name on his social media account, making it difficult to discover the defendant's identity.<sup>161</sup> Additionally, the First Amendment right to free speech includes the right to publish a statement anonymously.<sup>162</sup> The poster of a defamatory tweet or Facebook comment often hides behind a username that gives no indication of the person's real identity.<sup>163</sup> Further, Facebook and Twitter have certain protocols to protect the identities of their users and will reveal their information only in certain circumstances.<sup>164</sup> Therefore, when a person

*in the Twittersphere and Beyond*, 30 COMM. LAW. 1, 38–39 (2014).

159. *Id.*

160. *See infra* notes 161–67 and accompanying text.

161. *See infra* notes 163–67 and accompanying text. Courts have struggled to determine when plaintiffs have a right to discover the identity of anonymous posters in light of the right to speak anonymously. *See, e.g., In re Does 1–10*, 242 S.W.3d 805, 819–21 (Tex. App. 2007).

162. *McIntyre v. Ohio Elections Comm'n*, 514 U.S. 334, 342 (1995) (“Anonymity thereby provides a way for a writer who may be personally unpopular to ensure that readers will not prejudge her message simply because they do not like its proponent.”). This right to anonymity also applies on the Internet. *See, e.g., SaleHoo Grp., Ltd. v. ABC Co.*, 722 F. Supp. 2d 1210, 1213 (W.D. Wash. 2010). Individuals who have committed no wrongdoing should have the freedom to speak anonymously online to “facilitate[] the rich, diverse, and far ranging exchange of ideas” and avoid chilling free speech. *Id.* at 1213–14. However, when an online speaker does commit a wrongdoing—such as defaming another individual—he or she cannot seek protection under the right to anonymity. *See Ghanam v. Does*, 845 N.W.2d 128, 137 (Mich. Ct. App. 2014).

163. *Holschuh*, *supra* note 55, at 959. A user on Twitter may register using her real name or a pseudonym. *Twitter Privacy Policy*, TWITTER, <https://twitter.com/privacy?lang=en> (last visited Nov. 16, 2016).

164. Twitter's policy states:

Notwithstanding anything to the contrary in this Policy, we may preserve or disclose your information if we believe that it is reasonably necessary to comply with a law, regulation, legal process, or governmental request; to protect the safety of any person; to address fraud, security or technical issues; or to protect Twitter's rights or property.

*Id.* Facebook has a similar policy: “We may access, preserve and share your information in response

attempts to sue a social media user, one of the initial barriers to recovery will be properly naming the defendant.<sup>165</sup> The high costs associated with discovering the identity of the speaker may deter plaintiffs from filing suit.<sup>166</sup> The process of discovering the identity of an anonymous poster will often be determined by the specific approach of the court's jurisdiction.<sup>167</sup> Therefore, the defendant's anonymity may be an insurmountable problem for plaintiffs who seek to hold users liable for their defamatory online posts.<sup>168</sup>

### C. *Opinion or Actionable False Statement?*

The Constitution protects statements of opinion.<sup>169</sup> Distinguishing between actionable false statements and opinions is a challenging task, but is important in the social media context because many potentially defamatory statements made on social media consist of opinion.<sup>170</sup> The Supreme Court drew this important line in *Milkovich v. Lorain Journal Co.*<sup>171</sup> by setting forth a two-part test to determine when a statement is actionable rather than just mere opinion: (1) the statement is "provable as false"; and (2) the statement is capable of reasonably being interpreted as stating actual facts.<sup>172</sup>

As to the second prong of the *Milkovich* test, courts have often looked at

---

to a legal request (like a search warrant, court order or subpoena) if we have a good faith belief that the law requires us to do so." *Data Policy*, FACEBOOK, <https://www.facebook.com/about/privacy/other> (last visited Nov. 16, 2016).

165. Holschuh, *supra* note 55, at 959.

166. Lewis, *supra* note 64, at 953–54. If the plaintiff sues an anonymous Internet user, he or she "will file suit against 'John Doe' and then seek leave to subpoena Doe's ISP for Doe's identity." *Id.* at 954. Jurisdictional issues may arise in federal court if John Doe is unmasked and turns out to be residing in the same state as the plaintiff. *Id.*

167. *See generally* Ghanam v. Does, 845 N.W.2d 128 (Mich. Ct. App. 2014). The spectrum of standards includes the following: "good-faith basis to assert a claim, pleading sufficient facts to survive a motion to dismiss, showing of prima facie evidence sufficient to withstand a motion for summary disposition, and 'hurdles even more stringent.'" *Id.* at 137 (quoting Doe v. Cahill, 884 A.2d 451, 457 (Del. 2005)).

168. *See supra* notes 159–66.

169. *See infra* notes 170–72 and accompanying text.

170. *See supra* Section III.A.

171. 497 U.S. 1 (1990).

172. *Id.* at 19; *cf.* RESTATEMENT (SECOND) OF TORTS § 566 (AM. LAW INST. 1977) (stating that an opinion may be actionable "if it implies the allegation of undisclosed defamatory facts as the basis for the opinion").

the type of forum in which the statement was made.<sup>173</sup> In *Underwager v. Channel 9 Australia*, the Ninth Circuit further expanded upon how to determine whether an opinion implies a factual assertion.<sup>174</sup> The court first looked at the statement “in its broad context, which includes the general tenor of the entire work, the subject of the statements, the setting, and the format of the work.”<sup>175</sup> Second, the court looked specifically at the context of the statement itself, noting that there are certain online forums where an audience may reasonably expect posts to contain figurative or hyperbolic language.<sup>176</sup> Finally, the court looked at the statement itself, analyzing whether it was “sufficiently factual to be susceptible of being proved true or false.”<sup>177</sup>

A few courts have applied the *Milkovich* test in the social media context.<sup>178</sup> As an important preliminary note, a California Court of Appeal ruled that just because a statement is made on social media does not mean that speakers are completely insulated from liability by the First Amendment.<sup>179</sup> The court noted that “[w]hile courts have recognized that online posters often ‘play fast and loose with facts’, this should not be taken to mean online commentators are immune from defamation liability.”<sup>180</sup> Despite implicitly recognizing the informal nature of social media sites, the court firmly established that “[w]here specific false factual allegations are published and they cause damage, a defamation action will lie.”<sup>181</sup>

In another case dealing with an allegedly defamatory tweet, *Feld v.*

---

173. See, e.g., *Underwager v. Channel 9 Australia*, 69 F.3d 361, 366 (9th Cir. 1995) (“To determine whether a statement implies a factual assertion, we examine the totality of the circumstances in which it was made.”).

174. *Id.* In *Underwager*, a psychologist who had testified in 250 trials regarding the unreliability of children’s testimony of sexual abuse sued individuals who appeared on a television program that disputed his theories and credentials for conspiracy to defame him. *Id.* at 363–64. The court concluded that all but one statement in the program was an assertion of opinion and not actionable. *Id.* at 366.

175. *Id.* at 366.

176. *Id.*

177. *Id.*

178. See *infra* notes 179–97 and accompanying text.

179. *Sanders v. Walsh*, 162 Cal. Rptr. 3d 188, 194 (Ct. App. 2013).

180. *Id.* at 196 (citation omitted).

181. *Id.* at 196.

*Conway*,<sup>182</sup> the court considered whether a tweet was defamatory in the context of the entire conversation.<sup>183</sup> Feld, the plaintiff, arranged to have a horse breeder ship her horse to a farm to become a companion horse, but the horse was instead shipped to Canada and potentially slaughtered.<sup>184</sup> The mishap became the topic of an ongoing online debate.<sup>185</sup> Conway, the defendant, entered into the debate and posted the allegedly defamatory tweet.<sup>186</sup> Feld's professional success was dependent upon "public review and endorsement of her publications," and interested parties often searched her name on the Internet.<sup>187</sup> The allegedly defamatory tweet appeared in the search results on Internet search engines when users entered Feld's name, causing extensive harm to Feld's reputation.<sup>188</sup> In her defense, the defendant asserted that the statement was a constitutionally protected opinion.<sup>189</sup> The court ultimately found that the statement was protected because the tweet was a part of a heated Internet debate about the plaintiff's responsibility for the disappearance of her horse.<sup>190</sup>

Similarly, in *Finkel v. Dauber*,<sup>191</sup> a court again analyzed allegedly defamatory social media statements by looking at the conversation's context.<sup>192</sup> In *Finkel*, a social media user sued a group of adolescents when they posted comments about her on a secret Facebook page that stated she "was seen having sexual relations with a horse, contracted HIV from sharing needles with heroin addicts, contracted AIDS from a male prostitute, and

---

182. 16 F. Supp. 3d 1 (D. Mass. 2014).

183. *Id.* at 4 ("The tweet cannot be read in isolation, but in the context of the entire discussion.").

184. *Id.* at 2.

185. *Id.*

186. *Id.* The defendant's tweet stated: "Mara Feld aka Gina Holt—you are fucking crazy!" *Id.*

187. *Id.* at 3.

188. *Id.*

189. *Id.*

190. *Id.* at 4; see also Matthew E. Kelley & Steven D. Zansberg, *140 Characters of Defamation: The Developing Law of Social Media Libel*, 18 J. INTERNET L. 1, 10 (2014) (citing *Feld*, 16 F. Supp. 3d at 4) ("The court considered not just the tweet itself but related online postings involving the broader controversy into which the tweet fit in holding that it was understood as nondefamatory 'criticism—that is, as opinion—not as a statement of fact.'). The court further determined that the statement was intended to be a criticism rather than a statement of fact when considering how a reasonable person would interpret the tweet. *Feld*, 16 F. Supp. 3d at 4.

191. 906 N.Y.S.2d 697 (Sup. Ct. 2010).

192. *Id.* at 700.

transformed into the devil.”<sup>193</sup> The adolescents were officers of a secret Facebook group<sup>194</sup> called “Ninety Cents Short of a Dollar” that did not have public content or appear on any Facebook member’s profile.<sup>195</sup> In considering whether the plaintiff was defamed, the court took special note of the fact that the plaintiff never alleged that the secret group’s posts were “accessible to anyone outside the group” and that her name was never actually used in the posts.<sup>196</sup> Taken together, the court stated that: “The entire context and tone of the posts constitute evidence of adolescent insecurities and indulgences, and a vulgar attempt at humor. What they do not contain are statements of fact.”<sup>197</sup>

Courts should follow the *Finkel* court’s approach to ensure that tweets that are no more than “rhetorical hyperbole” are not held to be

---

193. Hunt, *supra* note 48, at 599. The following are some of the posts (spelling, punctuation, and typos as in original):

BTW the 11th cent, unbeknownst to many, acquired AIDS while on a cruise to Africa (with another member of the group who shall remain nameless). While in Africa she was seen fucking a horse. NICHTE NICHTE eleventh cent! I mean you know ... I kinda felt bad for the eleventh cent ... but then again I felt WORSE for the horse. ... (Leah Herz 1/29/07, 5:04pm)

In regards to the 7th cents comments,, it was not from an African cruise... it was from sharing needles with different heroin addicts, this led to cross “mojination” which caused the HIV virus... she then persisted to screw a baboon which caused the epidemic to spread. (Jeff Schwartz 1/29/07, 7:43pm)

I heard that the 11th cent got aids when she hired a male prostitute who came dressed as a sexy fireman apparantly .. she was lonely, because her friends no longer associated with her. her sexy fireman prostitute was her only company. in addition to acquiring aids, this nameless 11th cent aquired crabs, and syphillis. (Melinda Danowitz, 1/29/07, 8:00 pm)

also i heard that the stds (sexually transmitted diseases) she got were os bad that she morfed intot he devil in one of our pictures ... oops did i reviel the 11th cent!> (Michael Dauber, 1/29/07, 8:32 pm)

*Finkel*, 906 N.Y.S.2d at 700.

194. *Finkel*, 906 N.Y.S.2d at 700. A secret Facebook group is a group on Facebook where only members can see who is in the group and what members post in the group. *What Are the Privacy Settings for Groups?*, FACEBOOK HELP CTR., <https://www.facebook.com/help/220336891328465> (last visited Nov. 16, 2016). Additionally, only members of the group can find the group in a search. *Id.*

195. *Finkel*, 906 N.Y.S.2d at 700. The group had only six listed members. *Id.*

196. *Id.*

197. *Id.* at 702.

defamatory.<sup>198</sup> Conversely, courts should not be too quick to decide that any hastily constructed and posted tweet is necessarily an opinion and not a false statement of fact.<sup>199</sup> For social media backlash, the context is typically the same: an individual with an average social media presence posts some type of statement or photo onto his or her social media page, and the social media community responds by condemning the user and flooding the Internet with a false perspective of that individual's identity.<sup>200</sup> As was the case in *Feld v. Conway*,<sup>201</sup> these defamatory tweets or Facebook posts are typically searchable and available to the public through search engines.<sup>202</sup> Therefore, social media backlash implicates more reputational issues because the wider online community thinks less of the plaintiff rather than simply a small group of people.<sup>203</sup> It is easy to say that these types of statements are not "vulgar attempt[s] at humor"<sup>204</sup> because they are real responses to some distinct post.<sup>205</sup> The issue would be whether, in the context of strangers commenting about a social media user they have never met, the allegedly defamatory statements could be construed by a reasonable observer as conveying facts.<sup>206</sup>

The court's view of Twitter or Facebook will likely determine if the court will hold that a reasonable factfinder could conclude that the tweet or Facebook comment is a provable false statement of fact.<sup>207</sup> If the court believes social media sites are simply platforms for users to instantaneously vent and the reasonable viewer does not actually take these comments seriously, then it will likely rule that the allegedly defamatory tweet or Facebook comment is not based in fact.<sup>208</sup> However, if the court sees social

---

198. Hunt, *supra* note 48, at 599. The Supreme Court has held that statements that are simply "rhetorical hyperbole" are not to be considered actionable false statements. *Greenbelt Coop. Publ'g Ass'n v. Bresler*, 398 U.S. 6, 13–14 (1970).

199. *See supra* notes 179–81 and accompanying text.

200. *See supra* Part II.

201. 16 F. Supp. 3d 1 (D. Mass 2014).

202. *See* Allen, *supra* note 44, at 75–76; Marton et al., *supra* note 53, at 81.

203. *Cf.* Finkel v. Dauber, 906 N.Y.S.2d 697, 700 (Sup. Ct. 2010) (involving allegations of defamation in a small, "secret" Facebook group).

204. *Id.* at 702.

205. *See supra* notes 69–73 and accompanying text.

206. *See supra* notes 173–77 and accompanying text.

207. *See infra* notes 208–10 and accompanying text.

208. *See* D.C. v. R.R., 106 Cal. Rptr. 3d 399, 420 (2010) (looking at the content of the tweet itself).

media sites as valuable mediums for discourse,<sup>209</sup> it will be more likely to conclude that a defamatory comment would be taken seriously and therefore not be protected by the First Amendment.<sup>210</sup>

#### V. HOW THE PUBLIC-FIGURE TEST SHOULD BE EMPLOYED IN SOCIAL MEDIA BACKLASH SCENARIOS

Many social media users that have posted harassing and abusive comments have not been held accountable for the great harm they have caused.<sup>211</sup> If a target of social media backlash were to sue particular users for defamation and attempt to hold these speakers liable for their actions, one of the speakers' first arguments would likely be that the plaintiff is a public figure who has to prove by clear and convincing evidence that the statement was made with actual malice.<sup>212</sup> If the court were to make such a determination, it would pose a serious barrier to recovery.<sup>213</sup>

Commentators have taken divergent views on whether individuals defamed on social media should be considered public figures.<sup>214</sup> While it is tempting to argue that drastic changes in the media have rendered the public-

---

to decide that the actions of the defamers were deliberately false statements, rather than thoughtless venting).

The threat in this case was not merely a few words shouted during a brawl; it was a series of grammatically correct sentences composed at a computer keyboard over a period of at least several minutes. Even after the message was typed, the author still had to decide whether to send it. He placed his cursor over "send" and clicked the mouse or touchpad. Thus, the content of the message and its transmission show deliberation on the part of the author.

*Id.*

209. See Hunt, *supra* note 48, at 594.

210. See *id.*

211. See, e.g., Rebecca Phillips, Comment, *Constitutional Protection for Nonmedia Defendants: Should There Be a Distinction Between You and Larry King?*, 33 CAMPBELL L. REV. 173, 192 (2010).

212. See *supra* notes 103–08 and accompanying text.

213. See *id.*

214. See, e.g., Lat & Shemtob, *supra* note 112 (arguing that the public-figure test is obsolete considering the recent changes in media); O'Connor, *supra* note 104 (arguing that the access-to-media requirement should be reconsidered due to the proliferation of social media); Aaron Perzanowski, Comment, *Relative Access to Corrective Speech: A New Test for Requiring Actual Malice*, 94 CAL. L. REV. 833, 837 (2006) (recommending a balancing test that requires actual malice to "remain faithful to the rationale that informed the Court's balancing of reputational integrity and uninhibited debate," which "is needed to account for the cacophony of internet speech").

figure test obsolete, the Supreme Court has continuously affirmed its applicability, and lower courts thus far have continued to apply the test to Internet users.<sup>215</sup> Carefully analyzing the policy considerations behind the public-figure test will help courts make future determinations.<sup>216</sup> Helping courts apply the public-figure test to targets of social media backlash with *consistent* results aligns with Supreme Court precedent, which has counseled against an ad hoc determination of public figurehood and has encouraged developing “broad rules of general application.”<sup>217</sup>

Thus far, courts faced with defamatory comments on social media have applied traditional defamation principles.<sup>218</sup> While courts have grappled with how to fit emerging technologies into defamation’s existing framework, the Supreme Court has made clear that “whatever the challenges of applying the Constitution to ever-advancing technology, ‘the basic principles of freedom of speech and the press, like the First Amendment’s command, do not vary’ when a new and different medium for communication appears.”<sup>219</sup>

Lower courts have begun applying the limited-purpose public-figure test in social media defamation cases to determine whether a plaintiff must prove by clear and convincing evidence that the statement was made with actual malice.<sup>220</sup> Considering the existing framework as it stands—and where it has stood unmoved for over forty years—courts will have to decide whether

215. *See infra* notes 303–28 and accompanying text.

216. *See id.*

217. *See Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 343–44 (1974).

218. *See Kelley & Zansberg, supra* note 190, at 10. Of course, courts have already dealt with applying traditional defamation principles to changing mediums, such as the telegraph and the motion picture. *Id.* at 8; *see also* Cummins v. Bat World Sanctuary, No. 02-12-00285-CV, 2015 WL 1641144, at \*8 (Tex. App. Apr. 9, 2015) (“Neither the United States Supreme Court nor the Texas Supreme Court has seen the need to adjust defamation law in light of the changes in technology.”), *reh’g overruled* (Apr. 30, 2015), *review denied* (Aug. 28, 2015). Some commentators advocate that Twitter should be completely libel-free. *See* Thomas R. Julin & Henry R. Kaufman, *Twibel Tweak Needed for Tweeters*, CNA PRO NEWS (Apr. 26, 2011), <https://web.archive.org/web/20121222124219/http://www.cnapro.com/pdf/TwibelTweakNeededforTweeters%204-26-11.pdf>.

219. *Brown v. Entm’t Merchs. Ass’n*, 131 S. Ct. 2729, 2733 (2011) (quoting *Joseph Burstyn, Inc. v. Wilson*, 343 U.S. 495, 503 (1952)); *see also Kelley & Zansberg, supra* note 190, at 8 (“[T]he fact that Twitter, Vine, LinkedIn, and other yet-to-be-developed communications platforms are operated online or on mobile networks and are subject to various technological constraints and format conventions does not alter the fundamental principles of the First Amendment or its restraints on the tort law of defamation.”).

220. Lafferman, *supra* note 38, at 216.

the use of social media sites such as Twitter and Facebook convert a private individual into a public figure.<sup>221</sup>

This Comment does not argue that the public-figure test is unworkable in social media defamation,<sup>222</sup> but rather argues that the average social media user should not be deemed a public figure simply because she is exposing parts of her life on the Internet.<sup>223</sup> Because there are already so many other barriers to recovery,<sup>224</sup> courts must cautiously apply the public-figure test to targets of social media abuse to ensure that their harm can be remedied.<sup>225</sup> Specifically looking at average social media users who are not already public figures, the central question for the public-figure determination is whether their exposure or communication on social media thus far has put them so much into the public eye that defamatory speech targeting them should receive extra protections.<sup>226</sup> This Comment argues that the average social media user does not have the traditional characteristics of a public figure outlined by the Supreme Court, and thus should only have to prove that the defamatory speech was made negligently.<sup>227</sup>

*A. Targets of Social Media Backlash Do Not Have the Requisite Access to Media to Counteract False Statements*

In analyzing whether a plaintiff is a limited-purpose public figure, the Supreme Court evaluates whether the injured party has the ability to redress his or her reputational injuries in the media.<sup>228</sup> Public figures, like public officials, have “as ready access as ‘public officials’ to mass media of communication, both to influence policy and to counter criticism of their views and activities.”<sup>229</sup> According to at least one court, “[t]he inquiry into access to channels of communication proceeds on the assumption that public

---

221. Messenger & Delaney, *supra* note 117, at 4.

222. *But see* Lat & Shemtob, *supra* note 112 (arguing that media has changed so significantly since *Gertz* that the public-figure test is obsolete).

223. *See infra* notes 224–27 and accompanying text.

224. *See supra* Part IV.

225. *See supra* Part IV.

226. *See supra* Part III.

227. *See infra* Sections V.A–C.

228. *See supra* note 127 and accompanying text.

229. *Curtis Pub. Co. v. Butts*, 388 U.S. 130, 164 (1967).

controversy can be aired without the need for litigation and that rebuttal of offending speech is preferable to recourse to the courts.<sup>230</sup>

While access to the media is a valuable consideration when determining whether a plaintiff is a public figure,<sup>231</sup> the Supreme Court did not define the particular channels of communication a public figure should have access to.<sup>232</sup> At the time of the *Gertz* decision, the Justices were primarily referencing print and broadcast media.<sup>233</sup> But there is no doubt that the type of media referenced in *Gertz* has evolved.<sup>234</sup> Now, rather than simply print newspapers and television, there are many different forms of media, each with very distinct characteristics that impact the channel's effectiveness in counteracting false statements.<sup>235</sup> Surely, not all types of media are equal when considering whether the individual can properly counter another's false statements.<sup>236</sup> For targets of online public shaming, the central question for the access-to-media requirement is whether simply having access to social media sites affords them with a meaningful opportunity to rebut false, defamatory statements.<sup>237</sup>

An individual's access to media involves two characteristics: (1) the plaintiff must have a "realistic opportunity to counteract false statements,"<sup>238</sup> and (2) the channel must give the plaintiff regular and continuous access.<sup>239</sup> As to the second requirement, social media is accessible to any person with

230. *Reuber v. Food Chem. News, Inc.*, 925 F.2d 703, 708–09 (4th Cir. 1991). This view is in line with the self-help principle in *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 344 (1974).

231. *Cf. supra* notes 139–40 and accompanying text. Some courts do not apply the access-to-media requirement, but failing to analyze the characteristic in the social media context would be a misapplication of Supreme Court precedent. O'Connor, *supra* note 104, at 520.

232. O'Connor, *supra* note 104, at 514.

233. *Id.* The Court likely assumed that "it was in those media that defamation could be expected to originate, and it was in those same media that the plaintiff should seek to rebut such defamation." *Id.*

234. *See id.* at 521.

235. *See id.* at 531–32.

236. *See, e.g.,* Messenger & Delaney, *supra* note 117, at 5 ("Plaintiffs now have access to an almost endless amount of media sources, many of which lack effectiveness (for example, a webpage that is read by only a small number of people). Consequently, it is possible that lower courts will adapt their analysis to evaluate the access requirement with greater frequency.")

237. *See infra* notes 238–75 and accompanying text.

238. *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 344 (1974); Messenger & Delaney, *supra* note 117, at 5.

239. *Hutchinson v. Proxmire*, 443 U.S. 111, 136 (1979).

Internet access and will not be interrupted unless the platform has reason to suspend or terminate the user's account,<sup>240</sup> so a target of social media backlash will typically have regular and continuous access to social media.<sup>241</sup> Whether a social media site offers a "realistic opportunity to counteract false statements" is a more difficult question.<sup>242</sup>

Today, the average individual with computer access can post on a blog, upload a video to YouTube, send a tweet on Twitter, or update his or her status on Facebook and reach virtually anyone in the world.<sup>243</sup> Some urge that because of how easy it is to post on social media sites, "ordinary citizens have historically unprecedented access to effective communication channels" and therefore are able to sufficiently counter defamatory statements without needing a day in court.<sup>244</sup> However, simply having access to social media sites should not be the end of the inquiry.<sup>245</sup> Social media is much less effective in rebutting false statements than the traditional forms of print and broadcast media.<sup>246</sup>

An ability to counter false statements would logically "assume that an effective channel of communication would reach an audience of *similar composition* to the one that originally heard the defamation."<sup>247</sup> If the

240. See O'Connor, *supra* note 104, at 508.

241. See Messenger & Delaney, *supra* note 117, at 6.

242. See O'Connor, *supra* note 104, at 514 (quoting *Gertz*, 418 U.S. at 344).

243. See O'Connor, *supra* note 104, at 508. Note, however, that the ability to reach the public through use of social media platforms depends on the privacy settings of that user's particular account. See, e.g., *Protecting and Unprotecting Your Tweets*, TWITTER HELP CTR., <https://support.twitter.com/articles/20169886?lang=en> (last visited Nov. 16, 2016).

244. Lat & Shemtob, *supra* note 112, at 410–11. The authors emphasize that now the average citizen can refute false statements through a proliferation of online outlets, such as a blog or social networking site, and that "if false rumors started online, refuting them online may be the most effective response." *Id.* However, the authors do not take into account whether the access to these outlets is effective in rebutting false statements that have reached a widespread audience. See O'Connor, *supra* note 104, at 530–31 (arguing that *Gertz*'s definition of access to media includes defending oneself against the audience who read the defamatory statements). In discussing *Gertz*'s access-to-media inquiry, one court noted that "with the advent of the [I]nternet and the widespread use of social media, this assessment is less true than it once was." *Cummins v. Bat World Sanctuary*, No. 02-12-00285-CV, 2015 WL 1641144, at \*8 (Tex. App. Apr. 9, 2015), *reh'g overruled* (Apr. 30, 2015), *review denied* (Aug. 28, 2015).

245. See *infra* notes 246–75 and accompanying text.

246. See *infra* notes 247–75 and accompanying text.

247. Messenger & Delaney, *supra* note 117, at 5 (emphasis added); see also O'Connor, *supra* note 104, at 530–31 (arguing that *Gertz* set out a "relatively narrow definition of access to media: not one

attempt to counteract the false statement is heard by a much smaller audience, the injured plaintiff did not have an effective channel of communication.<sup>248</sup> *Gertz* presumably gave weight to a public figure's ability to access the media because he or she could issue a public statement in one of the few mediums of discourse and reach the same audience that heard the defamatory statement in the first place.<sup>249</sup>

In the social media backlash scenario, using social media to reach an audience of similar composition to the one that originally heard the defamatory statements would be nearly impossible.<sup>250</sup> Usually, thousands of tweets or Facebook posts contribute to a target's public shaming.<sup>251</sup> Beyond the users directly involved, others can see the defamatory comments either while the backlash is taking place through the trending topic feature on Twitter or even years later through a search engine.<sup>252</sup> It would be nearly impossible to reach the same thousands of people again—or even a majority of them—because the average plaintiff cannot send out a strong enough message on social media platforms to be heard.<sup>253</sup>

The court in *Hibdon v. Grabowski*<sup>254</sup> found that the plaintiff was a public figure in part due to his ability to access the media.<sup>255</sup> The plaintiff, Hibdon, owned a jet ski customizing business, advertised his business online, and published descriptions of his jet skis' speed on an Internet news

that encompasses any and all opportunities to be heard by all varieties of audiences, but rather the opportunity to defend oneself to the audience . . . that initially received the damaging information").

248. See, e.g., *Messenger & Delaney*, *supra* note 117, at 6.

249. See O'Connor, *supra* note 104, at 530; *supra* notes 135–38 and accompanying text.

250. See *infra* notes 251–52 and accompanying text.

251. See *supra* Part II.

252. See *supra* Part II.

253. See *Messenger & Delaney*, *supra* note 117, at 6–7. It is important to note at least one Justice on the Supreme Court has discounted that the access-to-media requirement must be realistic:

[D]ifficulty in reaching all those who may have read the alleged falsehood surely ought not preclude a finding that [the plaintiff] was a public figure under *Gertz*. *Gertz* set no absolute requirement that an individual be able fully to counter falsehoods through self-help in order to be a public figure. We viewed the availability of the self-help remedy as a relative matter in *Gertz*, and set it forth as a minor consideration in determining whether an individual is a public figure.

*Time, Inc. v. Firestone*, 424 U.S. 448, 486 (1976) (Marshall, J., dissenting).

254. 195 S.W.3d 48 (Tenn. Ct. App. 2005).

255. *Id.* at 59–62.

group called *rec.sport.jetski*.<sup>256</sup> After being featured in *SPLASH Magazine* on four different occasions, the defendants posted on *rec.sport.jetski* challenging the quality of Hibdon's work, so Hibdon sued for defamation.<sup>257</sup> The court found that Hibdon was a limited-purpose public figure.<sup>258</sup> In analyzing his access to the media, the court concluded that there was no doubt that Hibdon could effectively counteract the defendants' statements through the news group and through *SPLASH Magazine*.<sup>259</sup>

Assuming that a social media user can only have effective access to media if he or she can reach an audience of a similar composition,<sup>260</sup> the average social media user does not have an effective command of social media to be heard by thousands, unlike the plaintiff in *Hibdon* where access to a specific Internet news group allowed him to directly refute other users' statements.<sup>261</sup> One source estimates that the average Twitter user has 208 followers,<sup>262</sup> whereas, for example, a celebrity public figure like Taylor Swift has 81.8 million followers.<sup>263</sup> Twitter users do not have equal voices,<sup>264</sup> so courts should look to see the command of the channel that the plaintiff has by looking at his or her number of followers.

Justine Sacco's experience can be illustrative.<sup>265</sup> When Justine Sacco

256. *Id.* at 53.

257. *Id.* at 53–54, 59.

258. *Id.* at 62.

259. *Id.*

260. *See supra* notes 247–48 and accompanying text.

261. *See supra* notes 255–59 and accompanying text. In *Hibdon*, the harm was also not as widespread as in the social media backlash scenario because there was no mention of the comments being accessible through search engines and no evidence showing that persons outside of the group read the defamatory statements. 195 S.W.3d at 62; *supra* notes 255–59 and accompanying text.

262. Craig Smith, *By the Numbers: 170+ Amazing Twitter Statistics*, DMR, <http://expandedramblings.com/index.php/march-2013-by-the-numbers-a-few-amazing-twitter-stats/> (last updated Aug. 28, 2016).

263. Taylor Swift (@taylorswift13), TWITTER, <https://twitter.com/taylorswift13> (last visited Nov. 16, 2016).

264. Hunt, *supra* note 48, at 582 (“[P]osting a rebuttal on Twitter does not guarantee that anyone will read the reply.”).

265. *See infra* notes 266–72 and accompanying text. Note, however, that Justine was in fact approached by ABC News and did issue a public statement. Kami Dimitrova, *Justine Sacco, Fired After Tweet on AIDS in Africa, Issues Apology*, ABC NEWS (Dec. 22, 2013), <http://abcnews.go.com/International/justine-sacco-fired-tweet-aids-africa-issues-apology/story?id=21301833>. For purposes of this Comment, the Author has not taken into account the possibility that a target of social media backlash may be approached by the media. The access-to-media analysis in this

posted her tweet on Twitter, she only had 170 followers.<sup>266</sup> Justine, like the average social media user, simply does not have a strong enough social media presence to reach the relevant audience of the defamatory content.<sup>267</sup> Not only were hundreds of thousands of tweets sent about Justine, but between December 20th and the end of the month, she was googled 1.2 million times.<sup>268</sup> Even if Justine were to use social media to rebut the false statements made about her, it is unlikely that she could reach anywhere near the amount of people who believed she was a racist or an heiress to a \$4.8 billion fortune.<sup>269</sup> Another issue is that, when searching for Justine's name on a search engine, a user would have only seen the thousands of defamatory posts about her and not her rebuttal tweet.<sup>270</sup> Therefore, Justine would be unable to effectively rebut the statements made about her because the number of followers she has is drastically less than the number of people who read the defamatory comments.<sup>271</sup>

After the digital mob ran its course, Justine had no self-help remedy to reach the hundreds of thousands of people who read the false and defamatory remarks made about her, making her vulnerable to the personal attacks and potentially defamatory statements of others.<sup>272</sup> A person's vulnerability to another's statements regarding his or her reputation was specifically referenced in *Gertz*: the Court did not require private persons to prove statements were made with actual malice because the persons' only option for recourse would be through litigation and thus requiring proof of actual malice would be too high a bar.<sup>273</sup> Despite having social media sites at our fingertips, the increased access to these channels of communication does not make them effective.<sup>274</sup> The average social media user does not have enough followers to effectively rebut false statements made about

---

Comment will necessarily be limited to the channel of communication that the user already has access to—social media.

266. RONSON, *supra* note 1, at 68.

267. See Messenger & Delaney, *supra* note 117, at 6.

268. RONSON, *supra* note 1, at 71.

269. *Id.* at 77.

270. See Messenger & Delaney, *supra* note 117, at 6.

271. See, e.g., Hunt, *supra* note 48, at 582.

272. See *supra* notes 266–71 and accompanying text.

273. See *Phila. Newspapers, Inc. v. Hepps*, 475 U.S. 767, 789 (1986) (Stevens, J., dissenting).

274. See *supra* notes 251–71 and accompanying text.

them.<sup>275</sup> While there are many factors at play in the access-to-media analysis, the average social media user does not have the effective access to channels of communication contemplated by the Supreme Court and therefore should not be deemed a limited-purpose public figure.

*B. Mob Shaming of a Social Media User Should Not Be Considered a Matter of Public Concern*

If a target of social media backlash sues for defamation, courts will need to decide whether that social media user had voluntarily inserted herself into a public controversy such that other users should receive extra protection when speaking out against her.<sup>276</sup> But when does a public controversy exist? Whether the subject matter of a defamatory statement involves a public controversy can be gleaned from analyzing the actual text of the tweet or Facebook comment.<sup>277</sup> Courts already have difficulty distinguishing matters of public concern and private life,<sup>278</sup> with the Supreme Court providing only that courts should analyze the “content, form, and context” of the speech.<sup>279</sup> Speech on the Internet poses a unique issue: if a comment is posted on the Internet and is viewable by the public or if users are collectively discussing a particular topic, does that automatically convert the topic into a matter of public concern?<sup>280</sup> Of course, if a topic is already a public controversy in the general public or in traditional media, it will remain a matter of public concern if it is discussed on social media sites.<sup>281</sup> For example, criticizing an individual in public office is generally a topic of

275. Hunt, *supra* note 48, at 582 (“An average user would likely be hard-pressed to effectively rebut a defamatory statement posted by a user with a larger than average number of followers. Moreover, for the rebuttal to be ‘effective,’ it would have to reach at least some of the audience that had access to the original statement.”).

276. See *supra* Section III.A; see also Hunt, *supra* note 48 at 590 (showing voluntary injection into the controversy is a key element in determining if an individual is a limited-purpose public figure).

277. Hunt, *supra* note 48, at 591.

278. See *supra* Section III.C.2.

279. *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 761 (1985).

280. See *infra* 281–282 and accompanying text.

281. Hunt, *supra* note 48, at 591. The media channel where the topic is discussed does not affect the public figure analysis, and courts “should proceed in the same manner that it would if the defamatory statement were published in a different form.” *Id.*

public concern, and that designation does not change if the criticisms are made on the Internet.<sup>282</sup>

But the hundreds of thousands of tweets backlashing against individuals that first arise on social media platforms are not addressing topics that would individually be considered a matter of public concern.<sup>283</sup> For Twibel, courts should generally consider contemporaneously published tweets as part of the context of the expression.<sup>284</sup> At least one commentator has opined that it is possible that a topic can become a matter of public concern simply because it is discussed on Twitter.<sup>285</sup> This commentator argues that courts should gauge the specific online community's interest in a topic by searching the subject matter through the Twitter Search feature<sup>286</sup> or, if a user includes a hashtag relevant to the defamatory tweet, that the court should look to the amount of times other users used the same hashtag.<sup>287</sup> Both of these methods serve as good starting points for a court to discern if the topic is popular within a specific online community but should not be the end of the analysis.<sup>288</sup> Simply deciding that a subject is a public controversy because it has been frequently referenced on Twitter contradicts Supreme Court precedent: if all popular subjects on Twitter became public controversies, private individuals would automatically be transformed into public figures “just by becoming involved in or associated with a matter that attracts public attention.”<sup>289</sup>

282. See, e.g., *Ghanam v. Does*, 845 N.W.2d 128, 136 (Mich. Ct. App. 2014) (noting that the “First Amendment provides strong protections to those who use their freedom of speech to criticize public officials over public issues”).

283. See *supra* Part II.

284. See, e.g., *Hunt*, *supra* 48, at 593–95. This commentator further argues that fairness dictates this approach because the 140-character limitation on tweets restricts the court's ability to determine whether the expression is a matter of public concern. *Id.* at 593–94. The context of the expression matters because it follows the test set out by the Supreme Court. *Dun & Bradstreet, Inc.*, 472 U.S. at 761.

285. *Hunt*, *supra* note 48, at 563.

286. *Using Advanced Search*, TWITTER HELP CTR., <https://support.twitter.com/articles/71577?lang=en> (last visited Nov. 16, 2016). The advanced Twitter search helps a user find a tweet posted on a specific date or said by specific people. *Id.*

287. See *supra* notes 45–48 and accompanying text.

288. See *infra* note 289 and accompanying text.

289. *Wolston v. Reader's Digest Ass'n*, 443 U.S. 157, 167 (1979). The Court also warned against defining public controversies as “all controversies of interest to the public” because then too many individuals would be considered public figures. *Time, Inc. v. Firestone*, 424 U.S. 448, 454 (1976).

For Justine Sacco, one of the hashtags used in the abusive tweets was #HasJustineLandedYet, because thousands of social media users waited for Justine to find out she had lost her job.<sup>290</sup> Her social media demise was a top trending topic on Twitter and was known within the social media sphere and beyond.<sup>291</sup> If the court only looked to see if the hashtag was used with a high frequency, then condemning Justine as a racist would have been a matter of public concern.<sup>292</sup>

For targets of widespread shaming, courts should follow the reasoning in *D.C. v. R.R.*<sup>293</sup> and consider the Internet as an extension of the material world.<sup>294</sup> In *D.C.*, the court held that mere *access* to the Internet was insufficient to render someone a public figure.<sup>295</sup> “Thus, this court viewed the Internet as an extension to the material world, where participation in activities that are widely considered attributable to a private forum—talking and socially interacting with friends—is not considered sufficient in itself to turn someone into a public figure.”<sup>296</sup>

Just as social interaction among friends should not convert a topic into a

290. RONSON, *supra* note 1, at 70.

291. *See supra* note 7 and accompanying text.

292. *See supra* note 208 and accompanying text.

293. 106 Cal. Rptr. 3d 399 (Ct. App. 2010).

294. *See id.*; Lafferman, *supra* note 38, at 238.

295. *See id.* at 426, 428–29; *see also* Lafferman, *supra* note 38, at 238. In *D.C.*, a high school student created a website to assist in his pursuit of an entertainment career. 106 Cal. Rptr. 3d at 404. Several students made derogatory comments on the website about “his perceived sexual orientation and threaten[ed] him with bodily harm.” *Id.* at 404–05. The court cited to a few hateful comments:

One post read, “Faggot, I’m going to kill you.” Another read, “[You need] a quick and painless death.” One student wrote, “Fuck you in your fucking fuck hole.” Another commented, “Fucking ass clown. Nigga what?” One post announced, “You are now officially wanted dead or alive.” Another threatened, “I will personally unleash my manseed in those golden brown eyes.”

*Id.* at 406. The court sympathized with the target of such remarks, stating that “[t]he students who posted . . . threats sought to destroy D.C.’s life.” *Id.* Some of the injuries the student suffered included emotional harm, loss of income, and the cost of moving out of the community. *Id.*

296. Lafferman, *supra* note 38, at 238. The court emphasized that merely published information on a website “should not turn otherwise private information . . . into a matter of public interest.” *D.C.*, 106 Cal. Rptr. 3d at 426 (quoting *Du Charme v. Int’l Bhd. of Elec. Workers*, 1 Cal. Rptr. 3d 501, 509 (Ct. App. 2003)). In concluding that D.C. was not a public figure, the court noted that the threatening messages did not mention a public issue nor was D.C.’s personal life of widespread interest to the public. *Id.* at 428.

matter of public concern simply because it is discussed on the Internet,<sup>297</sup> a group of people collectively defaming a person's reputation should not become a matter of public concern simply because of the large number of users involved and the public nature of the defamation. For example, just as in *D.C.* where the posters were singularly out to destroy the high school teenager's life and not discussing a matter of public concern, Justine Sacco's defamers were not discussing an important social topic; they were trying to ruin her reputation and take away her job.<sup>298</sup> While her ruined reputation is of the utmost importance to herself, it is not a matter of grave concern to the public. Typically, targets of social media backlash are not part of an ongoing public discussion when they post their inflammatory comment.<sup>299</sup> However, a court might still decide a user's moral transgressions, such as Justine's poor attempt at humor in saying she could not get AIDS because she is white, became a topic of widespread public interest.<sup>300</sup> The court would then need to analyze whether the user became voluntarily involved in the controversy.<sup>301</sup>

*C. A Single, Isolated Social Media Post Is Not Enough to Become Voluntarily Involved in a Controversy*

A number of courts have considered whether a particular plaintiff was voluntarily assuming public scrutiny based on his or her Internet activity.<sup>302</sup> Some courts have found that because the Internet is a public forum, online speakers are public figures.<sup>303</sup> For example, in *Tipton v. Warshavsky*,<sup>304</sup> a website content provider brought a defamation claim against a website host.<sup>305</sup> The court found that Tipton was a limited-purpose public figure

---

297. *See supra* notes 295–96 and accompanying text.

298. *See supra* notes 7–8 and accompanying text.

299. *But see* Backlund v. Stone, No. B235173, 2012 WL 3800883, at \*5 (Cal. Ct. App. Sept. 4, 2012) (concluding that the plaintiff had voluntarily subjected himself to potential scrutiny by the public through his contributions to the ongoing public discussion of sextortion).

300. *Cf. D.C.*, 106 Cal. Rptr. 3d at 428 (noting that D.C.'s personal life was not a matter of widespread public interest).

301. *See supra* Section III.B.

302. Hunt, *supra* note 48, at 591.

303. *See infra* notes 304–08 and accompanying text.

304. 32 F. App'x 293 (9th Cir. 2002).

305. *Id.* at 294.

“because he voluntarily involved himself in public life by inviting attention and comment” on the website,<sup>306</sup> echoing *Gertz*, where the Court noted that public figures often invite attention and comment.<sup>307</sup> The court in *Tipton* seemed to be applying a more liberal interpretation of the voluntariness principle than the court in *Gertz*.<sup>308</sup>

However, other courts have found that the Internet is a private forum.<sup>309</sup> In *D.C. v. R.R.*,<sup>310</sup> the court treated the Internet as an inherently private forum<sup>311</sup> and seemed concerned that finding that the plaintiff was a public figure would convert millions of teenagers into public figures through their use of social media sites such as Myspace, Facebook, and YouTube.<sup>312</sup> By using a website to pursue an entertainment career, D.C. “did not, ipso facto, put him[self] in the public eye.”<sup>313</sup>

In contrast, the California Court of Appeals in *Backlund v. Stone*<sup>314</sup> found that a plaintiff who voluntarily chose to be featured on television and in print as an expert in cyber-harassment was a public figure because he “voluntarily subjected [himself] to inevitable scrutiny and potential ridicule by the public and the media.”<sup>315</sup> Additionally, the speaker contributed to an “ongoing public discussion about sextortion,” and therefore, the comments “were not private at all: they were a public comment about a publicly disseminated threat against her made by public figure Stone.”<sup>316</sup>

306. *Id.* at 295. The court further concluded that some of the statements made about Tipton on ourfirsttime.com raised a triable issue of fact as to whether the defendant acted with reckless disregard for the truth. *Id.*

307. See *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 345 (1974).

308. See *supra* notes 151–52 and accompanying text.

309. See, e.g., *Franklin Prescriptions, Inc. v. N.Y. Times Co.*, 267 F. Supp. 2d 425, 437 (E.D. Pa. 2003), *aff'd*, 424 F.3d 336 (3d Cir. 2005); *D.C. v. R.R.*, 106 Cal. Rptr. 3d 399, 428–30 (Ct. App. 2010); *Bieter v. Fetzer*, No. A04-1034, 2005 WL 89484, at \*4 (Minn. Ct. App. Jan. 18, 2005).

310. 106 Cal. Rptr. 3d 399 (Ct. App. 2010).

311. *Id.* at 428–30.

312. *Id.* “Millions of teenagers use MySpace, Facebook, and YouTube to display their interests and talents, but the posting of that information hardly makes them celebrities.” *Id.* at 428.

313. *Id.*

314. No. B235173, 2012 WL 3800883 (Cal. Ct. App. Sept. 4, 2012).

315. *Id.* at \*5 (quoting *Seelig v. Infinity Broad. Corp.*, 119 Cal. Rptr. 2d 108, 115 (2002)).

316. *Id.* “Stone became a limited public figure by operating a publicly accessible website that published lewd photos of minors, and by seeking the public eye when he appeared on television and in print media to discuss the topic of sextortion.” *Id.* at \*6.

Similarly, in *Franklin Prescriptions, Inc. v. New York Times Co.*<sup>317</sup> and *Bieter v. Fetzer*,<sup>318</sup> the courts were looking for specific, voluntary activity that would convert the plaintiff into a public figure on the Internet.<sup>319</sup> In *Franklin Prescriptions, Inc.*, the court found that a prescription company was not a limited-purpose public figure.<sup>320</sup> In the Third Circuit, courts employ a two-pronged limited public-figure test: first, the court must determine whether the defamatory statement involves a public controversy, and second, the court must “inquire into the nature and extent of plaintiff’s involvement within that controversy.”<sup>321</sup> The court, assuming there was a public debate regarding the Internet’s ability to make drugs more readily available and affordable, found that the plaintiff’s involvement in the controversy was not enough to convert him into a public figure: “[A] pharmacy that uses the Internet for information purposes only, and does not sell or take orders over the Internet, is not an ‘online’ pharmacy contributing to such a debate.”<sup>322</sup> The court found that “Franklin’s limited involvement with the Internet [was] too remote and tenuous to consider it a public figure.”<sup>323</sup>

Alternatively, in *Bieter v. Fetzer*,<sup>324</sup> a plaintiff claimed that a conspiracy theorist defamed him when the theorist published a newspaper article

317. 267 F. Supp. 2d 425 (E.D. Pa. 2003), *aff’d*, 424 F.3d 336 (3d Cir. 2005).

318. No. A04-1034, 2005 WL 89484, at \*4 (Minn. Ct. App. Jan. 18, 2005).

319. *Id.*; *Franklin Prescriptions, Inc.*, 267 F. Supp. 2d at 425.

320. *Franklin Prescriptions Inc.*, 267 F. Supp. 2d at 436–37.

321. *Id.* A public controversy has been defined as a dispute that “has received public attention because its ramifications will be felt by persons who are not direct participants.” *Waldbaum v. Fairchild Publ’ns, Inc.*, 627 F.2d 1287, 1296 (D.C. Cir. 1980).

322. *Franklin Prescriptions, Inc.*, 267 F. Supp. 2d at 437. The court considered the defendant’s argument that there was a controversy surrounding the Internet’s ability to make drugs available and affordable as a post hoc attempt to create a controversy. *Id.* at 436.

323. *Id.* at 437. The plaintiff was a neutral party that did not place itself into any online controversy and therefore was too remote a party to be required to meet any actual malice standard. *Id.*

Franklin did not inject itself into any controversy. Franklin merely provided an information only Website on the Internet and did not invite public attention, comment or criticism regarding the controversy of making drugs available via the Internet. Again, Franklin does not take or fill prescription orders online and does not allow for communication between the pharmacy and Internet users.

*Id.* at 440.

324. No. A04-1034, 2005 WL 89484 (Minn. Ct. App. Jan. 18, 2005).

asserting that some senators were involved in the sabotage of another senator's airplane, resulting in a crash that killed him.<sup>325</sup> To rebut the article, the plaintiff started an online chatroom of the theorist's claims.<sup>326</sup> In first analyzing whether a public controversy existed, the court held that an alleged assassination of a public official by a senator of the opposing political party is "a matter of grave concern for the entire nation."<sup>327</sup> The court found that it was clear that the plaintiff voluntarily thrust himself into the assassination controversy by forming the chatroom and inviting discussion.<sup>328</sup>

In the typical social media backlash scenario, a court likely could not point to sufficient conduct, like in the cases above, that would show that the plaintiffs voluntarily inserted themselves into a public controversy online.<sup>329</sup> Having a social media site does expose a user to the public.<sup>330</sup> Further, because individuals' lives are now so connected to their online selves on social media, individuals are not necessarily assuming the risk of public scrutiny.<sup>331</sup> The court should instead look to the plaintiff's post or social media history to determine whether the user had voluntarily inserted herself into a public controversy.<sup>332</sup>

In the public-shaming scenario, the user has typically posted a comment or picture that the public has found reprehensible.<sup>333</sup> For Justine Sacco, it was an ill-advised joke about AIDS in Africa.<sup>334</sup> Assuming this is a topic of public concern, Justine did not "rise to the forefront of a public controversy."<sup>335</sup> She was simply attempting a joke, and therefore, her involvement in the AIDS controversy was so tenuous that she was not necessarily assuming the risk of being scrutinized by the public for her

---

325. *Id.* at \*1.

326. *Id.*

327. *Id.* at \*3.

328. *Id.*

329. *See infra* notes 330–34 and accompanying text.

330. *See supra* notes 311–13 and accompanying text.

331. *Cf.* Elizabeth Atkins, *Spying on Americans: At What Point Does the NSA's Collection and Searching of Metadata Violate the Fourth Amendment?*, 10 WASH. J.L. TECH. & ARTS 51, 83 (2014).

332. *See infra* notes 333–38 and accompanying text.

333. *See supra* Part II.

334. *See supra* notes 69–73 and accompanying text.

335. *See supra* note 151 and accompanying text.

comment.<sup>336</sup> Therefore, looking at the plaintiff's initial posting that prompted the social media backlash will assist courts in determining whether the plaintiff truly assumes the risk of public scrutiny.<sup>337</sup> For the average target of social media backlash who posted one isolated comment that the public backlashed against, any involvement in the resolution of the controversy is likely too tenuous to find that the user was attempting to influence the resolution of an "important public question."<sup>338</sup>

## VI. CONCLUSION

With the advent of social media, defamatory remarks now can become worldwide trending topics in a matter of hours, devastating an individual's reputation in a manner not possible before the days of tweets and hashtags.<sup>339</sup> The sad truth is that it has become a regular occurrence for social media users to flood the Internet with false criticisms of others—most often strangers—when these targets have seemingly exhibited a trait the public finds distasteful.<sup>340</sup> Despite that many of these remarks are hastily constructed, they have had real-life consequences.<sup>341</sup> Victims of these digital mobs have suffered from depression and insomnia.<sup>342</sup> They have been scared to leave their homes.<sup>343</sup> Beyond emotional abuse, targets have consistently lost their jobs and livelihoods as a result of spontaneous public

---

336. *Cf.* *Backlund v. Stone*, No. B235173, 2012 WL 3800883, at \*5 (Cal. Ct. App. Sept. 4, 2012) (concluding that the plaintiff voluntarily subjected himself to potential public scrutiny through his contributions to the ongoing public discussion of sextortion). If the court applies a rule that analyzes whether the plaintiff had a "position of prominence in the public controversy," then simply one tweet on social media regarding the AIDS epidemic in Africa would in no way put Justine at the forefront of that controversy. *See Contemporary Mission, Inc. v. N.Y. Times Co.*, 842 F.2d 612, 617 (2d Cir. 1988).

Using another example, when Lindsay Stone posted her offensive photograph at the Arlington National Cemetery, a public controversy did not exist about honoring veterans who have lost their lives. *See supra* note 73. Even though there was public concern regarding her actions after the fact, the inquiry does not focus on a post hoc controversy. *See supra* note 322.

337. *See supra* note 320 and accompanying text.

338. *Curtis Publishing Co. v. Butts*, 388 U.S. 130, 155 (1967) (Warren, C.J., concurring).

339. *See supra* Part II.

340. *See supra* Part II.

341. *See supra* Part II.

342. *See supra* note 86 and accompanying text; *see also supra* Part II.

343. *See supra* note 86 and accompanying text; *see also supra* Part II.

outcry.<sup>344</sup> Users are helpless against these digital waves of defamatory remarks that result in devastating harm.<sup>345</sup>

The most viable remedy to help ease this harm is the law of defamation.<sup>346</sup> The average targets of social media backlash are not famous; they are simply private individuals with a modest social media presence.<sup>347</sup> However, some might argue that being the subject of a top-trending topic on Twitter would render these victims public figures, requiring them to prove the defamatory statements were made with actual malice.<sup>348</sup> But the well-established justifications for the public-figure doctrine simply do not fit.<sup>349</sup> First, social media does not provide the average user with a realistic ability to rebut the false statements that have reached thousands.<sup>350</sup> Further, group shaming an individual for a perceived reprehensible act should not be considered a matter of public concern.<sup>351</sup> Finally, a social media user's isolated comment is not enough to thrust that user into the forefront of a controversy.<sup>352</sup>

Unless defamation law is applied with care to social media backlash scenarios, mob justice will rule over legal justice.<sup>353</sup> These victims must remain private individuals distinguishable from public figures, which properly keeps defamation law as a viable remedy to ease the harm that stems from these now-common public shamings.<sup>354</sup>

Cory Batza\*

---

344. *See supra* Part II.

345. *See supra* Part II.

346. *See supra* Part III.

347. *See, e.g.*, notes 69–73 and accompanying text; *see also supra* Part V.

348. *See supra* Part V.

349. *See supra* Part V.

350. *See supra* Section V.A.

351. *See supra* Section V.B.

352. *See supra* Section V.C.

353. *See supra* Part II.

354. *See supra* Part V.

\* J.D. Candidate, Pepperdine University School of Law; B.S. in Business Administration, Pepperdine University. I would like to first thank the staff of *Pepperdine Law Review* for painstakingly checking my sources, Bluebook citations, and grammar for correctness. I would also like to thank Zach Tafoya for helping me choose the topic of this Comment during a short car ride and Lindsay Gochnour for giving me the confidence to see this Comment through to its completion. Finally, thank you to my father, David Batza, who inspired in me, among countless other things, a love for reading and writing.