

1 Joseph R. Saveri (SBN 130064)  
2 Steven N. Williams (SBN 175489)  
3 Kevin Rayhill (SBN 267496)  
4 Kate Malone (SBN 290884)  
5 Kyle Quackenbush (SBN 322401)  
6 **JOSEPH SAVERI LAW FIRM, INC.**  
7 601 California Street, Suite 1000  
8 San Francisco, CA 94108  
9 Telephone: (415) 500-6800  
10 Facsimile: (415) 395-9940  
11 jsaveri@saverilawfirm.com  
12 swilliams@saverilawfirm.com  
13 krayhill@saverilawfirm.com  
14 kmalone@saverilawfirm.com  
15 kquackenbush@saverilawfirm.com

*Attorneys for Plaintiff and the Proposed Class*

11 **SUPERIOR COURT OF THE STATE OF CALIFORNIA**  
12 **COUNTY OF SAN MATEO**

13 **JANE DOE**, individually and on behalf of all  
14 others similarly situated,

15 *Plaintiffs,*

16 v.

17 **YOUTUBE, INC.**

18 *Defendant.*

Civil Action No. 20-CIV-04023

**COMPLAINT AND DEMAND FOR JURY TRIAL**

CLASS ACTION

19 **NEED FOR ACTION**

20  
21  
22 1. Plaintiff JANE DOE seeks to protect herself and all others similarly situated from the  
23 dangers of psychological trauma resulting from exposure to graphic and objectionable content on  
24 YouTube, Inc.'s platform and YouTube's failure to provide a safe workplace for the thousands of  
25 contractors that scrub YouTube's platform of disturbing content.

26 2. Every day, YouTube users upload millions of videos to its platform. Millions of these  
27 uploads include graphic and objectionable content such as child sexual abuse, rape, torture, bestiality,  
28 beheadings, suicide, and murder. To maintain a sanitized platform, maximize its already vast profits, and

1 cultivate its public image, YouTube relies on people like Plaintiff—known as “Content Moderators”—  
2 to view those videos and remove any that violate the corporation’s terms of use.

3 3. Working at YouTube’s offices in California and offices of contract employers (“YouTube  
4 Vendors”) across the country, Content Moderators—including Plaintiff—witnessed thousands of acts  
5 of extreme and graphic violence and sexual assault. From genocide in Myanmar to mass shootings in Las  
6 Vegas and Christ Church to videos of children being raped and animals being mutilated, Content  
7 Moderators spend hours a day making sure that disturbing content like this never appears to YouTube’s  
8 users.

9 4. Content Moderators also face repeated exposure to conspiracy theories, fringe beliefs,  
10 and political disinformation—from false information about participating in the census, to lies about a  
11 political candidate’s citizenship status or eligibility for public office, to manipulated and/or doctored  
12 videos of elected officials, to denials that the Holocaust occurred, to suggestions that Covid-19 is a  
13 fraud. This type of content has destabilized society and often features objectionable content.

14 5. As a result of unmitigated exposure to highly toxic and extremely disturbing images  
15 viewed using YouTube’s proprietary “single review tool” (“SRT”), Plaintiff developed and suffers from  
16 significant psychological trauma including anxiety, depression and symptoms associated with PTSD.

17 6. To cultivate its image, YouTube (through its parent company Google, LLC) helped draft  
18 workplace safety standards to attempt to mitigate the negative psychological effects that viewing graphic  
19 and objectionable content has on Content Moderators. These safety standards include obtaining a  
20 candidate’s informed consent during the initial employment interview process; providing Content  
21 Moderators with robust and mandatory counseling and mental health support; altering the resolution,  
22 audio, size, and color of trauma-inducing images and videos; and training Content Moderators to  
23 recognize the physical and psychological symptoms of PTSD, anxiety, and depression.

24 7. Although these safety standards could not eliminate the risk that Content Moderators  
25 would develop negative psychological disorders after viewing graphic and disturbing content, these  
26 standards could have reduced the risk and mitigated the harm.



1 **PARTIES**

2 15. Plaintiff JANE DOE is a resident of Travis County, Texas. From approximately January  
3 16, 2018 until approximately August 24, 2019, Plaintiff worked as a Content Moderator, reviewing  
4 content for YouTube at an office located at 7700 West Parmer Lane, Austin Texas, 78717. During this  
5 period, Plaintiff was employed by Collabera, Inc., (“Collabera”). Plaintiff has been diagnosed with  
6 depression and suffers from symptoms associated with anxiety and PTSD. Because of the trauma JANE  
7 DOE suffered, and the very real threat that publicity from this case could exacerbate JANE DOE’s  
8 mental health problems, JANE DOE is using a pseudonym as provided for by California law.

9 16. Defendant YouTube, Inc., is incorporated under the laws of Delaware with its  
10 headquarters located at 901 Cherry Avenue, San Bruno, California. YouTube is a fully owned subsidiary  
11 of Google, LLC.

12 **FACTUAL ALLEGATIONS**

13 **A. Content moderators watch and remove some of the most depraved images on the internet to**  
14 **protect YouTube’s profits**

15 17. In fiscal year 2019, YouTube and Google combined made approximately \$150 billion in  
16 advertising revenue. In 2018, that number was nearly \$127 billion, and in 2017 that number was almost  
17 \$103 billion.

18 18. YouTube is attractive to companies and individuals that want to buy ads because of its  
19 immense user base. YouTube has over a billion monthly active users. These users value YouTube for its  
20 plethora of content and ability to share information—from learning how to ride a bike, to viewing the  
21 daily news, to watching funny home videos.

22 19. To generate this content, YouTube relies on users to upload videos to its platform.  
23 YouTube users upload 500 hours of video per minute, which works out to 30,000 hours per hour and  
24 720,000 hours per day. As Johanna Wright—Vice President of Product Management at YouTube—  
25 describes it, YouTube is the “video library for the world.”

26 20. Instead of scrutinizing content before it is shared with its users, YouTube relies on users  
27 to report inappropriate content. YouTube receives millions of user reports of potentially objectionable  
28 content on its platforms.

1           21.     Depending on how the content is flagged, YouTube directs the content to various  
2 queues, including “Violent Extremism” (“VE”), “Adult,” “Hate and Harassment,” “Child Sexual  
3 Abuse Imagery” (“CSAI”), “Redshift,” “Flagged,” and “Sexual Activity and Nudity.” Content  
4 Moderators then review the content to determine whether the content violates YouTube’s Community  
5 Guidelines. According to YouTube, these Content Moderators are “essential...because human  
6 judgment is critical to making contextualized decisions on content.”

7           22.     YouTube requires Content Moderators to review hundreds of thousands if not millions  
8 of potentially rule-breaking posts per week via YouTube’s SRT. For example, between June and  
9 December of 2017, YouTube’s Content Moderators reviewed nearly 2 million videos for violent  
10 extremist content alone.

11          23.     To tackle this immense amount of content, Susan Wojcicki—CEO of YouTube—  
12 pledged that YouTube would hire 10,000 Content Moderators in 2018. Plaintiff is informed and  
13 believes that there are thousands of Content Moderators that review YouTube’s content in the United  
14 States.

15           **B. Repeated exposure to graphic imagery can cause devastating psychological trauma,**  
16           **including PTSD, Anxiety, and Depression**

17          24.     It is well known that exposure to images of graphic violence can cause debilitating  
18 injuries, including PTSD, anxiety, and depression.

19          25.     In a study conducted by the National Crime Squad in the United Kingdom, seventy-six  
20 percent of law enforcement officers surveyed reported feeling emotional distress in response to  
21 exposure to child abuse on the internet. The same study, which was co-sponsored by the United  
22 Kingdom’s Association of Chief Police Officers, recommended that law enforcement agencies  
23 implement employee support programs to help officers manage the traumatic effects of exposure to  
24 child pornography.

25          26.     In a study of 600 employees of the Department of Justice’s Internet Crimes Against  
26 Children task force, the U.S. Marshals Service found that a quarter of the cybercrime investigators  
27 surveyed displayed symptoms of psychological trauma, including secondary traumatic stress.  
28

1           27.     Another study of cybercrime investigators from 2010 found that “greater exposure to  
2 disturbing media was related to higher levels of . . . secondary traumatic stress” and that “substantial  
3 percentages” of investigators exposed to disturbing media “reported poor psychological well-being.”

4           28.     The Eyewitness Media Hub has also studied the effects of viewing videos of graphic  
5 violence, including suicide bombing, and found that “40 percent of survey respondents said that  
6 viewing distressing eyewitness media has had a negative impact on their personal lives.”

7           29.     Whereas viewing or hearing about another person’s traumatic event used to be  
8 considered “secondary traumatic stress,” the current Diagnostic and Statistical Manual of Mental  
9 Disorders (American Psychiatric Association, 5<sup>th</sup> ed. 2013) (“DSM-5”) recognizes that secondary or  
10 indirect exposure to trauma, such as repeated or extreme exposure to aversive details of trauma through  
11 work-related media, meets the first diagnostic criterion for PTSD.

12           30.     While there is no way to eliminate the risk created by exposure to graphic and  
13 objectionable content, especially demanding job requirements or a lack of social support reduce  
14 resilience in the face of trauma exposure and increase the risk of developing debilitating psychological  
15 symptoms.

16           31.     Depending on many factors, individuals who have experienced psychological trauma  
17 may develop a range of subtle to significant physical and psychological symptoms, including extreme  
18 fatigue, dissociation, difficulty sleeping, excessive weight gain, anxiety, nausea, and other digestive  
19 issues.

20           32.     Trauma exposure and PTSD are also associated with increased risk of chronic health  
21 problems including cardiovascular conditions, pain syndromes, diabetes, and dementia.

22           33.     There is growing evidence that early identification and treatment of PTSD is important  
23 from a physical health perspective, as a number of meta-analyses have shown increased risk of  
24 cardiovascular, metabolic, and musculoskeletal disorders among patients with long-term PTSD.

25           34.     Psychological trauma and/or PTSD are also often associated with the onset or worsening  
26 of substance use disorders. Epidemiologic studies indicate that one-third to one-half of individuals with  
27 PTSD also have a substance use disorder. Compared to individuals without PTSD, those with PTSD  
28 have been shown to be more than twice as likely to meet the diagnostic criteria for alcohol abuse or

1 dependence; individuals with PTSD are also three to four times more likely to meet the diagnostic  
2 criteria for drug abuse or dependence.

3 35. PTSD symptoms may manifest soon after the traumatic experiences, or they may  
4 manifest later, sometimes months or years after trauma exposure.

5 36. An individual's risk of developing PTSD or associated symptoms may be reduced  
6 through prevention measures, categorized as primary, secondary, and tertiary interventions. Primary  
7 interventions are designed to increase resilience and lower the risk of future PTSD among the general  
8 population. Secondary interventions are designed to lower the risk of PTSD among individuals who  
9 have been exposed to trauma, even if they are not yet showing symptoms of traumatic stress. Finally,  
10 tertiary interventions are designed to prevent the worsening of symptoms and improve functioning in  
11 individuals who are already displaying symptoms of traumatic stress or who have been diagnosed with  
12 PTSD.

13 37. Individuals who develop PTSD or other mental health conditions following traumatic  
14 exposure require not only preventative measures but also treatment. Unlike prevention, treatment  
15 measures are aimed at symptom resolution and recovery from the disorder.

16 38. Preliminary screening is necessary to determine the types of prevention or treatment  
17 measures most appropriate for an individual.

18 **C. YouTube helped craft industry standards for mitigating the harm to Content Moderators**

19 39. Founded in 2006, the Technology Coalition was created "to develop technology  
20 solutions to disrupt the ability to use the Internet to exploit children or distribute child pornography."

21 40. Google (YouTube's parent company) was a member of the Technology Coalition at all  
22 times relevant to the allegations herein.

23 41. In January 2015, the Technology Coalition published an "Employee Resilience  
24 Guidebook for Handling Child Sex Abuse Images" (the "Guidebook").

25 42. According to the Guidebook, the technology industry "must support those employees  
26 who are the front line of this battle."  
27  
28

1           43.     The Guidebook recommends that internet companies implement a robust, formal  
2 “resilience” program to support Content Moderators’ well-being and mitigate the effects of exposure  
3 to trauma-inducing imagery.

4           44.     With respect to hiring Content Moderators, the Guidebook recommends:

- 5           a.     In an informational interview, “[u]se industry terms like ‘child sexual abuse  
6           imagery’ and ‘online child sexual exploitation’ to describe subject matter”;
- 7           b.     In an informational interview, “[e]ncourage candidate to go to websites [like the  
8           National Center for Missing and Exploited Children] to learn about the problem”;
- 9           c.     In follow-up interviews, “[d]iscuss candidate’s previous experience/knowledge with  
10           this type of content”;
- 11           d.     In follow-up interviews, “[d]iscuss candidate’s current level of comfort after  
12           learning more about the subject”;
- 13           e.     In follow-up interviews, “[a]llow candidate to talk with employees who handle  
14           content about their experience, coping methods, etc.”; and
- 15           f.     In follow-up interviews, “[b]e sure to discuss any voluntary and/or mandatory  
16           counseling programs that will be provided if candidate is hired.”

17           45.     With respect to safety on the job, the Guidebook recommends:

- 18           a.     Limiting the amount of time an employee is exposed to child sexual abuse imagery;
- 19           b.     Teaching moderators how to assess their own reaction to the images;
- 20           c.     Performing a controlled content exposure during the first week of employment with  
21           a seasoned team member and providing follow up counseling sessions to the new  
22           employee;
- 23           d.     Providing mandatory group and individual counseling sessions administered by a  
24           professional with specialized training in trauma intervention; and
- 25           e.     Permitting moderators to “opt-out” from viewing child sexual abuse imagery.



1           46.     The Technology Coalition also recommends the following practices for minimizing  
2 exposure to graphic content:

- 3           a.     Limiting time spent viewing disturbing media to “no more than four consecutive  
4           hours”;
- 5           b.     “Encouraging switching to other projects, which will allow professionals to get relief  
6           from viewing images and come back recharged and refreshed”;
- 7           c.     Using “industry-shared hashes to more easily detect and report [content] and in  
8           turn, limit employee exposure to these images. Hash technology allows for  
9           identification of exactly the same image previously seen and identified as  
10          objectionable”;
- 11          d.     Prohibiting Content Moderators from viewing child pornography one hour before  
12          the individuals leave work; and
- 13          e.     Permitting Content Moderators to take time off as a response to trauma.

14           47.     According to the Technology Coalition, if a company contracts with a third-party vendor  
15 to perform duties that may bring vendor employees in contact with graphic content, the company  
16 should clearly outline procedures to limit unnecessary exposure and should perform an initial audit of  
17 the independent contractor’s wellness procedures for its employees.

18           48.     The National Center for Missing and Exploited Children (“NCMEC”) also  
19 promulgates guidelines for protecting Content Moderators from psychological trauma. For instance,  
20 NCMEC recommends changing the color or resolution of the image, superimposing a grid over the  
21 image, changing the direction of the image, blurring portions of the image, reducing the size of the  
22 image, and muting audio.

23           49.     Based on these industry standards, some internet companies take steps to minimize  
24 harm to Content Moderators. For instance, at Microsoft, “[t]he photos are blurred, rendered in black  
25 and white, and shown only in thumbnail sizes. Audio is removed from video.” Filtering technology is  
26 used to distort images, and Content Moderators are provided with mandatory psychological counseling.

27           50.     At the UK’s Internet Watch Foundation, each applicant for a content moderator  
28 position is assessed for suitability by a psychologist, who asks about their support network, childhood

1 experiences, and triggers. Applicants are then interviewed about their work skills before proceeding to a  
2 final interview where they are exposed to child sexual abuse imagery. Candidates sit with two current  
3 Content Moderators and review a sequence of images getting progressively worse, working towards the  
4 worst kinds of sexual violence against children. This stage is designed to see how candidates cope and  
5 let them decide whether they wish to continue with the application process. Once they accept the job,  
6 Content Moderators have an enhanced background check before they start their six months' training,  
7 which involves understanding criminal law, learning about the dark web, and, crucially, building  
8 relevant trauma resilience.

9 **D. YouTube failed to implement the very standards it helped create**

10 51. YouTube failed to implement workplace safety measures that it (through its parent  
11 company Google) designed in the Guidelines and that other companies and non-profits have  
12 implemented.

13 52. During the hiring process, YouTube failed to properly inform prospective Content  
14 Moderators about the nature of the work or the effect reviewing graphic content can have on their  
15 mental health. Prospective Content Moderators are told they might be required to review graphic  
16 content, but they are not provided examples and they are not told that they would be required to review  
17 graphic content daily. They are also not asked about their experience with graphic content, they are not  
18 told that this content can have negative mental health impacts, they are not exposed to graphic content,  
19 they are not told to seek out other outside information, and they are not offered multiple days of  
20 interviews.

21 53. Before Content Moderators are exposed to any graphic content or receive any training,  
22 they are required to sign an employment contract and Non-Disclosure Agreement (“NDA”). Only  
23 after these documents are signed does the training begin.

24 54. During the training process, YouTube failed to train Content Moderators on how to  
25 assess their own reaction to the images, and YouTube failed to ease Content Moderators into review of  
26 graphic content through controlled exposure with a seasoned team member followed by counseling  
27 sessions.  
28

1           55.     Instead, Content Moderators are provided a two-week training where an instructor  
2 presents PowerPoints created by YouTube. The PowerPoints covered various categories of content,  
3 including graphic violence, child abuse, dangerous organizations, solicitation, porn, animal abuse,  
4 regulated products, fraud, and spam. Each category was covered by 60–80 slides. For each category, the  
5 PowerPoint began with a brief description of the applicable Community Guidelines, and then dozens of  
6 examples of content, applying the Community Guidelines.

7           56.     This content was extremely graphic. For example, during training, Plaintiff witnessed a  
8 video of a smashed open skull with people eating from it; a woman who was kidnapped and beheaded by  
9 a cartel; a person’s head being run over by a tank; bestiality; suicides; self-harm; children being rapped;  
10 births and abortions. As the example was being presented, Content Moderators were told that they  
11 could step out of the room. But Content Moderators were concerned that leaving the room would mean  
12 they might lose their job because at the end of the training new Content Moderators were required to  
13 pass a test applying the Community Guidelines to the content.

14           57.     During the three-week training, little to no time was spent on wellness and resiliency.  
15 Half-way through the training Plaintiff received—after Plaintiff was exposed to graphic content—two  
16 on-site Wellness Counselors spoke for an hour to the new Content Moderators. The Wellness  
17 Counselors told the Content Moderators where their offices were located, recommended that the  
18 Content Moderators get enough sleep and exercise, and reminded them that this job isn’t for everyone.  
19 The Wellness Counselors also told the Content Moderators that they could take breaks if they saw  
20 graphic content. However, as described below, the quantity and quality quotas Content Moderators  
21 were required to meet meant that these promised breaks were illusory.

22           58.     YouTube also failed to provide safeguards known to mitigate the negative effects of  
23 reviewing graphic content.

24           59.     Content Moderators are required to review hundreds of graphic and disturbing videos  
25 each week. To determine whether a video should be removed, YouTube created and continually revises  
26 hundreds of rules that Content Moderators must use to determine whether flagged content violates  
27 YouTube’s policies.  
28

1           60.     Despite these complex Community Guidelines, YouTube imposed strict quantity and  
2 accuracy quotas on Content Moderators. Content Moderators were required to review between 100 and  
3 300 pieces of content per day with an error rate of two to five percent. Supervisors often reminded  
4 Content Moderators of their quantity and accuracy quotas, telling the Content Moderators that “the  
5 Client [YouTube] isn’t happy with the amount of content that has been reviewed” and would tell  
6 Content Moderators how many posts they needed to review each day.

7           61.     To determine whether Content Moderators meet these metrics, YouTube audits  
8 Content Moderator’s work. To complete this audit, YouTube used two levels of reviewers above the  
9 first-level Content Moderators. Second-level Content Moderators audit first-level Content Moderators  
10 and specialize in certain areas of content, such as hate speech, foreign languages, and terrorism.  
11 Second-level Content Moderators are employed by YouTube Vendors. Third-level Content  
12 Moderators are employed by YouTube and audit second-level Content Moderators. If Content  
13 Moderators failed to meet the quantity and accuracy quotas, supervisors threat them with performance  
14 improvement plans (“PIPs”), which could lead to termination.

15           62.     YouTube was aware or should have been aware that its strict standards created  
16 stress and that such stress contributed to and exacerbated Content Moderator’s risk of  
17 developing psychological trauma.

18           63.     For many reasons—including low wages, short-term contracts, and the trauma  
19 associated with the work—many Content Moderators remain in the position for less than one year.

20           64.     Because of this high turnover and due to the immense amount of content that requires  
21 manual review, YouTube is chronically understaffed. To make up for this shortfall, Content  
22 Moderators are required to work long hours reviewing graphic content, despite YouTube’s own best  
23 practices described in the Guidebook and its claim in 2018 that Content Moderators would be limited to  
24 reviewing four hours of graphic content per day. In fact, Content Moderators routinely spend more  
25 than four hours a day reviewing graphic content, and some Content Moderators are given overtime pay  
26 to reduce backlogged queues.

27           65.     To review this content, YouTube designed and created the SRT, which it requires all  
28 Content Moderators to use, regardless of whether the Content Moderators are employed directly by

1 YouTube or by a YouTube Vendor and regardless of whether the Content Moderators are working at a  
2 YouTube facility or a facility operated by a YouTube Vendor.

3 66. YouTube monitors and is aware of the content of the videos the Content Moderators  
4 view, the number of videos the Content Moderators view per hour and per day, and the length of  
5 continuous content moderation sessions and breaks. YouTube controls how the videos are displayed  
6 (e.g., full screen versus thumbnails, blurred versus unblurred, etc.), how the accompanying audio is  
7 broadcast, and whether videos begin automatically upon completion of the prior video or whether the  
8 Content Moderator can catch his or her breath by controlling the start of the ensuing video.

9 67. YouTube failed to implement tooling safeguards in the SRT that would mitigate some of  
10 the harm caused by reviewing graphic and disturbing content, including changing the color or resolution  
11 of the video, superimposing a grid over the video, changing the direction of the video, blurring portions  
12 of the video, reducing the size of the video, and muting audio, although it knew that doing so was  
13 necessary to mitigate the harm to Content Moderators that was certain to result.

14 68. This failure is especially glaring considering the reasonably uncomplicated nature of  
15 many of the tooling changes. In 2017, a request was posted on Buganizer—YouTube’s internal  
16 reporting system for technical assistance—to implement tooling changes such as blurring images and  
17 videos. A Content Moderator in California commented on the request asking that a warning label be  
18 added to images and videos flagged as ultra-graphic violence. A YouTube engineer responded that this  
19 tooling change would take approximately half a day to implement. Suzanne French—Head of Global  
20 Vendor Operations at YouTube—commented that this tooling change was not a high priority and  
21 refused to implement the change.

22 69. When especially graphic content started being posted during and after the genocide in  
23 Myanmar, the Content Moderator again commented on the Buganizer request to see if YouTube would  
24 reconsider its decision. The Content Moderator’s request was ignored, and he was reprimanded for  
25 raising the issue. As of 2019, no tooling changes had been implemented.

26 70. YouTube also failed to provide psychological support to Content Moderators. YouTube  
27 purportedly offered Content Moderators “wellness” benefits, including a “Wellness Coach,” a  
28 hotline, and a human resources department.

1           71.     However, these support services were insufficient. Wellness Coaches were unavailable  
2 to Content Moderators that worked the evening shifts: 3:00 p.m. to 12:00 a.m. and 10:00 p.m. to 7:00  
3 a.m. And even those Content Moderators that had access to a Wellness Coach did not receive any on-  
4 site medical care because Wellness Coaches are not medical doctors and cannot diagnose or treat  
5 mental health disorders. Instead, Wellness Coaches would occasionally recommend that a Content  
6 Moderator see a licensed clinician but would not provide any information on how to find treatment.

7           72.     Wellness Coaches were also underqualified and undertrained, and consequently Content  
8 Moderators did not feel comfortable asking them for help. For example, in spring of 2018, Plaintiff met  
9 with a Wellness Coach to discuss upsetting videos she had witnessed that were particularly traumatic.  
10 The Wellness Coach recommended that Plaintiff take illegal drugs. The Wellness Coach did not  
11 provide any resiliency training or coping mechanisms (beyond self-medicating with an illegal  
12 substance).

13           73.     A few months after that, Plaintiff spoke with a fellow Content Moderator that had met  
14 with a different Wellness Coach. That Wellness Coach told Plaintiff's co-worker to "trust in God,"  
15 advice that was unhelpful.

16           74.     Content Moderators also believed that communications with Wellness Coaches were not  
17 kept confidential and feared that anything that was said to a Wellness Coach would be reported to  
18 management. And when Content Moderators tried to discuss the effect viewing graphic content had on  
19 their mental health, employees within the human resource department told Content Moderators they  
20 could not help them.

21           75.     YouTube also demands that Google Vendors require their employees to sign sweeping  
22 Non-Disclosure Agreements ("NDAs"). YouTube further requires YouTube Vendors to provide  
23 YouTube-developed training to all Content Moderators that instructs the Content Moderators not to  
24 speak about the content or workplace conditions to anyone outside of their review team, including  
25 therapists, psychiatrists, or psychologists retained by Content Moderators. By prohibiting Content  
26 Moderators from discussing their work or seeking outside social support, YouTube impedes the  
27 development of resiliency and increases the risk that Content Moderators will develop psychological  
28

1 trauma. Furthermore, by imposing NDAs in violation of California law, YouTube is estopped from  
2 asserting any statute of limitations defense to these claims.

3 **E. YouTube knows that exposure to graphic content can cause psychological trauma but**  
4 **seeks to silence whistle blowers and shield itself from liability**

5 76. In 2019, YouTube acknowledged that viewing graphic content could lead to  
6 psychological trauma. Well before that, YouTube engaged in an aggressive campaign to hide evidence  
7 and silence whistle blowers. In 2017, Content Moderators were told to stop talking or posting about the  
8 negative effects of reviewing graphic content. YouTube also purged its messaging systems of any of  
9 these reports, deleting old posts by Content Moderators that shed light on the trauma they were  
10 experiencing.

11 77. YouTube also sought to shield itself from liability. On December 20, 2019—four days  
12 after *The Verge* published an investigation into PTSD among workers at YouTube’s Content Moderator  
13 facility in Austin, Texas—YouTube responded by requiring its Content Moderators to sign a document  
14 acknowledging that performing the job can cause PTSD.

15 78. YouTube also required Content Moderators to acknowledge that “no job is worth  
16 sacrificing my mental or emotional health” and that “this job is not for everyone”—language likely  
17 drafted by lawyers and which is intended to suggest Content Moderators suffering from negative  
18 psychological health effects caused by exposure to graphic content might be terminated if they reported  
19 any negative impacts to their psychological health.

20 79. If a Content Moderator is fired or quits, they lose medical insurance and other  
21 healthcare benefits, as well as their income. Therefore, Content Moderators were left with a  
22 Hobbesian’s choice—quit and lose access to an income and medical insurance or continue to suffer in  
23 silence to keep their job.

24 80. From approximately January 16, 2018 until August 24, 2019, JANE DOE worked as a  
25 “Content Review Analyst” (i.e., Content Moderator), reviewing content for YouTube at an office  
26 located at 7700 West Parmer Lane, Austin Texas, 78717.

27 81. During this period, Plaintiff was employed solely by Collabera.

28 82. At all times relevant to this complaint, Collabera was a YouTube Vendor.

1 83. Collabera directly oversaw all human resources matters concerning Plaintiff.

2 84. Plaintiff has never been employed by YouTube in any capacity.

3 85. Plaintiff never received any wages from YouTube.

4 86. Plaintiff never received YouTube's employee benefits package (e.g., wellness benefits,  
5 paid time off, and parental financial assistance).

6 87. Plaintiff worked as a first-level Content Moderator.

7 88. During her employment as a Content Moderator, Plaintiff was exposed to thousands of  
8 graphic and objectionable videos, including graphic violence, sexual assault, and child pornography. For  
9 example, Plaintiff witnessed a video of: a smashed open skull with people eating from it; a woman who  
10 was kidnapped and beheaded by a cartel; a person's head being run over by a tank; a man eating the  
11 head off a rat; a fox being skinned alive; a man falling to his death off a roof that included audio of the  
12 impact of his body hitting the ground; school shootings included dead bodies of children; a politician  
13 shooting himself; backyard abortions; child abuse; and child sexual assault.

14 89. As a result of training for and providing content moderation services through YouTube's  
15 SRT and in accordance with YouTube's policies, Plaintiff developed severe psychological trauma  
16 including depression and symptoms associated with anxiety and PTSD.

17 90. PTSD and related syndromes caused by exposure to harmful content can be triggered by  
18 witnessing abuse; watching the news or seeing violence on television; hearing loud noises like gunshots,  
19 fireworks, cars backfiring, or objects falling; seeing ISIS members or paraphernalia; and seeing racially-  
20 discordant posts sowing political dissension in America. She has trouble sleeping and when she does  
21 sleep, she has horrific nightmares. She often lays awake at night trying to go to sleep, replaying videos  
22 that she has seen in her mind. She cannot be in crowded places, including concerts and events, because  
23 she fears mass shootings. She has severe and debilitating panic attacks. She has lost many friends  
24 because of her anxiety around people. She has trouble interacting and being around kids and is now  
25 scared to have children.

26 **CLASS ACTION ALLEGATIONS**

27 91. Plaintiff brings this class action individually and on behalf of all persons who performed  
28 content moderation work for YouTube in the United States at any time up until the present.



1           92.     Excluded from this definition are Defendant’s officers, directors, and management, any  
2 judicial officer presiding over this action and the members of his/her immediate family and judicial  
3 staff, any juror assigned to this action, and any Content Moderators that are employed directly by  
4 Defendant.

5           93.     The class is so numerous that joinder of all members is impracticable. Plaintiff does not  
6 know the exact size of the class since that information is within the control of YouTube and its Vendors.  
7 However, upon information and belief, Plaintiff alleges that the number of class members is in the  
8 thousands. Membership in the class is readily ascertainable from YouTube’s records such as those  
9 relating to its contracts with YouTube’s Vendors or to registered users of YouTube’s SRT.

10          94.     There are numerous questions of law or fact common to the class, and those issues  
11 predominate over any question affecting only individual class members. The common legal and factual  
12 issues include the following:

- 13           a.     Whether YouTube committed the violations of the law alleged herein;
- 14           b.     Whether viewing graphic and objectionable conduct in the manner which Content  
15               Moderators do for YouTube is an abnormally dangerous activity;
- 16           c.     Whether YouTube participated in and perpetrated the tortious conduct complained  
17               of herein;
- 18           d.     Whether Plaintiff and the class are entitled to medical screening, treatment, and  
19               damages;
- 20           e.     Whether YouTube should be ordered to implement and comply with industry  
21               guidelines for safety in content moderation.

22          95.     The claims asserted by Plaintiff are typical of the claims in that the representative  
23 plaintiff, like all class members, was exposed to highly toxic, unsafe, and injurious content while  
24 providing content moderation services for YouTube. Each member of the proposed class has been  
25 similarly injured by YouTube’s misconduct.

26          96.     Plaintiff will fairly and adequately protect the interests of the class. Plaintiff has retained  
27 attorneys experienced in class actions, complex litigation, California law, and issues involving content  
28

1 moderation. Plaintiff intends to vigorously prosecute this litigation. Neither Plaintiff nor their counsel  
2 have interests that conflict with the interests of the other class members.

3 97. Plaintiff and the class members have all suffered and will continue to suffer harm  
4 resulting from YouTube's wrongful conduct. A class action is superior to other available methods for  
5 the fair and efficient adjudication of the controversy. Treatment as a class action will permit a large  
6 number of similarly situated persons to adjudicate their common claims in a single forum  
7 simultaneously, efficiently, and without the duplication of effort and expense that numerous individual  
8 actions would engender. Class treatment will also permit the adjudication of claims by many members  
9 of the proposed class who could not individually afford to litigate a claim such as is asserted in this  
10 complaint. This action likely presents no difficulties in management that would preclude maintenance  
11 as a class action.

12 **FIRST CAUSE OF ACTION**  
13 **NEGLIGENCE**  
14 **(Abnormally Dangerous Activity)**

15 98. Plaintiff realleges and incorporates by reference herein all allegations above.

16 99. A company is strictly liable to individuals that are injured while the company engages in  
17 an abnormally dangerous activity.

18 100. An activity is abnormally dangerous if it (a) necessarily involves a risk of serious harm to  
19 the person, land or chattels of others which cannot be eliminated by the exercise of the utmost care, and  
20 (b) is not a matter of common usage.

21 101. Requiring Content Moderators to review graphic and objectionable content is an  
22 abnormally dangerous activity. Content Moderators risk serious and debilitating psychological trauma,  
23 including severe anxiety, depression and PTSD and there is no way to eliminate this risk. Content  
24 moderation is also not a matter of common usage. Only a handful of technology companies, non-profits,  
25 government agencies, and non-governmental organizations review content.

26 102. Strict liability for a defendant that engages in abnormally dangerous activity represents a  
27 social-policy determination that the defendant, while engaged in an enterprise tolerated by the law,  
28 must pay for the damage caused by its enterprise.

1           103. In fiscal year 2019, YouTube and its parent company Google made a combined  
2 approximately \$150 billion in advertising revenue. In 2018, that number was nearly \$127 billion, and in  
3 2017 that number was almost \$103 billion. YouTube and Google are some of the most successful  
4 companies in history based on revenue.

5           104. YouTube derives this vast wealth from providing a platform safe from graphic and  
6 objectionable content. YouTube relies on Content Moderators to ensure that its platform is free from  
7 graphic and objectionable content. Therefore, YouTube is required under the law to pay for the harm  
8 caused by requiring Content Moderators to review and remove graphic and objectionable content.

9           105. As a result of YouTube's tortious conduct, Plaintiff and the class are at an increased risk  
10 of developing serious mental health injuries, including, but not limited to, PTSD, anxiety, and  
11 depression.

12           106. To remedy that injury, Plaintiff and the class need medical monitoring that provides  
13 specialized screening, assessment, and treatment not generally given to the public at large.

14           107. The medical monitoring regime includes, but is not limited to, baseline screening,  
15 assessments, and examinations that will assist in diagnosing the adverse health effects associated with  
16 exposure to trauma. This screening and assessment will also inform which behavioral and/or  
17 pharmaceutical interventions are best suited to preventing or mitigating various adverse consequences  
18 of post-traumatic stress and other conditions associated with exposure to graphic imagery.

19           108. In particular, the medical monitoring regime includes (a) secondary preventative  
20 interventions, designed to reduce the risk of later onset of PTSD among class members who are not yet  
21 displaying symptoms of PTSD; (b) tertiary interventions, designed to reduce the worsening of  
22 symptoms among those who are already experiencing symptoms associated with post-traumatic stress  
23 or have a diagnosis of PTSD; and (c) evidence-based treatments to facilitate recovery from mental  
24 health conditions.

25           109. Monitoring, assessing, and providing preventative interventions and/or treatment to  
26 Plaintiff and the class will significantly reduce the risk of long-term injury, disease, and economic loss  
27 that Plaintiff and the class have incurred as a result of YouTube's unlawful conduct.  
28



- a. Requiring Content Moderators to use a YouTube-developed review platform that presented unmitigated traumatic content to Content Moderators according to YouTube-developed algorithms;
- b. Requiring that Content Moderators—through their employers—sign NDAs and undergo YouTube-developed confidentiality trainings that prohibited Content Moderators from discussing their work outside their review teams;
- c. Requiring that Content Moderators be interviewed and undergo training using YouTube-developed training materials and procedures; and
- d. Setting expectations as to the overall timeframe for and accuracy of content review, calculating the amount of time it should take a Content Moderator to review different types of posts, and deciding the overall number of hours required to meet the overarching timeframe and accuracy expectations.

120. Based on its exercise of retained control, YouTube has had at all relevant times a duty to exercise reasonable care with regard to the safety of Plaintiff and the class.

121. YouTube negligently exercised its retained control in a manner that affirmatively contributed to the injuries of Plaintiff and the class, including by exacerbating Plaintiff's and class members' risks of developing PTSD or other health issues. For example:

- a. YouTube failed to provide adequate technological safeguards to protect Content Moderators from risks associated with exposure to traumatic content via YouTube's SRT;
- b. YouTube's NDAs and confidentiality requirements diminished Content Moderators' social support networks and resilience by prohibiting Content Moderators from speaking about the content they reviewed or other related workplace conditions to anyone outside of their review teams;
- c. YouTube failed to provide Content Moderators with an interview process and training that met the standards it developed through the Technology Coalition's Guidebook; and

1 d. By setting demanding standards for review, both in terms of quantity and quality  
2 expectations, YouTube imposed stressful work conditions that served to further  
3 reduce Content Moderators' resilience to trauma.

4 122. YouTube was aware of the psychological trauma that could be caused by viewing graphic  
5 and objectionable content, including videos and/or images of child abuse, rape, torture, bestiality,  
6 beheadings, suicide, murder, and other forms of extreme violence through the SRT.

7 123. YouTube was also aware or should have been aware that the SRT could be made safer if  
8 proper precautions were followed, that requiring Content Moderators not to discuss their work or  
9 workplace conditions reduced their ability to deal with traumatic content, and that YouTube's overall  
10 quality and quantity standards had the effect of imposing intense workplace stress and, accordingly,  
11 increasing Content Moderators' risk of injury from psychological trauma.

12 124. YouTube breached its duty to Plaintiff and the class by failing to provide the necessary  
13 and adequate technological safeguards, safety and instructional materials, warnings, social support, and  
14 other means to reduce and/or minimize the physical and psychiatric risks associated with exposure to  
15 graphic imagery through YouTube's SRT.

16 125. YouTube continues to breach its duty to class members by failing to exercise its retained  
17 control with reasonable care; that breach continues to elevate class members' risk of injury from  
18 psychological trauma.

19 126. As a result of YouTube's tortious conduct, Plaintiff and the class are at an increased risk  
20 of developing serious mental health injuries, including, but not limited to, PTSD, anxiety, and  
21 depression.

22 127. To remedy that injury, Plaintiff and the class need medical monitoring that provides  
23 specialized screening, assessment, and treatment not generally given to the public at large.

24 128. The medical monitoring regime includes, but is not limited to, baseline screening,  
25 assessments, and examinations that will assist in diagnosing the adverse health effects associated with  
26 exposure to trauma. This screening and assessment will also inform which behavioral and/or  
27 pharmaceutical interventions are best suited to preventing or mitigating various adverse consequences  
28 of post-traumatic stress and other conditions associated with exposure to graphic imagery.



1           138. YouTube had a duty to exercise reasonable care to furnish a safe review platform to its  
2 contractors.

3           139. YouTube was aware of the psychological trauma that could be caused by viewing graphic  
4 and objectionable content, including videos and/or images of child abuse, rape, torture, bestiality,  
5 beheadings, suicide, murder, and other forms of extreme violence through its review platforms.

6           140. YouTube was aware or should have been aware that its review platforms could be made  
7 safer if proper precautions were followed.

8           141. YouTube nevertheless provided unsafe review tools to Plaintiff and the class that  
9 exposed Plaintiff and the class to unmitigated traumatic content.

10          142. YouTube breached its duty to Plaintiff and the class by failing to provide necessary and  
11 adequate technological safeguards, safety and instructional materials, warnings, and other means to  
12 reduce and/or minimize the physical and psychiatric risks associated with exposure to graphic imagery  
13 through YouTube's review platform.

14          143. YouTube continues to breach its duty to class members by failing to provide a reasonably  
15 safe review platform; that breach continues to elevate class members' risk of injury from psychological  
16 trauma.

17          144. As a result of YouTube's tortious conduct, Plaintiff and the class are at an increased risk  
18 of developing serious mental health injuries, including, but not limited to, PTSD, anxiety, and  
19 depression.

20          145. To remedy that injury, Plaintiff and the class need medical monitoring that provides  
21 specialized screening, assessment, and treatment not generally given to the public at large.

22          146. The medical monitoring regime includes, but is not limited to, baseline screening,  
23 assessments, and examinations that will assist in diagnosing the adverse health effects associated with  
24 exposure to trauma. This screening and assessment will also inform which behavioral and/or  
25 pharmaceutical interventions are best suited to preventing or mitigating various adverse consequences  
26 of post-traumatic stress and other conditions associated with exposure to graphic imagery.

27          147. In particular, the medical monitoring regime includes (a) secondary preventative  
28 interventions, designed to reduce the risk of later onset of PTSD among class members who are not yet



1 displaying symptoms of PTSD; (b) tertiary interventions, designed to reduce the worsening of  
2 symptoms among those who are already experiencing symptoms associated with post-traumatic stress or  
3 have a diagnosis of PTSD; and (c) evidence-based treatments to facilitate recovery from mental health  
4 conditions.

5 148. Monitoring, assessing, and providing preventative interventions and/or treatment to  
6 Plaintiff and the class will significantly reduce the risk of long-term injury, disease, and economic loss  
7 that Plaintiff and the class have incurred as a result of YouTube's unlawful conduct.

8 149. Plaintiff seeks medical screening and treatment to facilitate the screening, diagnosis, and  
9 adequate treatment of Plaintiff and the class for psychological trauma, including to prevent or mitigate  
10 conditions such as PTSD, anxiety, and depression.

11 150. Plaintiff also seeks compensatory damages for the injuries she and the class have  
12 suffered.

13 151. Plaintiff also seeks an award of attorney's fees.

14 **FOURTH CAUSE OF ACTION**  
15 **CALIFORNIA UNFAIR COMPETITION LAW**

16 152. Plaintiff realleges and incorporates by reference herein all allegations above.

17 153. Solely in the alternative and to the extent that this Court concludes that YouTube is not  
18 strictly liable for the harm caused by engaging in an abnormally dangerous activity, Plaintiff brings this  
19 fourth cause of action for violation of California Unfair Competition Law.

20 154. YouTube's negligent exercise of retained control of the content moderation work  
21 performed by Plaintiff and the class violates California common law.

22 155. YouTube's negligent provision of unsafe equipment to its independent contractors for  
23 use by Plaintiff and the class also violates California common law.

24 156. Plaintiff each suffered an injury in fact because of YouTube's negligent conduct and has  
25 lost money because of YouTube's conduct.

26 157. Specifically, Plaintiff paid out of pocket for medical treatment and therapy for her  
27 depression and symptoms of anxiety and PTSD, which was caused by YouTube's conduct.  
28



1 safety devices and safeguards reasonably adequate to render the employment and place of employment  
2 safe,” “adopt and use methods and processes reasonably adequate to render the employment and place  
3 of employment safe,” and “do every other thing reasonably necessary to protect the life, safety, and  
4 health of employees.” Cal. Labor Code § 6403

5 166. No employer can “require or permit any employee to go or be in any employment or  
6 place of employment which is not safe and healthful.” Cal. Labor Code § 6402.

7 167. YouTube failed to provide a safe working environment. YouTube routinely and  
8 repeatedly exposed Plaintiff and the class to content known to cause psychological trauma, including  
9 PTSD, anxiety, and depression. Even though YouTube knew of and could have reasonably  
10 implemented adequate safety measures, the corporation refused to implement necessary and adequate  
11 safety and instructional materials, trainings, warnings, and means to reduce and/or minimize the risks  
12 associated with exposure to graphic content.

13 168. YouTube’s failure to provide a safe workplace for Plaintiff and the class violates, *inter*  
14 *alia*, sections 6400, 6401, 6401.7, 6402, and 6403 of the California Labor Code.

15 169. In requiring Content Moderators to sign sweeping NDAs and instructing Content  
16 Moderators not to disclose information about working conditions—including the traumatic nature of  
17 the content, the intense stress from quantity and quality expectations, and the lack of training and safety  
18 measures to protect moderators from trauma exposure—YouTube further violates section 232.5 of the  
19 California Labor Code.

20 170. YouTube’s illegal conduct was and is willful and serious and has directly caused harm to  
21 Plaintiff and the class.

22 171. Plaintiff suffered an injury in fact because of YouTube’s conduct and has lost money  
23 because of YouTube’s conduct.

24 172. Specifically, Plaintiff paid out of pocket for medical treatment and therapy for her  
25 depression and symptoms of anxiety and PTSD, which was caused by YouTube’s conduct.

26 173. There were reasonably available alternatives to the conduct described herein that would  
27 further YouTube’s legitimate business interests.

28



- 1 or mitigate conditions such as PTSD, anxiety and depression—until it can be determined  
2 that psychological trauma is no longer a threat to their health;  
3 g. Award Plaintiff and class members their reasonable litigation expenses and attorneys’ fees;  
4 and  
5 h. Award any further relief that this Court deems just and equitable.

6 **DEMAND FOR JURY TRIAL**

7 Plaintiff hereby request trial by jury.

8  
9 Dated: September 21, 2020

Respectfully Submitted,

10 JOSEPH SAVERI LAW FIRM, INC.

11 By: 

12 Joseph R. Saveri (SBN 130064)  
13 Steven N. Williams (SBN 175489)  
14 Kevin Rayhill (SBN 267496)  
15 Kate Malone (SBN 290884)  
16 Kyle Quackenbush (SBN 322401)  
17 **JOSEPH SAVERI LAW FIRM, INC.**  
18 601 California Street, Suite 1000  
19 San Francisco, CA 94108  
20 Telephone: (415) 500-6800  
21 Facsimile: (415) 395-9940  
22 jsaveri@saverilawfirm.com  
23 swilliams@saverilawfirm.com  
24 krayhill@saverilawfirm.com  
25 kmalone@saverilawfirm.com  
26 kquackenbush@saverilawfirm.com

27 *Attorneys for Plaintiffs and the Proposed Class*  
28