

Attachment—Additional Questions for the Record

**Subcommittee on Communications and Technology and
Subcommittee on Consumer Protection and Commerce
Joint Hearing on
“Fostering a Healthier Internet to Protect Consumers”
October 16, 2019**

Mr. Steve Huffman, Co-Founder & CEO, Reddit, Inc.

The Honorable Anna G. Eshoo (D-CA)

- 1. Just a few years ago, Reddit was involved in many controversies, such as GamerGate, concerning the company’s content moderation practices. Yet you seem to have made changes to your content policies that have, to some degree, cleaned up some of the worst issues. What did you do in the last few years, and how did Section 230 play a role in that process?**

Response: Thank you, Congresswoman, for recognizing this, as it’s something we’ve worked hard at over the past several years and we’re proud of the progress we’ve made, while understanding of course that there is still more work for us to do here. Section 230 has been Reddit’s biggest tool in evolving our content moderation practices. In the past several years we have updated our policies on violent content, on involuntary pornography, on controlled goods, and on bullying and harassment, allowing us to be much more proactive in removing bad material from our site. We’ve also very intentionally grown and built out the teams and tools that we employ to address these problems, and we continue to evolve our approach. It is Section 230 that has allowed us to do all this without either the fear of retribution from bad actors angry that we’ve taken their content down, or of punishment in the form of liability for things that we might miss despite our good-faith efforts.

- 2. Dr. Citron’s idea for including a requirement that platforms employ reasonable content moderation practices to receive Section 230 immunity is an interesting idea. If it was signed into law, how would it affect the business of Reddit?**

Response: The difficulty with Dr. Citron’s new standard is that it turns every content moderation decision our teams make into a court case over whether we were “reasonable” or not in our content moderation practices. Our teams along with our volunteer community moderators have to make very difficult decisions every day. Some people will be upset that their piece of content came down. Others will be upset that another person’s content stayed up. Someone is likely going to be unhappy either way. Dr. Citron’s new standard merely emboldens those unhappy people to go to a trial lawyer and sue us. In order to avail ourselves of Section 230, we would need to prove in court that we have reasonable content moderation practices. I can guarantee you that my definition and your definition of “reasonableness” would be very different from that of a trial lawyer’s.

These lawsuits would spell ruin for smaller businesses like ours, while entrenching the largest players in the industry. Even in cases where companies may ultimately prevail in litigation, the legal expenses of defending the case could be existential for startup companies like ours, which are privately-held, still on venture funding, and not yet profitable.

- 3. Please tell us about how Reddit responded to the Allow States and Victims to Fight Online Sex Trafficking Act of 2017 (Public Law 115–164), better known as FOSTA-SESTA, being signed into law, including its effects, both intended and unintended, and how your company determined how to ensure compliance with the law.**

Response: FOSTA-SESTA has created a chilling effect not only on the Reddit platform but across the internet. Now let us be clear, it is our deeply held belief that content facilitating sex trafficking should never exist on our platform. It is important to note sex trafficking is illegal in the United States, and even before FOSTA-SESTA, Section 230 did not provide immunity regarding this or other criminal activity. What FOSTA-SESTA does do is create a great deal of ambiguity around when a platform has culpable knowledge of trafficking activity. This uncertainty forced Reddit, on the advice of counsel, to take down a number of safety and harm-reduction discussion communities run for, and by, those in the sex work industry, further marginalizing an already at-risk population. And while we understand that some may not consider sex workers to be a particularly sympathetic group, it is not a difficult leap to see the same choices forced around other vulnerable communities that have sparked 230 conversations, such as opiate addiction support and harm reduction communities.

The Honorable Kathy Castor (D-FL)

- 1. On June 19, 2019, The Verge published an investigation into one of Facebook’s content moderation sites in Tampa, FL, which is operated by the firm Cognizant. The article details allegations of appalling working conditions including sexual harassment, verbal and physical fights, theft, and general filthiness in addition to adverse mental health effects associated with the nature of their work.**
 - a. Operationally, how should tech platforms moderate their content? What role should human content moderators play? What role should technology play?**

Response: At Reddit, we’ve found that the only thing that scales with users is users. That is why we’ve empowered each Reddit user to also be a content moderator, and have structured our site in a way that is akin to neighborhoods—manageable units that are familiar to the people in them—so that no one individual or entity needs to have the burden of trying to police the entire site. This structure ensures that our moderation system is sensitive to social and cultural context, since it is community members themselves who voluntarily perform the moderation actions. This type of contextual awareness is something that machines will have a difficult time duplicating.

But where machines come in is in giving humans superpowers, and helping humans scale their work. For example, while our in-house content moderators (known as our “Anti-Evil Team”) review Reddit reports individually by hand, those reports are prioritized by machine learning, to help us ensure that the reports that are the most likely to be urgent go to the top of the pile, rather than just being buried chronologically.

Machines can also help cut down on the mundane. For example, we have a tool called “Automoderator,” which is essentially a helper-bot that we have built and made available for Reddit community moderators to use. They can set the Automoderator to do a number of things, such as, for example, banning certain words or slurs from appearing in the community. This is something that is customized from community to community by the volunteer moderators themselves, so that context and nuance is always present. It’s a great example of machines complimenting and amplifying human judgement at scale, reducing the burdens on individuals.

Finally, we also try and reduce the burden on individuals who do this work by always keeping their wellness at front of mind in our policies and practices. We provide them with benefits such as access to counselors, mandatory vacation time, and other measures to fight trauma or burnout. We are also mindful of their needs in the tools we build for them to do their jobs. For example, the interface of the programs that they use to review potentially disturbing content includes features such as progressively blurring, muting, and/or greyscaling images and video where human review is necessary.

b. What standard should a private company use to evaluate content? “Quasi constitutional”, a “community standard” established by the company along the lines of other private media, other?

Response: At Reddit we strongly believe that users themselves need to be empowered in setting appropriate standards. This is why we have created a federal system, wherein we at the corporate level set out topline principles, but beyond that, we empower our individual communities themselves to establish customized rules that are appropriate to their culture and context. We think that such a model is the only thing that scales while being flexible enough to accommodate the vast diversity of people who find their home on Reddit.

c. Given that private companies are not governed by standards that government would be when it decides not to post content, why do content moderators have to spend so much time reviewing and in such great detail evaluating explicit, violent, or hateful content? What value is there to society and the site owner to work to ensure that such explicit, violent, or hateful content is given every opportunity to be posted?

Response: We’d agree that there is little value to a great deal of violent, hateful, or otherwise objectionable content, which is why we are grateful for the way that Section 230 empowers us to exercise judgment to take it down, and take it down quickly. But in our experience, it’s not the obvious, black-and-white cases that take the most time. It’s the grey area content, where different, reasonable, well-intentioned people can disagree on whether something is

inappropriate or a violation. That is where we spend the most time. And we think that it is important that we be fair and consistent when these hard cases come up, because that fairness and consistency is what is going to maintain trust with our users, who range the entire political spectrum.

- d. This explicit, violent, or hateful content often is known to be inconsistent with the tech platform's content bylaws. Why do tech platforms, like Facebook, force content moderators to not only look at but also evaluate in great detail explicit, violent, or hateful content that is often inconsistent with the tech platform's bylaws?**

Response: I can't speak for other companies' practices. But at Reddit, the vast majority (more than 99%) of content moderation is done on a volunteer basis, by users themselves on their own schedule, in their own spaces, and we take their wellbeing seriously. We have an entire team at the company whose job it is to cater to moderator needs and build tools for them. Additionally, as part of our efforts to continually improve the speed and efficacy of our enforcement mechanisms, we are working toward being more proactive around bad content-- that is, taking action on it before a human ever encounters it. For example, like many of our peers, we employ content hashing against child exploitation content and foreign terrorist content, which means that this material comes down swiftly and automatically, without a human having to see it and potentially be traumatized.

- e. Should content moderators have more leeway to ban harmful content so they don't have to look at it over such lengthy time periods and evaluate the content in such detail?**

Response: Given that more than 99% of content moderation is undertaken on a volunteer basis by users themselves, typically in accordance with rules they themselves have created rather than ones handed down by us, content moderation on Reddit is fundamentally different than the approach other companies take, and has leeway and flexibility inbuilt. For the in-house employee teams that we have that also engage in content moderation, we take their wellbeing enormously seriously. We have in place benefits for them such as mandatory periodic vacation, access to counselors, and other measures to care for them. We also have taken their wellness into account in the tools we build for them to do their jobs. For example, review programs for potentially sensitive or disturbing content includes features such as progressively blurring, muting, and/or gray-scaling images and video where human review is necessary. On top of this, we employ hashing technology around abhorrent illegal content like child exploitation and foreign terrorist propaganda, so that a human never even has to be exposed to it.

- f. What should industry best practices be for treating content moderators? Should Congress play a role in ensuring worker rights in this unique industry? If so, how?**

Response: There are certain benefits that we give to our employees who handle sensitive content that we believe are a must. This includes mandatory periodic vacation, strictly limited working hours, access to counselors, mental healthcare benefits, and a comfortable, inviting workspace.

- g. Is it common practice among tech platforms to use contractors to conduct content moderation for their sites? Why do some tech platforms use contractors to conduct content moderation for their sites? Should tech platforms do this?**

Response: There is actually significant diversity in the industry in how companies approach content moderation, and recognizing that diversity is important. While some of the largest companies use contractors at an industrial scale, that's not really a viable option for smaller companies. We, for example, have chosen to rely on a community moderation model. Other companies do all of their content moderation in-house with small teams. The approach that each company takes largely depends on their individual resources and business model.

The Honorable Lisa Blunt Rochester (D-DE)

- 1. At the October 16, 2019, joint hearing, you provided commitments that Reddit will disclose information on diversity of your content moderators and issues with hiring diverse content moderator teams. Please provide that information to the Energy and Commerce Committee and my office.**

Response: As noted in our annual Transparency Report, over 99% of content moderation actions are done by volunteer moderators, who come from the Reddit userbase itself. Our company has a different business model from others in the industry, in that we don't collect large amounts of user data. In fact, we collect hardly any. This means that we actually know very little about the real-world identities of our users-- we don't know their age, we don't know their gender, we don't know their race, we don't know their religion. However, what we do know is that the moderators of any particular community are invested members of that community, and they are appointed by their peers. This gives us confidence that our moderators are attuned to the special context and needs of the communities that they help govern.

That said, when it comes to our actual employee base, we care deeply about ensuring that we have a broad and diverse team. We track the racial and ethnic makeup of our company, as well as inclusion sentiment amongst our employees, so that we can understand where we are and what further progress we need to make. To help us meet our goals, we recruit from historically black colleges and universities (HBCUs) as well as Hispanic-serving institutions (HSIs), and we sponsor tech events focused on engaging women and minorities. In the past year, these sponsorships have included (among others) AfroTech, /dev/color, Tech Inclusion San Francisco, the Lesbians Who Tech Leadership Summit, and the Grace Hopper Celebration of Women in Computing. We also have employee resource groups for our diverse employees, and tap their opinions on relevant product decisions to ensure that we have diverse viewpoints represented, even if those viewpoints come from different teams than those directly responsible for a particular product or feature.

- 2. What can the federal government do to improve the capacity and ability to effectively moderate online content, including technological research?**

Response: The single most important thing the federal government can do is to protect laws like Section 230 that make it possible for us to moderate in good faith without fear of retribution. Aside from that, the federal government should encourage diversity in approaches to online content moderation--whether through academic research into different models of online communities or through encouraging startups that approach content moderation differently from the central model dominant in the social media industry today.

The Honorable Tom O'Halleran (D-AZ)

- 1. Mr. Huffman, as written in statute, Section 230 has “good Samaritan” language to incentivize online platforms to take actions “in good faith to restrict access to or the availability of” harmful content.**

Many platforms have established content or use of service policies to specify what behavior is allowed by the service, while others employ artificial intelligence formulas to automatically filter user-generated content. Some platforms also hire human content moderators to review and remove content posted by users on its platforms that is considered harmful, violent, or graphic. These content reviewers often suffer from Post-Traumatic Stress Disorder (PTSD).

- a. What more can be done by the government and industry to ensure sufficient mental health services are made available to human content reviewers?**

Response: Mental health services are important not only to workers in our industry, but to everyone. But simply recognizing mental health as a priority is a first step. A second step is identifying the breadth of people who come across sensitive content. It's not only the people who directly review content. It's also the people who manage them. Management of sensitive teams is a very important skill, and sensitivity to worker needs is something that needs to permeate an entire organization. For this reason, we offer counseling and mental health benefits directly to the individual employees who undertake such work, and we also train their managers in the special needs that apply to managing such workers.

Additionally, it is important to reduce employee exposure to harmful content in the first place. We try to do this through the tooling we provide teams who fight this content. We have built in safety options that allow reviewers to put the content in grayscale or blur, or mute it by default.

Where government can be helpful in this regard, aside from prioritizing mental healthcare for everyone, is sharing experiences with industry. There are very large numbers of public servants who encounter sensitive or traumatic situations in their work every day, from members of our military to our first responders. These industries have each acquired learnings and tools to help protect the mental wellbeing of their frontline employees. Thinking of content moderators as a

type of first responder, and valuing them as such, could go a long way. Encouraging public-private knowledge sharing and best practices in this regard could be extremely productive.

The Honorable Greg Walden (R-OR)

- 1. At the hearing, Rep. Bilirakis asked EFF whether they have argued for including language mirroring legislation in trade deals explicitly for the purpose of “baking” language into an agreement to protect the statute domestically.**

For the record, Yes or No: Is including such 230-like language in trade agreements an attempt to preclude us – the committee of jurisdiction – from revisiting the statute?

Response: Respectfully, I’m not qualified to answer that. You’d have to speak to the people who wrote the language.

- 2. In your testimony you talk about the importance Reddit places on the downvote feature. Can you describe how this type of moderation applies to the more “edge-cases” of speech issues vs. the clear-cut illegal activity that may find its way on your platform?**

Response: Thank you, Congressman, for making this important distinction. As you’re no doubt aware, there are many things on Reddit, that, while they don’t actually break either the law or our rules, are lower quality and not particularly valuable. In most of these cases, it would be inappropriate overreach for us as a company to take the content down. Not only would it be heavy-handed, it would distract our limited resources from things that are actually dangerous or illegal. Community moderation, and the downvote specifically, is a solution that allows the community itself to say, “we don’t think this content is worthwhile,” and downrank it accordingly, limiting its visibility. Indeed, to think about it in a real-world context, social pressures dictate real-life decorum every day in our lives. Reddit tries to replicate that online.

- a. Do you think the “user-moderator” approach could be exported to other platforms? If yes, should it? If no, why not?**

Response: There is a wide diversity of site structures and business models in the tech industry, and it’s important to recognize this, and avoid the government dictating any single approach. This is particularly true given the reality of just how fast technology changes. For us and our site structure, at this time, the user-moderator approach has worked well. However, more important than this particular model, which may or may not be right for others in the industry, is the principle from which it draws—that users of a site should be intimately involved in its governance and have a stake in it. This is what makes Reddit different from other sites. We put the user at the center of everything, and actually empower them.

- 3. In your testimony, you state that medium, small, and startup-sized companies do not have the resources to moderate content “from the center.” Surely, you’d**

agree that some tech companies have the resources to devote to better, more effective moderation. At what point should we consider a company to be big enough to take more responsibility for their platform?

Response: Although smaller companies might not have the resources to moderate content in a centralized way, that doesn't mean that they don't have the resources to moderate content at all. It simply means that they might take a different approach, and we should recognize that and acknowledge that it is healthy for the industry for there to be a diversity of approaches. And to be sure, no matter what the size of a company or which model of content moderation they follow, we can all do better and are continually improving our tools and processes in this regard.

- 4. Some have argued that Section 230 was not intended to provide a liability shield to a nascent industry that needed protection from insurmountable legal fees. See, e.g., Letter from TechFreedom, to U.S. House of Representatives Energy & Commerce Committee, at 11 (dated Oct. 15, 2019). You are in a unique position as head of Reddit to understand the importance of Section 230. Based on your understanding and experience, do you believe Section 230 was designed with companies like Reddit in mind? Should Congress consider the impact any liability carve-outs could have on the ability of smaller platforms to thrive?**

Response: It's hard to look at the case law leading up to 230 and conclude that the drafters didn't have companies like Reddit in mind when they proposed Section 230. If you look at the original Prodigy case that prompted 230, it looks remarkably like Reddit -- a vast system of message boards with good-faith community moderation attempts. Companies were trying to do the right thing, and being punished for it with liability. The incentives were totally backward. If you want a safer internet, you need to ensure that companies are empowered to moderate without fear of liability. On top of that, there are real competition aspects to this question nowadays, particularly when we are talking about costly things such as litigation. Even to bring a case to dismissal costs money in legal fees. Just as an example, some of our competitor companies have more lawyers than we have total employees. They'll survive the onslaught of litigation that even small carve outs to 230 would bring. Companies like Reddit, and those smaller than Reddit, probably wouldn't.

The Honorable Bill Johnson (R-OH)

- 1. How does Reddit distinguish between "unpopular" posts that don't violate your rules from those that do violate the rules?**

Response: There is plenty of content on Reddit, that, while it doesn't break the rules, is unpopular with the community. This is where Reddit's user-based content curation comes in. Any Reddit user may vote on content on Reddit, either up or down. Unpopular content that receives a high amount of downvotes is deprioritized and hidden. All this happens without our direct involvement. It is purely a function of the Reddit community self-governing.

2. **You also mentioned in your testimony that you “evolve” your policies “to ensure they keep up with reality.” What kind of updates and how often are you having to make these changes to your policies to keep up with content against your rules and standards?**

Response: Since 2017, we’ve made numerous updates to our sitewide Content Policy. This includes revisions to existing policies, such as our policies against violent content, harassment or bullying, and involuntary pornography. It also includes the creation of entirely new policies, such as a prohibition on the sale of controlled goods. Our policy team is constantly reviewing these rules, including via user feedback, to determine where the gaps are and what else may need to be addressed.

For some specific examples of what this all means, take our most recent update, to our harassment and bullying policy, which we launched this past September. One of the changes that we made, based on feedback from users, was to allow bystanders to report harassment, a departure from our previous practice of accepting reports only from the victims themselves. We thought this was too much of a burden to place on victims, and so we made the change. As another example, we updated our involuntary pornography policy in February 2018 to specifically prohibit faked content. We made the change to address the new problem of pornographic deepfakes, which are nearly always created without the consent of the person depicted.

The Honorable Richard Hudson (R-NC)

1. **One of the best parts of my job is having the privilege of representing the brave men and women who are stationed at Fort Bragg. Additionally, my district represents one of the fastest growing veteran populations in the country. I take it as my responsibility to advocate for them in everything I do. As you all are aware, the opioid epidemic is something that has ravaged our country and disproportionately affected veterans. One of the underlying issues in this area is the availability of these drugs and how easy it can be for an individual to gain access to them. Unfortunately, we have seen that these drugs are often available through illegal online sales that help fuel this crisis.**

- a. **Can you please explain how your company monitors content on your platform to ensure the illegal sale of opioids does not occur?**

Response: Thank you for the question. Our approach to this content starts with our policies. Last year, we created a new policy explicitly forbidding the use of Reddit for the sale or exchange of controlled goods, which includes prescription opioids and other drugs. The enforcement of this policy is ongoing, and relies on a combination of human monitoring and reporting as well as automated tools. In this regard, we’ve been grateful for the leadership that the FDA has provided on this issue, and we’ve been in touch with them on multiple occasions. Their list of illegal pharmacy warning letters has been an incredibly important resource for us, in terms of providing an authoritative list of unauthorized online pharmacies. We monitor these warning letters, and

automatically block the posting of URLs leading to them. Additionally, when we find accounts attempting to post this material, we ban them. We can also apply special technical methods to fight against them from coming back to our site to continue their illegal activity.

On the advertising side, we have a strict advertising policy that has special rules governing the advertising of prescription drugs. All of these advertisements must be pre-approved by our policy team in house, and they must follow all FDA and industry guidelines, including practices such as side-effect labeling. We also prohibit the targeting of prescription drug advertisements to so-called sensitive communities, including communities for addiction recovery, or support communities for those with chronic conditions.

All that said, because of the high monetary stakes involved in the illegal drug trade, this is an ongoing battle, and we are working against a creative and adaptive adversary. These bad actors often use coded or obscure language to make their activity more difficult to detect. Nevertheless, we're committed to continuing our fight against them, while also ensuring that our approach is sensitive and thoughtful enough so as not to wrongly ensnare the many, many users who find support and community on Reddit for their recovery and harm reduction purposes.

b. When you find such content on your platform, do you engage law enforcement in addition to removing the content from your platform?

Response: We cooperate with federal, state, and local law enforcement, as outlined in our Guidelines for Law Enforcement, available publicly on our website. In serious cases of content involving real-world harm, we have made proactive reports to the Northern California Regional Intelligence Center (NCRIC). The NCRIC has proven to be an invaluable resource for escalating issues to law enforcement, especially given the complexity of the different law enforcement agencies that could be involved in an investigation regarding online activity. We hope that Congress, as it makes choices regarding resourcing, continues to support efforts like the NCRIC to better coordinate law enforcement reporting with adequate resources.