Appendix for *Proactive Blocking through the Automated Identification of Likely Harassers*

Ifat Gazia, Trevor Hubbard, Timothy Scalona, Yena Kang, Ethan Zuckerman*

a.   Types of Harassment

When asked to describe what their experiences using Twitter were, particularly as Kashmiri activists living in the United States, there was a clear burden the activists experienced, being more protected from the physically harmful forms of silencing within Kashmir, but subjected to other forms of harm, such as harassment and threats. IP1 referenced this theme clearly:

IP1: "And those people who do this type of work, especially if they are Kashmiris from the valley, are literally putting their lives on the line. Because there is a different harm that they face than even I would face because I have the privilege of sitting in the United States on a sunny day and tweeting from my phone. But there are actual Kashmiri activists who are doing this, who have family back in the valley, who will be targeted as a repercussion of their voice being elevated on these platforms. And if not their families, then they themselves are then targeted because India may or may not let them back into the country, let alone let them go back to their homes."

1) Physical Intimidation of Kashmiri activists by the Police and Military:

When asked to elaborate on whether they or their family members have been physically threatened as a result of their online activity, IP1 stated:

IP1: "I know, through the networks of activists that we're working with, both on Instagram and Twitter, that there were persons who were in Kashmir, who were helping to provide content, and the military absolutely showed up at their home. And they went silent for days and we'd be concerned about them and didn't know about their whereabouts."

2) Mass Reporting and Doxing

When asked what kind of online censorship they experience, IP1, IP2, and IP3 described being mass-reported on a regular basis, doxxed, and threatened with the deletion of their online accounts by Twitter and government institutions. IP2 and IP3 described the hate comments and threats they receive online.

IP1: "So, in my mind, given the current circumstances around Kashmir, the level of hate and the degree of organizing that exists in this in an Islamophobic way from a lot of Hindutva activists, I think that the reason I keep getting flagged is in a targeted way by either actual people that go out there and troll Twitter accounts or by bots that may have been created that look for my accounts like mine, to flag them."

IP2: "Because I speak a lot about Kashmir, and that's something that the Indian government doesn't like. They all the time send Twitter formal communications, asking them to pull down my account because my tweets violate India's Technology Act 2000. I do not even know what

that means. I never even bothered to Google it up and read about it what the Technology Act 2000 is. But these tactics of intimidation, these are not new."

IP2: "So many other activists and advocacy groups that work around Kashmir faced similar consequences that accounts get either directly suspended by Twitter, or they get mass reported by right wing Hindu trolls…I actually screenshotted this from Twitter, but there was this IT cell group that said that they had given me a case number and they had to get my account suspended. And there are these targeted, you know, efforts from all these right-wing groups that target accounts like mine to stifle our voices and to stifle minority dissident voices in general."

IP2: "The harassing they did wasn't enough, so they decided to dox me. There was this guy on Twitter with almost 100,000 followers who basically researched me and put all my personal information on Twitter. And then I received so many misogynist and very bad comments, basically, which were not politically harmful to me. But personally, they affected me so much. The guy had made sure to block me before he doxed me. But of course, like he had 100,000 followers, people saw it. And my friends tried not to send me the screenshots of it, so I had to log into a different account to see what was really happening"

IP3: "For example, there were a few scholars last year in 2019 through 2020, when their work on Kashmir really started picking up, when trolls basically put out personal information about these scholars, their pictures, emails, university departments, and telephone numbers. They urged their followers to write to their university departments and claim that these scholars were part of an Islamic terrorist organization and were there to harm innocent civilians."

IP3: "They DM'ed me, you know, hateful things and my friends who were speaking out for Kashmir, or they would send us like, you know, images, threatening us or saying that they would find us and they would find our families."

IP2: "When I woke up, I had like 300 disgusting comments on my wall. It is as if these cells collaborated and decided to comment on my Twitter together…they'll use the same comments literally as if it's a copy paste."

3) <u>Targeting of the friends and family of international Kashmiri activists</u>

IP2 went even further, describing the harmful implications of their online activism on their family members, who live in Kashmir, who are subjected to threats and increased surveillance.

IP2: "For example, if I put too many pictures on Twitter, if I put any picture with my brother, it is just going to bring him in more like into security agencies under that level of surveillance and threat. And it's not like we have not already suffered…my house has been raided; my brother has been taken by the military."

b. **Consequences**

When describing their experiences using Twitter, both personally and as an activist, IP1, IP2, and IP3 described being de-platformed; feeling unsafe, both for themselves and for their families; losing trust in Twitter; and losing communities.

1) <u>De-platforming</u>

IP1: "I have a love hate relationship with Twitter right now, because Twitter is definitely one of the oldest platforms that I joined as an activist back when I was in college, and unfortunately, I've been blocked from Twitter twice now, in the last the last time was in March, just this past March. And I was blocked because of my engagement as a Kashmiri activist on Twitter."

2) <u>Fear for safety and consequent self-censorship</u>

IP2: "Initially, of course, it [the cyber harassment] was extremely unsettling. I think I became used to it because it was a do or die situation for me. If I worried about the harassment, then I would not be able to do the work I really need to do. It does not mean that it doesn't affect me. There isn't a single second in my life that I don't think about the impact of my activism and the hate I receive because it's not just about me."

IP3: "Am I really comfortable? Am I really safe? I would not say so. But am I again, everyday scared to open up Twitter? No, it's not like that. And again, it is just probably because I don't really have a public presence. I have always worked behind the scenes. So not many people would know who I am anyway…But even I know a few people who have a public presence, but they have also opted to kind of remain unknown and have a random profile on Twitter, just because -- who wants to go through the harassment every time?"

IP2: "My own brother…he wrote a post on Facebook and was detained at the police station. So, if you write anything against the government in the diaspora, your account is suspended. But if you are in Kashmir, then they'll physically arrest you. As a result, so many people who were critical of the government and the occupation, they self-censored themselves. Now, only people who are like are people like me who are in the diaspora, they are not inside Kashmir. But they are trying their best to raise their voice against what is happening in Kashmir. But then we are not also saved from censorship and surveillance and all these suspensions."

3) <u>Distrust of Twitter</u>

IP1: "It has negatively impacted my perception of Twitter. I definitely did want to go off the platform, especially after this last time [I was banned] in March."

4) <u>Loss of social capital</u>

IP2: "The thing that I honestly would have been sad about was losing the over a decade's worth of connections and network building — all the work that I put in over a decade on Twitter. There is somewhat of like a legacy in the activism that I've done, and I was just so sad about losing that, that history and that social capital. And so, I thought I'd have to like start all over again somewhere. I'm not going to give up my activism, so I would have absolutely started over if I had to."

   c. Institutional Complicity

In describing their experiences with online activism and cyber-harassment stemming from such, IP1, IP2, and IP3 also discussed actors and institutions who were culpable. They discussed ways in which Twitter and governments were complicit.

1) <u>Twitter's Complicity</u>

IP1: "I was blocked erroneously. It was not because I had actually broken any engagement rules that Twitter claimed. In reality, I noticed because my Twitter had a lot of traffic at that time that was in relation to Kashmir that somebody or something was flagging my account and associating it with other accounts that were breaking Twitter rules...And then it took forever to get back active on Twitter again."

IP2: "What's ironic is that Twitter has no mechanism in place to verify these mass suspensions. My accounts were suspended not just because Twitter suspended them; it's because the Indian government wanted my accounts to be suspended. And they are directly linked with these trolls. Once they told Twitter to suspend my account…Twitter was sending me emails asking me to find legal counsel because they were going to suspend my account."

IP3: "I'm sure they're getting help from inside of Twitter as well because it doesn't make sense that advocacy and activist accounts are blocked and suspended, whereas accounts that are openly congratulating their followers for mass reporting an account, some that don't do anything but highlight what's happening in Kashmir, are not suspended. It tells you that Twitter and social media giants are really complicit when it comes to these actions. They basically support this kind of selective silencing or selective censorship."

IP3: "All of these corporations have a lot of investments in India -- whether it's human power, or whether it's actual investments."

IP3: "I understand that the whole freedom of expression is something that's used in a good way. And then of course, beaten up and used in the worst way to silence you. While, they say that they are a platform and they encourage collaboration, expressing ideas, regardless of what community or where you're from, that's really not true. The more I interact with individuals from the black and indigenous community, I realized that it's not only Kashmiris that have experienced this kind of digital silencing."

IP1: "I could not understand how something that again is based in content sharing and the principles of freedom of speech, could then do something like this — where they're taking the voices away from people who are putting their lives on the line."

2) <u>Government Complicity</u>

IP2: "In August 2019, when they took our autonomy away, the entire region was put into lockdown. And there was no form of communication between Kashmiris and the rest of the world. That was devastating."

IP1: "Obviously, the government that is oppressing an entire people. Their primary intent in censorship is to hide any of their faults, their biggest fault being that they're oppressing an entire

people. So, by default, by allowing their censorship, you are censoring an entire body of activists that are fighting for these people."

   d. **Solutions**

When asked what they wished would be done about the harassment and silencing they experienced as online activists, IP1, IP2, and IP3 stated:

IP2: "It's high time that big tech platforms like Facebook, Twitter, stop catering to demands of authoritarian governments like India. They need to communicate with targeted groups and develop transparent mechanisms of accountability. They should also have individuals or teams that are primarily focused on addressing such issues of separation of dissident voices. They need to put systems in place that can recognize when accounts are being mass reported by bots or right-wing groups. If they hire human moderators to do the job, these moderators should be well versed in all social justice issues, including Kashmir."

IP1: "There's some level of educating that needs to be had. I do not believe that people are inherently bad, I just think that there's a level of ignorance that exists right now. And someone on the inside just needs to understand the impact that they're having and in all the ways that we've just discussed so that they can improve their platform."

IP1: "I think just off the bat, I wish that they would honor freedom of speech. Being based in and bred in America, it is such a huge principle that we've all fought for and fought to have. It really irks me that Twitter gives into India's pressure to ban posts about Kashmir."

IP3: "There is a community that has their hopes that they will actually, you know, stick to their mission and provide that platform for collaboration and voice."

Basic information like how many respondents live in Kashmir v the diaspora, gender, age, etc.

Two respondents live in Kashmir. Four live in the US.

IP1: Female, diaspora, 32-35
IP2: Female, diaspora, 30-32
IP3: Female, diaspora, 35-37
IP4: Male, Kashmir, 33-35
IP5: Male, Kashmir, 33-35
IP6: Female, diaspora, 45-48