World Athletics publishes Online Abuse Study covering Tokyo Olympic Games

Female athletes were the target of 87% of all abuse

65% of all abusive posts warrant intervention from social media platforms

In line with its commitment to making athletics a safe and welcoming environment for everyone, World Athletics today publishes findings of a study conducted during the Tokyo 2020 Olympic Games to identify and address targeted, abusive messages sent to athletes via social media.

The study revealed disturbing levels of abuse of athletes, including sexist, racist, transphobic and homophobic posts, and unfounded doping accusations. It also unequivocally highlights the greater levels of abuse female athletes receive in comparison to their male counterparts.

These results, which follow the launch of World Athletics' Safeguarding Policy earlier this month, raise concerns that existing safeguarding measures on social media platforms need to be tougher to protect athletes. Online abuse may cause trauma for the affected individual and can heavily impact athletes' performances – both in training and during competition.

The study was carried out in collaboration with Threat Matrix, an initiative by data science company Signify Group Ltd and sports investigations company Quest Global Ltd, and was developed to understand the size, scale and gravity of online abuse targeted at Olympic athletes on Twitter. It builds the foundation for the action World Athletics is taking and provides a basis for World Athletics to work more closely with social media platforms to tackle this issue.

To gain an understanding of the level of online abuse in athletics, a sample of 161 Twitter handles of current and former athletes involved in the Tokyo Olympic Games (derived from a list of 200 athletes selected by World Athletics) was tracked during the study period, starting one week prior to the Olympic opening ceremony and concluding the day after the Olympic closing ceremony (15 July – 9 August 2021).

In this timeframe, 240,707 tweets including 23,521 images, GIFs and videos were captured for analysis. This included text analysis through searches for slurs, offensive images and emojis and other phrases that could indicate abuse. It also used AI-powered Natural Language Processing to detect threats by understanding the relationship between words (allowing it to determine the difference between "I'll kill you" and "you killed it", for example).

The study revealed:

- 132 targeted discriminatory posts from 119 authors, with 23 of the 161 tracked athletes receiving targeted abuse.
- Out of the 23 athletes who received abuse, 16 were women with 115 of the 132 identified abusive posts directed at female athletes.
- Female athletes received 87% of all abuse.
- 63% of identified abuse was directed at just two athletes – both black and female – while the two most common categories of abuse were of a sexist (29%) and/or racist (26%) nature, accounting for 55% of all identified abuse.

"When we published our Safeguarding Policy earlier this month, I said athletics clubs, schools and community sports environments should be safe and happy places for those in our sport," World Athletics President Sebastian Coe said.

"In a world where we share so much of our lives online, this must apply to the virtual, as well as the physical world. This research is disturbing in so many ways but what strikes me the most is that the abuse is targeted at individuals who are celebrating and sharing their performances and talent as a way to inspire and motivate people. To face the kinds of abuse they have is unfathomable and we all need to do more to stop this. Shining a light on the issue is just the first step."

Unfounded doping accusations made up 25% of abusive messages, while 10% consisted of transphobic (9%) and homophobic (1%) posts.

89% of racist abuse was targeted at US athletes, despite them representing only 23% of the study set.

Max Siegel, CEO of USA Track&Field, commented: "The entire USATF community is grateful to World Athletics for conducting this vital survey which has confirmed unfortunately what we have all known for a long time: US athletes are disproportionately targeted for abuse and hate on social media.

"Increasing evidence indicates that this is driven by a huge rise in prejudice against race, gender and social status. Simply put, this type of behaviour is disgusting and utterly unacceptable. USATF remains committed to working alongside World Athletics, our athlete and constituent community, social media proprietors, the US Center for SafeSport and law enforcement to eliminate abuse and make our sport safe and welcoming for all."

The analysis consisted of a triage process which involved assessing each of the flagged posts, analysing their content and context and providing a view as to which one of four tiers of required action they met.

65% of abusive posts were deemed gravely abusive, thus warranting intervention from the social channel. We work closely with Twitter and have notified them of the relevant posts and users and will continue to work with their team to take appropriate actions against the perpetrators.

Jonathan Hirshler, CEO of Signify Group, commented: "Working on this important study with World Athletics and our partners Quest Global, has allowed us to illustrate the tactics and nuances used by bad actors in this space – while also providing clarity on the scale of the problem. Threat Matrix is focused on proactive monitoring and informed assessment of targeted discriminatory abuse, and the identification of accounts sending these messages.

"Our goal is to shift the process from being victim-led to a proactive and preventative approach, making the athlete online experience safe, less threatening, and less toxic. We are grateful to World Athletics for taking a lead on this important issue."

Posts marked as Tier 3 made up 12%, with one serious case of erotic fixation targeting one female athlete and four cases of egregious racism. Where tweets have passed a criminal threshold, World Athletics has, together with their relevant Member Federation, passed over information to law enforcement in that country. We have also been in direct contact with respective Member Federations where abusive accounts have been identified to determine the individuals' level of affiliation to the sport in the region.

World Athletics will be conducting further research in this area and has used the findings of this survey to introduce an Online Abuse Framework for its own social media channels to ensure they are environments free from abuse.

The beauty of athletics is its diversity. We want our channels to not only represent, but to celebrate this diversity, free from harassment and abuse.

To make sure this is the case, World Athletics pledges to

- remove hate speech, including but not limited to sexist and racist language, bullying and other misconduct from the comments sections of our channels
- block people bringing abuse and hate to our channels
- report the most serious cases to relevant authorities
- ensure our channels continue to cover and celebrate diversity and equality in sport
- build on this promise to make our channels and the social media platforms used by our members a safer and more equal environment for everyone

World Athletics will also work closely with the Safeguarding Working Group on providing appropriate materials to Area Associations and Member Federations as part of its Safeguarding Policy resources to ensure the implementation of safeguarding measures covers not just physical, but also digital spaces.

Key findings of study

This initial report covers a representative sample of 200 athletes (focusing on 161 with active Twitter accounts). The output from this scaled set provides World Athletics with clarity over the size and scale of the issue and the tactics being used by bad actors.

The output of this report will further support World Athletics' own efforts in working with social media platforms to help protect athletes by supporting their moderation capabilities.

The primary focus of this study has been Twitter, due to the availability of public data on the platform. Instagram data has also been used in parts and has contributed to investigatory work and evidence gathering.

161	**240,707**	**23,521**
athletes out of 200 selected for monitoring with active Twitter accounts	Tweets captured for analysis	images, GIFs and videos analysed

The study revealed disturbing levels of abuse of athletes, including sexist, racist, transphobic and homophobic posts, and unfounded doping accusations.

Within the study, the following key findings were identified:

- 132 posts from 119 authors targeting 23 of the 161 tracked athletes that were flagged as abusive.

- The most common categories of identified abuse were sexist and racist – accounting for 29% and 26% of all identified abuse and more than half of the total abuse.

- Out of the 23 athletes who received targeted abuse, 16 were women.

- Of the 132 abusive posts identified, 115 were directed at women.

- Two athletes received more than three-fifths of all identified abuse between them.

132
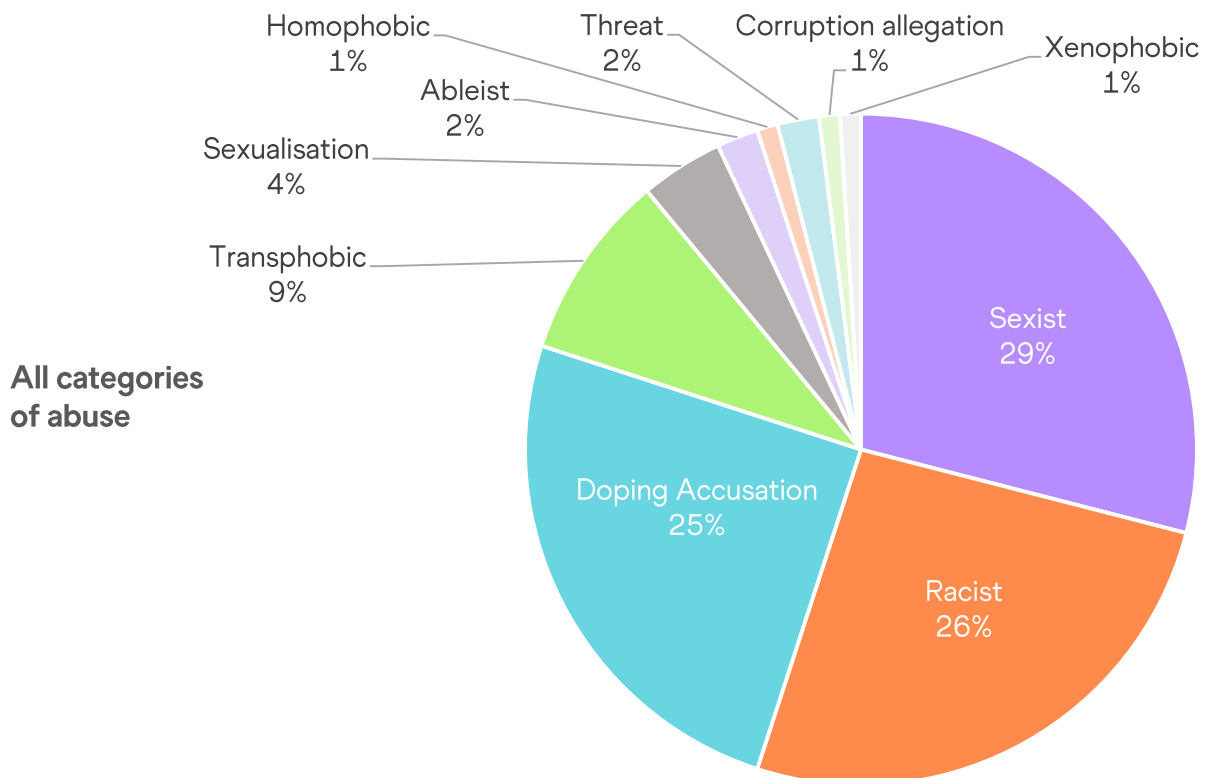abusive posts from 119 authors were identified

23
of the 161 tracked athletes received targeted abuse

55%
of identified abuse was sexist or racist in nature

Analysis of abuse

Across the study, the most frequent types of identified abuse were sexist or racist in nature. However, if comments currently listed as transphobic and homophobic were incorporated into the sexist abuse category, this would constitute the biggest single category, close to 40%.

All categories of abuse

Homophobic 1%
Threat 2%
Corruption allegation 1%
Xenophobic 1%
Ableist 2%
Sexualisation 4%
Transphobic 9%
Sexist 29%
Racist 26%
Doping Accusation 25%

The study also highlights the much higher level of abuse that female athletes face online compared with their male counterparts -70% of all abuse was targeted at women with 87% of all abusive posts directed at female athletes.

Recipients of abuse by gender

Men 30%
Women 70%

Men 17
Women 115

Number of posts

■ Women ■ Men
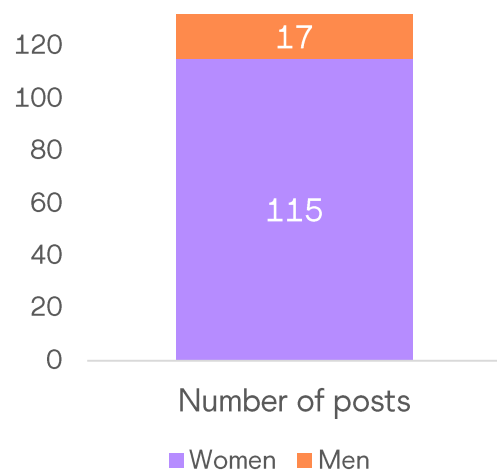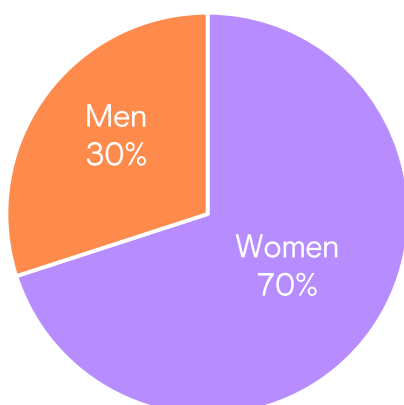
As part of the analysis, a detailed triage was applied to all flagged abuse, with the content and context of each message assessed and categorised into a set of tiers according to the gravity of abuse.

Analysis of flagged posts
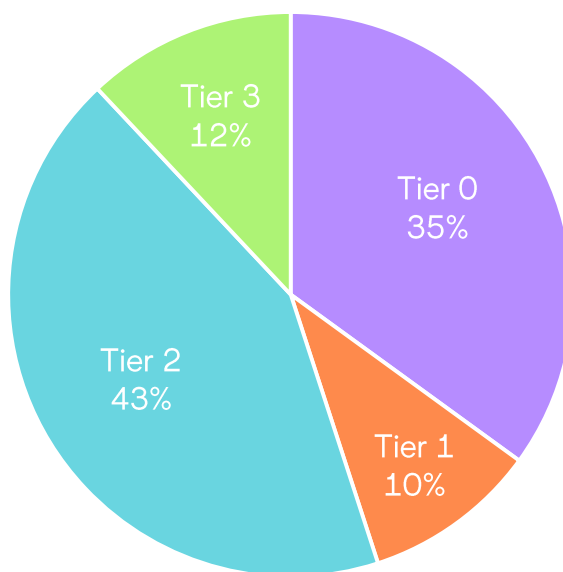
Of the posts flagged as being abusive:

- 35% were marked in Tier 0, meaning that there was abuse, but not of a level where we could expect external action.

- 10% were marked in Tier 1, meaning these comments are abusive enough to warrant intervention from the platform, but not from any other stakeholder. Twitter has to be notified of these cases.

- 43% were marked in Tier 2, with the comments warranting not only intervention from the platform, but these are cases where World Athletics is considering its options in keeping these individuals out of the sport.

- 12% were marked in Tier 3, with the comments warranting intervention from the platforms, World Athletics as the regulator and law enforcement.