DOES THE TRUTH ABOUT BLACK LIVES MATTER ON FACEBOOK?

Black Lives ALWAYS

matter

ROTEST

CUTS.

THE VIRAL DISINFORMATION ABOUT GEORGE FLOYD, BREONNA TAYLOR AND JACOB BLAKE



- Not for distribution -

CONTENT

January 2021 update

Introduction

<u>The disinformation narratives active on Facebook about George Floyd,</u> <u>Jacob Blake and Breonna Taylor</u> Other relevant observations

Is Facebook Applying its Policies?

Not "just" misinformation -- families are also facing abuse online

Why Correcting the Record is important for justice

Methodology

Document not for distribution. If this research is used, Avaaz must be informed and can be cited as follows, "Research from global civic organisation Avaaz shows/suggests." As part of Avaaz's ongoing non-partisan investigation into disinformation, our team will share snapshots of our findings that serve the public interest. Findings presented below may be updated as our investigation continues. After reviewing our reporting, Facebook may take moderation actions against the content described in this brief, such as flagging it or removing it from circulation. We provide archived links to every post we included in our analysis for this brief.

Note to journalists:

Some of the examples documented in this report have the potential to re-traumatise those directly affected. We urge caution and sensitivity regarding the reproduction of hateful content. If you include any of the screenshots below in your reporting we recommend putting a red line through screenshots of false posts in order to avoid spreading disinformation further -- see an example from the New York Times <u>here</u>.

January 2021 update

Before last week's violence on Capitol Hill, vulnerable communities across the nation were targets of violence, hate and disinformation -- organised online, often by the same actors. Throughout the 2020 election cycle, Avaaz's investigative team monitored and investigated the spread of disinformation. Our findings over the summer of 2020 indicated that violence throughout and after the election season was highly likely, because violence was already being glorified, instigated and practiced against minority communities across the country.

Communities of colour have been at the receiving end of disinformation and hate spread on social media platforms and are often disproportionately targeted and victimized by this content. The briefing below provides a small snapshot of the platforms' failure to clamp down on this harmful activity by focusing on three people whose deaths and injury inspired widespread outrage: George Floyd, Breonna Taylor and Jacob Blake. Our findings show that there were over 19 million views of misinformation content targeting these individuals between June and September, and that is only for a small sample of just 137 false and misleading posts. The overall numbers are potentially much higher.

Following the escalation and multiplication of violent confrontations targeting legislators, Congress was forced to further <u>ramp up security</u> to protect members. Most African American activists and families across the country who are targeted with violence at the hands of individuals and organized hate groups mobilizing online do not have access to this kind of protection. Meanwhile, these families are also forced to face the consistent threat of police brutality.

The research that forms the basis of this briefing originally took place in September 2020, not long after Facebook <u>publicly committed to Black lives</u>. Avaaz researchers then re-checked the results on 30 October, 1 December and 8 January. Facebook's actions on this content, despite promises to improve, remained insufficient and far removed from what experts had recommended.

Over time, with civil society and fact checkers reporting such content to Facebook, the acceleration of interactions on posts with false information on Facebook decreased, and more posts were removed. However, neither the speed nor the scope of Facebook's actions was sufficient to protect those being targeted with disinformation.

Only 25 of the 137 sample posts investigated, all of which targeted Floyd, Jacob, or Blake, were no longer available by 8 January -- it is unclear whether these were removed by the profiles who posted them, or by Facebook, or if something else happened. Facebook does not give external researchers, nor its users, transparency about how these changes and fluctuations come about. Many of these disinformation posts insinuate racist stereotypes against the victims of police brutality.

| ORIGINAL RESEARCH Sep 25, 2020 | | RECHECKS Oct 30, 2020 | | RECHECKS Dec 1, 2020 | | RECHECKS Jan 8, 2021 | |
|--------------------------------------|-----|--------------------------------------|-----|--------------------------------------|-----|--------------------------------------|-----|
| Number of Posts | 137 |
| Removed | 1 | Removed | 15 | Removed | 21 | Removed | 25 |
| False info note | 9 | False info note | 27 | False info note | 24 | False info note | 22 |
| Overlay | 60 | Overlay | 32 | Overlay | 31 | Overlay | 32 |
| Related Article | 1 | Related Article | 0 | Related Article | 0 | Related Article | 0 |
| TOTAL WITH MEASURE: | 70 | TOTAL WITH MEASURE: | 59 | TOTAL WITH MEASURE: | 55 | TOTAL WITH MEASURE: | 54 |
| TOTAL live on FB with no measure: | 66 | TOTAL live on FB with no measure: | 63 | TOTAL live on FB with no measure: | 61 | TOTAL live on FB with no measure: | 58 |

Dangerously, but unsurprisingly, many of these same pages and groups that had targeted victims of police brutality with disinformation moved on to amplify the election fraud claim. Some even mobilised their followers to attend the rally in Washington DC that transformed into an insurrection on Capitol Hill (see examples on the next page).



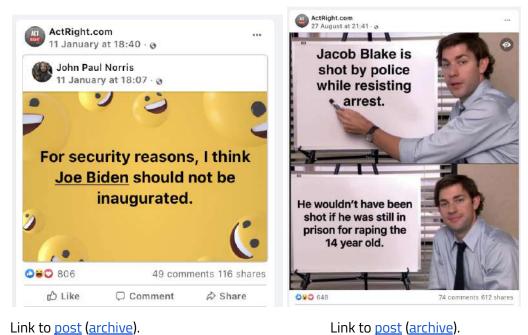




Link to post (archive)

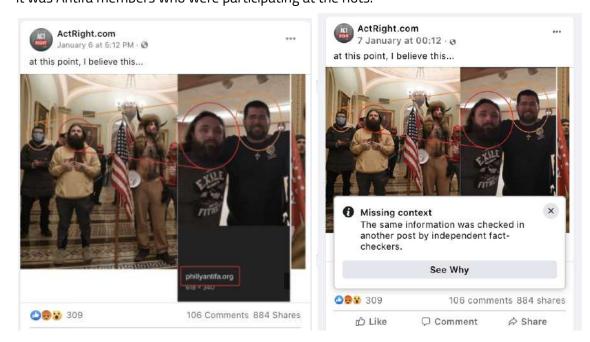
Link to post (archive).*

*The QR code on the screenshot to the right leads to Twitter, where one can RT 'JOIN US IN WASHINGTON D.C. ON JANUARY 6 AT THE NATIONAL MALL - THIS IS GOING TO BE THE DAY WE OFFICIALLY WIN AND START TO TAKE OUR COUNTRY BACK! #MILLIONMAGAMARCH #FIGHTFORTRUMP #STOPTHESTEAL #PENCECARD"



Other pages are still encouraging claims of election fraud and insinuating violence:

And some are sharing disinformation about what happened on Capitol Hill, including false claims that it was Antifa members who were participating at the riots:



Link to post (archive)



Link to post (archive)

Many of these pages could have been demoted by Facebook's algorithms for sharing disinformation, instead of amplified. Facebook could have stopped recommending these pages and groups to users as early as June, instead of accelerating their growth.

Facebook ignored expert advice and did not act as effectively as it could have against misinformation and hateful content targeting the most vulnerable in our society. It belatedly adopted better policies near and right after the election, only to roll them back again weeks before the insurrection occurred.

The erosion of truth and the polarisation, facilitated by the platform, put Americans and democracy at risk long before the attack on Capitol Hill and will continue to do so in the years to come, unless smarter policies are adopted. The highly publicised actions the platforms took against President Trump after the Capitol Hill incident were a band-aid on a bullet wound and will become meaningless if Facebook does not detoxify its algorithm and provide more transparency to users -- with retroactive corrections -- in the near future.

The potential of failure is already obvious just days after the platform promised to do more. As of 11 January, Avaaz found 90 groups with 166K total members pushing "Stop the Steal" and related content after Facebook claimed to suspend such pages. Of these, 50 are public groups that have received 200,572 total interactions in the last week alone.

Facebook must learn from the harm caused by delayed action and its refusal to implement systematic solutions to this problem. A first step in this direction, that would protect the legacy of victims of this harm, would be to retroactively correct the record on the misinformation content still spreading, highlighted in the report below, that defames George Floyd, Breonna Taylor and Jacob Blake, polarises American society and endangers all those unable to defend themselves against this tidal wave of disinformation.

Moreover, President-Elect Biden and Vice President-Elect Harris must prioritise regulating Facebook and all social media platforms, otherwise their administration's priorities may be sidetracked by the continued tidal waves of disinformation on social media that set the tone of political debate across the nation. As recent events show, social media's failure to defend the most vulnerable in our society could eventually put all of our communities at risk.

Introduction

What if your child was dead, but lies about them were kept alive on social media?

This is the situation facing the families of many victims of police brutality today -- disinformation is hurting families, defaming innocent victims and dividing the country.

In this disinformation spot check, Avaaz investigates false and misleading narratives about George Floyd, Breonna Taylor and Jacob Blake that are racking up nearly 1 million¹ interactions² (and 19 million³ in estimated views) on Facebook, despite being debunked by fact checkers. We also show the very human cost of these lies and how Facebook is failing to address them. Our research found claims that have been rated as false or misleading by reputable fact-checking organisations, often by Facebook's third party fact-checking partners, specifically:

- Disinformation targeting George Floyd, which received **168,921 total interactions**⁴
- Disinformation targeting Jacob Blake, which received **129,731 total interactions.**
- Disinformation targeting the case of Breonna Taylor, which received **661,430 total** interactions.

Our research further found:

 Facebook added false or partly false-information overlays or other fact-checking measures to 51% of content investigated⁵, but fails to detect and label copycat content (exact or slightly

¹ The total of 960,082 interactions was obtained from Crowdtangle's independent data figures for each post.

² Interactions are shares, reactions and comments.

³ The number for best estimated views is 19,575,338. More details on how this was calculated in the Methodology section of this briefing.

⁴ All interactions are for the past 12 months.

⁵ At the time of documentation, out of the 137 posts Avaaz investigated, 1 is no longer publicly accessible, 70 received a Facebook fact-checking measure, e.g. an overlay or an infobox under the post that indicate the post contains incorrect information and directs the user to articles by Facebook's third party fact checking partners, 66 had no such measure applied. Posts with a fact-check measure had 299,703 interactions while posts without had 660,320 interactions.

altered variations of false claims, see more detail below⁶). This unlabelled content investigated by Avaaz had over twice as many interactions than labelled posts.⁷

• Facebook is failing to prevent hate speech, bullying and abuse in its comment sections.

Ultimately our findings cast serious doubt on whether Facebook's policies and actions reflect its <u>claims</u> to *"stand against racism and in support of the Black community and all those fighting for equality and justice every single day."*

By allowing lies that amplify racist tropes and prejudice to go viral, while refusing to correct the record when fact-checks are available, Facebook is hurting the efforts of those fighting for equality across the country.

The content examined by Avaaz in this disinformation spot check included debunked claims that George Floyd died from an overdose⁸ and that he was a crisis actor.⁹ These claims are utterly untrue. We saw false claims that Jacob Blake raped a 14-year-old¹⁰ linked to a Twitter account impersonating a Black woman, also debunked by leading fact checkers. The false narratives about Breonna Taylor included some that were not overtly hostile but nonetheless false, like claims that Taylor was shot in bed while she was asleep¹¹ and was shot eight times¹² -- with one high-interaction post stemming from Kremlin-backed Redfish media. Avaaz also found coordinated behaviour by 15 pages controlled by The Daily Wire in posting a misleading article alleging that: "Medical Examiner Concluded George Floyd Likely Died Of Fentanyl Overdose, Court Docs Reveal".¹³

The attacking of victims and their families on social media in moments of trauma is not without precedent -- it has happened before, in the case of the Sandy Hook <u>shooting</u>, where parents

ebook-fact-checks and

⁶ Avaaz has highlighted this problem in other investigations -- more detail at <u>https://www.npr.org/2020/10/09/921791419/tiny-changes-let-false-claims-about-covid-19-voting-evade-fac</u>

https://edition.cnn.com/2020/10/09/tech/facebook-misinformation-loophole-avaaz/index.html.

⁷ Content investigated by Avaaz: labeled content has 299,703 interactions, while unlabeled content has 660,320 interactions.

⁸ Debunked by <u>AFP</u>.

⁹ Debunked by <u>Reuters</u> and <u>Factcheck.org</u>.

¹⁰ Debunked by <u>USA Today</u>.

¹¹ Debunked by <u>Courier Journal</u>.

¹² Debunked by <u>Courier Iournal</u>.

¹³ <u>Medical Examiner Concluded George Floyd Likely Died Of Fentanyl Overdose, Court Docs Reveal</u> (<u>https://archive.vn/fiuxS</u>), Daily Wire, Aug 26, 2020.

mourning the loss of their children also faced the spread of bizarre conspiracy theories claiming the shooting was staged and the children never existed. The families of Sandy Hook victims also continued to <u>endure threats and abuse online years after the shooting</u>. As Lenny Pozner, the father of a Sandy Hook Elementary School shooting victim, said in an <u>open letter to social media CEOs</u>: "I lost my beautiful 6-year old son Noah during the Sandy Hook shooting. That pain alone was unfathomable, yet I and my family were forced to relive that pain as viral fake news stories and hoaxes on Twitter, YouTube, Facebook and Instagram claimed we were actors who had staged the death of our own son."

The crisis actor narrative was a key false claim¹⁴ found by Avaaz on Facebook -- this is a common thread with other tragedies such as the <u>Parkland High School</u> shooting, when students were accused of being 'crisis actors' and <u>lies and conspiracy theories about them spread online</u>. The Pittsburgh synagogue shooting was also described as a "<u>false flag</u>" (an attempt to create sympathy, committed by Jews themselves) by some social media users.

Most recently, and connected to the topic of this investigation, the protests in <u>Kenosha</u>,¹⁵ Wisconsin, in the wake of the shooting of Jacob Blake, are a further example of how the spreading of <u>false or</u> <u>misleading claims targeting the victims of violence</u> and abuse against those closest to them happens at speed -- and without timely (retroactive) corrections.

The disinformation narratives active on Facebook about George Floyd, Jacob Blake and Breonna Taylor

Avaaz found false and misleading narratives debunked by fact checkers that continued to be shared on Facebook. The findings differed for George Floyd, Breonna Taylor, and Jacob Blake -- Avaaz found the most false content on George Floyd (65 debunked posts) and less for Jacob Blake (40 debunked posts) and Breonna Taylor (32 debunked posts).

It is important to note that our findings about the misinformation being shared are not exhaustive. We only look at fact-checked claims and only take a sample of the content available. We also use CrowdTangle, which does not provide a comprehensive analysis of what Facebook users see.

¹⁴ Debunked by <u>Reuters</u> and <u>Factcheck.org</u>.

¹⁵ Scroll down for Kenosha research by Avaaz.

False or misleading narratives observed by Avaaz included the following claims and variations on these themes:

George Floyd false claims (168,921 interactions, 3.44M estimated views)

- George Floyd's death was <u>staged</u> (archive)¹⁶
- George Floyd's death was filmed <u>before Covid-19</u> (archive)¹⁷
- George Floyd isn't dead (<u>archive</u>, content no longer available)¹⁸
- Those involved in Floyd's arrest are <u>"crisis actors"</u> (archive)¹⁹
- George Floyd has "<u>hair in one shot but is bald</u> in the shots of him laying on the ground" (archive)²⁰
- George Floyd <u>died of an overdose</u> (archive)²¹
- <u>Pelosi Gifts George Floyd's Brother</u> Folded American Flag Reserved Only For Fallen Veterans (archive)²²
- Photo of beaten woman is Aracely Henriquez, a pregnant woman who <u>George Floyd assaulted</u> during an armed robbery (<u>archive</u>)²³

¹⁶ Debunked by <u>PolitiFact</u>.

¹⁷ Debunked by <u>Factcheck.org</u> and <u>PolitiFact</u>.

¹⁸ Debunked by <u>Lead Stories</u> and <u>Reuters</u>.

¹⁹ Debunked by <u>Factcheck.org</u>, see also New York times <u>here</u>.

²⁰ Debunked by <u>PolitiFact</u> and <u>Lead Stories</u>.

²¹ Debunked by <u>AFP</u>.

²² Debunked by <u>FactCheck.org</u> and <u>PolitiFact</u>.

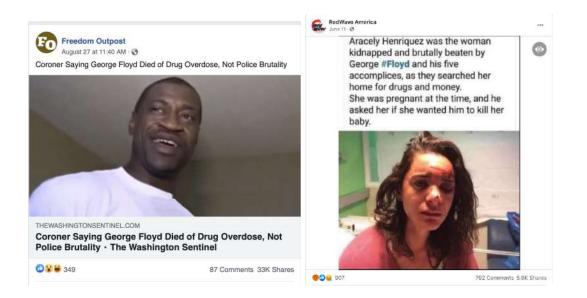
²³ Debunked by <u>FactCheck.org</u>, <u>AFP</u>, <u>Reuters</u> and <u>PolitiFact</u>.

Below is a breakdown of labelled vs unlabelled posts at the time of research²⁴ concerning George Floyd:

| Person | # of Posts | % of Total Floyd Posts | # of Interactions | Estimated views | % of Total Floyd Interactions |
|---|------------|---------------------------|-------------------|--------------------|-------------------------------------|
| Floyd posts WITH Facebook Measure | 40 | 62% | 129,553 | 2,641,487 | 77% |
| Floyd posts WITHOUT Facebook Measure | 24 | 37% | 39,309 | 801,480 | 23% |
| Floyd Posts NO LONGER AVAILABLE | 1 | 2% | 59 | 1,203 | Almost 0% |
| TOTAL | 65 | 100% | 168,921 | 3,444,170 | 100% |

For both George Floyd and Jacob Blake (see below), the majority of posts containing false information documented by Avaaz now have a Facebook measure applied. However, our findings and the high interactions on many of those posts suggest a lag in speed and, ultimately, the need for **retroactive corrections to reach the hundreds of thousands of users who have been exposed to the false claim before Facebook acted to apply the correction label.**

²⁴ Research period: September 14-24, 2020.



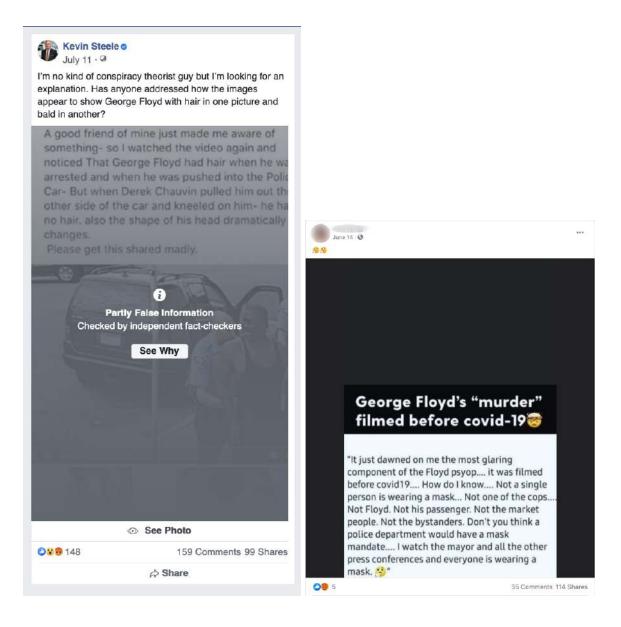
Link to post (archive)

Note: Facebook's *false information* label is now applied, but was not applied at time of research, see archived link. The hundreds of thousands of users who had seen this post before it was labeled never received a retroactive correction from Facebook. Note: **Interactions of post: 33,625**

Link to post (archive)

Facebook's *false information* label is applied. The users could have seen this high interaction post before it was labeled and would never receive a retroactive correction.

Interactions of post: 7,559



Link to <u>archive</u>, post no longer available. This appears to be the original <u>post</u> (<u>archive</u>).

Note: Facebook's *false information* label is applied. The users could have seen this high interaction post before it was labeled, and would never have received a retroactive correction.

Interactions of post: 406

Note: Facebook's *false information* label is not applied. **Interactions of post: 154**



This <u>post</u> (archive) with the <u>false claim</u> that Floyd "overdosed while resisting" also links to an <u>online</u> <u>apparel shop</u>, whose <u>Facebook page</u> has over 330,289 followers as of the writing of this briefing. Facebook's false information label is now applied. However, users who had seen this post before it was labeled never received a retroactive correction from Facebook.

Interactions of post: 3,128

Jacob Blake claims (129, 731 interactions, 2.64M estimated views)

- Jacob Blake <u>raped a 14-year-old</u> (archive)²⁵
- Jacob Blake has an <u>arrest warrant for having sex with a minor (archive</u>)²⁶
- Jacob Blake <u>declared he had a gun in the car and "brandished" a knife (archive</u>)²⁷

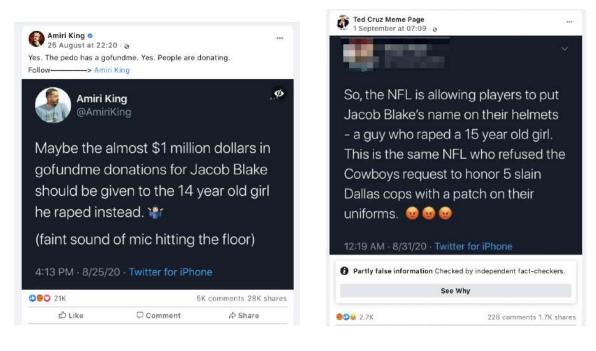
Below is a breakdown of labelled vs unlabelled posts at the time of research concerning Jacob Blake:

| Person | # of Posts | % of Total Blake Posts | # of Interactions | Estimated views | % of Total Blake Interactions |
|--|---------------|------------------------|-------------------|--------------------|-------------------------------------|
| Blake posts WITH Facebook Measure | 25 | 63% | 113,331 | 2,310,732 | 87% |
| Blake posts WITHOUT Facebook Measure | 15 | 38% | 16,400 | 334,383 | 13% |
| Blake Posts NO LONGER AVAILABLE | 0 | 0% | 0 | 0 | 0% |
| TOTAL | 40 | 100% | 129,731 | 2,645,116 | 100% |

²⁵ Debunked by <u>AP News</u> and <u>USA Today</u>. <u>Archive</u> of original post by Amiri King whose Facebook account is no longer available.

²⁶ Debunked by <u>Reuters</u>.

²⁷ Debunked by <u>USA Today</u>.



Link to post (archive).

Note: Facebook's *false information* label is applied. The users who had seen this post before it was labeled never received a retroactive correction from Facebook. **Interactions of post: 54,547.** Amiri King's Facebook account appeared to have been disabled by 8 October, after Avaaz's research for this briefing had concluded. Previous research that Avaaz submitted to Facebook had identified his page as a top misinformation-sharing page.²⁸

Link to post (archive).

Note: Facebook's *false information* label is applied. The users who had seen this post before it was labeled never received a retroactive correction from Facebook. **Interaction of post: 4,678**

²⁸ <u>www.vice.com/en/article/n7w8zd/how-facebook-allows-misinformation-to-spread-even-after-its-flagged</u>



.

Link to post (archive).

Note: Facebook's *false information* label is applied. The users who had seen this post before it was labeled never received a retroactive correction from Facebook. **Interactions of post: 3,906**

Still My President - Donald Trump 26 August at 01:18 - @ Kenosha, Wisconsin is burning because police shot a PEDOFILE RAPIST for resisting arrest! Our country is under attack from within its own borders! Look at who is supporting this criminal.... LeBron James and other high profile celebrities. Jacob Blake raped an underaged girl and these people are still supporting him Amiri King Maybe the almost \$1 million dollars in gofundme donations for Jacob Blake should be given to the 14 year old girl he raped instead. 📽 (faint sound of mic hitting the floor) 4:13 PM · 8/25/20 · Twitter for iPhone Partly false information Checked by independent fact-checkers. Foo Why

| 098 12K | | 1.7K comments 11K shares |
|---------------------|-----------|--------------------------|
| n ⁵ Like | C Comment | A Share |

Link to post (archive).

Note: Facebook's *false information* label is applied. The users who had seen this post before it was labeled never received a retroactive correction from Facebook. **Interactions of post: 26,282**



Link to post (archive).

Note: Facebook's *false information* label is applied. The users who had seen this post before it was labeled never received a retroactive correction from Facebook. **Interactions of post: 1,334**

Breonna Taylor claims (661,430 interactions, 13.48M estimated views):

- Her boyfriend was a <u>drug dealer trafficking drugs through her apartment</u> (archive)²⁹
- Police were at the <u>wrong apartment</u> (archive)³⁰
- Taylor was shot while she was asleep in bed (archive)³¹
- Taylor was <u>shot eight times by police</u> (archive)³²

²⁹ Debunked by <u>Courier Journal</u>; <u>USA Today</u>.

³⁰ Debunked by <u>Courier Iournal</u>.

³¹ Debunked by <u>Courier Journal</u>, see also New York Times article <u>here</u>.

³² Debunked by <u>Courier Journal</u>, see also New York Times article <u>here</u>.

• A <u>meme with 7 claims</u>, several of which were rated (partly) false, misleading, inconclusive, or in dispute, including that she was terminated for her job as EMT in 2017 (<u>archive</u>)³³

The 7-claim meme listed last, above, reached over 46,308 interactions³⁴ and was spotted by Avaaz researchers when running an additional search of Breonna Taylor's name on Crowdtangle between 23 September and 24 September 2020 following the <u>Grand Jury decision in the Taylor case</u>, to see how the news spike would impact this type of content shared on Facebook about Breonna Taylor.

As part of this post-announcement check, Avaaz also found a number of other high-interaction posts, some with overlays, some now deleted, pushing lines like "<u>Her boyfriend was a drug dealer trafficking</u> <u>drugs through the apartment</u>"³⁵, "<u>Her boyfriend was a career criminal</u>"³⁶ and others.

³³ Debunked by <u>LeadStories</u>, <u>USA Today</u> and (for some of the claims in the meme) <u>PolitiFact</u>. The meme appears to have originated from <u>Brandon Tatum</u>, a former police officer linked to TurningPointUSA and Prager University. The meme can no longer be found on Brandon Tatum's Twitter or Facebook. Brandon Tatum's <u>Facebook page</u> and <u>official website</u> also sell merchandise.

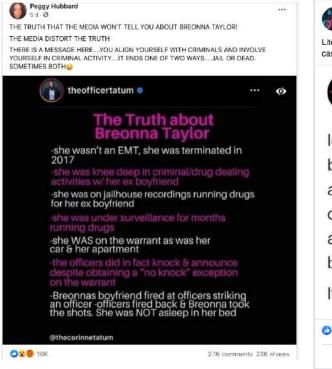
³⁴ Examples of where this meme was being shared: <u>Peggy Hubbard (archive)</u>, <u>Conservative Hangout (archive)</u>, <u>Thank you NYPD (archive)</u>, <u>Trump for President 2020 (archive)</u>. The meme was rated as "partly false information" on the Facebook overlay and <u>Lead Stories</u> (see also fact-checks by <u>USA Today</u> and <u>PolitiFact</u>) wrote: "The meme contains several claims that are false, misleading or in dispute. For example, she was not terminated from her job as an EMT in 2017. Her alleged involvement in the drug trade is in doubt -- no drugs or money were found at her home. Finally, whether or not police identified themselves at Taylor's door is in question. Officers have said that they did, though only one witness reportedly confirmed their account."

³⁵ Debunked by <u>Courier Journal</u> and <u>USA Today</u>.

³⁶ Debunked by <u>Courier Journal and USA Today</u>.

Below is a breakdown of labelled vs unlabelled posts at the time of research concerning Breonna Taylor:

| Person | # of Posts | % of Total Taylor Posts | # of Interactions | Estimated views | % of Total Taylor Interactions |
|---------------------------------------|------------|----------------------------|-------------------|--------------------|--------------------------------------|
| Taylor posts WITH Facebook Measure | 5 | 16% | 56,819 | 1,158,496 | 9% |
| Taylor posts WITHOUT | | | 50,015 | 1,130,430 | 2.18 |
| Facebook Measure | 27 | 84% | 604,611 | 12,327,556 | 91% |
| Taylor Posts NO LONGER AVAILABLE | 0 | 0% | 0 | 0 | 0% |
| TOTAL | 32 | 100% | 661,430 | 13,486,052 | 100% |





🛱 Share

Link to post (archive)

Note: Facebook's *false information* label is now applied, but the users who had seen this post before it was labeled never received a retroactive correction from Facebook. **Interactions of post: 36,876**

Link to post (archive)

Note: Facebook's *false information* label is not applied. **Interactions of post: 65,756**

Some of the false claims about Breonna Taylor were more innocuous or had a more sympathetic focus than for George Floyd or Jacob Blake. They were posted from a place of solidarity. Nonetheless, the false and misleading content shared, even if as a form of support, can increase polarisation and intensify confusion in an already fraught information environment.

One example of a solidarity post that contained misinformation and that had the highest number of interactions (over 140K at the time of research, according to Crowdtangle) in our study, was published by the <u>Kremlin-backed Redfish media start-up</u> which, while initially presenting itself as a "grassroots" community-based news organisation, is in fact a subsidiary of Ruptly, RT's video agency. The content, (<u>post</u>, <u>archive</u>) presented as a post sympathetic to Taylor, repeated a number of false claims including that police were at the wrong home, she was shot eight times and she was sleeping when she was shot. As of 11 October 2020, the post was still being shared, without a false-information overlay.

Avaaz cannot conclusively say that RedFish was the first to share this post, particularly since CrowdTangle only provides search results for content shared by public pages. It is important to note, however, that evidence from the 2016 and 2018 elections shows how Russia's IRA targeted African Americans with its <u>disinformation campaign</u>. A <u>report</u> commissioned by the Senate Intelligence Committee highlighted how many of the posts by actors connected to Russia used sophisticated strategies, often through claiming to be supporters of the Black Lives Matter movement, aimed at suppressing the African American vote and that sought to further polarise the nation through misleading content.

Approximately three hours after the RedFish post appeared, the popular progressive page <u>The Other</u> <u>98%</u> posted (<u>post</u>, <u>archive</u>) what appears to be a slightly edited version of Redfish's original, down to the same misspelling of "noone" in the final sentence. As of the writing of this document, The Other 98%'s version of the post has generated 53,546 interactions, according to Crowdtangle.

redish Russia stat

Russia state-controlled media - May 13 - 🕲

Her name was Breonna Taylor. She was only 26 when Louisville police raided the house where she was sleeping, looking for someone who did not live there, and shot her dead. Police fired 22 shots, 8 of which hit Taylor and the rest passed into other apartments.

Taylor was a healthcare worker on the frontlines as an EMT, saving lives during this pandemic. Police raided the house without knocking and killed her, even though she was unarmed and innocent. Not only were they at the wrong home, the person they were looking for was already in police custody.

Noone has been charged with her murder. #SayHerName #Justice4Breonna



Link to <u>post</u> (<u>archive</u>) Interactions of post: 142,299

The claim that police were at the wrong apartment was debunked by Courier Journal -- which obtained five search warrants and concluded that "the search warrant for Taylor and her home explicitly identified her and her address. The Louisville police were not there by mistake. They believed that Taylor had ties to Glover, one of the main suspects in the investigation".³⁷ The claim that Breonna Taylor was murdered "asleep in her bed" has also been confirmed as false by the Courier Journal (and written about in the New York Times). According to statements by Kenneth Walker to police, though

³⁷ Debunked by <u>Courier Journal</u>.

he and Taylor had been in bed "when they heard the banging on the apartment door at approximately 12:40 a.m (...) they got out of bed before police entered, and Taylor died on her hallway floor".³⁸

Similar false claims to the RedFish posts were then amplified by celebrities, who had the well-meaning goal of bringing attention to this case of police brutality, including Mariah Carey and Halle Berry. While those sharing these false claims are likely unwitting victims of a climate of misinformation surrounding these cases, they can contribute to a social media environment where it is increasingly hard to tell what is fact and what is not.



It has been 6 months since first-responder #BreonnaTaylor was murdered in her home, asleep in her bed. She still deserves justice - we have not forgotten, and we will not forget. Let's keep speak out...until our voices are heard!



Link to post (archive)

Note: Facebook's *false information* label is not applied. **Interactions of post: 60,764**

³⁸ Debunked by <u>Courier Journal</u>; also see <u>New York Times</u>.

Other relevant observations

While investigating the false and misleading narratives in this report, Avaaz made other relevant observations, which we are including here, but which are not included in the tally of interactions due to a lack of specific fact checks during the period of our investigation. We mention these observations in order to further fill in the picture of content spreading on Facebook about George Floyd, Jacob Blake and Breonna Taylor and encourage both the platform and fact-checking organisations to further investigate these findings.

Coordinated posting / "traffic for cash"

Nineteen posts by 15 pages³⁹ controlled by The Daily Wire shared an <u>article</u> alleging that "Medical Examiner Concluded George Floyd Likely Died Of Fentanyl Overdose, Court Docs Reveal"⁴⁰. A <u>Politifact</u> <u>fact-check</u> released on 25 September 2020 confirms that although George Floyd did have the substance Fentanyl in his bloodstream at the time of death, there was no question that the cause of death was the police brutality he suffered. It quotes the medical examiner's statement that "The cause of death, according to the medical examiner, was 'cardiopulmonary arrest complicating law enforcement subdual, restraint and neck compression". Further confirmation of the inaccuracy with which The Daily Wire wrote its misleading account of the Coroner's memos can be found in the <u>Wall</u> <u>Street Journal</u> analysis.

There are strong indications that these posts were made in a coordinated manner, with several being posted at exactly the same time -- see chart below. While several of the pages are branded Daily Wire properties, most disclose their relationship with Daily Wire only in the pages' "About" sections. Daily Wire has been warned about its <u>undisclosed coordinated behaviour before</u>, with Facebook stating that it ran "Pages [that] violate our policies against undisclosed paid relationships between publishers". It was reported that <u>Mad World News</u> was apparently paid for referrals and that Daily Wire used unbranded pages without disclosing their relationship to Daily Wire.⁴¹ One of Daily Wire's pages, Angry Patriot, and Mad World News were among the Facebook pages reportedly <u>followed by the "MAGA Bomber,</u>" Cesar Sayoc, before his arrest. This underlines the genuine, physical danger that

³⁹ Unique pages are: Ben Shapiro; Boycott; Daily Wire; Don't Mess With America; Donald Trump is My President; Fed-Up Americans; Lady Patriots; Matt Walsh; Michael Knowles; Pro-America News; Restless Patriot; The Angry Patriot; The Conservative; The Right News; The United Patriots.

⁴⁰ Prestigiacomo A.<u>Medical Examiner Concluded George Floyd Likely Died Of Fentanyl Overdose. Court Docs</u> <u>Reveal (https://archive.vn/fiuxS</u>), Daily Wire, Aug 26, 2020.

⁴¹ <u>https://popular.info/p/three-updates-and-one-new-hack</u>

disinformation poses to its targets, not only to public figures, but to anyone unfortunate enough to be caught up in its false news cycle. A list of the coordinated posts can be found on the next page.

| Page Name | Created | Post | Archive URL | Article | Total Interactions |
|-----------------|----------------------------|--|---|---|-----------------------|
| Ben Shapiro | 2020-08-26 21:17:07 EDT | https://www.faceb ook.com/officialbe nshapiro/posts/41 26519904052074 | <u>https://archive.vn/</u> wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? utm_source=faceb ook&utm_medium =social&utm_cam paign=benshapiro | 19,701 |
| Daily Wire | 2020-08-26 21:17:06 EDT | https://www.faceb ook.com/DailyWire /posts/26073075 39565061 | https://archive.vn/ wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? utm_source=faceb ook&utm_medium =social&utm_cam paign=dwbrand | 4,522 |
| Michael Knowles | 2020-08-26 21:17:06 EDT | https://www.faceb ook.com/michaelk nowlesshow/posts /3302196486525 182 | <u>https://archive.vn/</u> wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? utm_source=faceb ook&utm_medium =social&utm_cam paign=mjk | 3,006 |

| Matt Walsh | 2020-08-26 21:17:06 EDT | https://www.faceb ook.com/MattWals hBlog/posts/3581 356831897401 | https://archive.vn/ wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? utm_source=faceb ook&utm_medium =social&utm_cam paign=mattwalsh | 837 |
|-------------|----------------------------|--|-----------------------------------|--|---------|
| Ben Shapiro | 2020-08-26 11:31:34 EDT | https://www.faceb ook.com/officialbe nshapiro/posts/41 23769157660482 | https://archive.vn/ wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? utm_source=faceb ook&utm_medium =social&utm_cam | 219,102 |
| Daily Wire | 2020-08-26 11:31:34 EDT | https://www.faceb ook.com/DailyWire /posts/26067191 59623899 | https://archive.vn/. wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? utm_source=faceb ook&utm_medium =social&utm_cam paign=dwbrand | 19,285 |

| Matt Walsh | 2020-08-26 11:31:34 EDT | https://www.faceb ook.com/MattWals hBlog/posts/3580 022005364217 | https://archive.vn/ wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? utm_source=faceb ook&utm_medium =social&utm_cam paign=mattwalsh | 10,817 |
|---------------------------------|----------------------------|--|----------------------------------|--|--------|
| Michael Knowles | 2020-08-26 11:31:34 EDT | https://www.faceb ook.com/michaelk nowlesshow/posts /3300746043336 893 | https://archive.vn/ wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? utm_source=faceb ook&utm_medium =social&utm_cam paign=mik | 7,099 |
| Pro-America News | 2020-08-26 11:03:13 EDT | https://www.faceb ook.com/proameri canews/posts/325 4216297947854 | https://archive.vn/ wip/t3FYP | http://dlvr.it/RfNxlL | 94 |
| Donald Trump is My President | 2020-08-26 11:03:04 EDT | https://www.faceb ook.com/TheTrump FanPage/posts/19 72688212871606 | https://archive.vn/ wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=trumppre sident | 6,485 |

| The Angry Patriot | 2020-08-26 11:03:04 EDT | https://www.faceb ook.com/angrypatr iotmovement/post s/3181163548648 644 | https://archive.vn/. wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=angrypat riot | 2,223 |
|---------------------|----------------------------|--|---|---|-------|
| Restless Patriot | 2020-08-26 11:03:04 EDT | https://www.faceb ook.com/RestlessP atriots/posts/2931 385313634899 | <u>https://archive.vn/</u> wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=restlessp atriot | 541 |
| The United Patriots | 2020-08-26 11:03:04 EDT | https://www.faceb ook.com/TheUnite dPatriots/posts/26 88224031444587 | https://archive.vn/ wip/t3FYP | <u>http://dlvr.it/RfNxk</u> Q | 415 |
| The Conservative | 2020-08-26 11:03:04 EDT | https://www.faceb ook.com/the.conse rvative2/posts/35 06996612651832 | <u>https://archive.vn/</u> wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=tconserv | 90 |

| | | | | <u>ative</u> | |
|----------------------------|----------------------------|---|---|---|-----|
| Don't Mess With America | 2020-08-26 11:03:04 EDT | https://www.faceb ook.com/dontmess withamerica1776/ posts/115355536 8361682 | https://archive.vn/. wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=messwit hamerica | 58 |
| Lady Patriots | 2020-08-26 11:03:04 EDT | https://www.faceb ook.com/Patriotlad ies/posts/315742 3734352809 | <u>https://archive.vn/</u> wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=ladypatri ot | 18 |
| Boycott | 2020-08-26 11:03:03 EDT | https://www.faceb ook.com/BOYCOTT AETV1/posts/168 1123085382976 | https://archive.vn/ wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=boycott | 355 |

| Fed-Up Americans | 2020-08-26 11:03:03 EDT | https://www.faceb ook.com/angrymov ement/posts/1835 553506582896 | | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=fedup | 313 |
|------------------|----------------------------|---|---|--|-----|
| The Right News | 2020-08-26 11:03:03 EDT | https://www.faceb ook.com/therightn ews2/posts/2704 066949814197 | <u>https://archive.vn/</u> wip/t3FYP | https://www.daily wire.com/news/m edical-examiner-co ncluded-george-flo yd-likely-died-of-f entanyl-overdose- court-docs-reveal? %3Futm_source=fa cebook&utm_medi um=social&utm_c ampaign=trnews | 283 |

High-interaction disinformation in comments

Avaaz documented a potentially misleading <u>post</u> by Blue Lives Matter linking to a Police Tribune article⁴², which we believe would merit engagement from fact-checking organisations and which is in the same vein as the Daily Wire article above. False claims about George Floyd dying from an overdose have been debunked by <u>AFP</u> and Politifact as above. The key aspect Avaaz researchers observed here was the high interactions in comments on this post, where individual comments spreading disinformation debunked by official fact-checking organisations racked up hundreds of interactions. This points to a "dog-whistle"-like dynamic, where a post may be just ambiguous enough to escape false information treatment, but the resulting comments are rife with false claims. We have documented some examples in the section on comments needing corrections below.

⁴² <u>Newly-Released Evidence Shows Coroner Said George Floyd's Death Looked Like An Overdose</u>, Police Tribune.

Is Facebook Applying its Policies?

Avaaz researchers observed that 51% of the examples reviewed had "(partly) false information" labels and other measures and links to fact-checks. These 51% amount to only 31% *of total interactions* reviewed by Avaaz.⁴³ Avaaz researchers also found that Facebook is failing⁴⁴ at detecting identical or very slightly altered pictures containing misinformation and significantly failing at applying false-information overlays consistently, meaning these lies can continue to be spread without consequences. These findings raise questions about Facebook's content moderation algorithms and their apparent lack of basic image recognition capabilities and data sets of known disinformation posts, which should make flagging such copycat posts for demotion a simple task. Similarly, these findings indicate that better tools need to be built for fact-checkers to help them find all examples of existing disinformation content they flag.

Examples of such variations include:

- **Same exact text but different formatting**: Four identical posts⁴⁵ pushing the debunked claim that Jacob Blake raped a 14-year-old have false-information overlays and yet <u>this post</u>⁴⁶, with the exact same text but different formatting, does not.
- **Same picture of a tweet:** <u>all</u>⁴⁷ <u>of these</u>⁴⁸ <u>posts</u>⁴⁹, also on Jacob Blake, have false-information overlays, but <u>this post</u>⁵⁰ with the exact same tweet doesn't. The picture is possibly cropped slightly differently, but is still overall identical to the others.

⁴⁴ Avaaz has highlighted this problem in other investigations -- more detail at

https://www.npr.org/2020/10/09/921791419/tiny-changes-let-false-claims-about-covid-19-voting-evade-faceebook-fact-checks and

⁴³ The full list of examples reviewed by Avaaz does not represent the exhaustive list of *all* content for each false narrative, see methodology.

https://edition.cnn.com/2020/10/09/tech/facebook-misinformation-loophole-avaaz/index.html

⁴⁵ <u>http://archive.vn/DE6r4</u> (live link no longer available), <u>https://archive.vn/RwqEr</u> , <u>https://archive.vn/UEpHw</u>, <u>https://archive.vn/fADLW</u>.

⁴⁶ <u>https://archive.vn/x84kf</u>

⁴⁷ <u>https://archive.vn/RJQEw</u>

⁴⁸ <u>https://archive.vn/RIQEw</u>

⁴⁹ <u>https://archive.vn/SKKb7</u>

⁵⁰ <u>https://archive.vn/9rknd</u>

- Same picture of a tweet, sometimes with overlay and sometimes without: <u>These</u>⁵¹ posts⁵² relating to Breonna Taylor had false-information overlays at the time of research (see archived links in footnotes), but <u>these</u>⁵³ posts⁵⁴ with the exact same tweet didn't. The claim was rated false by <u>USA Today</u> and <u>Courier Journal</u>.⁵⁵ A more recent fact-check once again dismisses the friendly fire theory.⁵⁶
- Slightly altered text and photo relating to Jacob Blake, <u>one with a false-information overlay</u>⁵⁷ and <u>one without a false-information overlay</u>.⁵⁸

Not "just" misinformation -- families are also facing abuse online

In addition to the false and misleading narratives detailed above, relatives of victims of police brutality can also face significant abuse in the comment threads under their own public posts on Facebook.⁵⁹ Those comments can also contain some of the same disinformation narratives laid out above. Avaaz checked comments on public Facebook posts by Jacob Blake's father, Jacob Blake's sister, Breonna Taylor's sister and George Floyd's sister. The abuse was most intense on Jacob Blake's father's Facebook wall, where comments repeatedly spelled out the n-word, and violent imagery involving guns is also used. He has now deleted his Facebook account. Examples documented by Avaaz include:

⁵⁷ <u>https://archive.vn/XKpPv</u>

⁵¹ As of January 2021, the live post is no longer available, archive here: <u>https://archive.vn/OiLzT</u>

⁵² <u>https://archive.vn/LeQS8</u>

⁵³ <u>https://archive.vn/Lc5p2</u>

⁵⁴ <u>https://archive.vn/VZDJu</u>

⁵⁵ Archived Posts: <u>https://archive.is/OiLzT</u>, <u>https://archive.vn/LeQS8</u>, <u>https://archive.vn/Lc5p2</u>, <u>https://archive.vn/VZDJu</u>

⁵⁶ According to a <u>USA Today</u> from 8 October 2020: "The authorities said the four-page report dismisses the theory that Sgt. Jonathan Mattingly was hit by "friendly fire" from his own officers and supports Attorney General Daniel Cameron's conclusion that the 9 mm round that severed Mattingly's femoral artery came from the gun fired by Kenneth Walker."

⁵⁸ <u>https://archive.is/wip/XKpPv</u>, <u>https://archive.is/YAc9Z</u>

⁵⁹ Avaaz also looked for public profiles of the officers involved and could either find no verifiable public profiles on Facebook and Instagram, or only public profiles that were very likely fake.

Comments on Jacob Blake's father's Facebook profile

- "Too bad. Maybe an 8th shot woulda did in the pedophile rapist"
- Over multiple comments, repeatedly spelling out the n-word (see screenshot below)
- "Get your rapist son a pimped out wheelchair now. You gotta do right by him he followed in your shoes"
- "Congrats your family just won the ghetto lottery you POS"
- "Fuk ur paralyzed son hahahahah"
- "They need to castrate him while they are at it and convict him of the rape he has done don't feel sorry for him"
- "The whole pedos family is listed. It's gonna be a long night for me" (includes screenshot of family members' names listed on Facebook)

Comments on Jacob Blake's sister's Facebook profile

• "Her brother is a scumbag. He forced his way into a woman's home, sexual assulted her and stole her car n debit card. Classy great guy!! Karma got him now!!!"⁶⁰

Comments on George Floyd's sister's Facebook profile

- "He was a fucking criminal"
- "Frauds. Crisis actors. Period"

Comment on Breonna Taylor's mother's Instagram profile

• "If your daughter was not hanging with and doing all the things she shouldn't have been doing with her ex. Maybe we would not be here today"

⁶⁰ According to <u>USA Today, this is true</u>: "According to a criminal complaint, Blake allegedly entered the house of a woman he knew in the early hours of the morning, sexually assaulted her and then took a debit card and car keys before fleeing in her vehicle." We are including the comment here because it can still qualify as bullying/harassment.

| | LIKE · REDIV · ZW | |
|------|--|--------|
| | N () 22 | |
| - | Like - Reply - 2w | |
| | 🗣 🛑 🚺 🐂 🐂 replied - 3 Replies | |
| | 18 | |
| | Like - Reply - 2w | |
| | Y Provide the second se | |
| | G 🔿 🍃 19 | |
| | Like - Reply - 2w | |
| | 🗣 🛑 📰 👘 replied - 1 Reply | |
| | G 🔿 💱 17 | |
| | Like · Reply · 2w | |
| | 🗣 🛑 🔤 🔤 replied - 1 Reply | |
| | E 000 18 | |
| | Like · Reply · 2w | |
| | replied · 2 Replies | |
| | R ()) 19 | |
| 1000 | Like · Reply · 2w | |
| | how original | |
| | Like · Reply · 2w | |
| | now just to let this jogger know we are glad his son is paralyzed | |
| | Like Reply 2w Edited | 020 19 |

The comments above spell out the N-word, one letter per comment.





In the face of such false, divisive and hateful content that Facebook is allowing to poison comment sections, it is hard to believe that <u>the company</u> is doing what is required to respect its mission "to give people the power to build community and bring the world closer together" and its <u>emphasis</u> that users are expected to "respect the dignity of others and not harass or degrade others."

These messages are just a sample of what's festering on the platform and, even in this small sample, we count violations of at least four Facebook policies. The <u>Community Standards</u>, which also apply to comments, do not tolerate <u>bullying and harassment</u> including a prohibition to: maliciously post content about (victims of) a violent tragedy, that includes claims that a violent tragedy did not occur; claim that victims or survivors of violent tragedies are acting or being paid; maliciously target swearing at an individual; and call for death or disability. <u>Hate speech</u> is, at least on paper, not allowed on Facebook, although the comments above which are still visible on the platform, fit the following description: "Content that describes or negatively targets people with slurs". Some of the comments moreover breach <u>Facebook's policy</u> of not allowing content that glorifies violence or celebrates the suffering or humiliation of others. Finally, the platform claims that it <u>removes</u> "explicit attempts to mock victims" of serious harm.

From the comments easily accessible to Avaaz researchers at the time of research, there is a lack of enforcement of Facebook's own Community Standards when it comes to comments. **It is**

incomprehensible and unacceptable that Facebook does not protect victims of tragedies and their families from further harm, even though its policies are designed to do just that. Avaaz reports findings on hateful content and disinformation to Facebook when it is detected.

Why Correcting the Record is important for justice

Platforms like Facebook aren't powerless in the face of the lies and hate detailed in this briefing -there's already a solution, but leaders at Facebook and their counterparts at Google and Twitter are not urgently applying these steps. In Avaaz's work with <u>survivors of disinformation</u>, families who lost children at Sandy Hook told tech platform representatives that the constant conspiracy posts and accusations kept the trauma alive. Part of ensuring justice for the memory and families of George Floyd, Breonna Taylor and Jacob Blake, as well as countless other victims of police brutality, is to crack down on the injustice of lies still spreading on the platform, and to set the record straight for the victims.

"Correct the Record": Keeping users informed

To protect the victims of police brutality and their loved ones, Facebook must "Correct the Record" for harmful conspiracy theories and misinformation, which can stand in the way of justice, be used to radicalize and incite users against certain groups or individuals and can stand in the way of the grieving process of those whose loved ones have been killed or injured. "Correct the Record" would require platforms to retroactively distribute corrections from independent fact-checkers to every single person exposed to false or misleading information. Facebook's anti-disinformation policies currently require the application of labels to fact-checked content, but do not require the retroactive issuance of corrections to the hundreds of millions of people who have seen the initial falsehood in their News Feeds before these labels are applied.⁶¹

⁶¹ Facebook has begun to provide retroactive notifications to users who have engaged with **harmful Covid-19-related misinformation, but does not provide clear and direct corrections to this misinformation.** Instead Facebook has retroactively shown in certain Covid-19 related cases generic factual information from reputable sources, which -- while a positive step -- fails to ensure that the vast majority of users who saw the original misinformation will view the specific correction to that misinformation which may help them reconsider their own beliefs.

AN EXAMPLE OF HOW THE CORRECTIONS TESTED WOULD LOOK IS PRESENTED BELOW:

| Shana Grove | | | their notification centre, when the post is eventually fact-checked by an independent fact-checker: |
|--|---|------------------|--|
| 17 hrs - Q | | | |
| The measles outbre immigration | ak was directly caus | sed by illegal | It seems like you've been recently shown factually incorrect information. |
| | | Y. | Independent Fact-checks ● brought to you by Facebook 5 hrs • Ø No, illegal immigrants didn't cause the increase in U.S measles outbreaks. The measles outbreaks in the U.S were mainly due to unvaccinated American citizens w became infected when travelling to countries where large measles outbreaks were occurring, such as Isra Ukraine, and the Philippines. Fact-checked by Snoper and Healthfeedback See more |
| all outbreaks are c American children | reak is a massive fa aused by infected i | migrants, not | False Information Checked by independent fact-checkers |
| 09912 | 7 com | ments 307 shares | |
| | | | |

More examples can be found <u>here</u>.

"Correct the Record" would cut belief in misinformation by almost half

An academic study, commissioned by Avaaz and conducted by Dr. Ethan Porter of George Washington University and Dr. Tom Wood of Ohio State University, found that providing social media users who have seen false or misleading information with direct corrections from fact-checkers <u>can decrease</u> <u>belief in disinformation by an average of almost 50% and by as much as 61%</u>. Leading researchers from Cambridge University, MIT and 20 other top institutions from across the world have recently released <u>a "Debunking Handbook" that also reaffirms that corrections for disinformation work</u>.

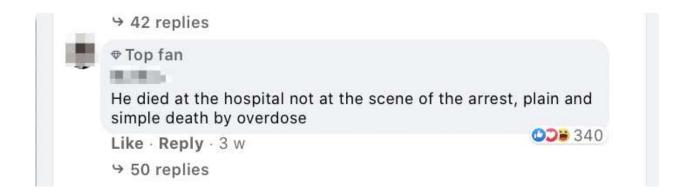
With protests against police brutality continuing, and fears of increasing violence and uncertainty in the context of elections, Facebook has a choice. The platform can continue to act reactively, apologizing for its "<u>operational mistake[s]</u>" while it is being weaponised to cause great harm to people already facing systemic injustices. Or, the platform can act proactively, applying the solutions above in addition to enforcing its existing policies and thus truly helping to protect and inform Americans at this tumultuous time.

"Correct the Record" -- Comments and Messenger need corrections, too

As our research has shown, disinformation doesn't only show up in posts, it also proliferates in comments. Facebook has a responsibility to tackle disinformation alongside hate speech, harassment and abuse in the comments; and retroactive corrections protecting the memory of George Floyd, Jacob Blake and Breonna Taylor should also be applied here.

In addition to the comments on family members' Facebook walls detailed above, which are in clear breach of Community Standards, Avaaz also found disinformation-spreading comments under some of the high-interaction false and misleading content examined for this briefing, where a single comment often reached hundreds of interactions itself.

From the comments under this <u>post</u>⁶² (encountered during research but excluded from overall tally of interactions) -- false claims about George Floyd dying from an overdose have been debunked by <u>AFP</u>:



⁶² <u>https://archive.vn/YTZVK</u>



And another comment repeating a false claim debunked by <u>USA Today</u> and <u>Factcheck.org</u> about Jacob Blake:



Finally, it's safe to assume (though impossible for external researchers to study systematically) that disinformation about George Floyd, Jacob Blake and Breonna Taylor, and attacks on their family members, also spread by Facebook Messenger. Facebook needs to urgently consider how it can apply its own Community Standards there and Correct the Record on disinformation spread this way.

Step 1 Identifying false content:

- Between 14 September and 24 September, 2020 Avaaz researchers searched the following fact-checking organisations for fact-checks regarding police shooting victims, using the names of those killed or injured in highly-publicised incidents over the last year:
 - Organisations: PolitiFact, FactCheck.org, AFP, Lead Stories, AP, Reuters, USA Today and Snopes.
 - Keywords: "Ahmaud Arbery", "Breonna Taylor", "George Floyd", "Jacob Blake", "Elijah McClain"
 - The following victims all had at least one piece of false information verified as such by the organisations listed above: Breonna Taylor, George Floyd and Jacob Blake.
- 2. We looked for further fact checks in local news sources and found more information on Breonna Taylor's case in the Courier Journal.
- 3. We selected key false and misleading narratives from step 1 and step 2 that specifically targeted Breonna Taylor, George Floyd and Jacob Blake.
- 4. We documented **matches of posts** that Avaaz researchers could find for each false or misleading claim identified in step 1 and step 2 and obtained overall article interactions from the CrowdTangle extension, when applicable. We searched public pages, groups and verified profiles for the last 12 months during the period of 14 September to 24 September 2020.
- 5. We ran an **additional check on Breonna Taylor only on 23 and 24 September**, following the Kentucky Attorney General's announcement on her case, to see if the media spike would coincide with a spike in false claims about her on Facebook.
- 6. Of the 137 posts that Avaaz investigated with a verifiably false claim, researchers documented how many did or did not receive a Facebook fact-checking label or other kind of measure, as well as how many were removed over the course of the investigation.

<u>Step 2 Estimating total views for content investigated about George Floyd, Jacob Blake and Breonna</u> <u>Taylor</u>

We estimated the views on Facebook of all the content documented above. We did this by:

1. Recording interactions with the 137 posts:⁶³

- Using the data provided by social media monitoring tool CrowdTangle, we calculated all interactions for the last 12 months during the period of 14 September to 24 September 2020.
- Cumulatively, this amounted to 960,082 best available interactions on CrowdTangle, including linked articles, memes or Facebook posts.

2. Calculating the views / interactions ratio:

- Facebook discloses the number of views for videos, but for posts containing an external domain weblink, the platform displays only the number of interactions (which are shares, reactions and comments).
- Therefore, in order to estimate the number of views of the 137 false or misleading posts, we designed a metric based on the publicly available statistics on CrowdTangle.
- Based on page data provided by CrowdTangle, we took into account the total number of views (1,109,990,000) for all videos uploaded or shared directly by the 84 Facebook pages spreading disinformation links between 24 September 2019 and 24 September 2020 and then divided it by the total number of interactions (54,440,000) for the same set of videos, which gives us a views/interaction ratio.
- The ratio applied to all images and text posts is **20.39** (1,109,990,00 / 54,440,000 = 20.39).

Total views calculation:

- All the137 posts from Facebook pages, groups and profiles between 14 September 2019 and 24 September 2020 received (a total) 960,082 interactions on Facebook.
- Multiplying the total 960,082 interactions across the 84 pages, 31 groups and 20 profiles by our views/interaction ratio of 20.39 we can extrapolate the final estimate of **19,575,338** views.

⁶³ The 137 posts were shared by 84 pages, 31 groups, 20 profiles. 2 pages that were documented during the research period are no longer available. All posts are documented and archived.

Our investigative team emphasises to reporters that these numbers are rough estimates based on the data available. But they remain estimates because Facebook has consistently refused to provide more transparency on the reach of misinformation content and the pages of misinformation superspreaders on its platforms. The platform also refuses to provide the relevant tools and access needed for researchers to calculate the exact number of views.

Documentation of abuse against families:

- 1. Identified the family members of victims of police violence through various news articles including from <u>CNN</u> (<u>CNN ex. 2</u>) and <u>NBC</u>.
- 2. Searched Facebook and Instagram for their profiles looking at those and similar names. Additionally reviewed friends/followers for connections to other family members.
- 3. Found public posts confirming relationships with victims.
- 4. Reviewed public posts and pictures from 15 September back to date of police violence for abusive comments and/or comments spreading disinformation regarding police violence.
- Documented most inflammatory comments and comments spreading disinformation that has been proven false by fact check and report them to Facebook as breaches of their Community Standards.