

# Facebook Censors Palestine

A Preliminary Analysis of Censored Palestinian Content by Facebook

Alysia Grapek

MA Cultural Studies and Communication (SDU)

## Abstract

Facebook has been deleting and censoring certain content from their platforms. An update set to go into effect on October 1st will tighten the Community Standards even further. Human rights activists and journalists have been targets of social media censorship at times where their voices have been imperative. This report aimed to build a profile that may be used for further studies and provides a hypothesis surrounding the discourse that may cause content to be censored. The findings and examples do not represent all censored content.

The analysis found that content surrounding the lives and experiences of the Palestinian people and the Israeli occupation of Palestine is more actively reviewed and censored by Facebook.

By analyzing blocked and restricted content against the Community Standards, this report briefly explored the mechanics of current Facebook censorship. One of the findings was that understanding what triggers the algorithm, as well as being aware of loaded keywords, terms, phrases and sentences is important to avoiding censorship on social media.

The study found that the Community Standards and enforcement of the guidelines indicate that subjective opinion and interpretation are a key part of what content is censored or allowed. These guidelines leave much up for interpretation and users agree to the terms that Facebook has the right to pick and choose.

The aim of the report was to understand the factors that lead to censorship in order to propose alternatives and use any findings for further studies. Based on the investigation of select discourse of the censored content, there may be measures a user can take should they want to lower the risk of having content flagged for being unintentionally anti-Semitic, xenophobic, or otherwise deemed in violation of the Community Standards. The report included some recommendations, as well as suggestions for further research.

## Background

Social media censorship of Palestinian voices is a repeated occurrence and it has been acknowledged and criticized by the UN<sup>1</sup>, among other groups. Calls for change, as well as petitions and boycotts of Facebook have been triggered in the past because of this censorship. According to 7amleh, the Arab Center for the Advancement of Social Media, the censorship had affected 2/3rds of young Palestinians who did not feel that they could express their views on social media<sup>2</sup> in 2019. According to their research, this was due to fear of oppression or reprisal.

Censorship of Palestinian voices on Facebook and Instagram have made international news and sparked attention worldwide. As recently July, Dutch-Palestinian model Bella Hadid had a post removed surrounding her Palestinian father. The Instagram story with a photo of her Palestinian father's passport was removed stating *Graphic violence, hate speech, harassment and bullying or nudity and sexual activity*<sup>3</sup> as the reason. After a statement from Hadid and international social media outrage at yet another example of censorship of Palestinians, Instagram restored the story. Instagram cited that the story contained a passport and was removed for privacy purposes, calling the removal a mistake<sup>4</sup>, despite stating other reasons prior.

In August, Facebook removed a post from the prominent American-Palestinian activist and human rights attorney Noura Erakat. In an Instagram post, Erakat shared a screenshot of her post, which had been written about the murder of her cousin Ahmed Erekat. This had been blocked as being in violation of the Community Standards for *harassment and bullying* and a warning was issued to her about consequences if future posts should be found in violation of Community Standards<sup>5</sup>.

The censorship has affected accounts, seemingly no matter the size- from celebrities, private user to organizations. Noura Erakat's accounts on Facebook and Instagram combined have around 53 thousand followers while Bella Hadid's accounts have around 8 million followers.

Though there is international recognition that Palestinian's right to free speech online is under attack through restrictions and limitations, there have not been actions taken to lessen these. The motivation of this study was to investigate what causes content relating to Palestine to be censored on social media. The report was approached with the intention of exploring technical and discursive elements of Facebook's censorship of Palestinian activists and advocates for Palestine internationally.

A research into whether there was Facebook censorship of Israeli voices, like with Palestinian voices, was conducted as well. However, the results did not support any current censorship of Israeli voices on social media.

## Methods

This report seeks to understand the discursive justification of the content that has been censored, based on a sample of posts and comments from Facebook and Instagram. This aims to build a profile that may be used for further studies. The findings and examples do not represent all censored content, nor are all comments with this language censored.

The analytical method used combined aspects of narrative and discourse analysis. The narrative analysis of the written content was based on Crang and Cook's (2007)<sup>6</sup> description of open coding. A

---

<sup>1</sup> Saraste, ICAHD; Times of Israel, 2012

<sup>2</sup> Fatafta, 2019

<sup>3</sup> Hadid, 2020

<sup>4</sup> Elle, 2020; Middle East Eye, 2020

<sup>5</sup> Erakat, 2020

<sup>6</sup> Crang and Cook, 2007

goal of the investigation was to analyze already censored material, which set the framework for the thematic coding. Once examples had been collected, the language of the censored or restricted comments was recorded. Any keywords or repeated themes were noted.

These were then used in the discourse analysis, which focused on the language used. This could include vocabulary, intention, interpretation and word placement. The background for the analysis is grounded in Foucault's approach to discourse analysis, which is influenced by the concept of power<sup>7</sup>. Political and societal power are connected to censorship of speech and communication, which was considered throughout the investigation.

For this report, examples of censored or deleted posts were collected from private users on Facebook, Instagram and Twitter, as well as articles documenting censorship of Palestinian voices. Written accounts of experiences of censorship from Facebook and Instagram were also included. The examples were noted and anonymized, however some users requested that documentation of their comments not be included in the report after facing consequences from Facebook. Research, content, keywords and hashtags are in English for the preliminary findings. Arabic and Hebrew content have not been included in this study; however, it would be relevant to include these in a more extensive study in the future.

Facebook's Community Standards are the framework for what is censored or permitted. These guidelines were documented, analyzed and included in the report<sup>8</sup>. Any keywords or repeated themes found in the text were noted. This formed the basis for the comparison and discourse analysis. The language from the censored content was then compared to the Community Standard's in order to understand what caused it to be a violation. This language was recorded and used to form recommendations that could be used to lower the risk of future censorship.

Facebook's terms of service also state that a user's information is shared among Facebook's platforms if a post is flagged or banned. These include Whatsapp and Instagram<sup>9</sup>. What was also noted during the investigation was, that users claimed Facebook does not always provide receipts or reasons for censoring material. The motives for this have been cause for speculation and have caused outrage in the past<sup>10</sup>. This also indicates a possible use of interpretation in the implementation of the Community Standards.

### Individual comments

It was a consideration that working from information provided by individual users may be biased or unreliable<sup>11</sup>. Additionally, some content was described through second-hand accounts by the users. This was noted because content that has been deleted or restricted may be unverifiable. The solution to this was not to attempt to legitimize or verify these claims, but rather to form a hypothesis and strategy for understanding the algorithm.

### Facebook Transparency

There is a clear political nature of social media censorship regarding censorship of Palestinian voices. Facebook has a history of complying with the Israeli government's requests to delete Palestinian's accounts, as well as block and filter content<sup>12</sup>. Facebook is transparent about their cooperation with foreign governments if there are requests for data or to remove data. In a report from

---

<sup>7</sup> Foucault, 1992

<sup>8</sup> Facebook, 2020

<sup>9</sup> New York Times, 2020

<sup>10</sup> Al Jazeera, 2018

<sup>11</sup> Crang and Cook, 2007

<sup>12</sup> The Intercept, 2017

2019, Facebook disclosed their cooperation with the Israeli government where it is possible to see how information is shared between the platforms and the government. In 2019, Facebook complied with 79% of the Israeli governments' requests for/ to remove user data<sup>13</sup>. Reasons for these requests are not disclosed.

In May 2020, Facebook appointed Emi Palmor, Israel's former Justice Ministry director-general to the Oversight board. This appointment led to further speculation about the censorship of Palestinians due to, in part, Palmor's alleged role in creating a "Cyber Unit" in 2016 which was responsible for censoring and limiting Palestinian accounts<sup>14</sup>. Factors such as the lack of Palestinian representation in Facebook's Oversight board and Israeli laws being the seemingly dominant authority over Palestinian's activity online highlight an imbalance of power. There was no indication in the research that Palestinian's have influence in Facebook's approach to speech online or protection of their freedom of speech.

**Disclaimer:** Judaism, Israel, Palestine and Islam are mentioned and discussed in user content and therefore included in the report.

The views, opinions and information expressed by individual users in the report are solely those of the individual users. The content does not represent the views of the researcher. This information is meant for the purposes of debate, education and research. The researcher strongly condemns anti-Semitism, xenophobia, racism and other forms of hate. In analyzing the discourse of the comments, one of the goals is to demystify the guidelines of the Community Standards and reason behind Facebook censorship. This, along with providing recommendations for future content, may help users who wish to lower the risk of having their content flagged as a violation of the Community Standards without their knowledge or intention.

### Upcoming Updates

As of September 2020, new updates have been made to the Instagram algorithm that is affecting audience reach more. This algorithm change on Instagram has not been studied at the time, however it would be a recommendation to continue this research with the new Instagram guidelines as well as Facebook's.

Also, in September, Facebook included a footnote in the updated Community Standards (2020) that will go into effect October 1<sup>st</sup>. This footnote provided context to the additions made surrounding specifically blackface and Jewish stereotypes that will be addressed and monitored from now on. This footnote reads:

*<sup>1</sup> While our hate speech policies apply equally to all people, we recognize that statements about specific groups of people may pose unique harm because of the way they've historically been used to attack, intimidate or exclude – for example, comparing black people to monkeys or farm equipment, or Jewish people to rats. In this vein, we are strengthening our policies to go beyond dehumanizing comparisons and also ban certain kinds of harmful stereotypes that have historically been used to attack, intimidate or exclude specific groups. We are in the process of identifying the specific stereotypes that most often show up on Facebook and Instagram so that we can establish a clearly defined list of harmful stereotypes that we will remove globally, and will begin by banning: 1) blackface, which is part of a history of dehumanization, denied citizenship and efforts to excuse and justify state violence; and 2) stereotypes about the power of Jews as a collective in the form of Jewish people running the world or controlling its major institutions, which reflect hatred toward Jews. This type of content has always gone against the spirit of our hate speech policies, but writing a policy to capture this content*

---

<sup>13</sup>Facebook Data Requests, 2019

<sup>14</sup> Middle East Monitor, 2020

*equitably and at scale has proven difficult. The challenges remain, but we are committed to doing the right thing and are attempting to close the gap between the letter and the spirit of our policy.*

The meaning of this note, along with the timing, likely comes as a direct result of the Anti-Defamation League's boycott on Facebook. ADL, along with other organizations participated in a campaign called #StopHateforProfit<sup>15</sup> which urged Facebook to tighten restrictions on hate. The ADL has moved for harsher guidelines in order to fight anti-Semitism. While speculation, the results of a survey published in January 2020 by ADL may provide context regarding the information and motive in the footnote. From the survey, two points stand out in particular regarding anti-Semitism and Facebook's updates and restrictions. These points (2020) were:

<sup>16</sup>**Stereotypes about Jewish control of business and the financial markets** are among the most pernicious and enduring anti-Semitic beliefs, with 15 percent finding Jews have too much power in the business world and 10 percent agreeing with the statement "Jews are more willing than others to use shady practices to get what they want." Nearly one-third of respondents (31 percent) say that Jewish employers go out of their way to hire other Jews, and 17 percent say that "the movie and television industries are pretty much run by Jews."

**Anti-Semitism also takes the form of exaggerated claims about the Israeli government.** Fourteen percent say the Jewish state sometimes "behaves as badly as the Nazis," and 16 percent agreed with the statement that Israel's "record on human rights is worse than most other countries."

## Analysis

The analysis involves the types of content that may be censored, information surrounding the updates to the Community Standards, as well as what content is permitted. The types of restricted content can be divided into two categories: the direct and indirect.

### Direct Content

The **direct content** can be defined as the content that directly triggers the algorithm. These are mainly keywords, hashtags and content coming from already flagged profiles. According to Facebook, there are certain words that are considered inappropriate. These are provided on their Community Standards page. Instagram's guidelines<sup>17</sup> are vaguer and less extensive than Facebook's. Content can be triggered by the algorithm or reported by independent users. The hypotheses formed are based on the available information and does not claim that all content with these keywords or themes is restricted.

### Coding

As personal content is often narrative, the material was approached from the lens of a narrative analysis. The material used in this study was thematically coded based on themes that were found to be repeated and more common in the posts. The main themes found in this process were:

#### *Direct Keywords*

---

<sup>15</sup> #StopHateforProfit, 2020

<sup>16</sup> ADL, 2020

<sup>17</sup> Instagram, 2020

*Generalizations**Anti-Semitism**Context and Commentary**Shadow Banning**Permitted Material*

These were further explored using discourse analysis in order to understand the implementation and extent of the Community Standards on user material.

## Keywords

Some main keywords were found to have been linked to removal of content. These were included from prior research on the censorship of Palestinian voices by Marwa Fatafta in 2019. The main keywords found to trigger review were<sup>18</sup>:

*Hamas**Jihad**Al Qassam**Al Saraya**Hezbollah**Martyr**Resistance**Terrorist/Terrorism*

These keywords, along with variations of the spelling of these words were found to be in connection to the counterterrorism initiatives and algorithms by Facebook. It is for that reason they are considered direct content as the words risk triggering a response from implemented counterterrorism measures<sup>19</sup>. While not every post using these hashtags is deleted, the account may be flagged or observed.

## Key terms for hate speech/ bullying/harassment

*Hate speech* was the most reported reason for content restrictions or removal. In order to understand how the discourse could be interpreted as *hate speech*, the Community Standards, guidelines, discourse, context and expression were all considered and analyzed. Statements made alone or in combination with other statements, as well as photo or video content have been deemed examples of hate speech and removed from the platforms.

Key terms included in the Community Standards are marked as direct content as well. The expressions found to be repeated or relevant to this study are italicized.

Sections of the Hate Speech Community Standards (2020) are listed below. The full list, including the third tier which was not included at this time, can be found at Facebook's Community Standards page<sup>20</sup>. Several additions to the previous version will be enacted in the October 1st update. These are highlighted in green in the text below.

**Tier 1- Hate speech**


---

<sup>18</sup> Fatafta, 2019

<sup>19</sup> Facebook, 2020

<sup>20</sup> Facebook, 2020

Designated dehumanizing comparisons, generalizations, or behavioral statements (in written or visual form)- that include:

Black people and apes or ape-like creatures

Black people and farm equipment

Caricatures of black people in the form of blackface

Jewish people and rats

Jewish people running the world or controlling major institutions such as media networks, the economy or the government

Muslim people and pigs

Muslim person and sexual relations with goats or pigs

Mexican people and worm like creatures

Women as household objects or referring to women as property or "objects"

Transgender or non-binary people referred to as "it"

### **Tier 2- Based on characteristics**

*Character traits culturally perceived as negative, including but not limited to: coward, liar, arrogant, ignorant*

*Derogatory terms related to sexual activity, including but not limited to: whore, slut, perverts*

*Other statements of inferiority, which we define as:*

*Expressions about being less than adequate, including but not limited to: worthless, useless*

*Expressions about being better/worse than another protected characteristic, including but not limited to: "I believe that males are superior to females."*

*Expressions about deviating from the norm, including but not limited to: freaks, abnormal*

*Expressions of contempt (in written or visual form), which we define as:*

*Self-admission to intolerance on the basis of a protected characteristics, including but not limited to: homophobic, islamophobic, racist*

*Expressions that a protected characteristic shouldn't exist*

*Expressions of hate, including but not limited to: despise, hate*

*Expressions of dismissal, including but not limited to: don't respect, don't like, don't care for*

*Expressions of disgust (in written or visual form), which we define as:*

*Expressions that suggest the target causes sickness, including but not limited to: vomit, throw up*

*Expressions of repulsion or distaste, including but not limited to: vile, disgusting, yuck*

Based on the tiers, examples of keywords based on hate speech in the Community Standards were determined in relation to censored posts. These words may be found independently of each other or in combination with others which could be flagged according to Facebook's guidelines. A sample list of keywords is:

*Jewish / Muslim*

*Coward /Liar / Arrogant /Ignorant*

*Useless/Worthless*

*Hate/Despise*

*Vile/Disgusting*

## Indirect Content

In this study, the **indirect content** can be defined as the banned, restricted or removed content that has less of an obvious or direct connection to the prohibited material. Using the framework

provided by the Facebook Community Standards and the coding of the censored material, discourse analysis was used to form hypotheses about the content.

Certain factors were noted that relate to discourse, as well as how the content was displayed.

### Generalizations

Posts with discourse that uses broad generalizations have been censored or removed, whereas posts with more specificity may have less of a likelihood of the same occurring. These generalizations were more often found with one or more other terms that went against community guidelines. When the generalizations involved Israel, the broad language could be interpreted as hate speech and anti-Semitism, whether intentional or not.

Examples of this found in deleted or restricted content were the term “*Israeli*” in combination with terms such as “*thieves*”, “*cowards*” or “*criminals*”. The language, if read selectively, can be interpreted as a generalization and not context specific. In this case, these terms used in a post may have been flagged as hateful because the terms “*thieves*”, “*coward*” or “*criminals*” are understood to be used to describe “*Israelis*”. This also suggests that the context of the language may not necessarily be a factor when content is flagged as hate speech.

### Anti-Semitism

As of September, 2020, Facebook does not have an official definition for what is anti-Semitic. Because of this, interpretation is a factor in what is censored under this category. One observation from the censored material was that content that used the word “*Nazi*” when used as a description for Israel, the Israeli government etc. could be marked as anti-Semitic. Some of these comments drew parallels between Israeli occupation/ army/ settlers and Nazis, which would indicate that these comparisons are viewed as anti-Semitic.

The available posts directly flagged as anti-Semitic did not show what could be directly interpreted as hate speech about Jews or harmful Jewish stereotypes. Some posts that were removed stating anti-Semitism as the reason included the shared keywords/themes of:

“*Illegal settlers*”  
 “*Illegal Orthodox Jewish settlers*”  
 “*Zionist*”

This posed the question if content surrounding Israel and Israelis is treated as though it were content surrounding Jews? This could be investigated in a further study. While these keywords do not directly go against community guidelines, the statements draw on generalizations that may be flagged for anti-Semitism. This shows a discrepancy between the guidelines provided, the lack of definition for anti-Semitism provided by Facebook and the actions taken.

Users shared that posts criticizing governments had also been taken down. These were found to be far less often; however, they were reported. One user had received a 3-day restriction for writing: “*US has freedom like Israel is the only liberal democracy in the Middle East*”.

The reasoning for this comment being removed was not disclosed, however drawing on the ADL comments about anti-Semitism<sup>21</sup>, one possibility could be that the removal of this comment may be tied to the tightening of exaggerated claims about the Israeli government being classified as anti-Semitic.

---

<sup>21</sup> ADL, 2020

## Shadow Banning

Shadow banning is when a user or page becomes far less viewable to the audience<sup>22</sup>. In this case, users who noted that they experienced that while they were not blocked or banned, their posts were hidden or extremely hard to find which lowered their engagement. This was noted as shadow banning. It is recommended to research this further, as it may have affected more users without their knowledge due to the stealthy approach to censorship.

In one example, one user reported his page of over 38 thousand followers has had very little reach when posting about Palestine throughout the last 10 years. Another user reported having their posts about Palestine regularly hidden. The discourse on posts leading to shadow banning could not be analyzed at this time to be included in the preliminary findings, however, the process is on-going.

## Context and Commentary

In a blog about Facebook's approach to hate speech, it is stated that context and intent of the content is considered when removing a post<sup>23</sup>. In another section regarding violence, it states:

We understand that people commonly express disdain or disagreement by threatening or calling for violence in facetious and non-serious ways. That's why we try to consider the language, context and details in order to distinguish casual statements from content that constitutes a credible threat to public or personal safety. In determining whether a threat is credible, we may also consider additional information like a targeted person's public visibility and vulnerability. (Facebook 2020)

While this may refer to many types of statements, the inclusion of the consideration of language and context may also be transferred to other content. From the sample of restricted posts, users reported being blocked or banned after providing their own commentary on posts or videos. One user shared that it was a "mistake" to comment on a video and from now on they will only provide the facts. This type of personal commentary on pictures, articles and videos indicates that it was not the content itself that had been censored, but that it may have been triggered by the inclusion of personal commentary.

Words such as "*vile*" "*hate*" or "*evil*" have been used in restricted content, which are tied to strong emotions. The hypothesis is that personal comments may be more emotional, which may include keywords that trigger the algorithm. Some emotionally charged words that are flagged by Facebook's Community Standards are:

*"Coward/liar/arrogant/ignorant"*

*"Useless/worthless"*

*"Hate/despise"*

Another possibility regarding certain banned posts is that the material, for example the video mentioned above, was restricted because of the lack of context that would make the material acceptable. Whether it is clarifying the intent or avoiding expressions of disdain, the language used in the commentary is found to be a key determinant.

---

<sup>22</sup> What is Shadow Banning, 2020

<sup>23</sup> Facebook, 2017

## Permitted Content

In order to include another perspective, comments directed towards the users were also included. Users shared comments they had experienced to be xenophobic, violent or hateful. According to the users, these statements were not removed by the platform, but were found to be acceptable by Facebook and Instagram. Examples of posts that were not found to violate community guidelines were:

*“Death to Palestine”*

*“The people in Gaza should be exterminated with insecticides”*

*“Every Muslim is a dead terrorist”*

*“Proud Kahanist”*

These were included to analyze the enforcement of these guidelines from different sides of the spectrum. While hate speech against Muslims is prohibited in the Community Standards for example, the comments above were not deemed to be violations. A passage in the Community Standards about speech that incites violence states:

“In some cases, we see aspirational or conditional threats directed at terrorists and other violent actors (e.g. Terrorists deserve to be killed), and we deem those non credible absent specific evidence to the contrary.”

Another passage states:

“We allow people to debate and advocate for the legality of criminal and harmful activities, as well as draw attention to harmful or criminal activity that they may witness or experience as long as they do not advocate for or coordinate harm.” (Facebook 2020)

These passages indicate that subjective opinion and interpretation are a key part of what content is censored or allowed. Whether this is done manually or through an algorithm is not disclosed. These guidelines leave much up for interpretation and users agree to the terms that Facebook has the right to pick and choose.

One possible explanation is a selective enforcement of the guidelines where certain comments are deemed “non credible” and therefore not considered a violation. A thorough study of permitted and banned content based on the given guidelines was not the main goal of the report, however these results open up the possibility for further investigation into if the standards and disciplinary measures are being approached and enforced equally among users.

## Findings

One of the findings was that understanding what triggers the algorithm, as well as being aware of loaded keywords, terms, phrases and sentences is important to avoiding censorship on social media. These were compiled using the Community Standards, censored posts, as well as using prior research regarding the censorship of Palestinian voices online.

The study found that the Community Standards and enforcement of the guidelines indicate that subjective opinion and interpretation are a key part of what content is censored or allowed. The implementation of the Community Standards and lack of definition for key terms such as “anti-Semitic” leave much up for interpretation. This is something that users agree to by using the platform and it is still in the process of being updated according to Facebook’s most recent update.

Another conclusion of the report is that content surrounding the lives and experiences of the Palestinian people and the Israeli occupation of Palestine is more actively reviewed and censored by

Facebook. This is based on the user data, information provided by Facebook's transparency policy, the lack of censorship for Israelis, the discourse of the Community Standards and other available factors that highlighted the difference in approach to Palestinian content. Even in cases where content was restored, the response to ban or remove content deemed as a violation indicates an imbalance in enforcing the Community Standards.

While it cannot be verified at this time, on the basis of the analysis of discourse, one hypothesis is that context and word formulation play a role in which posts are censored. The Community Standards include factors such as public visibility, language, details, intent and context, which may play a role in the selective censorship. Furthermore, the combination of key terms in certain posts were found to likely be a determining factor in whether the content was prohibited or permitted.

The aim of the report was to understand the factors that lead to censorship in order to propose alternatives and use any findings for further studies. Based on the discourse analysis of the censored content and framework provided by Facebook, a number of recommendations were formed with the intention of providing possibilities to potentially lower the risk of having content flagged for being unintentionally anti-Semitic, xenophobic, or otherwise deemed in violation of the Community Standards.

## Recommendations

The recommendations provided in this section were developed using the findings of the report, the updated Community Standards on Facebook and the aforementioned external factors found during the research process. The text presented in the Community Standards is simultaneously vague and specific which may lead to difficulty understanding and adhering to the framework. The intention of these recommendations is to provide information which may aid organizations, content creators and individual users in staying informed and aware of the changes being made to social media platforms. This can help maintain their presence and ensure their voices are heard. Below are recommendations for posting on social media that could potentially lower the chance of being flagged or censored. Following these recommendations is not a guarantee that the content will not be restricted or removed.

### **Consider content and language**

When sharing content surrounding Palestine, any language, imagery or video may be reviewed, restricted or removed. Staying aware of language in particular may lower the chance of being restricted. This could be something like when using keywords that may trigger a response, know that the post is more likely to be scrutinized. Changing the format, word choice or providing context may help.

### **Be more specific when speaking about subjects**

Most censored posts studied were revealed to use a type of generalization. Avoid generalizations based on things such as country, religion and protected characteristics like those mentioned in Facebook's Community Standards. Adding a little more specificity may keep the content from being restricted. Some examples of this could be:

*(We listened to) a group of French politicians vs. The French*  
*(I bought something from) that businessman over there vs. businessmen*

In both of these examples, the extra context makes the statement less of a generalization and provides context.

### **Steer clear of stereotypes and stereotypical comparisons**

Along with the generalizations, the findings recommend avoiding the use of stereotypes. Facebook and Instagram have an active effort to crack down on these, which may affect your content. In order to lower the chance of your content being taken down, it could be helpful to avoid including language or imagery of stereotypes or comparisons, as they may be deemed harmful or hateful. Jokes and memes are included as in the Community Standards, which means they can also be restricted if they include a stereotype.

### **When writing about something that is unclear or not legally proven, include terms like “alleged” “debated” or “suspected”.**

Facebook includes that “unverifiable rumors” are in violation of their Community Standards. While an action in a photo or video may seem clear and obvious, certain statements may be perceived as accusations or rumors by Facebook. Without legal merit, these comments can be labeled or interpreted as hate speech and censored. Using terms such as “*alleged*” “*debated*” or “*suspected*” is an acceptable option for describing an individual/group and a claim without the statement becoming a rumor.

### **When quoting other comments or debating controversial content, add a line of context**

Facebook’s Community Standards shared that this context is necessary in their review of content. When sharing something that is in violation of the Community Standards, adding a line of context can distance you from the content. For example, if sharing a hurtful comment that you or someone you know received, include that they are not your words but that they were sent to you. This context separates you from the statement you shared. In doing this, according to the Community Standards, Facebook is less likely to penalize you for the content. If the post was removed and then disputed, the description will provide context and motive that explains the use of the language during the review process.

### **Avoid negatively charged emotions or words when writing about a group or subject**

According to the Community Standards, many of the keywords that are found to be associated with hate speech, are also tied to negative emotions. This recommendation is not to mask or hide emotions, but rather to attempt to provide an explanation for the censorship of some content. The emotionally charged vocabulary was found to be an issue mainly in combination with generalizations about groups or countries.

Another recommendation could be that when commenting or posting about something specific (such as a person, place or event), use vague language rather than specific language. Unlike in the recommendation about generalizations, this is meant to lessen the likelihood of misunderstandings when commenting on others’ content, videos and pictures. For example, if adding commentary about a video, do not directly address the object or subject in the video. A direct comment sharing disgust about a cat in a video might be:

*“This is a filthy cat... EW!”*

The statement is directly connecting the language tied to disgust to the cat, which could potentially make it a problem. A vague comment about the same video might be:

*“EW! That is filthy”*

Removing the subject has made the language vaguer, which does not tie the statement to the cat in the video. This may avoid statements unintentionally being flagged as hate speech.

#### **Check the Community Standards for when there are updates**

Stay in the loop of any changes that are made to the guidelines by Facebook and Instagram. This information will be helpful to keep reaching your audience and having your voice heard. Along with understanding the language and choices made by Facebook and Instagram, it can be helpful to keep up to date with technical updates, political measures and boycotts. These can all be influential in explaining what content is censored by the platform. If your content is censored, there is usually an option to dispute the choice and some users have had their content restored.

#### **Share this information with other organizations, activists, journalists and pages**

Help get the word out on these changes. If there is an update, make noise about it so more people can be aware before it happens and have the option to prepare for them. So many are dependent on social media to share the news about Palestine, as well as for Palestinians to tell their stories and document experiences. People all over the world rely on these platforms and Palestinian voices to learn, share and stay informed. Preserving Palestinian's right to free speech and combatting social media censorship of Palestinians is crucial.

## Sources

- Akpan, Paula. “What Is Shadow Banning & Why Does It Deserve Our Attention?” *Bustle*, Bustle, 2 Aug. 2020, [www.bustle.com/life/what-is-shadow-banning-how-does-it-work](http://www.bustle.com/life/what-is-shadow-banning-how-does-it-work).
- Al Jazeera. “Censored and Surveilled: The Digital Occupation of Palestinians.” *Palestine | Al Jazeera*, Al Jazeera, 25 Aug. 2018, [www.aljazeera.com/programmes/listeningpost/2018/08/censored-surveilled-digital-occupation-palestinians-180825122608591.html](http://www.aljazeera.com/programmes/listeningpost/2018/08/censored-surveilled-digital-occupation-palestinians-180825122608591.html).
- “Anti-Semitic Stereotypes Persist in America, Survey Shows.” *Anti-Defamation League*, 2020, [www.adl.org/news/press-releases/anti-semitic-stereotypes-persist-in-america-survey-shows](http://www.adl.org/news/press-releases/anti-semitic-stereotypes-persist-in-america-survey-shows).
- “Community Standards.” *Facebook*, 2020, [www.facebook.com/communitystandards/](http://www.facebook.com/communitystandards/).
- Crang, M. and Cook, I. 1995, 'Doing Ethnographies.', *Norwich*: Geobooks.
- “Facebook Appoints Ex-Israel Official to Its Oversight Board.” *Middle East Monitor*, 23 June 2020, [www.middleeastmonitor.com/20200623-facebook-appoints-ex-israel-officials-to-its-oversight-board/](http://www.middleeastmonitor.com/20200623-facebook-appoints-ex-israel-officials-to-its-oversight-board/).
- “Facebook Transparency Report: Gov Data Requests.” *Facebook Transparency Report | Gov Data Requests*, 2019, [transparency.facebook.com/government-data-requests/country/IL/jul-dec-2019](https://transparency.facebook.com/government-data-requests/country/IL/jul-dec-2019).
- Foucault, Michel, and A. M. Sheridan Smith. *The Archaeology of Knowledge*. Routledge, 1992.
- Gostoli, Ylenia. “Palestinians Fight Facebook, YouTube Censorship.” *News | Al Jazeera*, Al Jazeera, 20 Jan. 2018, [www.aljazeera.com/news/2018/01/palestinians-fight-facebook-youtube-censorship-180119095053943.html](http://www.aljazeera.com/news/2018/01/palestinians-fight-facebook-youtube-censorship-180119095053943.html).
- Greenwald, Glenn. “Facebook Says It Is Deleting Accounts at the Direction of the U.S. and Israeli Governments.” *The Intercept*, 30 Dec. 2017, [theintercept.com/2017/12/30/facebook-says-it-is-deleting-accounts-at-the-direction-of-the-u-s-and-israeli-governments/](https://theintercept.com/2017/12/30/facebook-says-it-is-deleting-accounts-at-the-direction-of-the-u-s-and-israeli-governments/).
- “Hard Questions: What Are We Doing to Stay Ahead of Terrorists?” *About Facebook*, 11 Nov. 2019, [about.fb.com/news/2018/11/staying-ahead-of-terrorists/](https://about.fb.com/news/2018/11/staying-ahead-of-terrorists/).
- “Hard Questions: Who Should Decide What Is Hate Speech in an Online Global Community?” *About Facebook*, 11 Nov. 2019, [about.fb.com/news/2017/06/hard-questions-hate-speech/](https://about.fb.com/news/2017/06/hard-questions-hate-speech/).
- “Instagram Community Guidelines FAQs.” *Instagram*, 2020, [about.instagram.com/blog/announcements/instagram-community-guidelines-faqs](https://about.instagram.com/blog/announcements/instagram-community-guidelines-faqs).
- Isaac, Mike. “Zuckerberg Plans to Integrate WhatsApp, Instagram and Facebook Messenger.” *The New York Times*, The New York Times, 25 Jan. 2019, [www.nytimes.com/2019/01/25/technology/facebook-instagram-whatsapp-messenger.html](http://www.nytimes.com/2019/01/25/technology/facebook-instagram-whatsapp-messenger.html).
- Marwa Fatafta December 4, 2019 | Edit, et al. “'Incitement' and 'Indecency': How Palestinian Dissent Is Repressed Online.” *972 Magazine*, 4 Dec. 2019, [www.972mag.com/censorship-online-palestinians/](http://www.972mag.com/censorship-online-palestinians/).

O'Malley, Katie. "Bella Hadid Accuses Instagram Of 'Bullying' After 'Removing' Photo." *ELLE*, ELLE, 8 July 2020, [www.elle.com/uk/life-and-culture/culture/a33244063/bella-hadid-accuses-instagram-censorship-palestinian-passport/](http://www.elle.com/uk/life-and-culture/culture/a33244063/bella-hadid-accuses-instagram-censorship-palestinian-passport/).

Saraste, Anne "Media Regulation and Censorship in Occupied Palestine" *ICAHD*, University of Tampere.

staff, TOI, et al. "UN Criticizes Censorship in Palestinian Territories and Israel." *The Times of Israel*, 19 June 2012, [www.timesofisrael.com/un-criticizes-censorship-in-palestinian-territories-and-israel/](http://www.timesofisrael.com/un-criticizes-censorship-in-palestinian-territories-and-israel/).

*Stop Hate for Profit*, 2020, [www.stophateforprofit.org/](http://www.stophateforprofit.org/).

"UN Censors UN Watch Submissions on Palestinian Abuse of Child Rights." *UN Watch*, 28 May 2020, [unwatch.org/un-censors-un-watch-submissions-on-palestinian-abuse-of-child-rights/](http://unwatch.org/un-censors-un-watch-submissions-on-palestinian-abuse-of-child-rights/).

"An Update on Combating Hate and Dangerous Organizations." *About Facebook*, 15 May 2020, [about.fb.com/news/2018/11/staying-ahead-of-terrorists/about.fb.com/news/2020/05/combating-hate-and-dangerous-organizations/](http://about.fb.com/news/2018/11/staying-ahead-of-terrorists/about.fb.com/news/2020/05/combating-hate-and-dangerous-organizations/).

## Appendix 1

### Facebook Comments



I notice mine don't get the exposure of my regular posts. I haven't had it blocked yet, but hidden.



Top Fan

Don't use Twitter or Instagram, but on occasion my comments have been erased when I check back. Especially when I call the Zionist Israeli army and settlers =NAZIS



I posted a video of a settler repeatedly running over a Palestinian until he was dead while a soldier looked on. I made the mistake of commenting. 7 days. Now I just give facts without my own views



I was blocked for hate speech after commenting on a post about the demolition of Palestinian homes and taking of the land by Israelis, leaving several families homeless. My comment - "Israeli thieves".



I was blocked because I likened some site to that of Adolf.....



Been banned from telling the truth about the crimes committed against the Palestinian people by Zionist Israeli criminals.  
FREE PALESTINE.

Like · Reply · 6d



Always my Facebook account was blocked for 3 days for calling Netanyahu and Emir of UAE evil



I was blocked for 24 hours because of making a posting about Hebron/Al Khalil with a photograph of the empty streets of Al Khalil. Of course the picture was blocked.

I have posted the photograph with this text:  
„One of my hardest experiences, was a journey to the occupied city Hebron/ Al Khalil. It is a Palestinian City in the Heart of the so-called Westbank. It is occupied by illegal orthodox Jewish Settlers, which are Month by month occupying more and more Palestinian Land with the help of the Israel Army Forces”

Ps

Few years ago, we've had a Facebook German political page with over 30.000 Followers. The topics were pretty diverse. About the Gaza Wars and the IOF using Phosphorus, about illegal settlers, about racism in Germany, about CIA hidden actions like Operation Ajax (coup d'etat) , political scandals like Enron and so on....

From one day to another the page was deleted by Facebook. Their reason: Antisemitism.  
We have never written something bad about Jews. It was always a politically thing .



Canadians Against Islamophobia is a Facebook page that fights Islamophobia. They do actually share many of your posts. Facebook banned them 3 years ago from boosting any post for money after racist islamophobes complained against them.



This is interesting. In the past week, I reported one comment on a JVP post for rambling about Jewish control of the media, & another on Al-Jazeera for celebrating the deaths of people in the Beirut explosion ('every dead Muslim is a dead terrorist'). Facebook found that neither comment violated community guidelines or the terms of service. It's insane.

 **Top Fan**  
  
I run a page with 38k followers. Whenever I post about palestine or post anything critical of Israel. My posts get absolutely no reach. And I've been doing it for 10 years

   
On another FB page, I jokingly/mockingly mentioned how America has "freedom" like how Israel is "the only liberal democracy in the Middle East" and it seems that post was taken down altogether.

Support Message

You anonymously reported **Pravda**'s photo for displaying hate speech.

UPDATED

WHAT YOU CAN DO

- Delete message  
Delete this message from your inbox?
- Block **Pravda**  
You won't be able to see or contact each other.

ACTIVITY

Read 3 Previous Messages

**f** **Thanks for your feedback**  
Today at 8:49 PM  
Thank you for requesting a review. A specialist from our team reviewed the photo again and confirmed that it doesn't go against any of our Community Standards, including hate speech. We base our Community Standards on input from the people who use Facebook around the world, as well as experts.

Delete

(Along with Image of profile)



## Appendix 2- Instagram

 I had been active on Facebook for ten years, educating people about the situation in Palestine. I never once advocated violence, and often challenged anti-Jewish hatred. Suddenly, a couple of months ago, I ceased to exist. Without any warning, or way too appeal, Facebook disabled my account. They refused my requests to show me my first gender. My Facebook account was . I had several thousand friends and followers.

2w 29 likes Reply

[View 5 replies](#)

 I stopped going on Facebook because I kept getting kicked out for speaking out against Israels occupation. I speak out against Saudi Arabia and they don't say anything lol

2w 26 likes Reply

 Yes. I wish I had screen shot my reply to someone. I was talking about how it is European colonization just like it was in South Africa, Australia, North America, and South America. I went into detail how it is White Supremacy. Instagram removed my comment and said it was "Hate Speech". I made the mistake of clicking "Okay" instead of clicking on Instagram's "Think we made a mistake". It was not hate speech at all! To put it to the test, I reported a comment where somebody said, "Death to all Palestine". I quickly for an Instagram review response saying it does not go against their community guidelines for hate speech or violence. That is all the proof I need to know who is controls these comments.

 I'm surprised they're not censoring this. I called their suppressors terrible horrible people, using their country name and my comment was immediately banned

3d 1 like Reply

 [REDACTED] was blocked for 24 hours because of making a posting about Hebron/Al Khalil with a photograph of the empty streets of Al Khalil. Of course the picture was blocked. 

I have posted the photograph with this text:

„One of my hardest experiences, was a journey to the occupied city Hebron/ Al Khalil. It is a Palestinian City in the Heart of the so-called Westbank. It is occupied by illegal orthodox Jewish Settlers, which are Month by month occupying more and more Palestinian Land with the help of the Israel Army Forces“

2w Reply

 I was censored in a facebook group because I said I dont like Wonderwoman cause she is a Zionist IDF supporter. This was considered hate speech. And I was blocked from posting there for two weeks. 

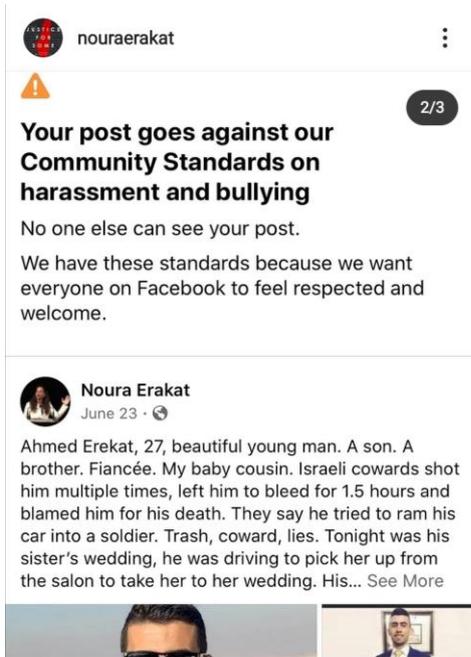
 [REDACTED] I have once reported a comment on Facebook saying that people in Gaza are insects and should be exterminated by insecticides. And guess what! They answered me that it doesn't go against the community guidelines!! 

2w 1 like Reply

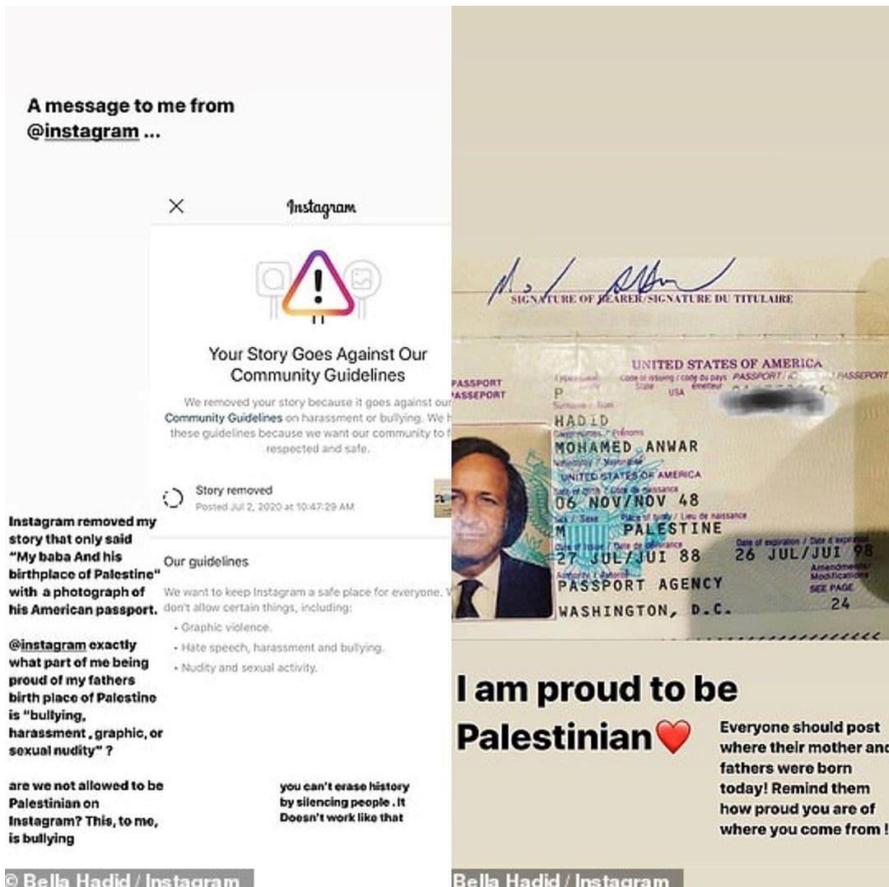
 [REDACTED] No, but once my comment has been banned because I cursed on the Israeli government LMAO 

 [REDACTED] Yep. I submitted for them to review again and they put my post back up. 700 

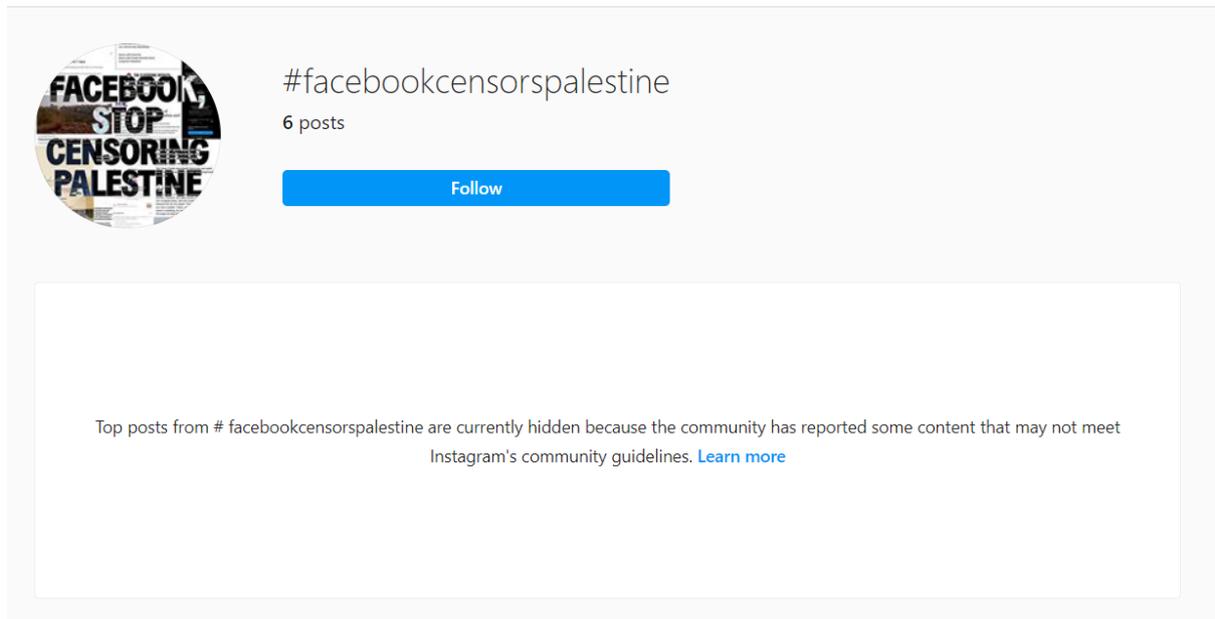
## Appendix 3- Noura Erakat / Bella Hadid



(Instagram)



## Appendix 4



September 2020