

## Methodology

To conduct this analysis, researchers collected every Facebook post and tweet created between Jan. 1, 2015, and June 14, 2020, by every official and unofficial account maintained by each voting member of the 116th U.S. Congress in office during the weeks following the killing of George Floyd (May 25 through June 14, 2020). The complete dataset contains nearly 1.3 million Facebook posts from 1,040 congressional Facebook accounts and over 2.8 million tweets from 1,021 congressional Twitter accounts, from a total of 530 members. With the exception of one finding based on all members of Congress – which includes every lawmaker who served at any point from January 2015 through June 2020 – this analysis focuses on just those members who served between May 25 and June 14, 2020. The broader data collection process is described [here](#).

This analysis includes all text from these Facebook and Twitter posts, including image captions and emojis. Photo and video posts were not included in this analysis unless the post also contained meaningful text, such as a caption. Text that appeared only within images was not included in the analysis. Posts by nonvoting representatives were also excluded, as were any posts produced by politicians before or after their official terms in Congress.

The few independent members of Congress who do not officially belong to the Democratic or Republican parties are treated as members of the party that they caucused with from May 25 to June 14 (i.e., Sen. Bernie Sanders is considered a Democrat). Rep. Justin Amash of Michigan, a Libertarian, is excluded from the analysis.

The Black Lives Matter movement and the #BlackLivesMatter hashtag started in 2013 as a response to the acquittal of George Zimmerman in the killing of Black teenager Trayvon Martin. Therefore, the Center’s collection of Facebook posts and tweets – which started on Jan. 1, 2015 – does not capture congressional social media conversation during the early years of the movement.

### Identifying posts mentioning “Black lives matter”

To identify posts about the Black Lives Matter movement, organizations and protests, researchers used a case-insensitive regular expression – a pattern of keywords and text formatting – that consisted of the terms “Black lives matter” and “#BlackLivesMatter”.

This pattern identified 2,285 unique Facebook posts and tweets that made explicit references to the movement, organizations, or protests.

## Distinct party keywords

To identify the different conversations between parties in the weeks following George Floyd’s death, researchers conducted a distinct keywords analysis using the complete set of 68,783 Facebook posts and tweets created by members of Congress from May 25 through June 14, 2020, including posts and tweets that did not mention “Black lives matter” explicitly.

Text from each document (post) was converted into a set of features representing words and phrases. To accomplish this, researchers applied a series of pre-processing functions to the text of the posts. First, researchers removed 3,059 “stop words” that included common English words, names and abbreviations for states and months, numerical terms like “first,” and a handful of generic terms common on social media platforms like “Facebook” and “retweet.” The text of each post was then converted to lowercase, and URLs and links were removed using a regular expression. Common contractions were expanded into their constituent words, punctuation was removed and each sentence was tokenized using the resulting white space. Finally, words were lemmatized (reduced to their semantic root form) and filtered to those containing three or more characters. Terms were then grouped into two- and three-word phrases.

Distinctive keywords and phrases used by each party’s members of Congress on each platform (Facebook and Twitter) were identified using [pointwise mutual information](#). Researchers then calculated the proportion of party members who mentioned each distinct term (phrase). Terms mentioned by fewer than 10 members of either party are excluded. Researchers then used the proportions to calculate a ratio of differences in mentions between parties for each term. The most distinctive party keywords were defined as those terms with the largest ratio difference between the parties.

Finally, researchers consolidated phrases: removing those that had a word in common with any other phrase that was associated with a larger difference (e.g., “Black lives matter” is not shown because “Black lives” was associated with an even larger party difference). Terms are displayed in their standardized form (e.g., “Black life” instead of “Black lives”) and have been edited slightly in some cases for readability (e.g., “the House passed” instead of “house passed”). Words that appeared in retweets are included in this analysis, even if the member who retweeted them did not create the original tweet.

## Racial and ethnic group labels for members of Congress

To assign race and ethnicity to each member of the 116th Congress, researchers used lists of the [Black](#), [Asian and Pacific Islander](#) and [Hispanic](#) members of Congress that are maintained as part of a [collaborative project](#) between the Office of the Historian and the Clerk of the House’s Office of

Art and Archives. In addition, according to the [Congressional Research Service](#), there are [four Native American members](#) in the 116th Congress. Researchers used the [Biographical Directory of the United States Congress](#) to confirm the identity of these four members.

After compiling the list of Black, Asian, Hispanic and Native American members in the 116th Congress, researchers assigned all other members to the white racial group. Multiracial members such as Sen. Kamala Harris appeared in multiple lists (in her case, Black and Asian). Researchers included these members in all applicable racial or ethnic categories when calculating the statistics reported here.