

Searching Racism after George Floyd

Christopher Barrie

¹University of Oxford

September 13, 2020

Abstract

On May 25, 2020, George Floyd, a forty-six year-old African American man, was killed by a white police officer in Minneapolis, Minnesota. His death led to US-wide protest that would eventually cross borders, spanning multiple continents. The transnational diffusion of protest and movement frames is a recognized phenomenon in the study of contention. Less well documented are the consequences of such protest for political awareness—absent survey data, this is difficult to measure. One alternative is to use Internet search data. This visualization uses Google Trends to document variation in search interest for “racism” across 101 countries and 32 languages during the period before and after George Floyd’s killing. In the majority of countries, we see significantly increased search interest that is sustained for at least two weeks. The findings suggest transnational protest diffusion may lead to global increases in political awareness.

Keywords— Racial attitudes; Political awareness; Digital trace; Heatmap; Cross-national time series

On May 25, 2020, George Floyd, a forty-six year-old African American man, was killed by a white police officer in Minneapolis, Minnesota. Harrowing video footage capturing his death circulated widely on social and broadcast media, prompting demonstrations across the United States [2]. Protest coalesced under the banner of Black Lives Matter (BLM), a movement founded in 2013 to fight racial injustice and police brutality in the US. But protest was not confined to the US or the BLM movement. Floyd’s death came to stand in for ongoing racial injustice worldwide, and protests would eventually span multiple continents.

Analysis to date has focused on protest incidence and social media activity related to George Floyd and BLM [6, 5]. But did the protests lead to wider political learning or interest in issues

connected to racism? Survey evidence suggests protest and protracted movement activity can lead to enhanced political awareness [1, 3]. Aside from surveys, another way of measuring generalized interest in topics of political interest is to use data derived from internet search engines [9]. These data are readily accessible and provide a way to gauge interest across multiple languages and countries over time.

This data visualization uses data from Google Trends. The data was collected by iteratively searching through country keyword combinations for different translations of the search term “racism.” For this, I used the Python package “pytrends” [8] alongside Google Sheets to translate the word “racism” into target languages.¹ Country keyword (“racism”) combinations were taken from an initial list of 244 country-keyword pairs. Data was ultimately accessible for 118 pairs, comprising 101 countries and 32 languages. Google does not provide raw search numbers; instead, search interest data is normalized before release. The normalization procedure involves taking the number of searches for the keyword in question denominated by total searches. This is then indexed on 1-100 scale and normalized relative to the location and date range specified (here, 2020-01-01 to 2020-08-22 for 101 countries). Thus, a value of 100 represents the day, for the specified country-keyword combination, on which search intensity was at its peak, relative to other days in the same country.
²

The visualization demonstrates that Floyd’s killing led to a pronounced uptick of interest in racism, and this is true across the overwhelming majority of country-keyword pairs. Individual tiles show daily search interest for each country on the normalized 1-100 scale, where darker colours indicate increased search intensity. To determine the effective duration of increased search interest after Floyd’s killing, I also visualize—in the upper time series—the duration of time after May 25 when (standardized) mean daily search interest exceeded overall mean search interest, as well as one and two standard deviations from the overall mean. This shows that increased search interest effectively lasted around four weeks, peaking over the period June 1 to June 4.³

¹See Appendix for full details of method.

²For further information, see <https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8>.

³Search interest is standardized by subtracting the mean and dividing by the standard deviation. Mean daily search interest is then calculated as the sum of daily means for each country-keyword pair denominated by the total number of country-keyword pairs.

Existing scholarship points to the potential for movements to diffuse across borders and achieve transnational resonance [10]. This data visualization suggests that episodes of mass unrest may lead to enhanced interest in, and awareness of, issues of injustice globally. Future scholarship should investigate the attitudinal and political consequences of such movement-led shifts in issue awareness.

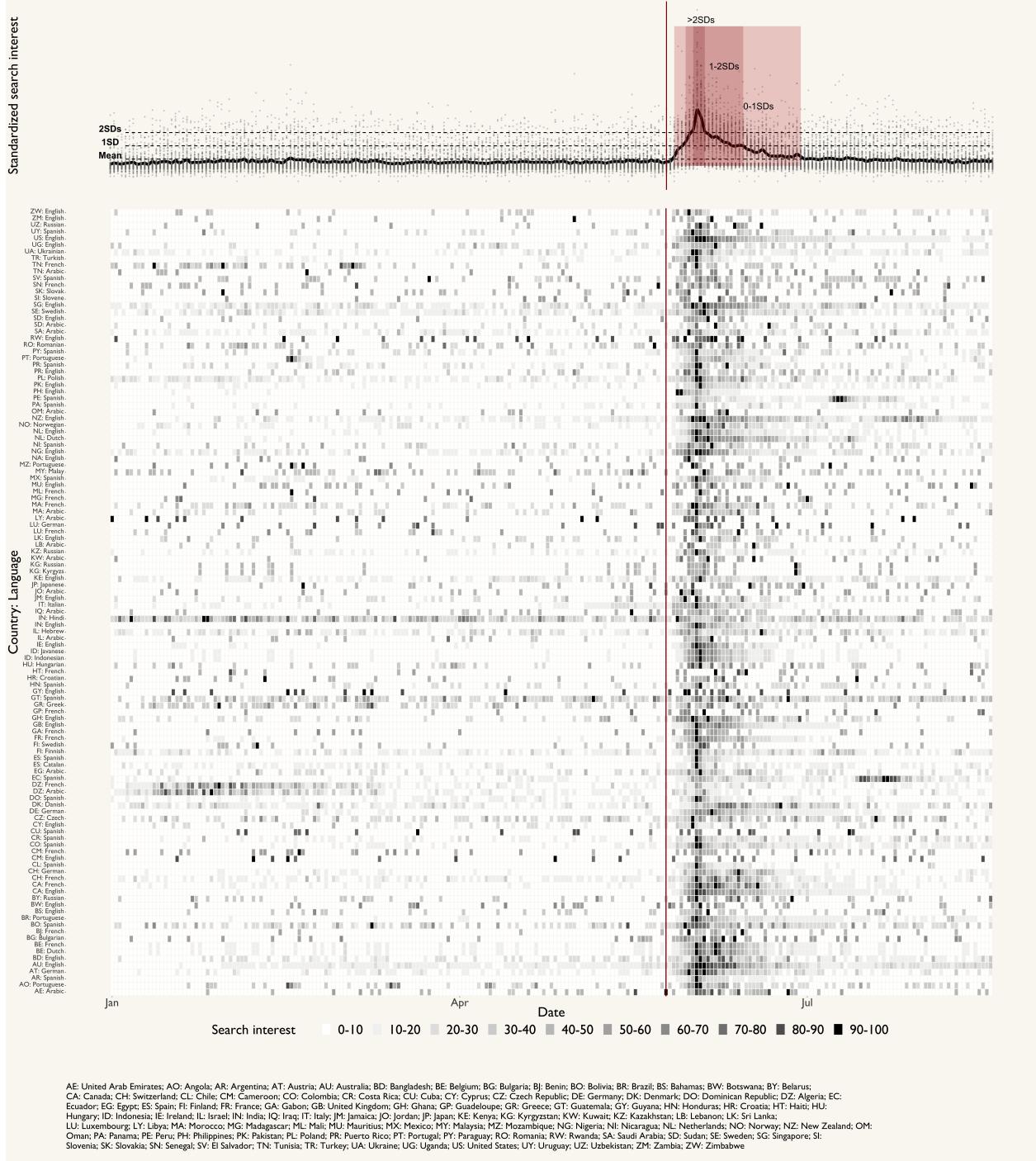


Figure 1: Top: Time series of mean standardized search interest averaged across all countries. Red shaded areas show duration of increased interest on standard deviation scale. Bottom: Tiled heatmap of cross-national search interest in racism, 2020-01-01—2929-08-18. Vertical red line indicates date of Floyd's killing.

References

- [1] R. Branton, V. Martinez-Ebers, T. E. Carey, and T. Matsubayashi. Social Protest and Policy Attitudes: The Case of the 2006 Immigrant Rallies. *American Journal of Political Science*, 59(2):390–402, Feb. 2015.
- [2] L. Buchanan, Q. Bui, and J. K. Patel. Black Lives Matter May Be the Largest Movement in U.S. History. *The New York Times*, July 2020. ISSN 0362-4331.
- [3] J. J. Castle, S. Jenkins, C. D. Ortbals, L. Poloni-Staudinger, and J. C. Strachan. The Effect of the #MeToo Movement on Political Engagement and Ambition in 2018. *Political Research Quarterly*, pages 1–16, June 2020.
- [4] CIA. CIA - The World Factbook. Technical report, Central Intelligence Agency, 2020.
- [5] T. Economist. George Floyd is remembered around the world. *The Economist*, 2020. ISSN 0013-0613.
- [6] M. Haddad. Mapping anti-racism solidarity protests around the world. *Al-Jazeera*, 2020.
- [7] J. Hogue and B. DeWilde. Pytrends: Unofficial API for Google Trends. Technical report, Python library version 4.7.3, 2020.
- [8] P. Massicotte and D. Eddelbuettel. gtrendsR: Perform and Display Google Trends Queries. Technical report, R Package Version 1.4.6., 2020.
- [9] L. T. P. Nghiem, S. K. Papworth, F. K. S. Lim, and L. R. Carrasco. Analysis of the Capacity of Google Trends to Measure Interest in Conservation Topics and the Role of Online News. *PLOS ONE*, 11(3):1–12, Mar. 2016. ISSN 1932-6203.
- [10] S. Tarrow and D. McAdam. Scale Shift in Transnational Contention. In D. Della Porta and S. Tarrow, editors, *Transnational Protest and Global Activism*. Rowman & Littlefield Publishers, Oxford, July 2004. ISBN 978-1-4616-6672-1.

A Supplementary Methods Appendix

A.1 Data sources and preparation

The sole data source for online search data was Google Trends. To arrive at a list of country-keyword pairs, I first took a list of country code top-level domains (ccTLDs) for the Google search engine and their associated main search languages available here <https://preview.tinyurl.com/ccTLDsgoogle>. I checked this list against details of main languages spoken by country from the CIA World Factbook (see CIA 4). As languages, I retained those spoken by a significant proportion of the population (defined as at least 30%) or when the language was used as a de facto lingua franca for administrative and official affairs—often the case with English. When this information was not available in the CIA World Factbook, I cross-referenced through internet searches. I also supplemented the list of ccTLDs and associated languages with several languages that I determined should be included based on these criteria but were missing. These were: German for Belgium; French for Cambodia; English for Cameroon; Turkish for Cyprus; Tigrinya for Eritrea; French for Laos; English for Bangladesh. Several countries were also missing from this list and were therefore added, including: Albania; Eritrea; Eswatini; Myanmar. This resulted in an initial set of 244 country-keyword pairs.

I then translated the word “racism” into each of the 109 languages included in the list and paired these with their associated country. This was achieved by using Google Sheets. Google Sheets allows the user to translate into a target language by specifying the origin word and target language with the associated two-letter language code (see Figure A.1).

A.2 Google Trends data extraction

To obtain search interest data, I used the “pytrends” library in the Python programming language [7]. The “pytrends” library interfaces with Google Trends, allowing the user to specify country-keyword pairings and download search interest for a specified time period for each country. I iteratively specified each of the country-keyword pairs as parameters in a query, thereby obtaining per-country search interest for the date range 2020-01-01 to 2020-08-18. I also tested the “gtrendsR” package in the R programming language to obtain search interest data [8]. I found pytrends to be more stable; when using “gtrendsR” the principle function used to grab search interest data—`gtrends()`—often failed with certain country-keyword pairs that returned data with pytrends. The full list of country-keyword pairs with the translation generated can be anonymously accessed at: https://docs.google.com/spreadsheets/d/1WTnvNDLaf5Y0za3xm1P254V-mEy5u10zJe28C_MQG1g/edit?usp=sharing. Only a subset of country-keyword pairs returns data. For others, insufficient data is available, meaning that the query returns no data. I then manually verified that there was insufficient data for the country-keyword pair by using the Google Trends “Explore” web interface. Finally, I retain only countries with at least twenty days of non-zero values.

A screenshot of a Google Sheets document titled "translate". The menu bar includes File, Edit, View, Insert, Format, Data, and a dropdown. Below the menu is a toolbar with icons for back, forward, print, and zoom (100%). The formula bar shows the formula =GOOGLETRANSLATE("racism", "en", B2). The main content is a table with three columns: A, B, and C. Column A lists words in various languages, column B lists their corresponding codes, and column C lists the language names. Row 2 is selected, showing "rassisme" in A2, "af" in B2, and "Afrikaans" in C2. The table has 10 rows, indexed from 1 to 10.

	A	B	C
1	racism	en	English
2	rassisme	af	Afrikaans
3	racizëm	sq	Albanian
4	ክራሲዝ	am	Amharic
5	عنصرية	ar	Arabic
6	ռասիզմ	hy	Armenian
7	ırqçılık	az	Azerbaijani
8	arrazakeria	eu	Basque
9	расизм	be	Belarusian
10	স্বাজাতিকতা	bn	Bengali

Figure A.1: Screenshot from Google Sheets translations of “racism.”