

Project proposal : Behavioural Consequences of Group Threat – Ethnic Derogation on YouTube in the Aftermath of Threatening Events

In order to conduct a successful project during the summer school, we propose several modular projects with a more or less overarching thematic structure. Our reasoning for doing so is that modular projects are well-suited to deal with the uncertainty of who applies to the summer school. That way, we are flexible enough to react to the individual strengths and preferences that junior researchers bring into the project group.

In effect, our project proposal includes direct extensions or follow-up research questions based on the ethnic abusive commenting in YouTube project which was shortly summarized in the “Motivation” section. In short, the main finding was that empirical patterns suggesting social contagion of ethnic abusive swearing are primarily a result of social selection effects. **The first project module** aims to test the generalizability of the general approach, research design as well as its findings by adopting a cross-nationally comparative research design. Germany is certainly not the only European country that faced exceptionally high immigration rates and terrorist attacks/sexual assaults. The situation in Belgium, Finland, France, Great Britain, Italy, Spain and Sweden was similar though varying in intensity both with respect to the rate of immigration and the exposure to violence. Depending on the language competences of the junior researchers (although native-speaking abilities are probably essential), we plan to replicate the analyses in one of the aforementioned countries. To do so, we would have to generate dictionaries of swear words and ethnic markers, construct a terror/sexual assault database, identify those national political talk shows with the widest audience reach and scrape the relevant comment data from YouTube. A replication provides insights into the extent to which the results for Germany may have been driven by it being a unique context – more than half of all refugees fleeing to Europe settled in Germany and Germany experienced the largest number of individual violent attacks/assaults. As argued by Kaufmann (2014), sudden increases in out-group size are related to more negative sentiment towards immigrants. Moreover, the few studies on the effect of terrorist attacks or sexual assaults on attitudes towards minorities also point towards a negative shift in those attitudes (Legewie 2013; Czymara & Schmidt-Catran 2017). Combining large immigrant influx with increased terrorist activity may thus have generated a specific climate in Germany that is highly conducive to voicing and the spread of ethnic abusive swearing. Replicating this study in any of the countries mentioned above would enable us to essentially vary immigration rate and/or terrorist activity in order to explore the robustness of the findings in the German context.

Apart from assessing the robustness of national findings, a cross-national extension would enable us to focus more thoroughly on the contribution of social selection mechanisms to the use and spread of ethnic abusive swearing. To do so, **the second project module** would explore the role that political mobilization plays in driving social selection effects in online commenting. In the German case, there was a massive increase in ethnic abusive commenting in YouTube videos of political talk shows leading up to the 2017 nation election which seemed to be associated with mobilization for the right-wing populist party AfD (Note: anecdotal evidence from skimming comment data). This module will thus integrate a party political perspective by investigating whether the increased frequency (and possibly intensity) of ethnic abusive swearing can be linked to increasing references of political parties and their programs in the lead-up to national elections or referendums. During the 2015-2018 time frame, several important elections or referendums qualify as test cases: the Spanish elections and Brexit in 2016; Austrian, French and German elections in 2017 or the Swedish elections in 2018. In many if not all of these elections, one or more sizable parties campaigned in opposition to immigration thereby allowing us to vary important drivers of social selection effects in political mobilization such

as party member count or party incumbency status. Another possibility would be to use the Twitter API or scrape news pages to gather information on party activities, like controversial statements by politicians concerning ethnic groups, and link this to activities in comment networks by the date the comment was posted. Such an approach could shed light on the claim that anti-immigrant rhetoric from politicians might strengthen anti-immigrant movements. Note that the general advantage of a cross-national framework, namely exploiting cross-country variation in societal conditions such as changes in immigration rates also benefits this project module. Data sources for this project module can again be comment data from YouTube videos of influential national political talk shows.

The third project module still focuses on exploring the social selection mechanisms but abandons the implicit assumption that comment data from YouTube videos can be treated as a fairly closed system. In reality, social selection could well be driven by individuals sharing YouTube videos of political talk shows on other social media platforms such as Facebook, twitter or reddit and encouraging network members implicitly or explicitly to react to the content in YouTube comment section. As stated above, much – but not all - of the relationship between previous' comments content and ego's likelihood of using ethnic abusive swearing can be explained by social selection. The main idea for this project module is to investigate two research questions: first, does cross-platform sharing lead to an increase in commenting volume? This research question would thus be concerned with exploring the contribution of other social media platforms to social selection effects. The second research questions asks whether accompanying message to the shared YouTube videos influence the content of YouTube comments over and above the comment content already populating YouTube videos. For example, consider the mobilizing potential of two (fictional) tweets referencing the video of a political talk show on YouTube: tweet one may simply state "Interesting discussion of the likely impact refugees have on social security systems: LINK TO VIDEO". A second tweet may phrase it like this: "Proof that refugees will suck our social security systems dry: LINK TO VIDEO". Our guiding hypotheses would be that the second tweet will elicit a much stronger mobilizing impetus while also priming for much more negative commenting behaviour due to the implied material threat that refugees pose.

This project has societal relevance in times when social media is portrayed to be a main driver of societal polarization and pleas for action against hate speech are getting more and more pronounced (see for example the European Commission's #Noplace4hate campaign). Should we replicate the results that the group of online swearers have certain political motives *before* the actual societal debate and that they are therefore not "triggered" into ethnic abusive swearing by other users; a large part of the proposed means that are currently discussed to lower the amount of hate speech might lead astray with potentially detrimental consequences to free speech in general. For example, the deletion (i.e., de-facto censoring) of online comments is unsatisfactory as it will not change the opinions of those who swear heavily in social networks and the comments themselves do not trigger more hate speech.

We feel that these projects provide a great fit with the overarching thematic as well as methodological goal of the BIGSSS Summer School in Computational Social Science: they provide a strong potential to advance our understanding of intergroup relations in online communication while at the same time remaining manageable enough to successfully complete during the comparatively short duration of the summer school.

Lecture or hands-on workshop portfolio

Again, depending on the composition of the expert and junior research and thus on the prevailing preferences, we are open to several possibilities.

First, we would be happy to present first results of a current project studying **immigrant selectivity** in Western European societies. Selectivity refers to the classic idea, that immigrants do not represent a random sample of the population at origin (Lee 1966). Instead, immigrants tend to be selective in a variety of dimensions: they tend to be younger, healthier, more often male, more ambitious and portray distinct educational profiles. The lecture would cover descriptive patterns for educational selectivity over origin groups, destination countries and time as well as provide in-depth empirical analyses of the relationship between educational selectivity and core dimensions of immigrant integration (such as language acquisition, labour market outcomes or host-society identification) relying on published and unpublished work (see https://chspoerlein.shinyapps.io/select_plot/ for an interactive description of educational selectivity of immigrants in German). Alternatively, we offer thematic lectures about the extent and reasons of ethnic residential segregation in Germany and the impact of mass media or terrorist attacks on attitudes towards immigration.

Second, we can also offer one of three hands-on workshops: one on scraping of YouTube, twitter and reddit comments and comment networks or one on text mining – both using R. **Scraping comments** introduces participants to registering with APIs, connecting to the APIs using R, saving data, visualizing networks and extracting network information as well as the basics of working with textual data (manipulating strings, regular expressions, etc.) The **text mining** work-shop aims to give a much deeper introduction to work with text as data. The first part will deal with the basics of working with textual data, placing a stronger focus on more advanced methods such as generating text corpuses and document-term matrices, tokenizing text, stemming words, stop word removal compared to the workshop of scraping comments. The second part will give an overview over common analysis tools for textual data: sentiment analysis, word frequencies, n-grams, topic models and – time permitting – graph-based text analysis which combines network analysis and graph theory to describe the relationship between words. Third, we would be happy to offer a hands-on workshop on **longitudinal causal analysis**. This course makes use of causal diagrams (Directed Acyclic Graphs) to formalize theoretical causal models (a model to capture our assumption about the world) to motivate certain statistical methods when the researcher's aim is to estimate causal effects. One important section of this course discusses which causal model legitimates the use of either Fixed-Effects estimators or other methods for the estimations of effects of time-varying treatments (like Inverse Probability of Treatment Weighting).

Literature

- Czymara, C.S. and Dochow, S. 2018. Mass Media and Concerns about Immigration in Germany in the 21st Century: Individual-Level Evidence over 15 Years. *European Sociological Review* 34:381-401.
- Czymara, C.S. and Schmidt-Catran, A.W. 2017. Refugees Unwelcome? Changes in the Public Acceptance of Immigrants and Refugees in Germany in the Course of Europe's "Immigration Crisis". *European Sociological Review* 33:556-577.
- Fryer, R.G. 2007. Guess who's been coming to dinner? Trends in interracial marriage over the 20th century. *Journal of Economic Perspectives* 21:71-90.
- Gordon, M.M. 1964. *Assimilation in American Life: The Role of Race, Religion, and National Origin*. Oxford: Oxford University Press.
- Kaufmann, E. 2014. "It's the Demography, Stupid": Ethnic Change and Opposition to Immigration. *The Political Quarterly* 85:267-276.
- Lee, E.S. 1966. A Theory of Migration. *Demography* 3:47-57.
- Legewie, J. 2013. Terrorist Events and Attitudes towards Immigrants: A Natural Experiment. *American Sociological Review* 118:1199-1245.
- Schlueter, E. and Scheepers, P. 2010. The relationship between outgroup size and anti-outgroup attitudes: A theoretical synthesis and empirical test of group threat and intergroup contact theory. *Social Science Research* 39:285-295.
- Spörlein, C. 2015. The Destination Choices of Recent Pan-American Migrants: Opportunities, Costs and Migrant Selectivity. *International Migration Review* 49:523-552.
- Spörlein, C. 2018. How educational systems structure ethnic inequality among young labour market participants in Europe: occupational placement and variation in the occupational status distribution. *Research in Social Stratification and Mobility* 55:109-119.
- Spörlein, C. & Kristen, C. 2018. Educational Selectivity and Language Acquisition of recently arrived Immigrants in Germany, Great Britain and Ireland. *International Migration Review*.
<https://doi.org/10.1177/0197918318798343>.
- Spörlein, C. & Kristen, C. Unpublished. Why We Should Care about Regional Origins: Educational Selectivity among Refugees and Labor Migrants in Western Europe. Submitted at *Frontiers in Sociology*.
- Spörlein, C. & Schlueter, C. 2018a. How educational systems shape cross-national ethnic inequality in math competence scores: moving beyond mean differences. *PLOS ONE* 13:e0193738.
- Spörlein, C. & Schlueter, C. 2018b. How to best examine segregation: A statistical and conceptual multilevel approach. *Methods, data, analyses* 12:211-232.
- Spörlein, C. & van Tubergen, F. 2014. The Occupational Status of Immigrants in Western and non-Western Societies. *International Journal of Comparative Sociology* 55:119-143.
- Spörlein, C., Schlueter, E. & van Tubergen, F. 2014. Ethnic Intermarriage in Longitudinal Perspective: Testing Structural and Cultural Explanations in the United States from 1880 to 2011. *Social Science Research* 43:1-15.