

Extracting and Analyzing User Comments to Expose How Americans Respond to Gun Violence: A Computational Social Science Study by Erika Fox '21

Introduction/Significance:

Framing is the process of selecting some aspects of reality and making them salient, while directing attention away from other aspects. This process has implications for audience members' understanding of and opinions about the world. This framing process is present in the American press system; therefore, it is important to study framing in news articles in order to understand if or how journalists steer their audience members' opinions on certain (often political) topics. One of the topics that is heavily framed by the American press is mass shootings. I discovered this in the Fall 2018 semester when I conducted a content analysis on a sample of 50+ New York Times articles that covered the Las Vegas Massacre that occurred on October 1st, 2017 as a part of my Media and Politics class with Professor Simmons. The frames in that sample of articles revealed a wide array of perspectives within the New York Times platform. For example, I discovered a divide in how journalists attributed the responsibility for the massacre: some articles solely hold the gunman Stephen Paddock responsible for the killing of 58 people, while others put more emphasis on the idea that mass shootings such as Vegas are a problem in the United States because of the country's lackadaisical gun laws. I speculated about potential implications of these frames as a part of that analysis, but I have not yet had the opportunity until now to see how many of them actually hold, and to what extent. Luckily, the New York Times website allows its audience members to comment directly on its articles, which essentially provides a sea of data about how people feel about what each article is saying. In this summer's project, I will use computational social science methods to automatically extract and analyze these comments and find out how they differ across articles with different frame profiles

for two types of shootings: The Las Vegas Massacre (10/1/2017) and the Parkland High School Shooting (2/14/2018). This study aims to expose how framing impacts public opinion as evident in the comments on online news resources.

Mentors: Professor Simmons and Professor Hay

Dates: June 10th, 2019- August 2nd, 2019

This study is designed to occur over ten weeks under the mentorship of Professor Simmons.

However, since this project spans both sociology and computer science concepts, Professor Hay will serve as a secondary mentor for the computer science aspects of the project.

Research Question: How are the implications of news framing reflected in the comments of New York Times articles on the Las Vegas Massacre and the Parkland High School Shooting?

The above is the main question of this project, but here it is broken down:

What is the nature of users' comments about online news articles about mass shootings?

- What topics are discussed? (e.g. victims, gun control, the suspect)
- What is the tone? (e.g. anger, calls to action, mourning)

How does the population of comments vary across case characteristics?

- Do the topics/tone of topic vary depending on the shooting?
 - Las Vegas 10/2/2017 vs Parkland 2/14/2018
 - Country Music Festival vs. High School
 - Stephen Paddock vs. Nikolas Cruz
 - Age of Victims- older concert goers vs. young students

Do comments by individuals vary by article?

Tasks (& approximate timeline):

1. Literature Review (Week 1: 6/10-6/14)
 - a. The study will begin with an analysis of any existing literature that should inform the study.
 - i. What has been written about online vocality?
 - ii. What has been written about the implications of news framing?
 - iii. What has been written about how people react to mass shootings specifically?
 - iv. What has been written about the use of NLP (Natural Language Processing) for similar causes?
2. Condensed analysis of frames in the coverage of the Las Vegas Massacre and Parkland Shooting (Weeks 2 and 3: 6/17-6/28)
 - a. New York Times content analysis of frames: In order to properly understand the potential implications of the framing in the news articles that I am looking for in the comments, I will need to first break down the most prominent frames in articles from the days following both the Las Vegas Massacre of 10/1/2017 and the Parkland High School Shooting of 2/14/2018.
 - i. Resources: New York Times, Excel, MAXQA
3. Extracting data (Weeks 4-6: 7/1-7/19)
 - a. This is the stage in which I will develop an algorithm to extract the comments from the comment sections of the articles in my sample.

- i. Resources: Java Programming Language, Git Hub, Professor Hay and various computer science papers on NLP. Towards the end of this stage I will be able to begin analyzing the data that I am collecting.

4. Analysis of data (Weeks 6-7: 7/15-7/26)

- a. This is the stage in which I will evaluate the success of my algorithm and interpret its results.

5. Research Paper (Weeks 1-8)

- a. Over the course of this 8-weeks study I will be compiling my findings into one research paper, with a section dedicated to all of the other components of the project: Literature Review, Analysis of Frames, Extracting data (Methods), Analysis of data, interpretation of the data and a conclusion. Even though I will be working on this paper over the entire project duration, the 8th week will be dedicated only to the completion of this document.