Does Disinformation Influence Election Outcomes?

To get a sense of the literature, I collected data from Google Scholar on peer-reviewed articles with “disinformation” in the title, 1980-present. The query and data collection process was done programmatically with R, aka data scraping. The query was set to return only articles, not patents or citations. Data scraping is an imperfect process. The fewer the pieces of information being sought in a query, the more efficient is the process. In the chart below, you can see how many pieces of information I my query sought by the number of variable. 27% of my table has missing values (see [Disinformation Tables.xlsm](https://usbe-my.sharepoint.com/:x:/g/personal/jared_walker_schools_utah_gov/EdZ0MGwbwnpIiB9cKKP0lLwBVpA7sBw3YNtp3lDoMN30Aw?e=JLzONj)). A superior approach would seek two pieces of information—citations and article title.

Figure 1: Missingness Visualizations

|  |
| --- |
| Chart, bar chart  Description automatically generated |
| Chart  Description automatically generated |

A

Chart, line chart

Description automatically generatedOver the last six years, researchers have shown an increasing interest in disinformation—what it is, the risk it poses to democratic institutions, and how to manage the problem. “Disinformation” is one of a constellation of terms including “fake news,” “misinformation,” “propaganda,” etc. which some scholars are referring to as “problematic information” (Freelon and Wells, 2020). As a burgeoning body of literature, there is little consensus on a conceptual framework for studying disinformation. But there is broad agreement on the definition of “disinformation.” According to the European Union’s High Level Expert Group on Fake News and Online Disinformation, **“Disinformation . . . includes all forms of false, inaccurate, or misleading information designed, presented, and promoted to intentionally cause public harm or for proﬁt” (High Level Expert Group, 2018, p.** 3). There are three critical components: 1) deception, 2) potential for harm, and 3) an intent to harm. Other deceptive messages which lack the potential and intent to harm constitute *misinformation* (Freelon and Wells, 2020)

After collecting data from Google Scholar, I used a function in R to get a frequency count of each word in the titles of articles I collected. Table 1 can be interpreted as a list of the most common themes in the disinformation literature. A superior analysis would be to map a semantic network based on common words (Vale Cunha, et.al. 2020)

As a body of literature, disinformation research is getting to know itself. Just in the last four years, six systematic literature reviews on disinformation have been published. Another five develop taxonomies for the study of disinformation.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Most Common Words in Article Titles   |  |  |  | | --- | --- | --- | | **Rank** | **Word** | **Frequency** | | 1 | *disinformation* | 2329 | | 2 | *fake news* | 212 | | 3 | *COVID-19* | 173 | | 4 | *information* | 167 | | 5 | *social media* | 164 | | 6 | *misinformation* | 152 | | 7 | *campaign(s)* | 120 | | 8 | *Russian* | 113 | | 9 | *election(s)* | 91 | | 10 | *propaganda* | 81 | | 11 | *twitter* | 62 | | 12 | *democracy* | 55 | | 13 | *pandemic* | 55 | | 14 | *2016* | 51 | | 15 | *research* | 35 | |

Just in the last four years, six systematic literature reviews on disinformation have been published. Another five developed taxonomies for disinformation. As Freelon and Wells point out, none of this is entirely new, but it is newly salient, most notably because of the Digital Revolution and the information age it gave rise to.

As Freelon and Wells point out, none of this is entirely new, but it is newly salient, most notably because of the Digital Revolution and the information age it gave rise to.

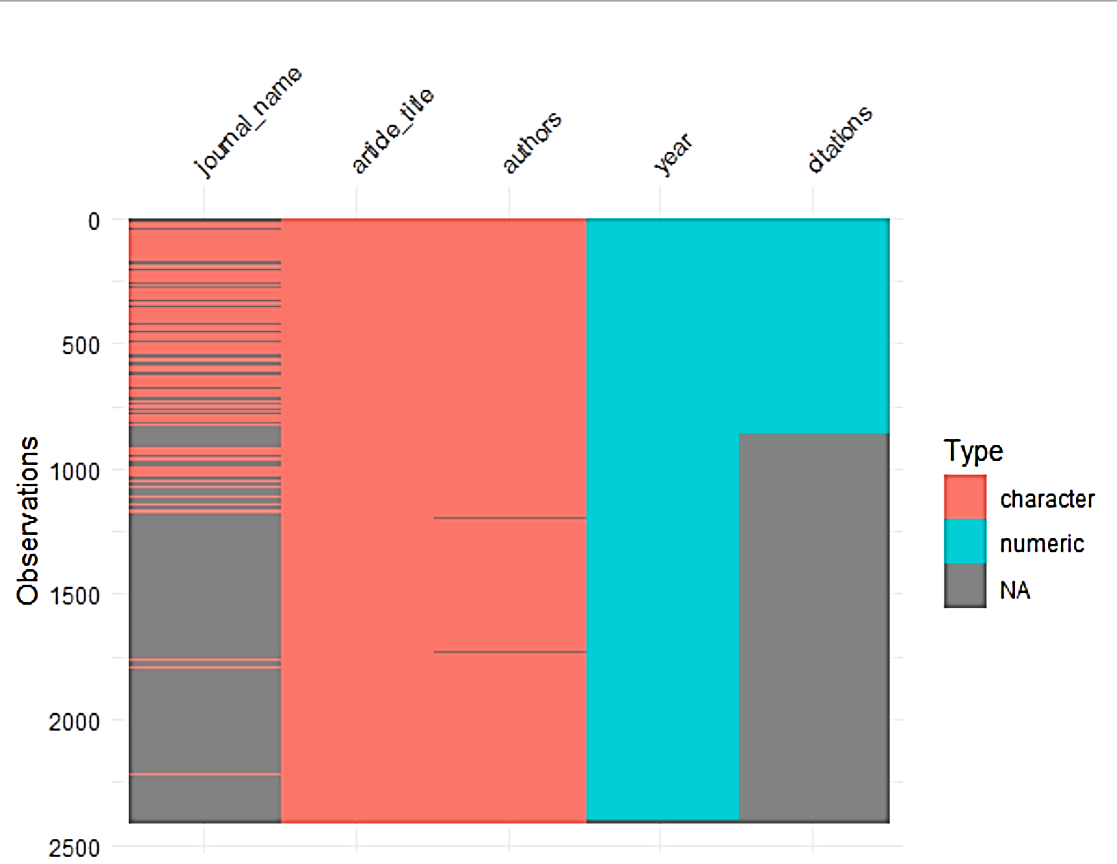
2415 papers published with ‘disinformation’ in the title since 1980. The

[Disinformation Tables.xlsm](https://usbe-my.sharepoint.com/:x:/g/personal/jared_walker_schools_utah_gov/EdZ0MGwbwnpIiB9cKKP0lLwBVpA7sBw3YNtp3lDoMN30Aw?e=JLzONj)

The query and data collection process was done programmatically with R, aka web scraping. The query was set to return only articles, not patents or citations. Web scraping is an imperfect process. mperfect process. The more information, the more problems.

2415 papers published with ‘disinformation’ in the title since 1980. The

If I did that from the start, data would have less missingness



|  |
| --- |
|  |
|  |
|  |
|  |
|  |

vvvvv

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | **Rank** | **Word** | **Frequency** | | 1 | *disinformation* | 2329 | | 2 | *fake news* | 212 | | 3 | *COVID-19* | 173 | | 4 | *information* | 167 | | 5 | *social media* | 164 | | 8 | *misinformation* | 152 | | 9 | *digital* | 137 | | 10 | *campaign(s)* | 120 | | 11 | *Russian* | 113 | | 12 | *election(s)* | 91 | | 13 | *political* | 81 | | 14 | *propaganda* | 81 | | 15 | *European* | 66 | | 16 | *twitter* | 62 | | 17 | *democracy* | 55 | | 18 | *pandemic* | 55 | | 19 | *2016* | 51 | | 20 | *war* | 51 | | 21 | *policy* | 47 | | 22 | *EU* | 43 | | 23 | *literacy* | 43 | | 24 | *warfare* | 41 | | 25 | *countering* | 40 | | 26 | *impact* | 39 | | 27 | *detection* | 38 | | 28 | *era* | 38 | | 29 | *fight* | 35 | | 30 | *research* | 35 | |  |  |  | |  |  |  | | | **12075**  **>**  **> sum(is.na(df))/sum(is.na(df))+sum(!is.na(df))**  **[1] 8811**  **>**  **>**  **>**  **> sum(is.na(df))/(sum(is.na(df))+sum(!is.na(df)))**  **[1] 0.2703934**  **> sum(!is.na(df))/(sum(is.na(df))+sum(is.na(df)))**  **[1] 1.349158**  **> sum(!is.na(df))/(sum(is.na(df))+sum(!is.na(df)))**  **[1] 0.7296066** |
|  |  | |
| |  |  |  | | --- | --- | --- | | **Rank** | **Word** | **Frequency** | | 1 | *disinformation* | 2329 | | 2 | *fake news* | 212 | | 3 | *COVID-19* | 173 | | 4 | *information* | 167 | | 5 | *social media* | 164 | | 6 | *media* | 157 | | 7 | *online* | 153 | | 8 | *misinformation* | 152 | | 9 | *digital* | 137 | | 10 | *campaign(s)* | 120 | | 11 | *Russian* | 113 | | 12 | *election(s)* | 91 | | 13 | *political* | 81 | | 14 | *propaganda* | 81 | | 15 | *European* | 66 | | 16 | *twitter* | 62 | | 17 | *democracy* | 55 | | 18 | *pandemic* | 55 | | 19 | *2016* | 51 | | 20 | *war* | 51 | | 21 | *policy* | 47 | | 22 | *EU* | 43 | | 23 | *literacy* | 43 | | 24 | *warfare* | 41 | | 25 | *countering* | 40 | | 26 | *impact* | 39 | | 27 | *detection* | 38 | | 28 | *era* | 38 | | 29 | *fight* | 35 | | 30 | *research* | 35 | | |  |

|  |  |
| --- | --- |
|  |  |
|  |  |
| Chart, bar chart  Description automatically generated |  |

|  |  |
| --- | --- |
| ITEM | NEEDED |
| Books | 1 |
| Magazines | 3 |
| Notebooks | 1 |
| Paper pads | 1 |
| Pens | 3 |
| Pencils | 2 |
| Highlighter | 2 colors |
| Scissors | 1 pair |

Binary threshold research disconitntiy design

|  |  |  |  |
| --- | --- | --- | --- |
| **College** | **New students** | **Graduating students** | **Change** |
|  | *Undergraduate* |  |  |
| Primary Candidate |  | 103 | 7 |
| Disinformation Statement(s) | Dummy variable | 214 | 9 |
| Disinformation Statement(s) | Count | 120 | 77 |
| Pine College | 134 | 121 | 13 |
| Fraction of Vote | Continuous |  |  |
|  |  |  |  |
|  |  |  |  |
| Oak Institute | 202 | 210 | -8 |
|  | *Graduate* |  |  |
| Cedar University | 24 | 20 | 4 |
| Elm College | 43 | 53 | -10 |
| Maple Academy | 3 | 11 | -8 |
| Pine College | 9 | 4 | 5 |
| Oak Institute | 53 | 52 | 1 |
| **Total** | **998** | **908** | **90** |

Enrollment in local colleges, 2005

|  |  |  |  |
| --- | --- | --- | --- |
| College | New students | Graduating students | Change |
|  | Undergraduate |  |  |
| Cedar University | 110 | 103 | +7 |
| Elm College | 223 | 214 | +9 |
| Maple Academy | 197 | 120 | +77 |
| Pine College | 134 | 121 | +13 |
| Oak Institute | 202 | 210 | -8 |
|  | Graduate |  |  |
| Cedar University | 24 | 20 | +4 |
| Elm College | 43 | 53 | -10 |
| Maple Academy | 3 | 11 | -8 |
| Pine College | 9 | 4 | +5 |
| Oak Institute | 53 | 52 | +1 |
| Total | 998 | 908 | 90 |

Source: Fictitious data, for illustration purposes only

|  |  |
| --- | --- |
|  |  |

|  |  |  |
| --- | --- | --- |
| **c** | **Word** | **Frequency** |
| 1 | *disinformation* | 2329 |
| 2 | *fake news* | 212 |
| 3 | *COVID-19* | 173 |
| 4 | *information* | 167 |
| 5 | *social media* | 164 |
| 6 | *media* | 157 |
| 7 | *online* | 153 |
| 8 | *misinformation* | 152 |
| 9 | *digital* | 137 |
| 10 | *campaign(s)* | 120 |
| 11 | *Russian* | 113 |
| 12 | *election(s)* | 91 |
| 13 | *political* | 81 |
| 14 | *propaganda* | 81 |
| 15 | *European* | 66 |
| 16 | *twitter* | 62 |
| 17 | *democracy* | 55 |
| 18 | *pandemic* | 55 |
| 19 | *2016* | 51 |
| 20 | *war* | 51 |
| 21 | *policy* | 47 |
| 22 | *EU* | 43 |
| 23 | *literacy* | 43 |
| 24 | *warfare* | 41 |
| 25 | *countering* | 40 |
| 26 | *impact* | 39 |
| 27 | *detection* | 38 |
| 28 | *era* | 38 |
| 29 | *fight* | 35 |
| 30 | *research* | 35 |

|  |  |  |
| --- | --- | --- |
| **Rank** | **Word** | **Frequency** |
| 1 | *disinformation* | 2329 |
| 2 | *fake news* | 212 |
| 3 | *COVID-19* | 173 |
| 4 | *information* | 167 |
| 5 | *social media* | 164 |
| 6 | *media* | 157 |
| 7 | *online* | 153 |
| 8 | *misinformation* | 152 |
| 9 | *digital* | 137 |
| 10 | *campaign(s)* | 120 |
| 11 | *Russian* | 113 |
| 12 | *election(s)* | 91 |
| 13 | *political* | 81 |
| 14 | *propaganda* | 81 |
| 15 | *European* | 66 |
| 16 | *twitter* | 62 |
| 17 | *democracy* | 55 |
| 18 | *pandemic* | 55 |
| 19 | *2016* | 51 |
| 20 | *war* | 51 |
| 21 | *policy* | 47 |
| 22 | *EU* | 43 |
| 23 | *literacy* | 43 |
| 24 | *warfare* | 41 |
| 25 | *countering* | 40 |
| 26 | *impact* | 39 |
| 27 | *detection* | 38 |
| 28 | *era* | 38 |
| 29 | *fight* | 35 |
| 30 | *research* | 35 |