

PolitiFact Analysis

Sofia Bliss-Carrascoas

9/28/2022

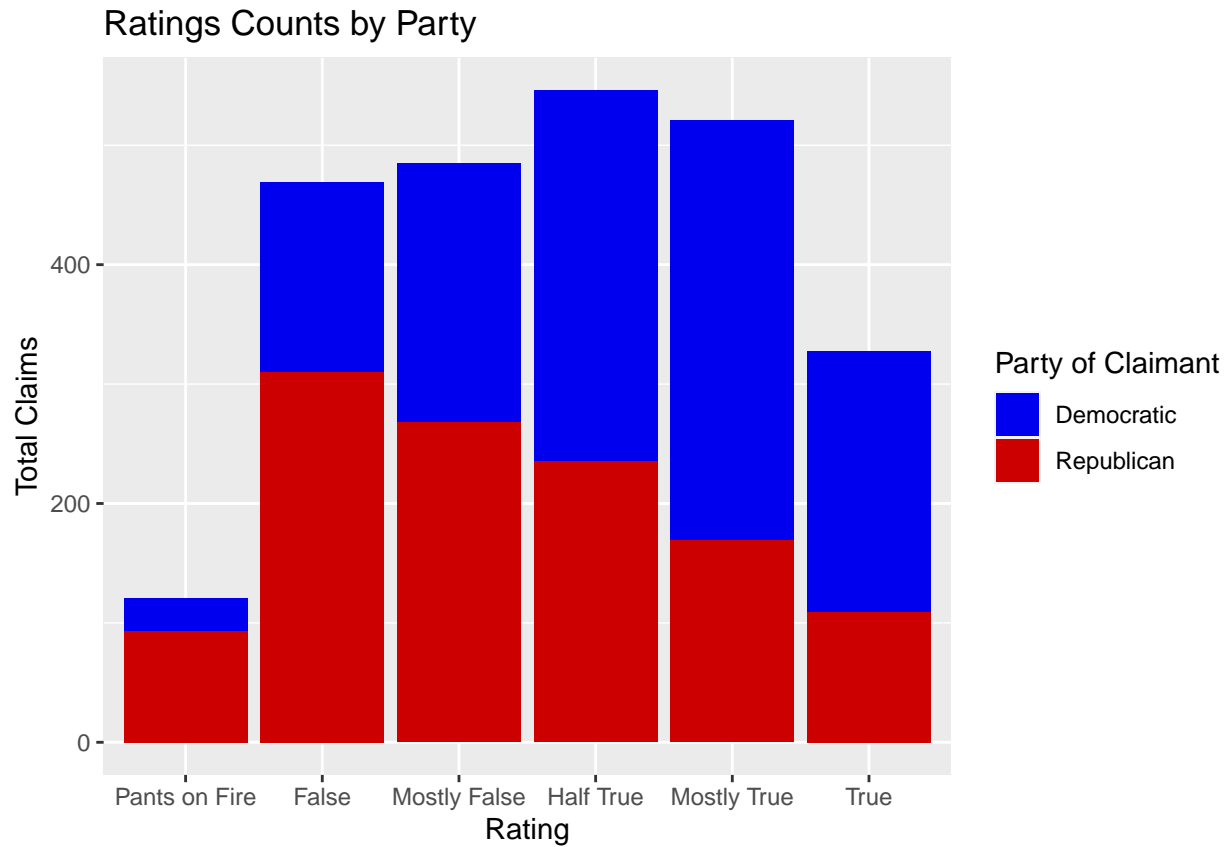
Having cleaning, sorted, and tagged all of our data, we conducted some initial data analysis to answer some of our research questions.

Which Party Lies More?

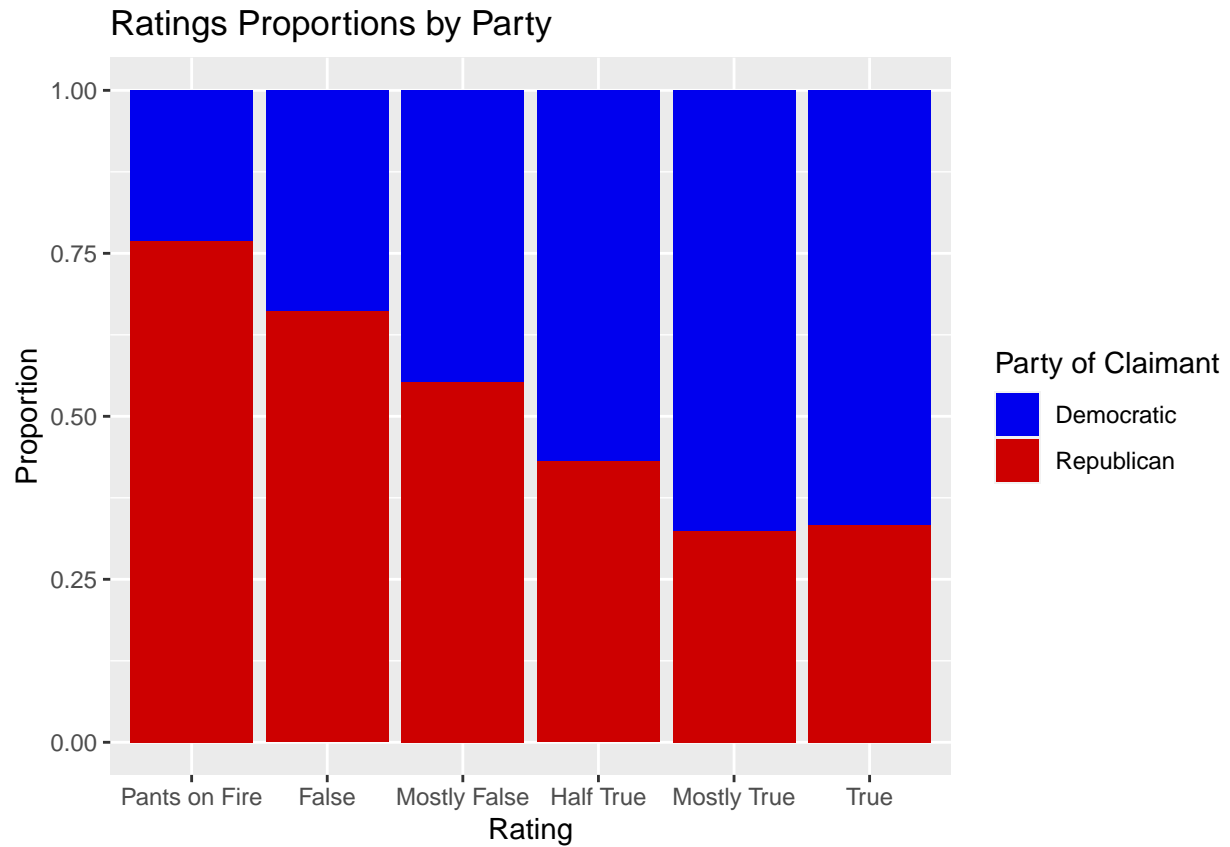
Our first research question was relatively simple. Which party lies more? While it is an impossible question to answer definitively, we decided to look at each rating tier and see the breakdown by party to get a sense of which party receives more false ratings.

Table 1: Claims by Party Sorted by Rating

| textualRating | Democratic | Republican | Total |
|---------------|------------|------------|-------|
| Pants on Fire | 28 | 93 | 121 |
| False | 159 | 310 | 469 |
| Mostly False | 217 | 268 | 485 |
| Half True | 311 | 235 | 546 |
| Mostly True | 352 | 169 | 521 |
| True | 218 | 109 | 327 |

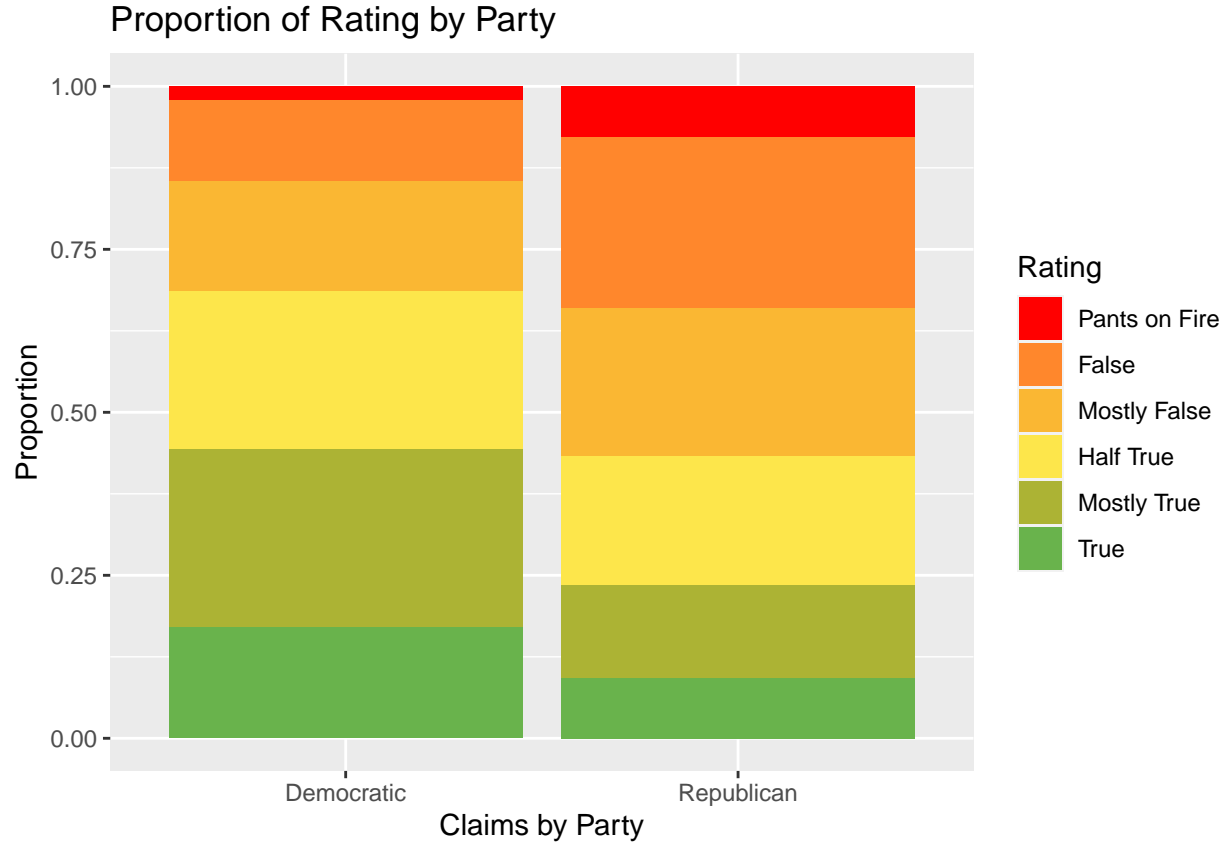


To account for the different samples sizes, we also viewed the data proportionally. Here we can see that Republican claimants are responsible for over 76.9% of the Pants on Fire claims, 66% of False claims, and 55.2% of Mostly False claims. Democrats have a majority of truthful claims responsible for 57% of the Half True claims, 67.6% of the Mostly True claims, and 66.7% of True claims.



Severity of Claims – by Party

Our second research question was, does one party lie worse. For this, we looking within each party's claims and reviewed the distribution of ratings.



Here we can see that 7.9% of all Republican claims are rated Pants on Fire as opposed to 2.2% of Democratic claims. 26.2 % and 22.6% of Republican claims are rated False and Mostly False respectively. 12.4% and 16.9% of Democratic claims are rated False and Mostly False.

Democrats also have a greater proportion of their claims rated True and Mostly True. To see a full percentage breakdown, review the table below.

Table 2: Percentage of Claims by Textual Rating

| textualRating | Democratic | Republican | Total |
|---------------|------------|------------|-------|
| Pants on Fire | 2.2 | 7.9 | 4.9 |
| False | 12.4 | 26.2 | 19.0 |
| Mostly False | 16.9 | 22.6 | 19.6 |
| Half True | 24.2 | 19.8 | 22.1 |
| Mostly True | 27.4 | 14.3 | 21.1 |
| True | 17.0 | 9.2 | 13.2 |

What do they Lie About?

Next we reviewed the distribution of our subject tags. A reminder that claims can be tagged with multiple tags each and only Mostly False, False, and Pants on Fire claims were tagged by subject.

Listed in our table below is the number of falsehoods tagged with each of the four macro tags, by party and in total. For each of tags (rows) the columns show the # of democratic claims tagged with that subject tag, the # of Republican claims tagged with that subject tag, and the total # of claims tagged with that subject tag. The third and fourth columns also shows what percentage of a given parties' falsehoods were assigned

that tag. For example 17.88 % of Republican falsehoods were tagged as Legislation. 15.59 % of Democratic falsehoods were tagged as Self/Personal Record. These percentages are valuable to compare the two parties because there is a greater number of Republican False(ish) claims overall, so reviewing the proportions helps to compare across parties.

Table 3: MacroTag Subject Counts and Percentages by Party

| MacroTag | Democratic | Republican | TotalCount | DemocraticPercentage | RepublicanPercentage |
|----------------------------|------------|------------|------------|----------------------|----------------------|
| Fear | 90 | 208 | 298 | 22.28 | 31.00 |
| Legislation | 67 | 120 | 187 | 16.58 | 17.88 |
| Self/Personal Record | 63 | 69 | 132 | 15.59 | 10.28 |
| Opponent/Opponent's Record | 153 | 287 | 440 | 37.87 | 42.77 |

We created a similar table to review each of the 32 micro subject tags. For each of tags (rows) The table shows the # of democratic claims tagged with that subject tag, the # of Republican claims tagged with that subject tag, and the total # of claims tagged with that subject tag. the table is the count of claims assigned that tag split by part and in total. Again the third and fourth columns show what percentage of a given parties' claims were assigned that tag. For example, 11.88 % of Democratic falsehoods were tagged as labor and employment compared to only 5.66% of Republican falsehoods receiving that tag.

This table has an additional column called "Percent Disparity." This column is a subtraction of DemocraticPercentage - Republican Percentage. This is the variable used to order the table. The lower the value (negative) signifies that falsehoods about that subject make up a greater proportion of Democrat lies than Republicans. The greater the value in Percent Disparity means that falsehoods about that subject make up a greater proportion of Republican lies than Democrats. This allows us to see where the parties diverge in subject.

Democrats are more likely than Republicans to have their falsehoods be about Labor and Employment, Economic Well-Being and Domestic Commerce, Education, Guns and Race.

Republicans on the other hand are more likely than Democrats to tell falsehoods that focus on Immigration and Refugee Issues, Law Crime and Policing, Voting/Elections, International Affairs and Foreign Aid, Environment, and History.

Table 4: Microtag Subject Counts and Percentages by Party

| MicroTag | Democratic | Republican | TotalCount | DemocraticPercentage | RepublicanPercentage | PercentDisparity |
|--|------------|------------|------------|----------------------|----------------------|------------------|
| Labor and Employment | 48 | 38 | 86 | 11.88 | 5.66 | -6.22 |
| Economic Well-Being and Domestic Commerce | 48 | 41 | 89 | 11.88 | 6.11 | -5.77 |
| Education | 33 | 28 | 61 | 8.17 | 4.17 | -4.00 |
| Guns | 31 | 25 | 56 | 7.67 | 3.73 | -3.95 |
| Race | 20 | 13 | 33 | 4.95 | 1.94 | -3.01 |
| Taxes | 38 | 47 | 85 | 9.41 | 7.00 | -2.40 |
| Foreign Trade | 13 | 9 | 22 | 3.22 | 1.34 | -1.88 |
| Health (non-care) | 43 | 59 | 102 | 10.64 | 8.79 | -1.85 |
| Record/Candidate Biography/Campaigns and Personal Behavior | 181 | 295 | 476 | 44.80 | 43.96 | -0.84 |
| Women | 10 | 12 | 22 | 2.48 | 1.79 | -0.69 |
| Agriculture | 5 | 7 | 12 | 1.24 | 1.04 | -0.19 |
| Community Development and Housing | 8 | 13 | 21 | 1.98 | 1.94 | -0.04 |

| MicroTag | Democrat | Republican | Total Count | Democrat Percent | Republican Percent | Percent Disparity |
|--|----------|------------|-------------|------------------|--------------------|-------------------|
| Media and Communications | 11 | 20 | 31 | 2.72 | 2.98 | 0.26 |
| Healthcare | 32 | 55 | 87 | 7.92 | 8.20 | 0.28 |
| National and State Macroeconomic Issues | 12 | 22 | 34 | 2.97 | 3.28 | 0.31 |
| LGBTQ | 4 | 11 | 15 | 0.99 | 1.64 | 0.65 |
| Civil Rights, Minority Issues, and Civil Liberties | 18 | 35 | 53 | 4.46 | 5.22 | 0.76 |
| Transportation | 6 | 16 | 22 | 1.49 | 2.38 | 0.90 |
| Defense/Military | 21 | 41 | 62 | 5.20 | 6.11 | 0.91 |
| Technology Science Space | 9 | 24 | 33 | 2.23 | 3.58 | 1.35 |
| Social Welfare | 34 | 67 | 101 | 8.42 | 9.99 | 1.57 |
| Energy | 7 | 23 | 30 | 1.73 | 3.43 | 1.70 |
| Religion | 2 | 15 | 17 | 0.50 | 2.24 | 1.74 |
| Abortion | 3 | 19 | 22 | 0.74 | 2.83 | 2.09 |
| Government Operations | 66 | 125 | 191 | 16.34 | 18.63 | 2.29 |
| Terrorism | 11 | 34 | 45 | 2.72 | 5.07 | 2.34 |
| History | 20 | 49 | 69 | 4.95 | 7.30 | 2.35 |
| Environment | 9 | 32 | 41 | 2.23 | 4.77 | 2.54 |
| International_affairs and Foreign Aid | 31 | 75 | 106 | 7.67 | 11.18 | 3.50 |
| Voting/Elections | 27 | 72 | 99 | 6.68 | 10.73 | 4.05 |
| Law and Crime and Policing | 40 | 101 | 141 | 9.90 | 15.05 | 5.15 |
| Immigration and Refugee Issues | 22 | 84 | 106 | 5.45 | 12.52 | 7.07 |

We determined the top 15 most common subject tags by count, excluding **Record/Candidate Biography/Campaigns and Personal Behavior**. We knew from experience that during specific time periods, certain subjects would be lied about at a disproportionate rate. For example, the spike in lies relating to Health at the outset of the pandemic. To visualize this phenomenon, we made the following graph displaying these 15 tags and how the frequency of lies progressed over time. It is of note that we only have six months of data in 2021, which is why counts are disproportionately lower.

