

Women's representation and voice in media coverage of the coronavirus crisis

Laura Jones

Laura Jones is a Research Associate at the Global Institute for Women's Leadership at King's College London. She would like to thank Jonathan Rankin for invaluable support and assistance on this project.

Executive summary

This research note summarises the results of a big-data analysis of 146,867 articles related to coronavirus published between 1 March and 31 July 2020 and drawn from 15 sources across the UK, Australia and the US.

Using natural language processing techniques, we searched each article for well-known individuals featured and quotations. Through this analysis, we established (1) the gender of well-known experts, (2) the proportion of female voices by article category and topic and (3) the visibility of politicians by gender.

Our key findings:

- Women are a minority among mentions of prominent economists and STEM experts (those famous enough to have their own Wikipedia pages) in articles on the Covid-19 crisis. For every mention of a prominent female STEM expert, there are 19 mentions of a male counterpart. For every mention of a prominent female economist, there are five mentions of a male economist.
- Only a third of quotations in articles concerning the pandemic were attributable to women, but with large variation by article category and topic. 37% of quotations in the health coverage were attributable to women, compared with 27% in business articles, and 24% in science and politics articles.
- When looking at coverage of specific sub-topics linked to the coronavirus crisis, there was a clear divide in female voice between traditionally feminine and masculine issues. Women make up over half of those quoted in articles related to childcare and domestic violence, and just over a third in those related to education – but less than a sixth of those quoted on topics related to finance and the economy.
- Female politicians make up just four of the top 20 most mentioned domestic politicians in articles related to coronavirus in the UK, Australia and the US, confirming that in each country the political response has had a largely male face.

Introduction

The coronavirus pandemic is a crisis on many fronts – political, economic, and scientific – with ramifications that touch every member of society. Yet the response to the pandemic has had an overwhelmingly male face. Just two of the 27 members of the White House Coronavirus Task Force are women,¹ and in the UK almost half of the government's daily press briefings featured an all-male line-up.²

Beyond looking at female representation in public bodies or formal institutional settings, quantifying the number of women appearing in media coverage of the pandemic provides another way of understanding how female voices shape the debate and potentially the policy response, too.

Early research in this area suggests that male dominance in the world of elected representatives and official bodies is mirrored in the voices represented in news media coverage of the Covid-19 crisis. Three times as many male experts than female experts were featured on the UK's flagship TV and radio news programmes in March, as reporting focused on the early political handling of the crisis.³ This imbalance was largely due to the numerical dominance of men among politicians and government spokespeople.

Until now, most research in this area has relied on manual analysis of a small number of media sources. To understand whether these findings are applicable more generally, we have applied big-data processing techniques to 146,867 articles on the topic of coronavirus from 15 sources across three countries – the UK, the USA and Australia – collected between 1 March and 31 July 2020.

Using natural language processing techniques we were able to analyse the content of each article in order to (1) quantify the gender of well-known STEM experts and economists⁴; (2) establish the gender breakdown of quotations in different types of coverage, to get a sense of female voice by article category and topic; and (3) extract every mention of a politician to generate a picture of the visibility of politicians by gender.

Sample

Our initial sample consisted of 248,741 articles on the coronavirus pandemic published between 1 March and 31 July 2020 and was constructed by pulling every article which mentioned the words “coronavirus” or “Covid-19” from the top five online news outlets (according to global Alexa rankings) in the UK, Australia and the US, using the API provided by eventregistry.org, an AI-driven media intelligence platform.

Many articles mentioned coronavirus only very tangentially, without it being a focus of the article. To refine our sample, we therefore carried out additional filtering. Eventregistry.org assigns a weight to each article according to how strongly the article is related to coronavirus. As we were primarily interested in articles with a substantial focus on the pandemic, we kept only those with a strong coronavirus weight. This resulted in a sample of 146,867 articles.

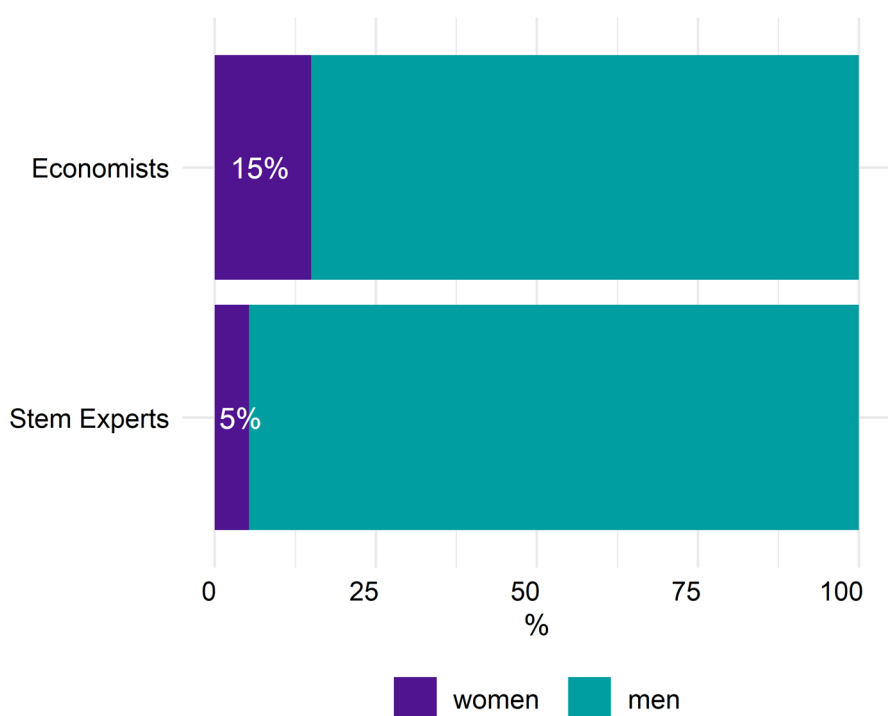
1. Gender of well-known STEM experts and economists

In this section we analyse the representation of women among well-known people mentioned in the coverage of coronavirus within two key categories, STEM experts and economists. These categories were chosen to reflect the economic and scientific dimensions of the crisis.

It is important to note that the analysis in this section is based only on mentions of people well-known enough to have their own Wikipedia pages – this is because it is only for these people that we are to pull information on their occupations. On the one hand, this is a disadvantage, in that it cannot tell us about the representation of female voices at a more junior level; but on the other hand, it gives us an idea of the voices within the most visible (and likely most powerful) section of these groups.

Our news content was sourced from the API provided by eventregistry.org, an AI-driven media intelligence platform. Eventregistry.org scans each article and uses the content of the text to annotate articles with links to relevant Wikipedia articles, including to the pages of any individuals mentioned who are well-known enough to have their own entry.⁵ Using this and the Wikidata Query API, we were able to link information on gender and occupation for 54,636 unique people who were mentioned 666,681 times. Occupations that appeared at least 25 times were manually coded if they belonged to one of three focal categories: STEM experts and economists (analysed in this section) and politicians (analysed in section 4). For more detailed information, please see the methodology endnote.

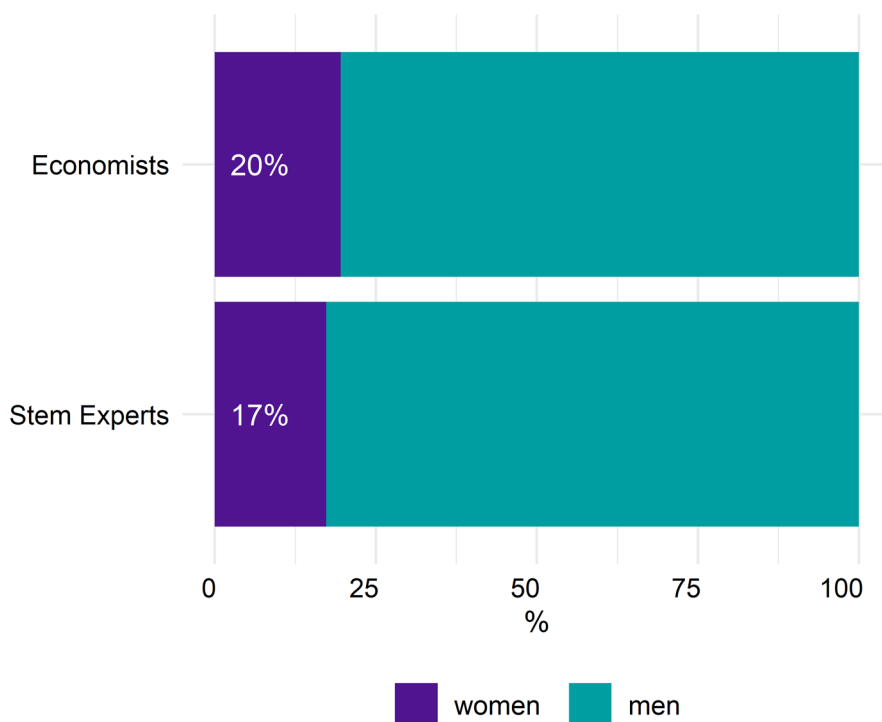
Figure 1: Gender split among mentions of STEM experts and economists in media coverage of coronavirus



Based on analysis of 146,867 articles referring to coronavirus from 15 sources drawn from the UK, Australia and US, 01/03/20–31/07/20.

We found that women were a small minority in both categories. Just 5% of mentions of well-known STEM experts were of women. This means that for every mention of a prominent female STEM expert in a news story concerning the virus, there were 19 mentions of prominent male STEM experts. This poor result was despite the fact that women made up 17% of unique prominent STEM experts mentioned (see figure 2) and is driven by the fact that the most mentioned STEM experts – such as Anthony Fauci and Chris Whitty – received a very large share of the citations.

Figure 2: Gender split among unique individuals mentioned in media coverage of coronavirus



Based on analysis of 146,867 articles referring to coronavirus from 15 sources drawn from the UK, Australia and US, 01/03/20–31/07/20.

15% of mentions of well-known economists were of women, meaning that for every mention of a prominent female economist there were five mentions of a prominent male economist. This category includes anyone who has the occupational label “economist” on Wikipedia, so includes, among others, government advisors, senior officials and academics.

2. Female voice by type of coverage

To complement the analysis above, which focuses on the representation of selected groups of individuals and the ratio of male to female mentions, we expanded the research to include all individuals and understand not only whether they were mentioned, but whether they received space to express their thoughts. To do this we analysed the quotations within each article.

To differentiate between a woman being merely mentioned and receiving space to express her opinions in the article, we extracted every instance where a gendered pronoun was followed by a verb. We manually tagged every verb that appeared at least 100 times as indicative of speech (e.g., “he said”, “she argued”) or not (e.g., “he went”).

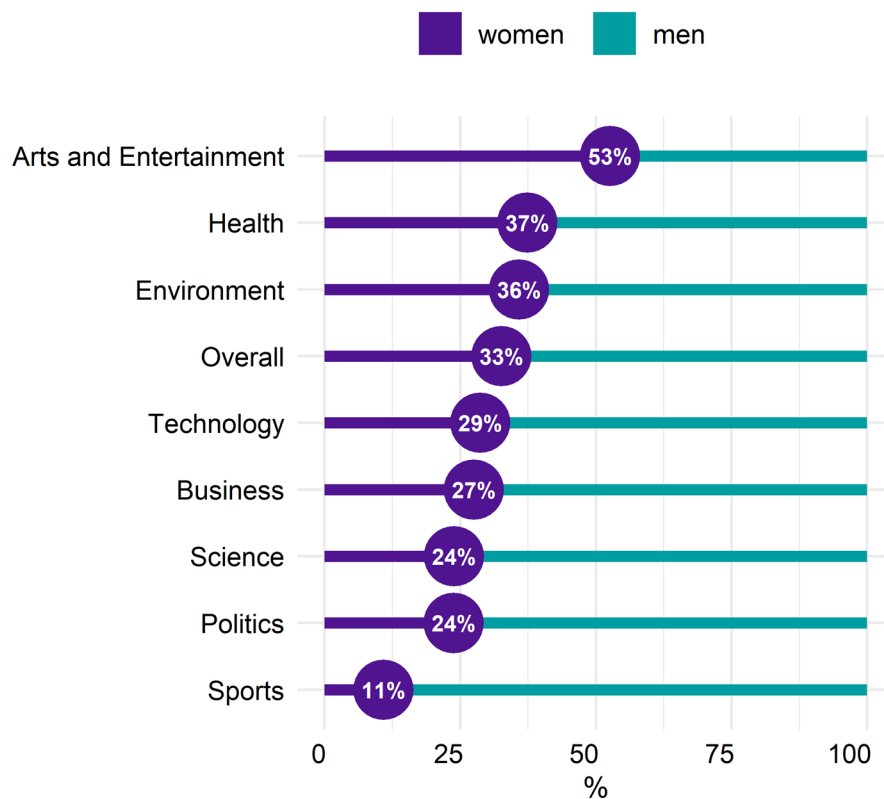
Because articles mentioning coronavirus are extremely diverse in their content, and previous research suggests large differences in the representation of women across article categories,⁶ the articles in our sample were classified into eight pre-defined categories using their linguistic content.⁷

Sample headlines from each article category:

- Arts and Entertainment: “Can the music business survive coronavirus?”
- Health: “Health workers going door-to-door in Covid-19 hotspots NOT given masks”
- Environment: “Jane Goodall says humanity is doomed if we don’t change after this pandemic”
- Tech: “Technology companies can help fight Covid-19”
- Business: “Australian economy will take biggest dive since 1930s, the IMF says”
- Science: “Before a vaccine, we’ll need treatments. Is interferon beta-1a a game changer?”
- Politics: “‘Very angry’: China pressured EU to drop COVID disinformation criticism”
- Sports: “Premiership: Cancelled season would be ‘frustrating’, says Umaga”

We then looked at the share of female voices by article category. The results are displayed in figure 3 below.

Figure 3: Percentage of quotations attributable to women in articles related to coronavirus, by article category



Based on analysis of 146,867 articles referring to coronavirus from 15 sources drawn from the UK, Australia and US, 01/03/20–31/07/20.

Overall, a third of the quotations in our sample of coronavirus coverage were attributable to women, but with large differences by article category. In the arts and entertainment category, over 50% of quotations were attributable to women, falling to just over a third in the health and environment categories, just over a quarter in the business and technology categories, and less than a quarter in the science and politics categories.

As we were interested in establishing the extent to which female voices were present in the discussion in areas of policy relevance, we selected four particularly salient article categories – health, science, politics and business – and looked to see whether there were any significant differences between countries. The results are displayed in figure 4 below.

Figure 4: Percentage of quotations attributable to women in articles related to coronavirus, by article category and country



Based on analysis of 146,867 articles referring to coronavirus from 15 sources drawn from the UK, Australia and US, 01/03/20–31/07/20.

The results were remarkably consistent across the three national contexts for most categories. Ranging from 36% to 39% female, the gender balance of quotations was most equal in the health category. Articles in this category included, for example, stories of individual cases of coronavirus and reporting from hospitals and care homes, and quoted doctors, nurses and pharmacists alongside members of the public and politicians.

In the business category, 27% to 28% of quotations were from women. Articles in this category reported on issues such as the economic impacts of the pandemic, business closures and sick leave, among others.

Articles in the science category had the widest national spread, with 21% (UK) to 31% (Australia) of quotations attributable to women. Articles in this category included those on coronavirus vaccine research, testing and treatments.

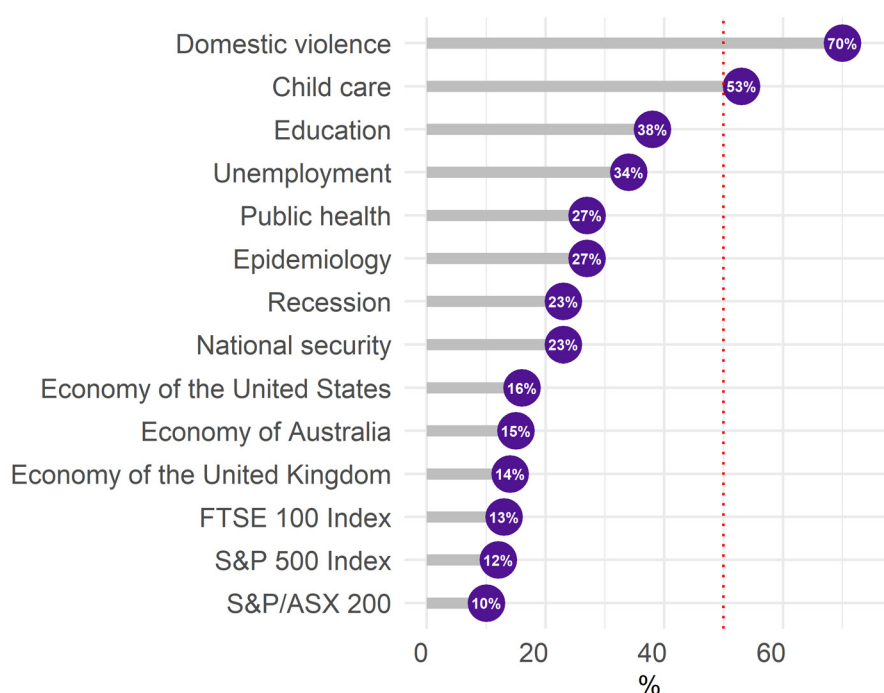
While the results in the science and business categories are lower than women's representation in the general population, they are much higher than women's share among the well-known experts reported in section 1, suggesting that women's share of voice is larger among those who are either being quoted in a non-expert capacity or who are less publicly prominent experts.

With the second lowest female representation overall after sport (see figure 3), approximately a quarter of the quotations in political articles in each country were attributable to women. Articles in this category included those discussing the government's handling of the pandemic and policy responses.

3. Female voices by topic

Looking beyond broad article categories, we were able to compare the representation of female voices in articles focusing on key topics of interest. The results are displayed below in figure 5.

Figure 5: Percentage of quotations attributable to women in articles related to coronavirus



Based on a subsample of 13,541 articles referring to coronavirus and these topics, part of a larger sample of 146,867 articles from 15 sources drawn from the UK, Australia and US, 01/03/20–31/07/20.

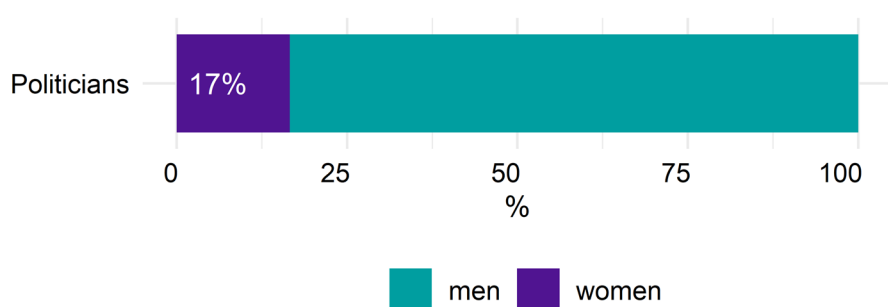
There appears to be a clear split between the representation of female voices in what are traditionally seen as feminine policy areas, such as domestic violence, childcare and education, and the policy areas traditionally seen as masculine, such as finance, defence and the economy,⁸ with the scientific topics of public health and epidemiology somewhere in the middle.

At one extreme, women are in the majority of those commenting in articles discussing domestic violence and childcare, but make up just one in five, or less, of those commenting in articles discussing economic and finance topics.

4. Visibility of politicians by gender

In this final section we examine the visibility of female politicians, first by repeating the analysis from section 1 with politicians as our focal group (figure 6), and then by looking at the gender breakdown of the top 20 most mentioned politicians from each country (figures 7-9).

Figure 6: Gender split among mentions of politicians in media coverage of coronavirus

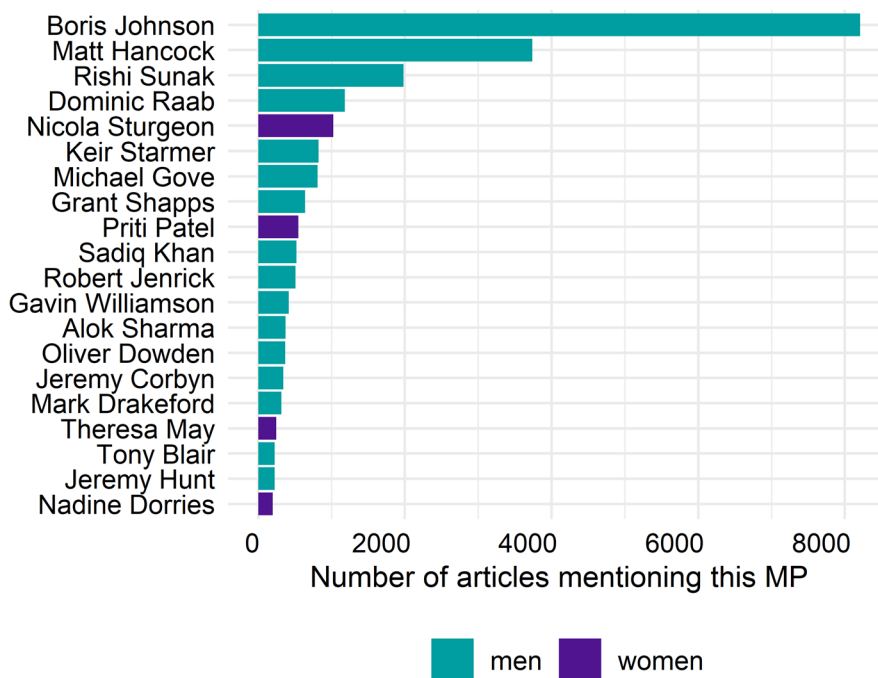


Based on analysis of 146,867 articles referring to coronavirus from 15 sources drawn from the UK, Australia and US, 01/03/20–31/07/20.

When looking across the whole sample we found that 17% of references to politicians were to women, meaning that for every mention of a female politician in an article about the coronavirus crisis, there were five mentions of a male politician. When we exclude references to the (male) national leaders of each country in our sample – Boris Johnson, Donald Trump and Scott Morrison – we still find that there are four mentions of a male politician for every mention of a female one.

We dug deeper into female representation in the political domain by extracting every mention of a politician and then plotting the top 20 most mentioned for each country separately (including only politicians from that country). The results are found in figures 6 to 9 below and confirm that the political face of the coronavirus crisis has been male-dominated in all three countries.

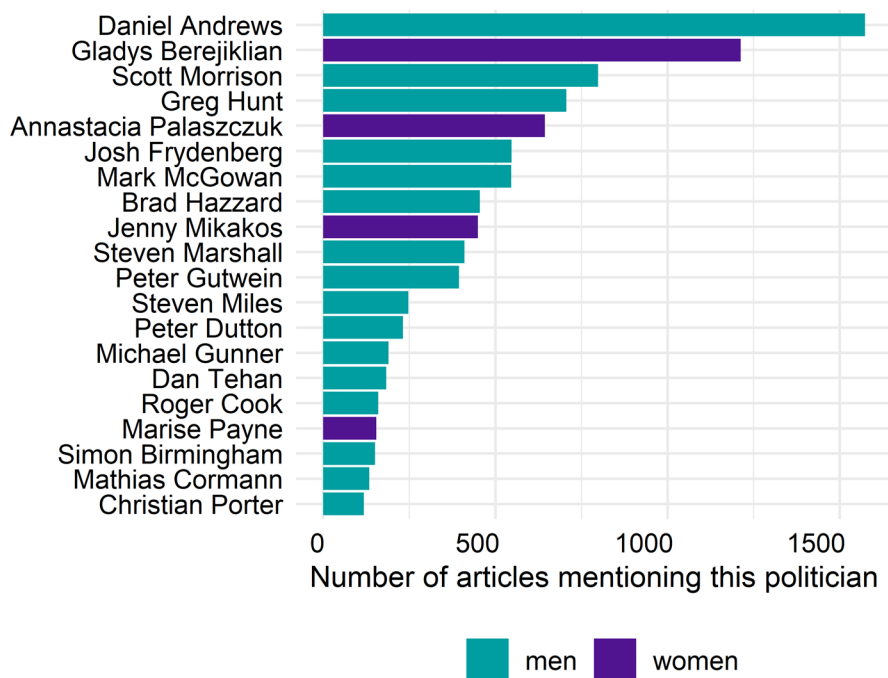
Figure 7: Top 20 most mentioned politicians – UK



Based on analysis of 77,137 articles from five UK sources, 01/03/20–31/07/20.

In the UK only four of the top 20 most mentioned British politicians are female. This is despite the fact that, at 34% female, the UK has the most gender-balanced parliament on record and reflects the numerical dominance of men in Boris Johnson's cabinet. Only Scottish First Minister Nicola Sturgeon, Home Secretary Priti Patel, former Prime Minister Theresa May and Health Minister Nadine Dorries make it into the top 20.

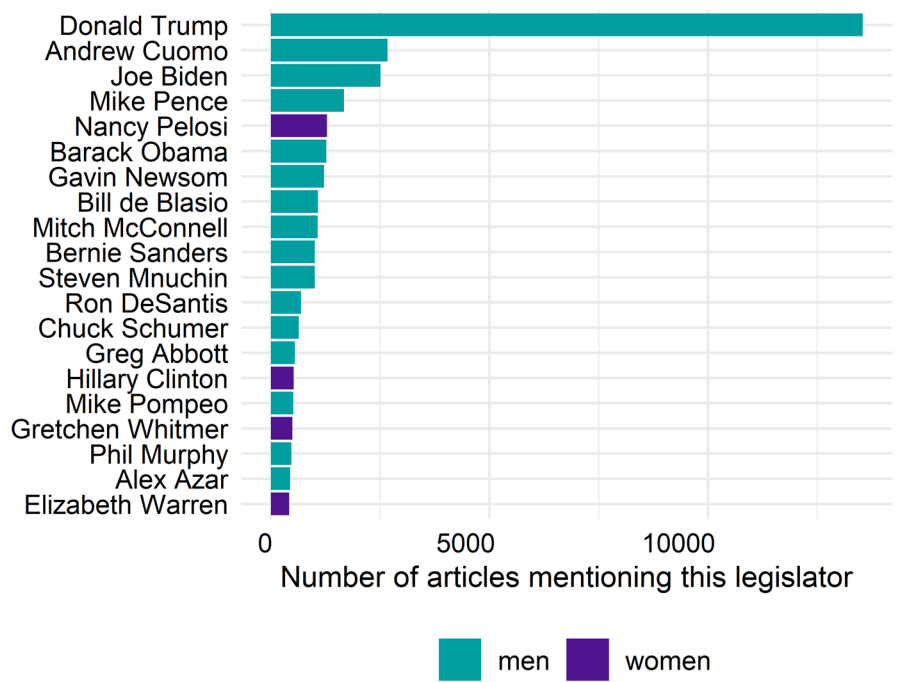
Figure 8: Top 20 most mentioned politicians – Australia



Based on analysis of 26,134 articles from five Australian sources from 01/03/20–31/07/20.

It is a similar story in Australia, where only four of the top 20 most mentioned Australian politicians are female. State Premiers Gladys Berejiklian and Annastacia Palaszczuk, as well as former Victorian Health Minister Jenny Mikakos and Minister for Foreign Affairs Marise Payne, are the only women to make the top 20.

Figure 9: Top 20 most mentioned politicians – US



Based on analysis of 43,596 articles from five American sources from 01/03/20–31/07/20.

In the US, the media coverage is dominated by Donald Trump, who received five times more mentions than the next most mentioned politician, Andrew Cuomo. It is interesting to note that of the four women who make the top 20, all are Democrats.

Conclusion

The coronavirus crisis has placed the role of women in leadership positions firmly on the agenda. This report has offered a complement to analysis which looks at women's numerical representation in decision-making positions by examining the extent to which this participation translates into "voice" in media coverage of the coronavirus crisis in three countries.

Providing a detailed breakdown of the categories and topics where women are more or less represented, we found that male voices and expertise dominate in almost every area. Only in articles classified as "arts and entertainment" or touching on the topics of childcare and domestic violence did women make up at least half of those quoted. As the voices become more prominent and well-known, the female share in them decreases.

Like previous epidemics, coronavirus, and the policy response to it, has had and continues to have gendered impacts, exacerbating existing structural inequalities between women and men. Early evidence suggests that in the UK, US and Germany, women have been more likely to either lose their jobs or experience reductions in income, largely due to their overrepresentation in precarious and hard-hit sectors, at the same time as caring responsibilities have intensified for many.⁹ Meanwhile, there are reports that domestic violence has increased under lockdown.¹⁰

Properly mitigating the impacts of these structural gender inequalities means incorporating gendered perspectives and female voices into emergency and long-term strategies. What is clear is that as the continuing crisis shapes our lives for the foreseeable future, female voices in the public debate are a minority.

Methodology endnote

Data

Our initial sample consisted of 248,741 articles on the coronavirus pandemic published between 1 March and 31 July 2020 and was constructed by pulling every article which mentioned the words “coronavirus” or “Covid-19” from the top five online news outlets (according to global Alexa rankings) in the UK, US and Australia, using the API provided by eventregistry.org, an AI-driven media intelligence platform.

The sources were:

UK: theguardian.com, bbc.co.uk, telegraph.co.uk, independent.co.uk, dailymail.co.uk

USA: cnn.com, nytimes.com, cnbc.com, usatoday.com, foxnews.com

Australia: news.com.au, abc.net.au, smh.com.au, 7news.com.au, sbs.com.au

Many articles mentioned coronavirus only very tangentially, without it being a focus of the article. To refine our sample, we therefore carried out additional filtering. Eventregistry.org assigns a weight to each article according to how strongly the article is related to coronavirus. As we were primarily interested in articles with a substantial focus on the pandemic, we kept only articles with a strong coronavirus weight. This resulted in a sample of 146,867 articles. We ran all the analysis on the full sample as well as the filtered sample, and this did not change the substantive results.

Methodology

To analyse the representation of women within these sources, we applied a variety of techniques drawn from the fields of machine learning and natural language processing.

Entity linking

Our news content was sourced from the API provided by eventregistry.org, an AI-driven media intelligence platform. Eventregistry.org scans each article and uses the content of the text to annotate articles with links to relevant Wikipedia articles, including to the pages of any individuals mentioned who are well-known enough to have their own entry. To check the accuracy of these annotations, we manually checked 100 references and found that 90% were accurate.

Using this and the Wikidata Query API, we were able to link information on gender and occupation for 54,636 unique people who were mentioned 666,681 times. Occupations that appeared at least 25 times were manually coded if they belonged to one of three focal categories: STEM experts, economists and politicians. As we were interested in the contribution of contemporary experts, we excluded references to those born before 1930.

Article categorisation and topics

To group the articles into categories we used the article category tags provided by eventregistry.org, which classified articles into a set of eight fixed categories using a technique called topic classification. An algorithm is trained to classify articles based on their linguistic content using a training dataset of pre-categorised articles.

Article topic was determined using the concept tags provided by eventregistry.org. These are applied using the same technique as the tags related to individuals discussed in the entity linking section. As well as applying a concept tag (such as “childcare”) eventregistry.org provides a weight, according to how central that concept is the article (measured from 1-5). We selected only articles where the concept was weighted 4 or 5. This decision was made after an manual check of 100 articles with the “childcare” tag with weights 3-5. We found that of those articles, only two didn’t mention childcare at all, but 21 only mentioned it very tangentially. Among the articles weighted 4-5 all mentioned childcare and only 5% mentioned it tangentially.

Quotations

Each article was analysed using the Stanford core NLP java library. This processes sentences into their constituent parts and tags words according to their parts of speech. We then extracted every instance where a gendered pronoun was followed by a verb. We manually tagged every verb that appeared at least 100 times as indicative of speech (e.g., “he said”, “she argued”) or not (e.g., “he went”).

References

1. <https://www.npr.org/sections/goatsandsoda/2020/06/24/882109538/where-the-women-arent-on-coronavirus-task-forces>
2. Smith, J.C., (2020). Where Are the Women? Descriptive Representation and Covid-19 in U.K. Daily Press Briefings. *Politics & Gender*, 1–10. <https://doi.org/10.1017/S1743923X20000513>
3. city.ac.uk/news/2020/may/male-experts-on-uk-news-programmes-peak-during-march-2020
4. Those notable enough to have their own Wikipedia page.
5. For more information on this, see Brank, J., Leban, G. and Grobelnik, M., (2018). Semantic annotation of documents based on wikipedia concepts. *Informatica*, 42(1).
6. Jia, S., Lansdall-Welfare, T., Sudhahar, S., Carter, C., Cristianini, N. (2016). Women Are Seen More than Heard in Online Newspapers. *PLOS ONE* 11, e0148434. <https://doi.org/10.1371/journal.pone.0148434>
7. Categorisation based on eight general categories used to categorise news content: business, politics, technology, arts and entertainment, sports, health, science, and environment.
8. Krook, M. & O'Brien, D.Z. (2012). All the president's men? The appointment of female cabinet ministers worldwide. *Journal of Politics* 74(3): 840–885.
9. Adams-Prassl, A., Boneva, T., Golin, M. and Rauh, C. (2020). Inequality in the impact of the coronavirus shock: Evidence from real time surveys.
10. <https://www.theguardian.com/society/2020/apr/12/domestic-violence-surges-seven-hundred-per-cent-uk-coronavirus>
11. For more information on this, see Brank, J. et al. (2018), op. cit.

The Global Institute for Women's Leadership

The Global Institute for Women's Leadership at King's College London works towards a world in which women of all backgrounds have fair and equal access to leadership.

Chaired by Julia Gillard, the only woman to have served as Prime Minister of Australia, the institute brings together rigorous research, practice and advocacy to break down the barriers to women becoming leaders, while challenging ideas of what leadership looks like.

giwl.kcl.ac.uk
@GIWLkings
#womensleadership
giwl@kcl.ac.uk