Do Hero Images Perpetuate Gender Bias?

Anika Fuloria
Stanford University
afuloria@stanford.edu

Abstract

Hero images, or the thumbnail images that accompany online news articles, are often the first visual cue readers encounter, yet their potential to perpetuate media bias remains underexplored. While most existing studies of bias focus on textual content, this paper proposes a multimodal approach that evaluates bias embedded in hero images through a novel computer vision pipeline. My method takes as input a dataset of over 13,000 UK-based news articles, each including a hero image, headline, and summary. I apply BLIP-2 for image captioning, conduct sentiment analysis across all modalities, and use zero-shot classification to detect five types of sociopolitical bias. I find that image captions, especially those associated with women, are significantly more likely to exhibit negative sentiment and receive "gender bias" or "stereotyping" labels, even when the article text remains neutral. By comparing similar headlines across sources and examining outlet-level trends, I show that visual framing varies systematically by both gender and news source. My findings highlight the need for greater scrutiny of editorial image selection and demonstrate the value of combining vision-language models with interpretability tools to detect media bias at scale.

1. Introduction

While media bias has traditionally been studied through textual content, visual bias remains relatively underexplored. In particular, the bias of hero images, or the thumbnail images shown alongside online news articles, is understudied. These images are often the most salient visual elements on a news page and may encode racial, gender, or emotional biases that reinforce or contradict the article's framing. Such visual signals, while subtle, can influence public perception and perpetuate stereotypes, especially in high-traffic digital environments.

This project is motivated by the growing recognition that multimodal media (media combining text and images) plays a critical role in how audiences engage with and interpret news. Yet, most bias detection methods focus exclusively on language, leaving a significant gap in our understanding of visual rhetoric and framing.

To address this, I develop an end-to-end computer vision pipeline to detect potential bias in hero images. The input to my system is a set of online news articles, each containing an associated hero image, headline, and summary. Using this data, I apply a combination of image captioning models (BLIP-2), natural language inference (NLI) with zero-shot classifiers, and sentiment analysis tools to generate structured representations of visual content. The output of my system is a set of predicted bias scores related to gender, race, and emotional tone for each image. I then analyze these scores in aggregate to identify patterns of bias across news sources and demographic categories.

My goal is not only to reveal previously invisible forms of media bias but also to demonstrate how vision-language models can be repurposed to evaluate framing in visual journalism. Ultimately, this project contributes to broader efforts in AI ethics, fairness, and explainability by operationalizing normative concerns about media representation into quantifiable metrics.

2. Related Work

Previous studies of media bias have primarily focused on textual content, using methods such as ideological labeling, framing analysis, and natural language processing (NLP) to identify slant. For example, Budak et al. [2] modeled political bias through differences in article selection and phrasing, while Hamborg et al. [5] developed tools like NewsBird to detect biased reporting through structured annotation pipelines. These works demonstrate that media outlets influence public opinion through subtle narrative techniques like omission, emphasis, and framing. However, they concentrate almost exclusively on language.

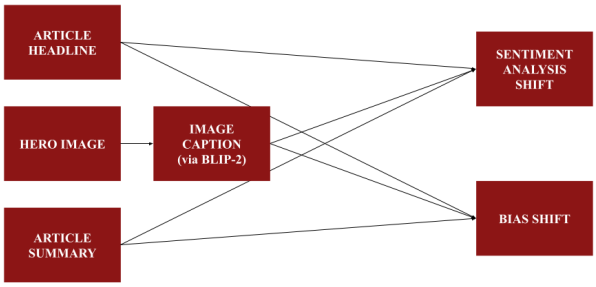
Meanwhile, growing interest in multimodal sentiment analysis has motivated new approaches that integrate text and image processing. Chaubey et al. [4] show that users frequently express opinions through image-text pairs on social media and propose a joint transformer-based model that improves sentiment classification performance across both modalities. Similarly, You et al. [8] and Majumder et al.

[6] have developed image-text fusion techniques to capture emotional valence more accurately. Zadeh et al. [9] also propose tensor-based models that enable fine-grained alignment of visual and linguistic signals. However, most of these approaches are trained on user-generated datasets like social media posts, which may not generalize well to curated editorial imagery in journalism.

In parallel, computer vision researchers have explored fairness and representation in facial detection and classification systems. For instance, Buolamwini and Gebru [3] revealed that commercial gender classifiers perform poorly on darker-skinned and female faces, highlighting the risk of biased training data. More recently, works like Steed and Caliskan [7] have shown that large vision-language models may replicate gendered and racial stereotypes when generating or captioning images. These concerns inform my methodological choices, especially around image captioning and attribute extraction. Similarly, Birhane et al. [1] analyzed large vision-language datasets like LAION and found evidence of racial and gender biases embedded in both captions and image distributions. Zhao et al. [10] show that models trained on image-caption pairs often amplify gender stereotypes, such as over-associating women with cooking or cleaning tasks. These findings inform my decision to explicitly audit gender representation in image captions.

Despite these advances, little work has been done to quantify bias in editorial images used by mainstream news outlets. While projects like Media Bias/Fact Check and AllSides provide ideological labels for news sites, they do not evaluate how images visually encode sentiment, identity, or emotion. My work builds on this foundation by offering a scalable, automated method for evaluating image bias in journalism. By adapting recent advances in captioning like BLIP-2, natural language inference like BART-based zero-shot classification, and sentiment analysis, I provide new tools for understanding how media representation operates not just in words, but in pixels.

3. Methods



My analysis begins with the hero image accompanying each news article in the dataset. These images are typically the primary visual representation associated with the article and often contain one or more people. To extract semantic meaning from these images, I apply the **BLIP-2** model which generates captions that describe image content. The primary aim of this step is to obtain descriptive captions that reflect both the gender and actions of any individuals present in the image.

Following image captioning, I conduct sentiment analysis on three distinct components of each article: the headline, the summary, and the image caption. This allows me to assess whether there are notable differences in emotional tone across these modalities. For example, do the headline and summary frame a story more negatively or positively than the visual representation does?

I then apply a zero-shot classification approach using Facebook's **bart-large-mnli** model to evaluate potential biases in each of the three textual components. Specifically, I classify each headline, summary, and caption against a predefined set of possible bias categories: `["gender bias", "racial bias", "political bias", "stereotyping", "neutral"]`. This step enables me to identify instances where different parts of the article may exhibit diverging framing, particularly in terms of implicit or explicit bias.

Finally, to explore the consistency of bias representation, I compare pairs of articles with highly similar headlines, using cosine similarity of TF-IDF vectors as the metric. I then evaluate whether these similar headlines are associated with consistent or divergent sentiment and bias classifications across their respective summaries and image captions. This step helps determine if subtle textual variations in framing contribute to measurable shifts in perceived bias or sentiment.

4. Dataset and Features

The dataset used for this study consists of over 13,000 UK-based news articles and was sourced from Kaggle. Each entry in the dataset includes several components: a headline, a brief article summary, a direct link to the full article, the publication date, and a hero image in JPEG format. The dataset is particularly valuable because it pairs textual information with a corresponding visual element, allowing for multimodal analysis of bias and sentiment.

As an example, one row of the dataset is:

```
What's gone wrong at Royal Mail?,
"Lost letters, big losses and
calls for reform:  BBC Panorama
speaks to customers and insiders.",
https://ichef.bbci.co.uk/ace/standard/240/
cpsprodpb/0553/production/_132736310
_postie-gettyimages-1232863080.jpg,
"Mon, 26 Feb 2024 00:01:07 GMT",
```

https://www.bbc.co.uk/news/business-6838228 IMAGE_2.jpg.

IMAGE_2.jpg (the corresponding hero image) is shown in 1. The caption accurately describes the contents of the image and identifies the word ”woman” as the subject of the image.

Figure 1. BLIP-2 generates the caption “a woman is standing in front of a mailbox with a letter in her hand.” This demonstrates the model’s capacity to extract contextual semantics from hero images.

Notably, the image files are relatively low-resolution, which presents a challenge for visual models like BLIP-2 that rely on fine-grained details to generate accurate captions. Despite this, the dataset’s breadth and real-world relevance (as it spans many major UK news outlets) make it well-suited for studying how media representational bias can manifest across both text and imagery. The presence of multiple text modalities (headline, summary, and image caption) per article enables a comprehensive comparison of how different channels may convey distinct biased narratives.

5. Results

5.1. Overall Results

To investigate visual bias across gender groupings, I calculated the ”bias shift” (as measured by the difference between sentiment scores for text and the image caption) for each article. Figure 2 visualizes the overall distribution of these shifts by gender group. The largest average negative shift is observed in images featuring only women, suggesting that image captions may more frequently convey negative emotional tones when women are shown. In contrast, images without people exhibit the smallest negative shift, highlighting the influence of human representation on perceived tone.

Figure 3 shows the distribution of bias type ”shifts” (as predicted by a zero-shot classifier) across gender groups. Captions for images featuring women are significantly more likely to be labeled with ”gender bias” or ”stereotyping,”

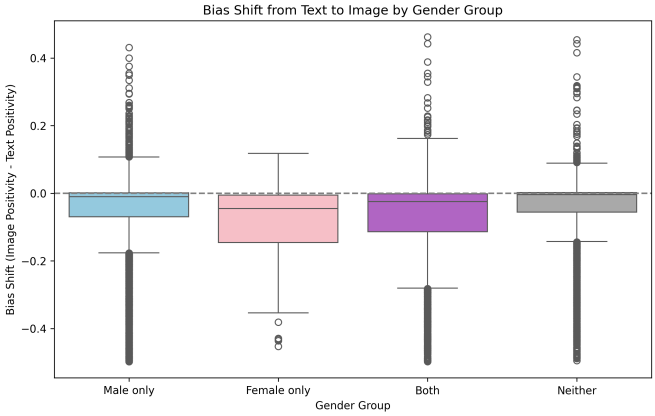


Figure 2. Overall, articles about women show a larger ”bias shift” than articles about men or about both genders.

while captions for men are more frequently labeled ”neutral.” This is different from how the article text is biased in about 50% of articles. This pattern supports the hypothesis that visual framing differs not just in sentiment but also in the kind of sociopolitical framing invoked.

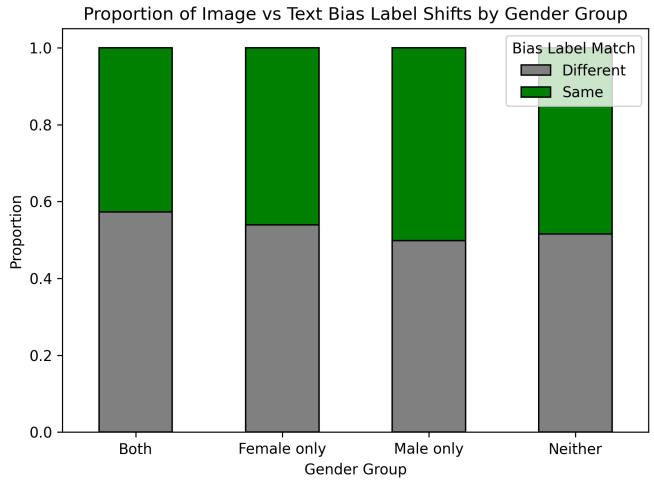


Figure 3. Bias labels change between titles/summaries and images for all gender groups.

Figure 4 further breaks down the sentiment distribution of image captions. The modal sentiment for female-only images skews negative, while male-only and both-gender images show slightly more balanced distributions. Together, these figures suggest consistent patterns in how news media visually frame gender.

Figures 5, 6, and 7 compare bias label predictions across the three modalities (headline, summary, and caption), disaggregated by gender. Headlines and summaries tend to be overwhelmingly classified as ”neutral,” whereas captions, especially for images featuring women, exhibit higher proportions of ”stereotyping” and ”gender bias.” This suggests

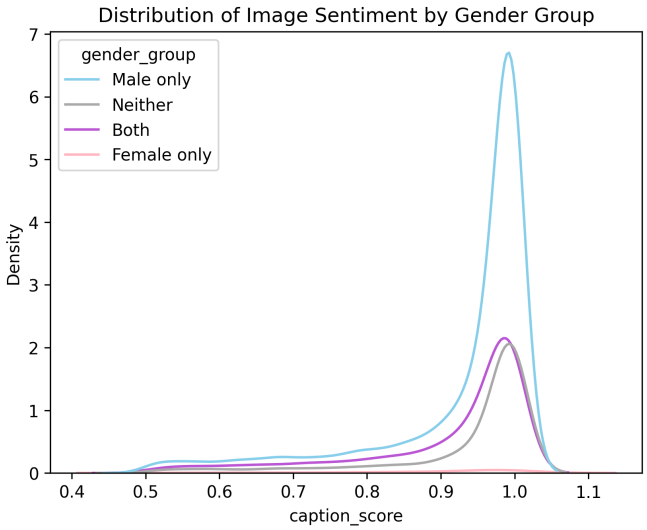


Figure 4. Men are overrepresented in the dataset, so their density graph has a higher peak. However, it seems as though all gender groups have a similar average caption score (from sentiment analysis).

that the framing shift introduced by images is both measurable and substantial.

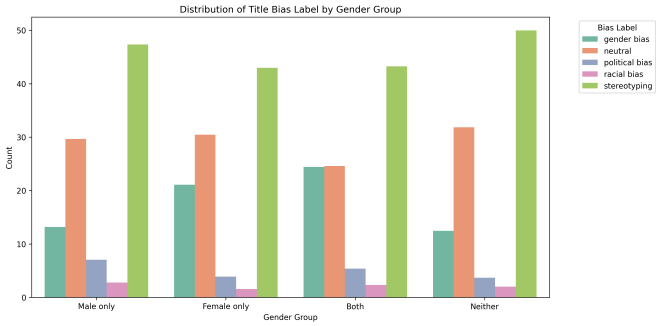


Figure 5. Article titles that are biased tend to have stereotyping bias.

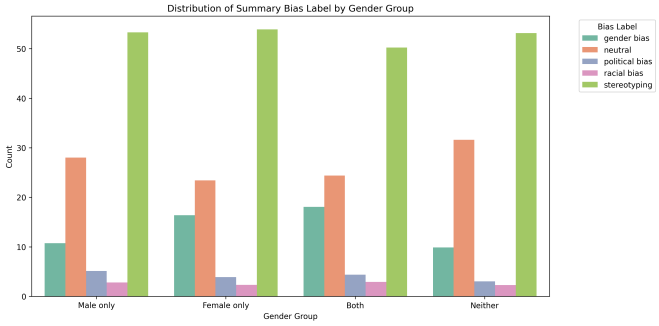


Figure 6. Article summaries that are biased tend to have stereotyping bias.

To better understand how this bias varies across news

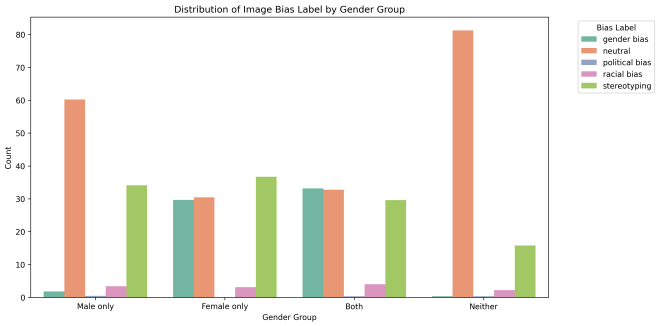


Figure 7. Images are more likely ot be unbiased when neither gender is present in the image.

sources, I calculate the mean bias shift per gender group for each outlet. For instance, bbc.co.uk and mirror.co.uk show consistently strong negative shifts for the "Both" and "Female only" categories. Some outlets, such as wired.co.uk and necn.com, show exceptionally large shifts for female-only images, although their sample sizes are small and may require cautious interpretation. These outlet-level averages illustrate the sociotechnical interaction between editorial practices and gendered visual framing.

5.2. BBC: A Case Study

BBC articles demonstrate a consistent negative sentiment shift, especially when both men and women appear in hero images. As shown in Figure 8, the shift is quantitatively more negative than in other gender groups. Figure 9 shows that the captions for these articles often reflect gender-related bias labels, even when the accompanying text remains neutral. This suggests an editorial pattern where visual content adds a biasing layer of interpretation.

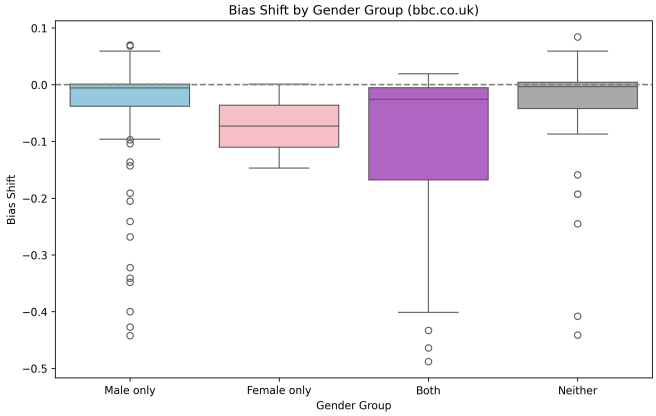


Figure 8. The BBC shows a significant bias shift in articles about women.

5.3. The Daily Mail: A Case Study

As illustrated in Figures 10 and 11, the Daily Mail shows particularly stark sentiment shifts and frequent classifica-

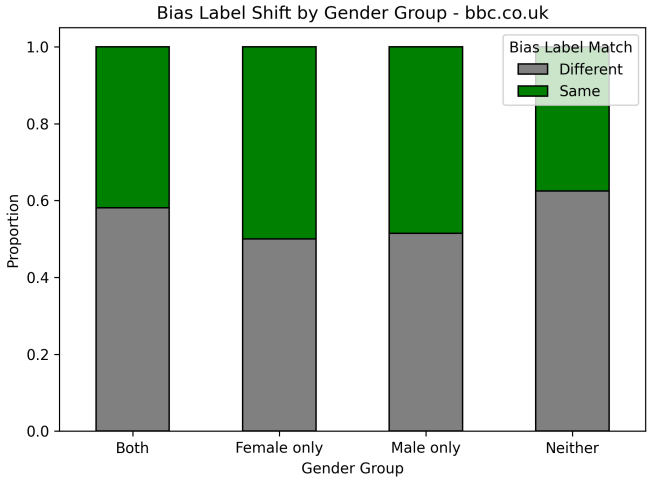


Figure 9. The BBC bias label changes are similar to all other news sources.

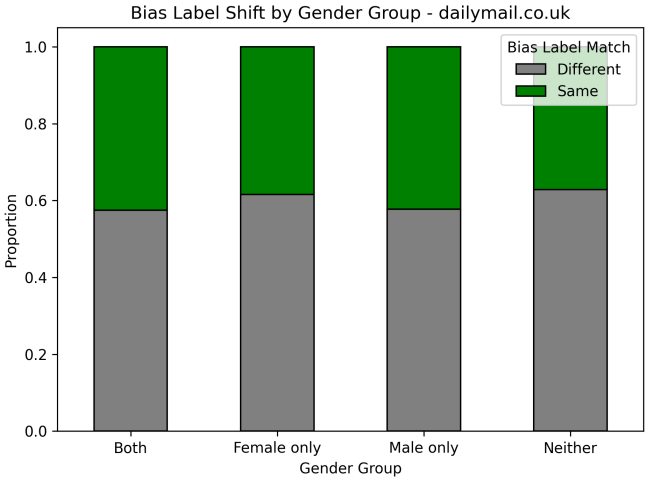


Figure 11. The Daily Mail bias label changes are similar to all other news sources.

tion of captions as biased, especially in male-only and both-gender image groups. This aligns with the publication's reputation for sensationalist and emotionally charged coverage, where images often heighten the article's framing rather than merely accompany it.

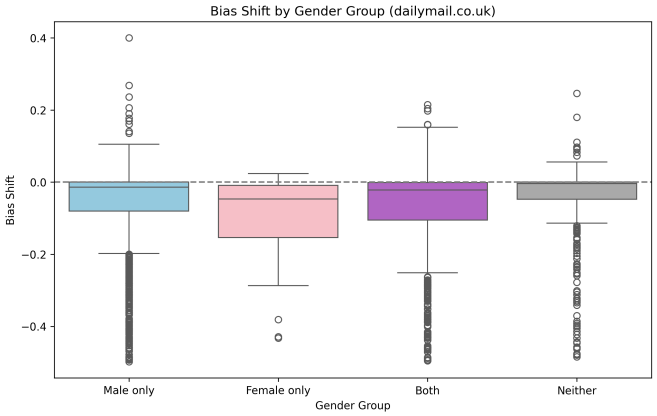


Figure 10. The Daily Mail shows less bias against women than the BBC. This is somewhat surprising as The Daily Mail is a well-known tabloid. One explanation could be that the articles themselves could be more biased so the biased images are not that different than the article content.

5.4. Similar Headlines

To explore how small textual differences might interact with visual framing, I identified near-duplicate headlines using cosine similarity (similarity > 0.8). Figure 12 shows caption score differences between matched articles. For some articles (e.g. "Wasp season"), captions and sentiment are nearly identical across outlets, suggesting consistent framing. In others (e.g. celebrity health or Oscars coverage), small changes in wording correspond to major shifts in image sentiment and gender representation, as seen in Figure 13.

Headline Pair #1 Headline pair similarity: 1.00
Headline 1: Wasp season: How to keep wasps out of your home this year
Sources: theargus.co.uk (Caption Score: 1.0) | Gender Group: Male only
Headline 2: Wasp season: How to keep wasps out of your home this year
Sources: theboltonnews.co.uk (Caption Score: 0.996) | Gender Group: Male only

Headline Pair #2 Headline pair similarity: 0.94
Headline 1: Good Morning Britain star rushed to hospital for emergency surgery
Sources: walesonline.co.uk (Caption Score: 0.796) | Gender Group: Male only
Headline 2: ITV Good Morning Britain star rushed to hospital for 'emergency surgery'
Sources: birminghammail.co.uk (Caption Score: 0.998) | Gender Group: Male only

Headline Pair #3 Headline pair similarity: 0.81
Headline 1: Oscars 2024: Here's How To Watch This Year's Academy Awards Live In The UK
Sources: huffingtonpost.co.uk (Caption Score: 0.999) | Gender Group: Both

`Headline 2: Academy Awards: How to watch the 2024 Oscars tonight in UK Sources: yorkpress.co.uk (Caption Score: 0.985) | Gender Group: Male only`

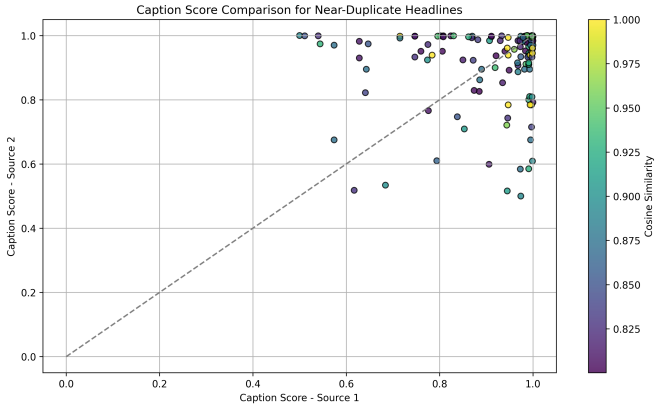


Figure 12. For similar headlines, image captions have varying bias content.

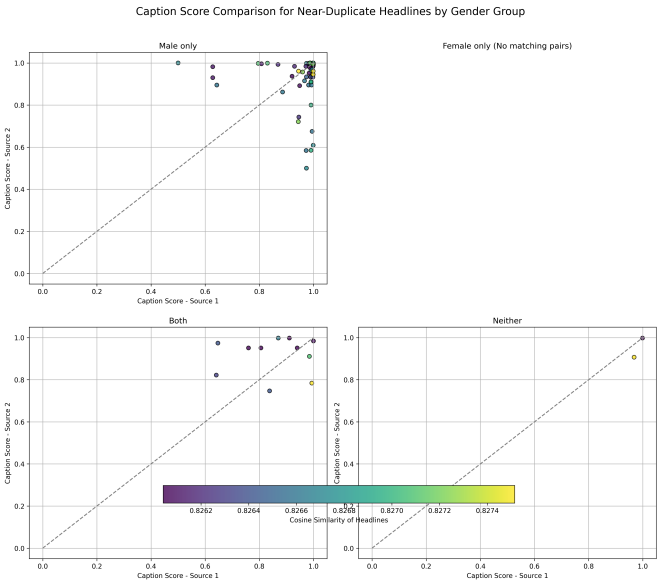


Figure 13. For similar headlines, image captions have varying bias content for different gender groups. Women did not appear in the "near-duplicate" caption subset, so there is not enough data to plot.

These examples reinforce the idea that seemingly minor textual variations or source-specific image choices can lead to meaningful differences in how bias manifests across articles, underscoring the sociotechnical complexity of media bias in multimodal contexts.

6. Discussion

The results suggest that hero images in news media are not neutral accompaniments to text, but active participants in constructing meaning. Across thousands of articles, hero images, particularly those featuring women, are more likely to introduce negative sentiment or bias not present in the article's headline or summary. This is especially pronounced in outlets like the Daily Mail, where editorial choices around imagery appear more deliberate and emotive.

These findings have implications for both journalism and machine learning. In journalism, they highlight the importance of critically evaluating image selection practices as part of editorial standards. In ML, they reinforce the value of multimodal analysis and the need for tools that can flag bias beyond text. By demonstrating that modern image-captioning and NLI tools can uncover these patterns at scale, this work contributes to the growing body of computational media studies that blends fairness, ethics, and AI.

However, limitations remain. The image captioning model (BLIP-2) may reflect biases in its training data, which could propagate into the final analysis. Similarly, zero-shot classifiers rely on subjective label sets and thresholding.

7. Conclusion and Future Work

This paper presents a novel, scalable pipeline for analyzing visual bias in hero images from news articles, integrating image captioning, sentiment analysis, and bias classification to compare visual and textual framing. I find strong evidence that images often diverge in tone and bias type from their accompanying text, particularly when women are visually represented.

Future work could expand image captioning to include markers like race and age. In addition, I could refine the sentiment and bias classifiers with news-specific fine-tuning to make these models better at analyzing biases.

Ultimately, this work demonstrates that computational methods can illuminate latent patterns in visual journalism, offering both diagnostic tools and ethical prompts for how news media shapes perception.

References

[1] A. Birhane, V. U. Prabhu, and E. Kahembwe. Multimodal datasets: misogyny, pornography, and malignant stereotypes. 2021.

[2] C. Budak, S. Goel, and J. M. Rao. Fair and balanced? quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, 80(S1):250–271, 04 2016.

[3] J. Buolamwini and T. Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on Fairness, Accountability, and Transparency*, pages 77–91. PMLR, 2018.

[4] P. K. Chaubey, T. K. Arora, K. B. Raj, G. R. Asha, G. Mishra, S. C. Guptav, M. Altuwairiqi, and M. Alhassan. Sentiment analysis of image with text caption using deep learning

techniques. *Computational Intelligence and Neuroscience*, 2022(1):3612433, 2022.

[5] F. Hamborg, K. Donnay, and B. Gipp. Automated identification of media bias in news articles: an interdisciplinary literature review. *International Journal on Digital Libraries*, 20(4):391–415, 2019.

[6] N. Majumder, D. Hazarika, A. Gelbukh, E. Cambria, and S. Poria. Multimodal sentiment analysis using hierarchical fusion with context modeling. volume 161, pages 124–133, 2018.

[7] R. Steed and A. Caliskan. Image representations learned with unsupervised pre-training contain human-like biases. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, page 701–713. ACM, Mar. 2021.

[8] Q. You, H. Jin, Z. Wang, C. Fang, and J. Luo. Image captioning with semantic attention. 2016.

[9] A. Zadeh, M. Chen, S. Poria, E. Cambria, and L.-P. Morency. Tensor fusion network for multimodal sentiment analysis. 2017.

[10] J. Zhao, T. Wang, M. Yatskar, V. Ordonez, and K.-W. Chang. Men also like shopping: Reducing gender bias amplification using corpus-level constraints. 2017.

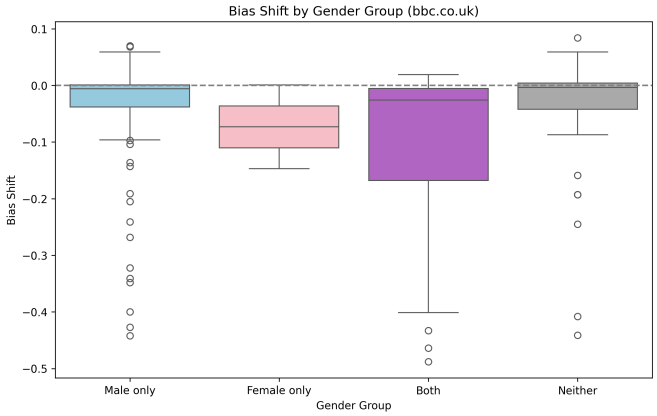
8. Appendix

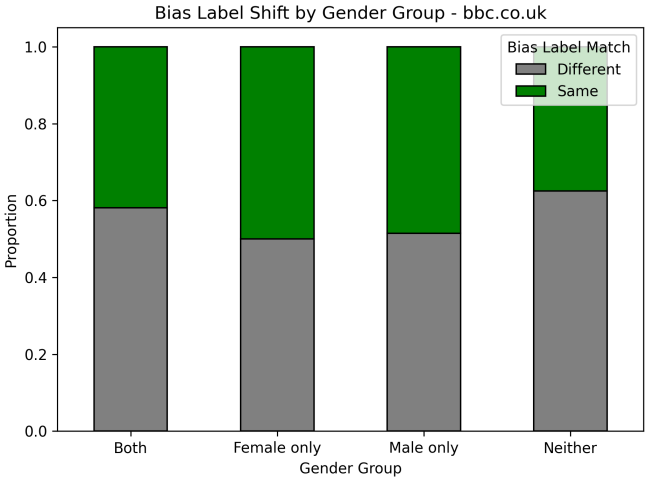
Table 1. Overall Bias Shift by Gender Group

Gender Group	Count	Mean	Std	Min	25%	50%	75%	Max
Both	3327	-0.078	0.127	-0.498	-0.1140	-0.025	-0.002	0.462
Female only	128	-0.092	0.120	-0.453	-0.1455	-0.045	-0.006	0.118
Male only	7393	-0.058	0.118	-0.498	-0.0700	-0.010	0.001	0.431
Neither	2439	-0.050	0.117	-0.495	-0.0560	-0.004	0.002	0.454

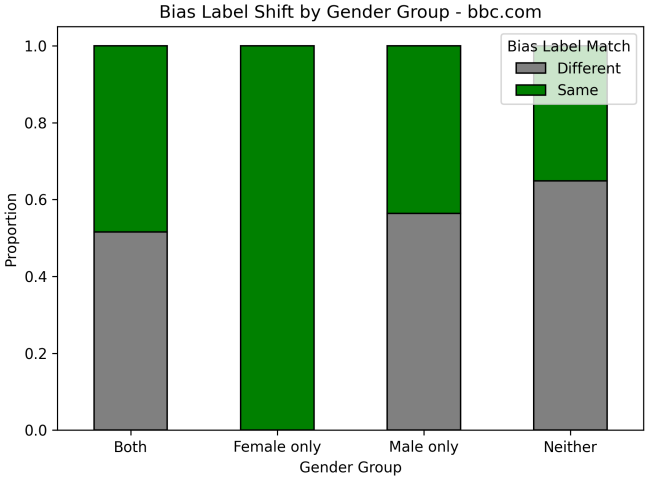
Table 2. Mean Bias Shift per Source by Gender Group

News Source	Both	Female only	Male only	Neither
bbc.co.uk	-0.105	-0.073	-0.045	-0.052
bbc.com	-0.080	-0.134	-0.066	-0.034
belfastlive.co.uk	-0.075	-0.010	-0.063	-0.042
birminghammail.co.uk	-0.057	-0.105	-0.063	-0.041
cambridge-news.co.uk	-0.061	-0.013	-0.049	-0.057
dailymail.co.uk	-0.077	-0.095	-0.064	-0.050
dailyrecord.co.uk	-0.090	0.038	-0.066	-0.056
express.co.uk	-0.014	NaN	-0.012	-0.195
glasgowtimes.co.uk	-0.075	-0.087	-0.049	-0.040
grimsbytelegraph.co.uk	-0.064	-0.179	-0.059	-0.047
heraldscotland.com	-0.089	-0.114	-0.049	-0.050
huffingtonpost.co.uk	-0.089	-0.136	-0.076	-0.048
irishmirror.ie	NaN	NaN	0.165	NaN
manchestereveningnews.co.uk	-0.079	-0.083	-0.058	-0.055
mirror.co.uk	-0.081	-0.136	-0.069	-0.074
nbcsportsboston.com	-0.076	0.095	-0.058	-0.110
necn.com	-0.089	-0.292	-0.060	-0.044
theargus.co.uk	-0.066	-0.099	-0.042	-0.044
theboltonnews.co.uk	-0.078	-0.151	-0.054	-0.055
thenorthernecho.co.uk	-0.064	-0.034	-0.050	-0.049
thepinknews.com	-0.098	NaN	-0.064	-0.001
walesonline.co.uk	-0.079	-0.101	-0.070	-0.055
wired.co.uk	-0.203	NaN	-0.069	-0.058
yorkpress.co.uk	-0.084	-0.057	-0.060	-0.052

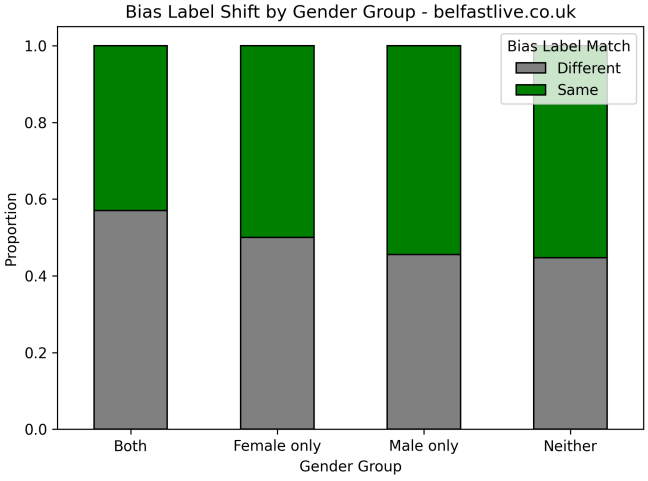

Bias Shift by Gender Group (bbc.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

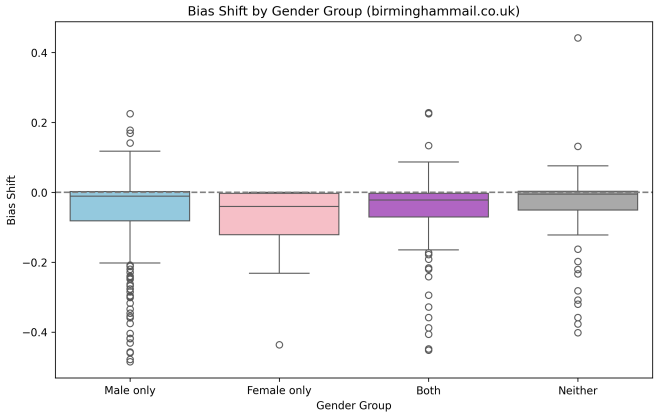

Bias Label Shift by Gender Group - bbc.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

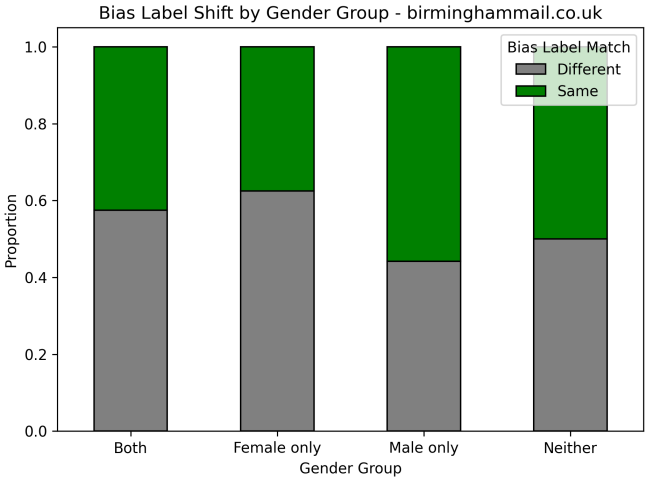

Bias Shift by Gender Group (bbc.com)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

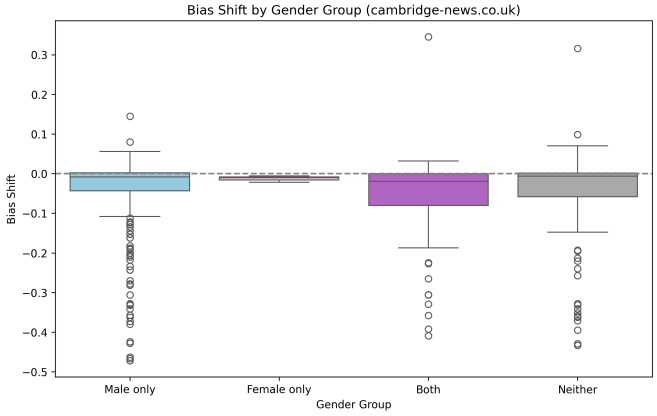

Bias Label Shift by Gender Group - bbc.com
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

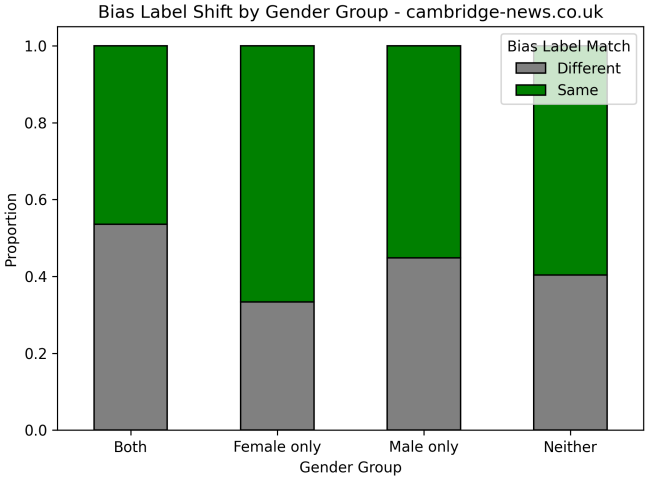

Bias Shift by Gender Group (belfastlive.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

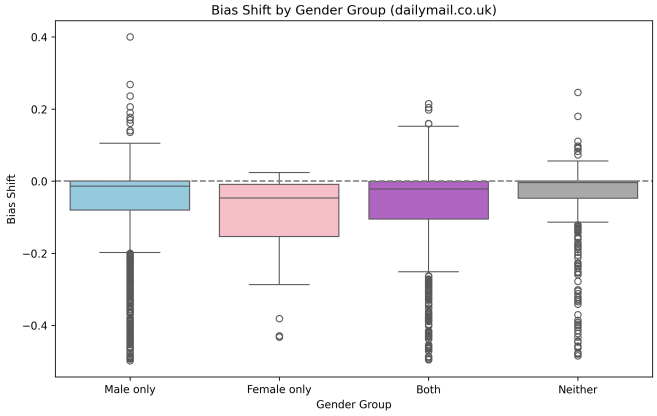

Bias Label Shift by Gender Group - belfastlive.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

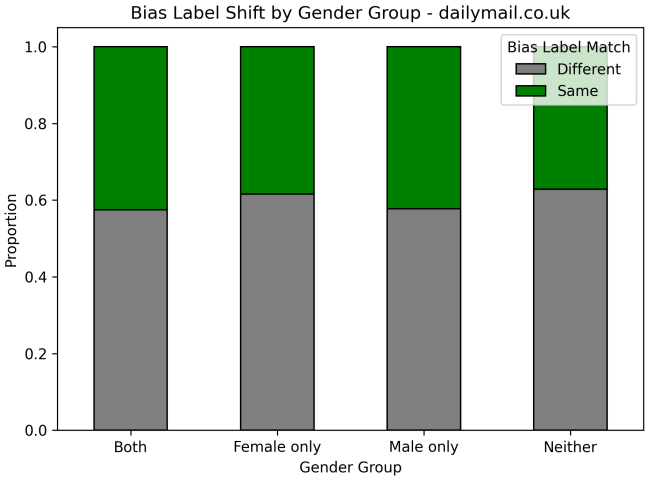

Bias Shift by Gender Group (birminghammail.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

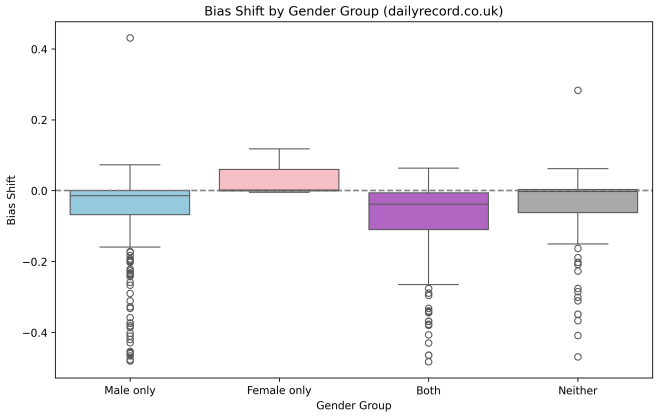

Bias Label Shift by Gender Group - birminghammail.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

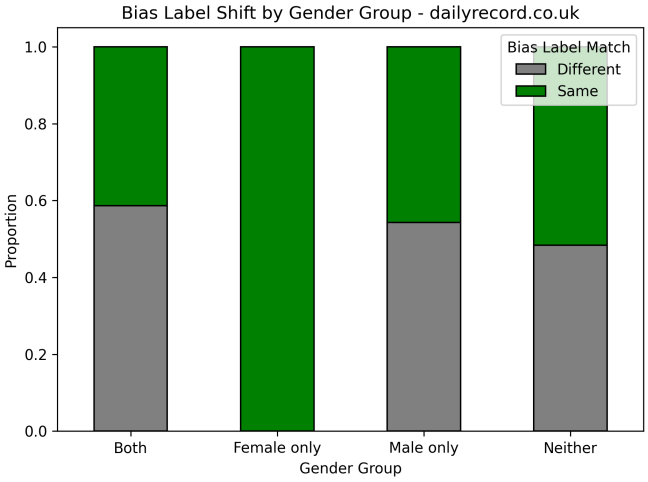

Bias Shift by Gender Group (cambridge-news.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

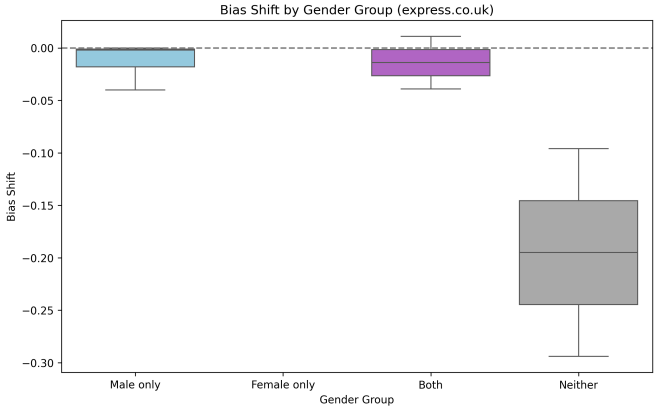

Bias Label Shift by Gender Group - cambridge-news.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

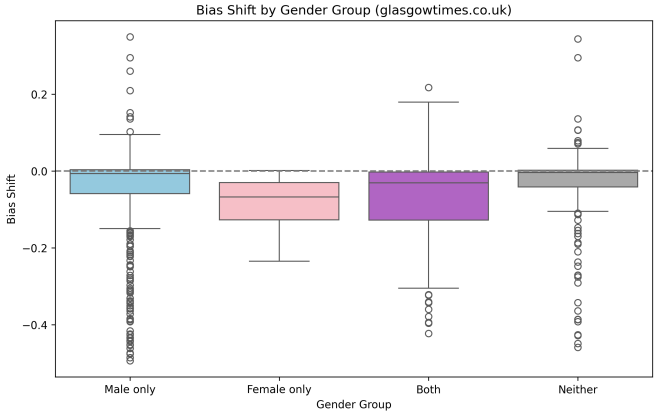

Bias Shift by Gender Group (dailymail.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

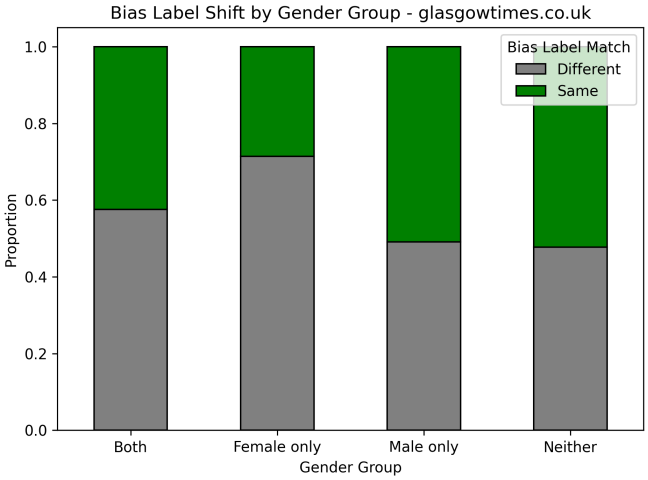

Bias Label Shift by Gender Group - dailymail.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

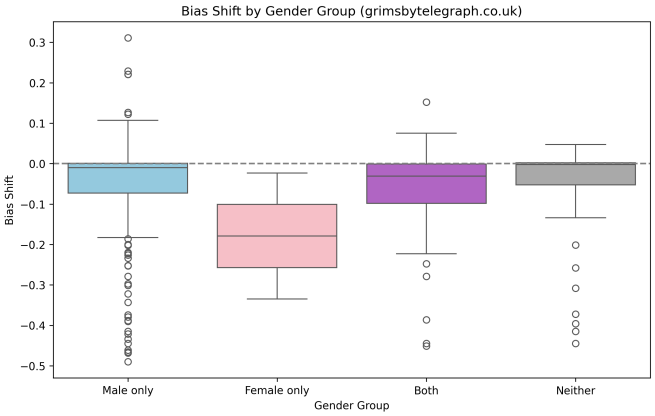

Bias Shift by Gender Group (dailyrecord.co.uk)
Bias Shift
0.4
0.2
0.0
−0.2
−0.4
Male only
Female only
Both
Neither
Gender Group

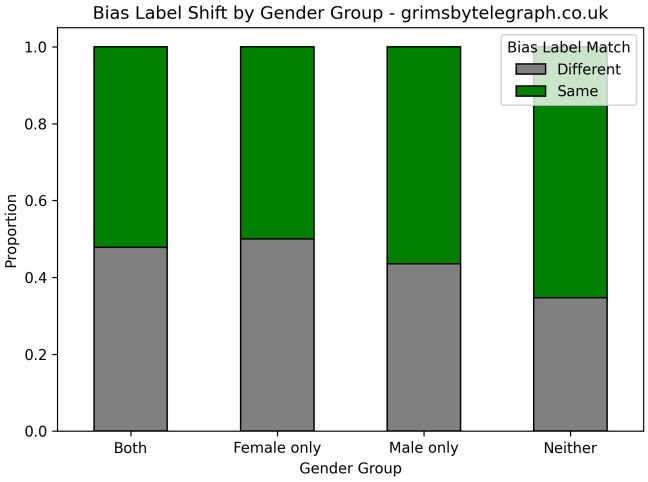

Bias Label Shift by Gender Group - dailyrecord.co.uk
Bias Label Match
Different
Same
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
Both
Female only
Male only
Neither
Gender Group

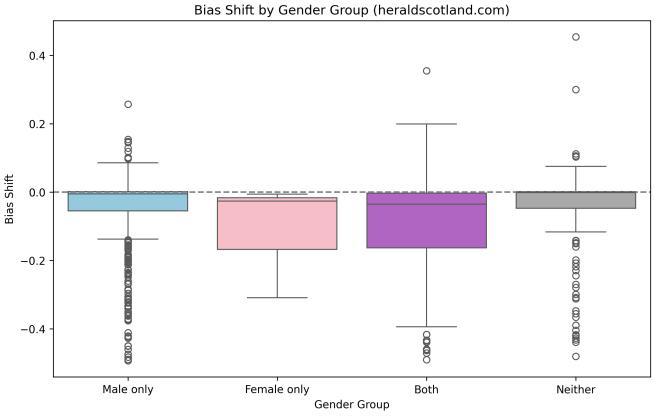

Bias Shift by Gender Group (express.co.uk)
Bias Shift
0.00
−0.05
−0.10
−0.15
−0.20
−0.25
−0.30
Male only
Female only
Both
Neither
Gender Group

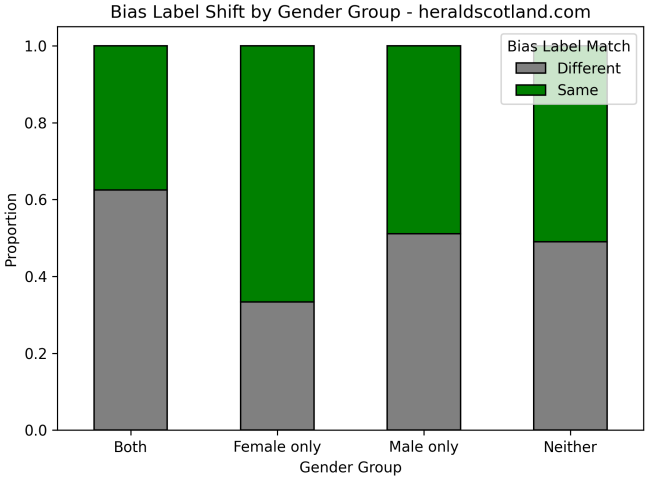

Bias Shift by Gender Group (glasgowtimes.co.uk)
Bias Shift
0.2
0.0
−0.2
−0.4
Male only
Female only
Both
Neither
Gender Group


Bias Label Shift by Gender Group - glasgowtimes.co.uk
Bias Label Match
Different
Same
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
Both
Female only
Male only
Neither
Gender Group

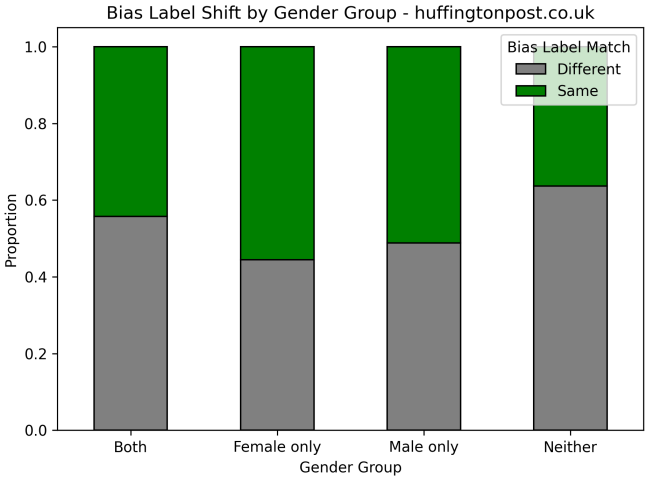

Bias Shift by Gender Group (grimsbytelegraph.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

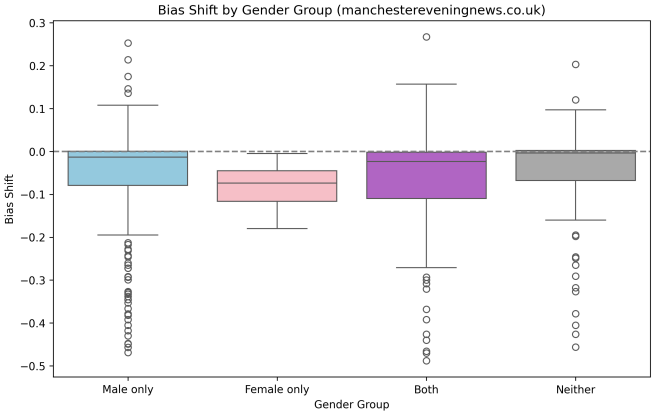

Bias Label Shift by Gender Group - grimsbytelegraph.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

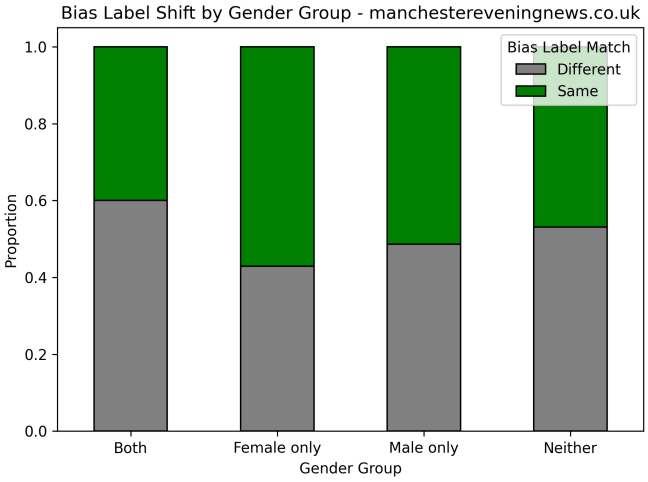

Bias Shift by Gender Group (heraldscotland.com)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

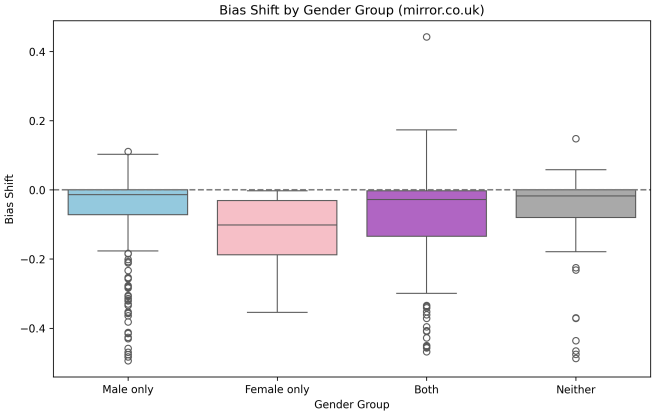

Bias Label Shift by Gender Group - heraldscotland.com
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

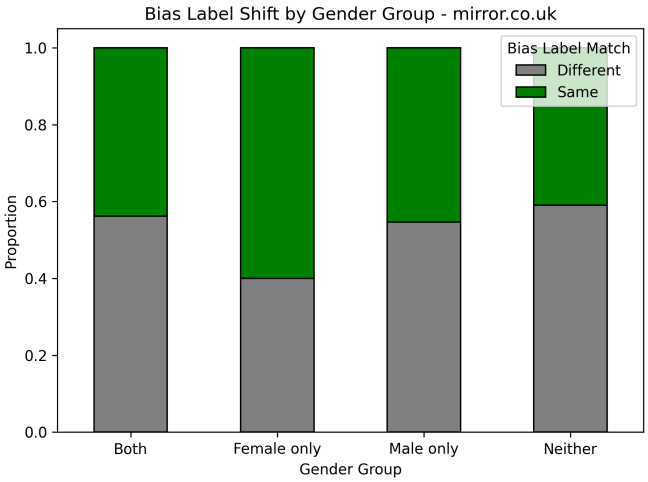

Bias Shift by Gender Group (huffingtonpost.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

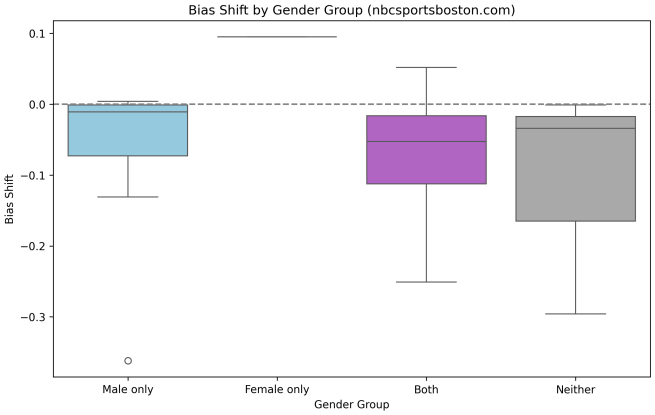

Bias Label Shift by Gender Group - huffingtonpost.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

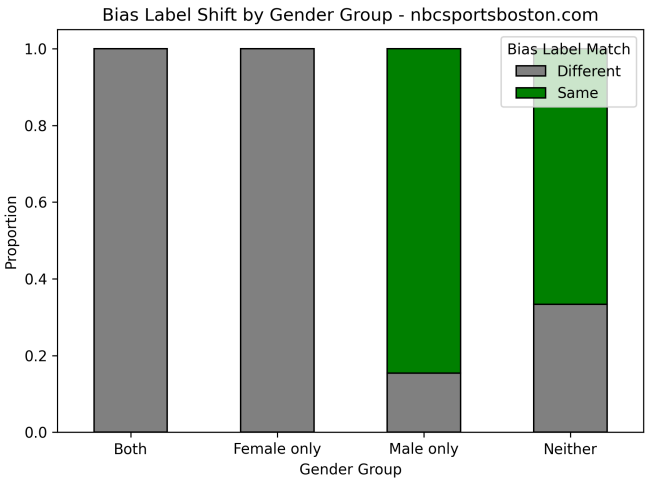

Bias Shift by Gender Group (manchestereveningnews.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group


Bias Label Shift by Gender Group - manchestereveningnews.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

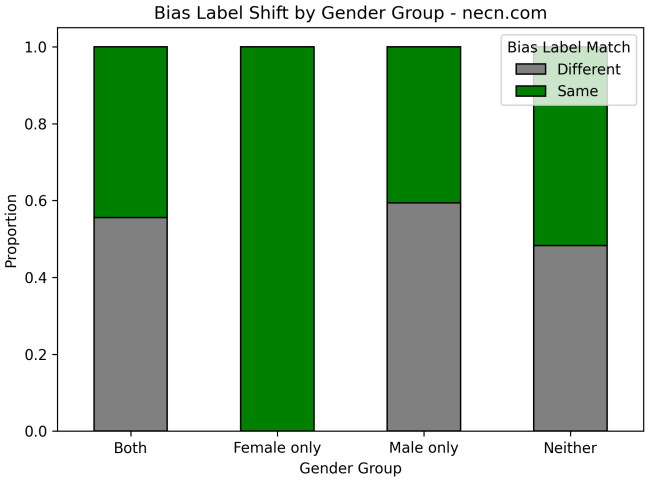

Bias Shift by Gender Group (mirror.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

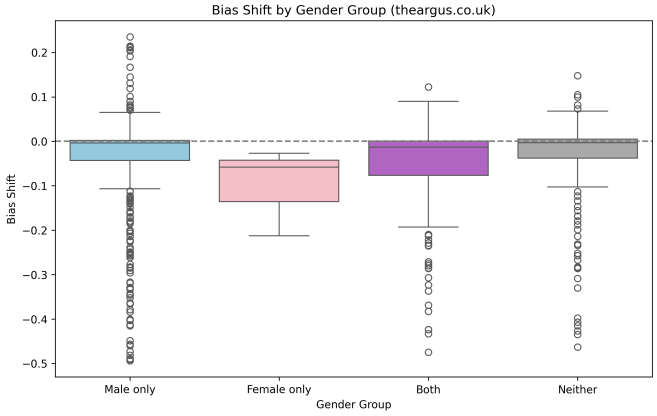

Bias Label Shift by Gender Group - mirror.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

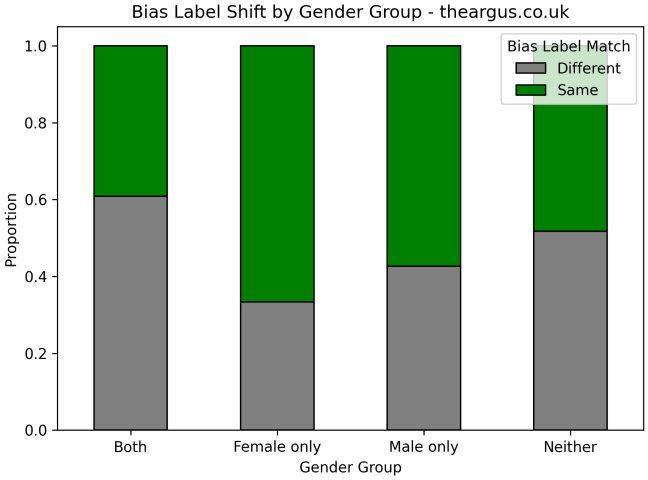

Bias Shift by Gender Group (nbcsportsboston.com)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

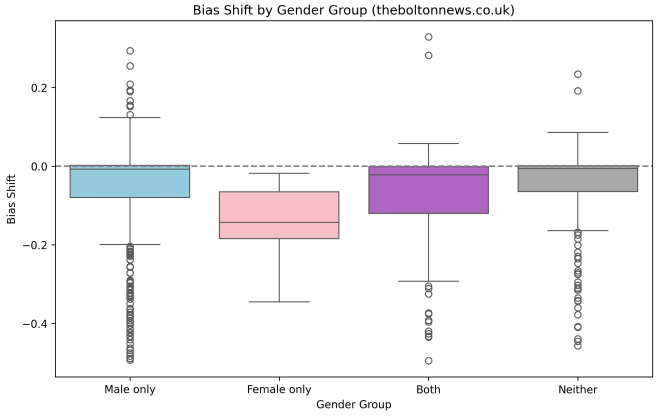

Bias Label Shift by Gender Group - nbcsportsboston.com
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

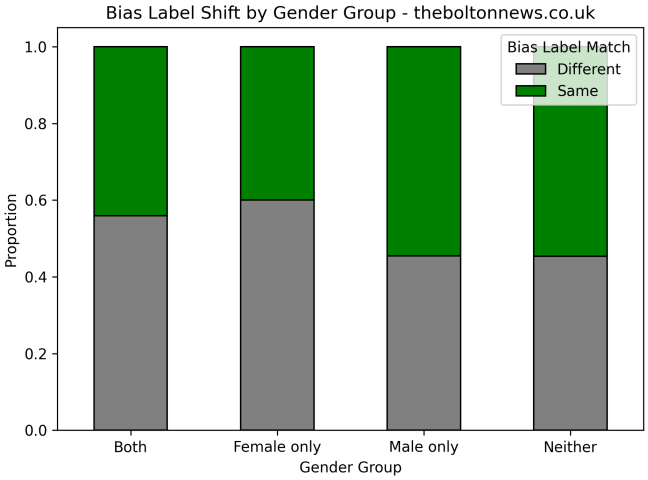

Bias Shift by Gender Group (necn.com)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

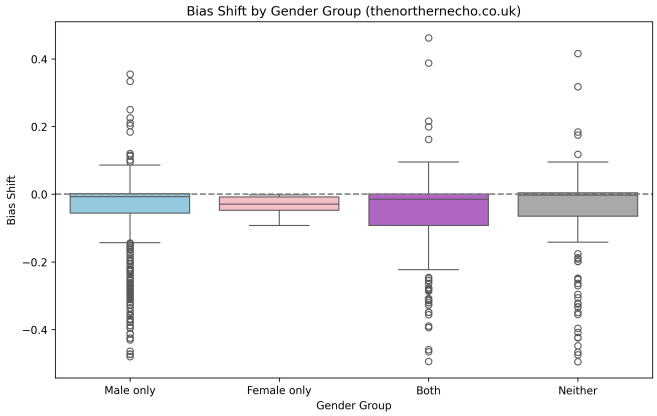

Bias Label Shift by Gender Group - necn.com
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

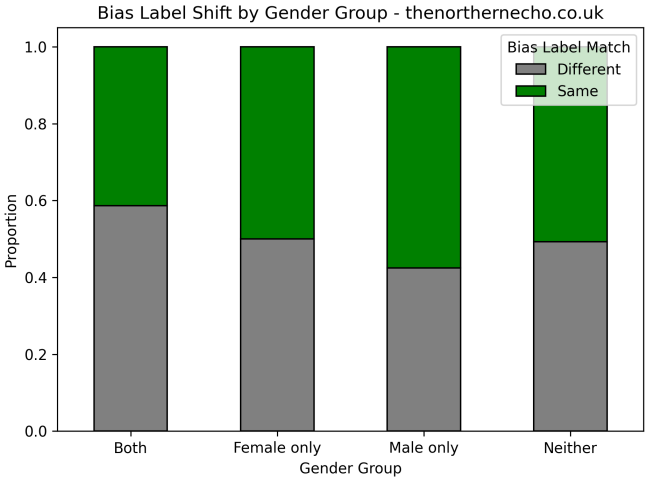

Bias Shift by Gender Group (theargus.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

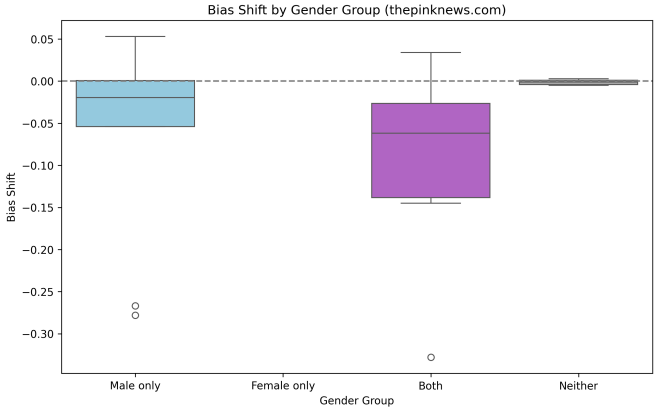

Bias Label Shift by Gender Group - theargus.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

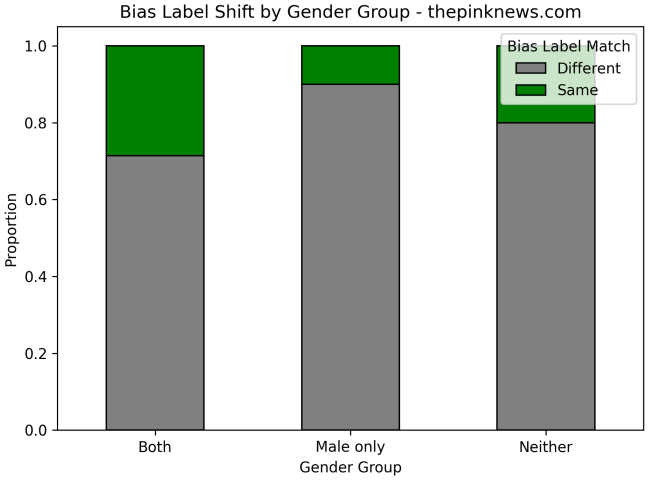

Bias Shift by Gender Group (theboltonnews.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

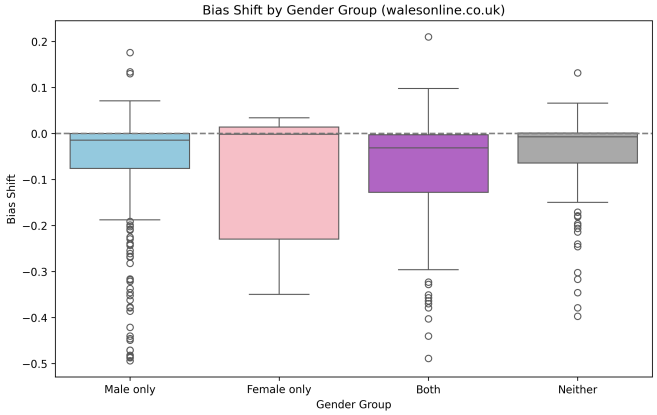

Bias Label Shift by Gender Group - theboltonnews.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

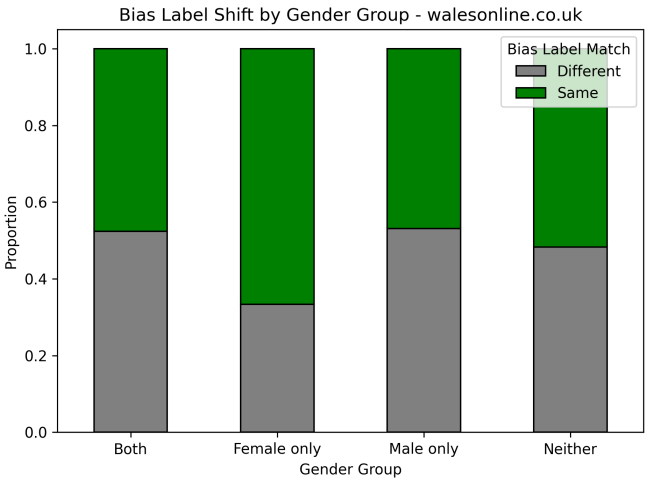

Bias Shift by Gender Group (thenorthernecho.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

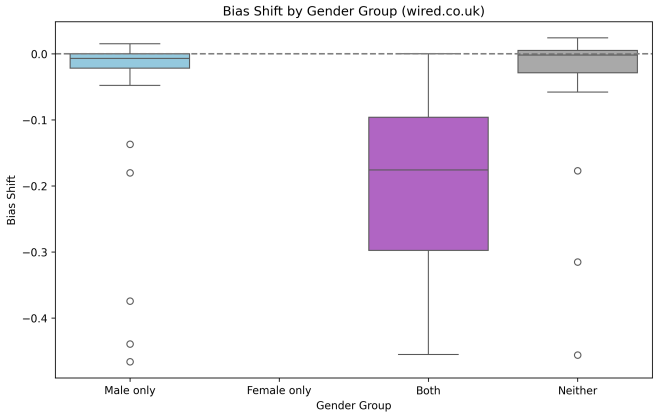

Bias Label Shift by Gender Group - thenorthernecho.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group

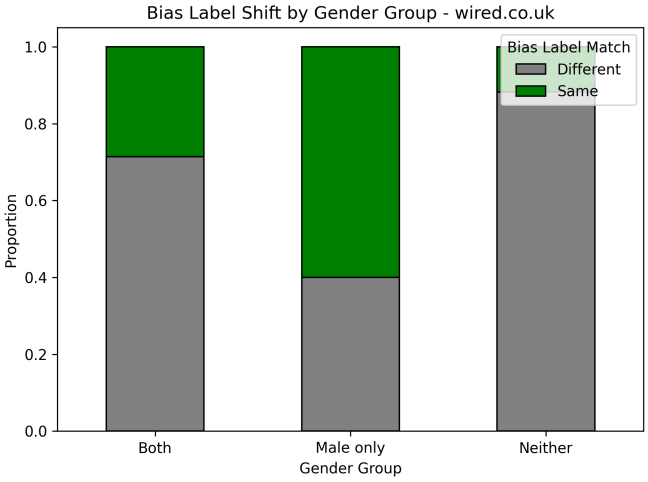

Bias Shift by Gender Group (thepinknews.com)
Bias Shift
Male only
Female only
Both
Neither
Gender Group

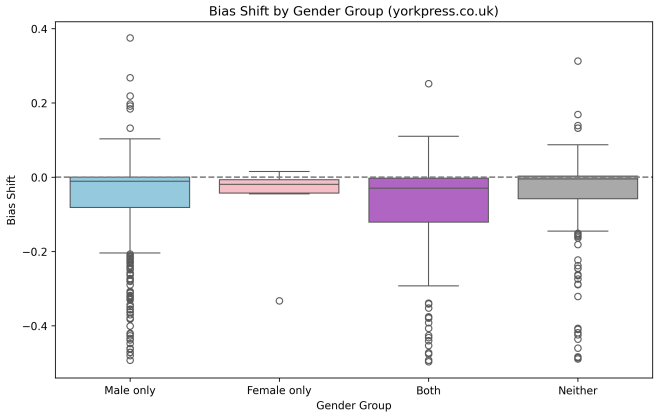

Bias Label Shift by Gender Group - thepinknews.com
Bias Label Match
Different
Same
Proportion
Both
Male only
Neither
Gender Group

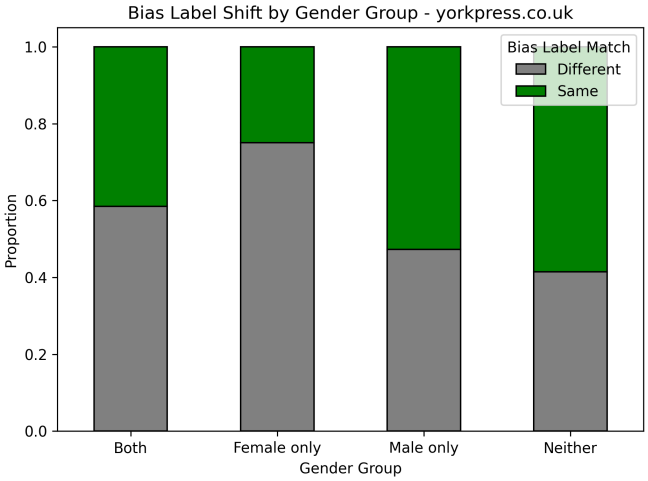

Bias Shift by Gender Group (walesonline.co.uk)
Bias Shift
Male only
Female only
Both
Neither
Gender Group


Bias Label Shift by Gender Group - walesonline.co.uk
Bias Label Match
Different
Same
Proportion
Both
Female only
Male only
Neither
Gender Group


Bias Shift by Gender Group (wired.co.uk)
Bias Shift
0.0
−0.1
−0.2
−0.3
−0.4
Male only
Female only
Both
Neither
Gender Group


Bias Label Shift by Gender Group - wired.co.uk
Bias Label Match
Different
Same
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
Both
Male only
Neither
Gender Group


Bias Shift by Gender Group (yorkpress.co.uk)
Bias Shift
0.4
0.2
0.0
−0.2
−0.4
Male only
Female only
Both
Neither
Gender Group


Bias Label Shift by Gender Group - yorkpress.co.uk
Bias Label Match
Different
Same
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
Both
Female only
Male only
Neither
Gender Group