

Women who break the glass ceiling get a “paper cut”: Gender, fame, and media sentiment

Eran Shor

McGill University

Arnout van de Rijt

European University Institute

Vivek Kulkarni

Stanford University

Abstract

Past *quantitative* studies have shown that most media coverage is of men. Here we ask if the scarce coverage that women get is *qualitatively* different from that of men. We use computer-coded sentiment scores for 14 million person names covered in 1,323 newspapers to investigate the three-way relationship between gender, fame, and sentiment. Additional large-scale data on occupational categories allow us to compare women and men within the same profession and rank. We propose that as women’s fame increases their media coverage becomes negative more quickly when compared to men (a “paper cut”), because their violation of gender hierarchies and social expectations about typical feminine behavior evokes disproportionate scrutiny. We find that while overall media coverage is much more positive for women than for men, this difference disappears and even reverses at higher levels of fame. In encyclopedic sentiment data we find no biographic basis for women’s disproportionate decline in media coverage sentiment at high fame, consistent with the conjectured double standard in media discourse.

1. Introduction

The sentiment of media coverage has significant career and life consequences for individuals. Negative coverage of politicians may translate into fewer campaign donations and negatively affect their ability to draw voters (Aaldering, van der Meer and van der Burg 2018, Heldman, Carroll and Olson 2005, Kahn 1994, Schlehofer et al. 2011). Similarly, negative coverage of businesspeople, in particular entrepreneurs and business owners, may hurt their ability to develop their business or maintain its public image and commercial stance (Baker, Aldrich and Nina 1997). For managers, it might hurt promotion prospects and future employability. Finally, for athletes, artists, book authors, and various entertainers, negative media coverage might limit employment opportunities and reduce salaries, royalties, commercial opportunities, and the sale of their products.

Research on women in the workforce has long established the existence of a glass ceiling, where in various fields women are upwardly mobile up to a certain rank below the top at which they stagnate (Alessio and Andrzejewski 2000, Cotter et al. 2001, Ridgeway 2011). Here we ask what happens to the ways in which the media covers women when they do manage to break through this glass ceiling and reach positions of power and high status. That is, we probe the three-way relationship between gender, fame, and coverage sentiment.

We examine this question by analyzing two unique large datasets. The first was generated by Lydia—a computerized text-analysis system, which collected both frequency and sentiment data on millions of unique person names from 1,323 US newspapers over a period of six years (Bautin et al. 2010). We manually assembled the second dataset from Wikipedia categories on about 50,000 well-known men and women in prominent social and occupational domains, including politics, business, entertainment, sports, science, and crime. Combining these two datasets allows us to compare the sentiment toward women and men at different coverage frequencies (i.e. degrees of fame).

We argue that when women's fame increases, rather than celebrating their achievements with favorable coverage, media scrutinize them more closely, ready to find blemishes and faults in their performance. This happens because successful and prominent women pose a threat to traditional gendered status hierarchies and stereotypical ways of thinking about femininity and gender roles. This supposition, in turn, leads us to expect an interaction effect of gender and fame on sentiment: As women become more famous, their media coverage trends negative more quickly when compared to men.

We refer to the disproportionately negative media coverage that famous women receive after breaking the glass ceiling as a "paper cut". Such paper cuts can be deep and surprisingly painful, with substantial repercussions for careers and personal lives. A paper cut originates from newspaper journalists and editors who subscribe to common social perceptions, stereotypes, and hierarchies. A paper cut is to be distinguished from negative media coverage that reflects a "glass cut," when media merely mirror real-world events, reporting on unfavorable career and personal life consequences that result from women breaking the glass ceiling.

We argue that famous women receive this paper cut in all major domains of media coverage, despite obvious differences in the media coverage contexts of politicians, businesspeople, criminals, and athletes. The media coverage of female politicians or businesswomen will become more negative as they move up the political and corporate ladder because this increases the unconventionality of their candidacy or position in the firm and puts their femininity under intense scrutiny. Women who commit serious crimes that render them infamous, such as murder, are perceived as doubly deviant because they transgress both the law and standard feminine behavior. And for elite female athletes, their sexual normativity and commitment to motherhood and family are often questioned.

Our analyses confirm the prediction of a paper cut, although the strength of the evidence varies by domain. We find that overall women receive much more positive coverage when

compared to men, but this difference disappears and even reverses when they become famous. By contrast, women's biographies are more positive when compared to those of men across the board, including when they are very famous. This contrast between biographic and media sentiment suggests that famous women's more negative media coverage is not driven by them engaging in negative actions, experiencing negative events, or selecting into occupations or situations of a negative nature, i.e. a glass ceiling cut. The evidence instead points to famous women in most domains paying a media premium, i.e. a paper cut.

2. Gender, fame, and media coverage: Theoretical expectations

Anecdotal evidence suggests that successful women in various fields may be receiving relatively negative coverage when compared with less successful women or with equally successful men. Women who reach professional success would often be painted in negative and unflattering ways, as the media carefully scrutinize their looks and actions, trivializing or undermining them. Working women's behaviors and demeanors are often painted as "unladylike" and irrational and they are described as "catty", "bossy", "bitchy", "ball-busting" "hysterical" and "demanding" (Barden 1996, Falk 2010, Fowler and Lawless 2009, Goddu 1999, Joo 2002, Saner 2014, Sanghani 2014). According to Halvorson (2015), successful women are often portrayed in the media as either competent but cold (also "bossy", "bitchy", "pushy", "frigid", and "ball-busting"), or as warm but incompetent, illogical, and irrational; the doormat who no one takes seriously (also "ditsy", "silly", "airhead", or "emotional").

However, while these accounts may resonate with common public perceptions, there is a lack of systematic large-scale research on the media coverage sentiment of women and men, in particular as they become increasingly successful and well-known. How might fame be associated with the coverage sentiment of women? We are unaware of any previous study that has theorized the complex relationship between gender, fame, and media sentiment, carefully identifying the

possible mechanisms that may shape it and testing them systematically. We therefore offer a new theoretical account, identifying two potential mechanisms that may generate differential media sentiment toward successful women. First, gendered coverage sentiment patterns may be due to media covering events or achievements differently when they happen to a woman than when they happen to a man (section 2.1). Specifically, women may receive disproportionately negative media coverage – a paper cut – when they become very famous, because of gender norms held by media actors and audiences. Second, different events may happen to women and men, with the differential media coverage merely mirroring those different events (section 2.2).

2.1. A paper cut: Women’s success as a threat to gendered status hierarchies and norms of ideal femininity

Differential media treatment of successful women may directly originate from the media itself and to common social perceptions that may affect media coverage. However, it is not obvious whether to expect news coverage of women to be relatively more or less favorable than that of equivalent men. On the one hand, it is possible that at least some journalists and editors practice what Glick and Fiske (1996) call “benevolent sexism”—viewing women stereotypically and in restricted roles, but in ways that are subjectively positive in feeling tone, emphasizing behaviors and traits typically categorized as prosocial, such as helping, showing compassion, or intimacy-seeking. While such tendencies for “benevolent sexism” may be countered by more traditional sexism, or what Glick and Fiske call “hostile sexism”, we suggest that they may nevertheless lead to relatively positive overall coverage for women in the news.

Journalists may also wish to celebrate the relative achievements of women, seeking to show that women can make it despite encountering multiple difficulties. Thus, they would tend to focus on flattering and positive features and stories, while disregarding or downplaying negative elements in successful women’s life events, personalities, or behaviors. Some have argued that even journalists

who hold a bias against women might be inclined to soften judgmental and unflattering descriptions, trying to prove that they are in fact not biased or being sensitive to critiques about their gendered coverage tendencies (Smith 1997). These arguments suggest the following net expectation:

H1. Women receive more positive media coverage when compared to men.

At the same time, media sociologists and mass communication scholars have emphasized the strength and persistence of masculine norms in newsrooms and editorial boards. The historical dominance of men in editorial positions has led to coverage norms that favor men, their actions, and their opinions. (Mills 1997, Rodgers and Thorson 2003, Ross 2007, Ross and Carter 2011). Traits perceived as “masculine”, such as competitiveness, assertiveness, and dominance, are often celebrated and commended in fields like sports, politics, and business. These masculine norms of journalistic practice have come to be regarded simply as professional routines, to which all journalists are expected to subscribe (Ross 2009, Ross and Carter 2011, Van Zoonen 1988). Consequently, these coverage norms persist when women reach lead editorial positions in media outlets (De Swert and Hooghe 2010, Mills 1997, Shor et al. 2015). We argue that such persistent masculine norms in media organizations may lead to disproportionately negative media coverage of women *when they become successful and famous*. Psychologists and sociologists of gender have argued that masculine norms lead women in positions of leadership and notoriety to be devalued relative to men in top positions. We argue that this devaluation could then lead journalists to comment more negatively on famous women.

According to Eagly and Karau (2002) there is a perceived incongruity between female gender roles and leadership roles. This gap, in turn, results in public and journalistic perceptions of women as less suitable for leadership roles and in less favorable performance evaluations of

women as leaders. (See also the meta-analysis by Koenig et al. (2011), suggesting a common stereotypical perception of most leadership roles as masculine).

Ridgeway (1997, 2009, 2011, 2013) has further argued that status biases affect individuals' willingness to evaluate women's actions, words, efforts, and accomplishments in a positive way. Consequently, men are often perceived as simply better and more capable of performing valued social tasks and serving in high status positions than are women. This, in turn, could translate into media practices that significantly exacerbate and artificially magnify existing inequalities between men and women (De Swert and Hooghe 2010, 2010, Zoch and Turk 1998).

According to some, the glass ceiling is not merely a transparent barrier that prevents women from advancing. As women move up the ranks, they experience other disadvantages as well (Baxter and Wright 2000, Morgan 1998, Prokos and Padavic 2005). For the domain of politics, Fowler and Lawless (2009) argue that when women run for office, they break gender stereotypes and norms about both women and politicians. Meeks (2012) has argued that media coverage becomes more negative as women move up the political ladder, because this increases the unconventionality of their candidacy and puts their femininity under intense scrutiny. More generally, women who manage to reach the very top positions in politics, business, science, literature, sports, and the arts may fail to receive coverage that is as respectful and positive as that of their male counterparts, mainly because their career advancement poses a challenge to social status hierarchies (Ridgeway 2013, Rudman et al. 2012). Negative perceptions should be especially likely when successful women are perceived as too assertive or domineering, as such "masculine" behaviors breach gender roles and status hierarchies, threatening the gendered status quo. In contrast, men who reach high status positions would likely receive more positive coverage because they are perceived as merely fulfilling their social role and enforcing these same status hierarchies. Ridgeway (2013) further argues that when gendered status hierarchies and distinctions appear to be challenged (e.g. following reports that women may be closing the

gaps in some domains), this increases anti-feminist sentiments. We argue that if media actors are susceptible to these stereotypical ways of thinking about successful women, then this should translate into more negative coverage of women once they become successful and famous.

In sum, the combination of masculine newsroom norms and social norms about gender hierarchies and ideal femininity should lead to worse media sentiment for women who threaten these norms by having career success and, consequently, becoming famous. While women's overall coverage may be more positive (H1), as they become famous the sentiment of their media coverage should decline more rapidly than men's:

H2: There is an interaction effect of gender and fame on *media* sentiment, with women's coverage becoming relatively more negative at higher levels of fame.

2.2. Media as a mirror: Personal and organizational factors

An alternative source of gender differences in media sentiment could be differences in the actual careers and public lives of male and female individuals. Differential sentiment in coverage of women and men may accurately reflect real differences in women's and men's accomplishments, evaluations, and the events that occur in their lives stemming from personal and organizational factors. With regards to the overall difference in sentiment between women and men, these factors would predict more positive media coverage for women. First, women are less commonly involved in criminality and are more likely to occupy professions that garner generally positive sentiment, such as teaching or providing care for children and the elderly. Women with successful careers arguably must be more talented, ambitious, and hard-working than equivalent men as they must overcome institutional barriers against women's advancement. For example, women in politics may be subjected to more stringent selection and promotion processes (Jalalzai 2008, Palmer and Simon 2008). It therefore stands to reason that as they move up the

organizational hierarchy, those women who remain in the pool of potential candidates for promotion will be more qualified than their male counterparts. Recent studies on female politicians show that on average they indeed tend to be of higher quality, work harder, and perform better than their male colleagues (Bauer 2020, Fulton 2012, Lazarus and Steigerwalt 2018).

Some recent research further suggests that women in leading executive positions in business may offer advantages to their firms, including improved firm performance, though evidence is not unequivocal (Dezso and Gaddis Ross 2012, Khan and Paolo Vieito 2013, Weber and Zulehner 2010). In other words, women who manage to reach high-level occupational positions may be exceptionally talented, charismatic, determined and hard-working. The traits and behaviors that help them overcome a hostile male-dominated environment and lead their organizations to better achievement might, in turn, also be reflected in more positive media coverage.

These arguments about personal and organizational factors lead to the same prediction as the one derived from arguments of media bias (section 2.1) that women would enjoy more positive media coverage (H1). At the same time, if the media act merely as a mirror, then women's public lives truly contain more positive events. This greater positivity should then also be reflected in women's biographies, such as those written on their Wikipedia page.

H3. Women's biographies are more positive when compared to those of men.

At the same time, organization scholars have suggested that women who reach high-end positions in various organizations appear to be less competitive, ambitious, confident, and assertive when compared to similarly positioned men (Castagnetti and Rosti 2013, Fox and Lawless 2014, Manning and Saidi 2010). Gendered socialization, prevailing gender norms, and status beliefs would keep women from self-promotion, bragging about their achievements, or deliberately putting themselves in the limelight. Women also learn that they are held to a double standard, as self-promoting behaviors

could be construed as non-feminine. Consequently, even when women do reach elite social and occupational positions, they might fail to promote themselves as much as men or they might lack (or suppress) qualities such as overt competitiveness, confidence, and assertiveness. Since these qualities are often viewed as necessary and beneficial for such high-end positions, this might consequently result in the assessment that the women are not sufficiently qualified for these positions. This would suggest that while media coverage of women overall may be more positive, this is less the case for women in top organizational positions.

Organizational scholars have also reported on another common pattern that may result in worse media coverage for relatively successful women. Studies have found that women and minorities are more likely to be promoted to risky and precarious leadership positions, with higher turnover rates and professional instability, a phenomenon that some have referred to as a “glass cliff” (Cook and Glass 2013, O'Brien 2015, Ryan and Haslam 2005). When organizations are in crisis, for example in the form of a scandal, very poor financial performance, or a political party losing its seat share, they often look for a change of pace. Consequently, they may be more likely to promote someone who is not their typical executive, i.e. not a white man. The argument is that women and minorities, possibly because they feel like this is their only shot at reaching an executive position, may accept the offer even when it is risky, while qualified white men are more likely to turn it down and look for safer opportunities. Consequently, women’s likelihood of being associated with failure and reputational risk would be higher, as they would be more often handed a mess to clean up. With the cards stacked against them, they fail to save the day, which, in turn, would lead to worse media coverage.

These arguments then lead to the same prediction as the one produced by the paper cut argument: an interaction effect of gender and fame on media sentiment (H2). At the same time, if the mechanism driving this interaction effect is the media acting as a mirror, then famous women’s worse media sentiment does not stem from media bias but rather reflects a disproportionate frequency of negative events occurring in famous women’s public lives. These negative events are

part of their biographies such as those on their Wikipedia page. For example, on July 5, 2021, former UK Prime Minister Theresa May's Wikipedia biography included negative phrases describing her failures to win a vote of confidence in parliament, her Brexit deal defeat, and her resignation. Similarly, Marissa May became the CEO of a troubled Yahoo! and her Wikipedia biography documents the continued difficulties the company experienced during her tenure from 2012-2017. Thus, if women indeed select (or are being selected into) relatively tough jobs at top levels—consistent with the concept of a glass cliff—then we should also see an interaction effect of gender and fame on sentiment in encyclopedic biographies:

H4: There is an interaction effect of gender and fame on *biographical* sentiment, with women's biographies becoming relatively more negative at higher levels of fame.

3. Gender, fame, and media coverage: Existing evidence

In this section we review shortly former research findings on the coverage of successful women in key media domains, including politics, business, entertainment, crime and sports.

Researchers have devoted substantial attention to the media representation of successful women in *politics*. Most of this research has focused on a handful of prominent political figures, analyzing their coverage and comparing it to the coverage of male politicians in countries such as New Zealand, Australia (van Acker 2003), and Canada (Gidengil and Everitt 2003). This research suggests that male politicians who appear commanding, competitive, or resolute are commended for these traits, while female politicians who show similar traits are portrayed in a negative light. On the other hand, when women displayed traits often perceived as feminine, such as vulnerability and non-competitiveness, they also failed to receive positive coverage (see also Henderson 1999). Heldman et al. (2005), who examined Elizabeth Dole's presidential campaign in the US, found that negative stereotypes are most obvious and harmful for women

who enter a presidential race (see also Huddy and Terkildsen 1993, Kahn 1992, Kahn 1994). Hillary Clinton's media coverage similarly became increasingly negative with her growing political involvement and her move from First Lady, to Senator, to Secretary of State, and to the Democratic candidate for presidency (Ryan 2013, Scharrer 2002). Studies in New Zealand have likewise found that when women reach top political positions their lives become heavily scrutinized and sensationalized by the media, more so than the lives of men (Comrie 2006, McGregor 1996, Trimble and Treiberg 2010).

More recently, a content analysis of Dutch national newspapers provided some larger scale evidence for the field of politics, reporting that male politicians received more media coverage on leadership traits when compared with female politicians. The latter received more coverage of their appearance and personal life, more negative coverage of their political viability, and more stereotypical trait coverage (Aaldering and Van der Pas 2020, Van der Pas and Aaldering 2020). In another recent cross-national study, Shor and Miltsov (2020) found that growth in the parliamentary representation of women was associated with more negative coverage to women in various countries.

Another domain in which women appear to receive particularly negative coverage is *crime*. While all criminals naturally tend to receive unfavorable coverage, the literature suggests that coverage for women who commit crimes is particularly negative (Collins 2016, Naylor 1990). Women who kill are portrayed by newspapers as an aberration of true womanhood (Creed 1996); as 'bad' (i.e. monsters), 'mad' (i.e. insane), or 'sad' (i.e. weak and helpless) (Cavaglioni 2008, Eastal et al. 2015). Common depictions in such cases include describing the women as irrational, immature, and emotionally unstable (Barnett 2005, Cavaglioni 2008, Huckerby 2003), man-hating, unattractive, unfeminine, lesbian vampires, sexual deviants (Berrington and Honkatukia 2002, Creed 1996, Farr 2000, Naylor 2001), evil, manipulative, cold-blooded monsters (Berrington and Honkatukia 2002, Hinds and Stacey 2010, Wilczynski 1991), and

inadequate/unnatural/bad mothers and wives (Barnett 2005, Huckerby 2003, Morrissey 2003). These depictions are particularly salient when women commit serious crimes (as opposed to more minor crimes) and seem particularly prominent in high-profile murder cases that receive substantial media attention (Grabe et al. 2006, Weimann and Fishman 1988). Again, however, evidence is mostly small-scale and qualitative, raising the question whether these findings hold up in systematic analysis of comprehensive newspaper data.

The representation of successful *businesswomen* has received somewhat less attention by media scholars, with most scholars focusing on the volume of coverage, rather than its tone (Grandy 2013, Greenwald 1990, Shor et al. 2014b, Shor et al. 2015, Shor, Van de Rijt and Fotouhi 2020). The few studies that did examine coverage tone have reported mixed results, with successful businesswomen on the one hand presented as heroic entrepreneurs and charmers, but on the other hand depicted as women who have transgressed women's "natural place" and as mad and cunning (Czarniawska 2004, Czarniawska 2008). As with successful politicians, coverage of businesswomen and female CEOs tends to focus on their family, marital status, and attire rather than on their position and work status (Bjursell and Backvall 2011, Krefting 2002, McGregor 2000). Yet, these depictions often carry a rather positive tone, such as when a woman is complemented on being able to balance family and work commitments, or presented as more nurturing, caring, and compassionate, in life and in business (Bjursell and Backvall 2011). Most recently, Bishop Smith, Chown and Gaughan (2021) found that the appointments of female CEOs do not receive more negative media coverage than those of male CEOs, except for when the appointment was given a lot of media attention, which is consistent with our paper cut hypothesis.

Finally, with regard to *sports* qualitative studies have found that whereas reporting on male athletes generates excitement, reporting on female athletes is often matter-of-fact and dull (Cook and Glass 2013, Cooky, Messner and Musto 2015, Fink and Kensicki 2002, Shugart 2003, Vincent et al. 2007). Female athlete's sexuality, sexual normativity, and their commitment to their

family are also more likely to come under scrutiny, in particular when they practice sports that are considered as less gender-appropriate (Fink and Kensicki 2002, Shugart 2003, Waldon 2005).

While the studies reviewed above advance our knowledge about the media representation of successful women, nearly all rely on anecdotal case studies, examining a small number of famous individuals (often only one or two). Such research designs raise concerns about selection bias and generalizability. In addition, much of the previous research has focused only on women, without systematically comparing media reports on men to those on women. Thus, it is hard to ascertain whether the suggested negative relationship between fame and coverage tone is indeed only (or especially) true for women, as some have suggested (Ridgeway 2013, Rudman et al. 2012) or whether this is also the case for men.

We therefore offer here both a comprehensive theory of the relationship between gender, fame, and media coverage tone, accounting for the various mechanisms that might drive this relationship, and a systematic large-scale test of this theory for millions of men and women across occupations and fields. While the literature above does not yield a strong hypothesis about the relationship between gender, fame and coverage sentiment in each domain, we argue that regardless of field, successful or prominent women pose a threat to traditional gendered status hierarchies and to stereotypical ways of thinking about femininity and gender roles. We therefore expect the interaction predicted in our H2, where more famous women should receive disproportionately negative coverage, to hold across domains.

4. Data and analytical strategy

4.1. Newspaper sample – Our primary data source for analyzing newspapers’ coverage volume and sentiment is the Lydia text analysis system (Bautin, Vijayarenu and Skiena 2008, Bautin et al. 2010, Shor et al. 2014a, Shor 2018). Lydia provides time-stamped records of occurrences of person-names for a large number of newspapers, magazines and online news sources from 2004

to 2009. This period allowed a unique opportunity for researchers to access large-scale longitudinal data from the online editions of a very large number of American newspapers, before most of these newspapers placed their content behind paywalls.

We follow work on the sociology of fame in operationalizing fame as the quantity of public attention, measured through mentions in the media. This operationalization measures both the status of individuals and the visibility they enjoy (Ferris 2007, van de Rijt et al. 2013). For each name occurrence, Lydia provides the timestamp, newspaper, and sentiment score (positive or negative) associated with the name. It computes sentiment scores by tracing the co-occurrence of a name and a sentiment word in the same sentence. While this process might produce some errors in complex sentences, the volume of text processed by Lydia and the aggregation over all usages of a name generates overall accurate sentiment scores (Godbole, Srinivasaiah and Skiena 2007), as demonstrated in the online Appendix 1. Lydia is also able to determine the gender of a named person through anaphora resolution with high accuracy. In the online *Appendix 2* we provide a detailed explanation of our gender classification process, as well as a table showing classification results for a random sample of names.

In the present analysis, we examine data from 1,323 US news outlets, including nationally-distributed newspapers, such as USA Today and the New York Times, as well as local newspapers, such as the Tulsa World and the Wichita Eagle). Lydia collected data on the sentiment associated with 4 billion mentions of 26 million person names of individuals who appeared in these newspapers over the six-year period of data collection. The comprehensiveness of the dataset is especially advantageous when attempting to test the relationship between coverage volume and coverage sentiment, as it allows a comparison of the coverage of a very large number of individuals who were mentioned only a few times in the news with the coverage of those mentioned hundreds of thousands of times.

Our empirical focus is on five domains: politics, business, crime, sports, and entertainment. These domains follow the news categories in the academic literature reviewed above and are also consistently among the most prominent sections in most newspapers. Additionally, these are domains for which Wikipedia categorizations offer comprehensive data and often provide exhaustive lists of various sub-domains, thus allowing us to examine each of these domains systematically.

4.2. Wikipedia data on individual names by field – While the Lydia system includes data on virtually all individuals who appeared in the news over the period of data collection, it is unable to tell us much about the characteristics of these individuals, beyond their coverage volume and the general section in which they appeared (e.g. news or entertainment). Moreover, names such as Donald Duck and Elizabeth Arden are falsely classified as person names, but cannot be pruned because of the scale of the dataset. We therefore manually assembled and coded a large complimentary database comprised of the names of individuals from Wikipedia, whom we are confident are real people. These individuals mostly enjoy at least a moderate level of public or professional recognition in various fields. Online *Appendix 3* shows examples of these names by level of fame.

In order to be able to compare men and women from equivalent occupational categories, we collected data from the Wikipedia categorization databases.¹ These databases classify individuals into categories and sub-categories. While for many categories, lists are not exhaustive, they nevertheless tend to capture the large majority of men and women who have made the most significant imprint in specific domains. Importantly, previous research has found that Wikipedia editors are predominately white and male (Antin et al. 2011, Collier and Bear 2012). This has the potential to introduce racial and gender underrepresentation into the contents of Wikipedia. However, while some studies suggest that women are underrepresented in certain categories of

¹ See <https://en.wikipedia.org/wiki/Wikipedia:Categorization>

Wikipedia, such as sociologists (Adams, Brückner and Naslund 2019) and engineers (White 2018), others report that relative to their share in various occupational domains, women are not underrepresented on Wikipedia and may in fact be slightly overrepresented (Wagner 2015, Wang, Pappu and Cramer 2021). While such potential misrepresentation may affect some of our analyses, for some categories, such as senators, lists are exhaustive and hence representative.

In order to collect our index of names, we first generated an initial list of larger social and occupational domains, based in part on common newspaper categorizations. These domains include politics, business, entertainment, sports, science, and crime. We then devised a list of important sub-domains within each of these larger categories. For example, within the domain of entertainment we identified the following sub-categories: Actors (TV and film), directors, singers, and dancers. Finally, for some domains, we identified more specific sub-categories, in which individuals are particularly likely to attract media attention (e.g. US senators in politics, Oscar nominees in entertainment, and Pulitzer Prize winners in literature).

Next, we merged this domain-specific data with our Lydia data, which provided the number of mentions and the coverage sentiment for the Wikipedia names. We analyzed the dataset of all 42,862 individuals who were both included in a relevant Wikipedia category and mentioned by the exact same name in newspapers. For each of these persons we have a newspaper occurrence frequency, a newspaper sentiment score, and an encyclopedic sentiment score. This dataset also allowed us to exclude common names from the analysis (e.g. Michael Smith), as these appear in Wikipedia with the area in which they became famous (e.g. poet) marked in parentheses. This exclusion is important, because in such cases it is often impossible to ascertain the identity of the individual to whom the newspaper article refers. In a small number of cases, individuals appeared in multiple categories, e.g. Arnold Schwarzenegger as politician and actor, and we analyzed such individuals accordingly in both categories. Our main results do not change when instead of double counting them we randomly choose one category.

4.3. Measuring the sentiment toward individuals mentioned in newspaper articles

Following standard natural language processing (NLP) procedures², The Lydia system assigns scores for thousands of adjectives with positive (+1) or negative (-1) sentiment to each distinct entity (name) in a text corpus, newspaper articles in our case (For more details on this process, see Godbole, Srinivasaiah and Skiena 2007). Cooccurrences with positive adjectives (such as “good” or “successful”) within the same sentence are counted as positive sentiment, while cooccurrences with negative terms (such as “ugly” or “failure”) are counted as negative sentiment (for further information on our sentiment analysis tool, see Bautin, Vijayarenu and Skiena 2008, Shor, van de Rijt and Miltsov 2019). Each name then receives an average sentiment score calculated across all co-occurrences in all news sources over the period of the study, ranging from -1 (for individuals who appeared only in co-occurrence with negative adjectives) to +1 (for those who appeared only in co-occurrence with positive adjectives). The online *Appendix I* demonstrates the ability of this sentiment tool to trace negative and positive events in the lives of a group of well-known celebrities and to accurately capture the temporal dynamics of coverage sentiment associated with the lives of these famous individuals. The appendix shows that sudden shifts in sentiment do indeed co-occur with these negative or positive life events, providing support for the validity of our sentiment measure.

It could be argued that a sentiment analysis that simply measures “positive” vs. “negative” references is too crude when trying to capture subtle differences in media coverage, in particular differences between women and men. For example, some scholars have argued that news reports on women often tend to overemphasize their physical appearance or motherly traits, while belittling their professional competence or intellectual skills (Byerly and Ross 2006,

² A simple application programming interface (API) explaining common natural language processing (NLP) and the procedure we followed here can also be found in the following link: <https://textblob.readthedocs.io/en/dev/index.html>

Heldman, Carroll and Olson 2005, Jia et al. 2016, Tuchman 2000, van Zoonen 1994). Our computerized text analysis system does not differentiate between positive adjectives such as “beautiful”, “attractive”, and “nurturing” on the one hand and adjectives like “competent”, “able”, “successful”, or “strong” on the other, which may be more beneficial for economic or political outcomes. However, the online *Appendix 4* demonstrates that the most common adjectives used in conjunction with female and male names are in fact very similar, suggesting that gender differences in qualitative dimensions of coverage sentiment may be more modest than previously suggested. Still, it is prudent to recognize that minor differences indeed exist and, accordingly, to interpret findings about sentiment differences with caution. We further elaborate on this issue in our concluding discussion and limitations sections.

We additionally use Wikipedia as a source of *biographic* sentiment data. For each Wikipedia person page in our dataset we calculated a biography sentiment score by aggregating positive and negative qualifications on these biographic pages. This additional sentiment data serves as a baseline against which to compare media sentiment, with discrepancies suggesting media bias. This approach is admittedly both conservative and naive. The honest intentions of the editors and editorial policy enforcement on Wikipedia notwithstanding, what is written on Wikipedia pages is inevitably influenced by the news cycle. To take one example, during the 2016 presidential election campaigns, Hillary Clinton’s email and private server mishaps were arguably negligible compared to Donald Trump’s many involvements in unethical behavior. Still, the mass media’s continued attention to the Clinton email scandal may have resulted in a disproportional discussion of this scandal on Clinton’s Wikipedia page, thereby depressing her biographic sentiment score. We argue that such potential spillover effects of the news cycle render the discrepancy between the two sentiment measures a conservative estimate of media bias.

4.4. Analytical Strategy:

Table 1 shows descriptive statistics for all variables used in our analyses. We compare the sentiment toward women and the sentiment toward men in our Wikipedia sample of 42,862 names using two-tailed independent sample t-tests. We then examine how fame mediates the relationship between sentiment and gender, by calculating mean media sentiment by gender at varying binned levels of fame, with corresponding confidence intervals of means. Next, we examine whether similar patterns can be found in encyclopedic sentiment, which captures positive and negative events in the lives of these famous individuals. We do so by calculating mean biographic sentiment by gender at varying levels of fame as well as by regressing biographic sentiment on fame, gender, and their interaction in OLS models with robust standard errors.

Table 1 about here

5. Findings

In **Table 2** we present comparisons between the sentiment toward women and the sentiment toward men in our Wikipedia sample. The final column shows the results of two-tailed independent sample t-tests for a difference in mean sentiment between women and men. The table shows that, with the obvious exception of criminals, both men and women receive more positive than negative coverage. Strikingly, in all major domains (politics, business, entertainment, crime, sports, and science), the coverage tone for women is significantly more positive than the coverage tone for men. Woman's coverage advantage can also be detected in most of the more specific categories, such as US House representatives and Oscar and Emmy nominees. In other subcategories – such as tennis Grand Slam winners and Nobel Prize winners – there is no statistically significant difference. These results provide strong support for H1 that women's media coverage overall is more positive than men's.

Table 2 about here

We next examine the role of fame in producing divergent sentiment for men and women. In **Figure 1**, we present two panels showing the interaction between gender, fame and coverage

sentiment. Panel 1 is based on the data from our larger sample of nearly 14 million person names (the Lydia newspapers sample), showing results for both well-known individuals and relatively obscure ones, who have appeared in the news only once or twice during the period of the study. Panel 2 is based on the smaller sample we collected from Wikipedia (N = 42,862), including individuals who are all well-known enough to have a Wikipedia entry. Note that there are relatively few individuals on Wikipedia with very little coverage, as indicated by the widening confidence intervals at lower levels in panel 2, but not in panel 1. The analyses of both samples show a similar pattern: at low levels of fame (1 to 10 yearly mentions), women receive coverage that is substantially more positive than that of equally renowned men (a 10% to 20% difference in coverage tone). However, as the number of mentions grows, the coverage tone associated with men remains fairly stable and even slightly improves, while the coverage tone for women becomes increasingly negative, resulting in an eventual elimination, and even reversal of sentiment differences. Indeed, among the most famous individuals, those who received in the order of one million mentions, the coverage of men is more positive than that of women.

Figure 1 about here

While Figure 1 supports the notion that women, unlike men, are more heavily scrutinized when they are famous, it leaves important questions about the origins of the effect unanswered. In particular, one may wonder whether this pattern holds for all women and men, regardless of their public status or their social and occupational domains. A second concern is that the pattern might reflect differential membership of women and men in occupational categories with different levels of sentiment and fame. For example, most politicians are men and the average politician receives both greater coverage and more negative coverage than the average entertainer, who is more likely to be a woman.

In **Figure 2**, we present results for coverage tone by gender and fame for men and women who were classified into six major social and occupational domains by the Wikipedia

categorization pages. Panels 1 through 6 of the figure present results for politicians, businesspeople, entertainers, criminals, athletes, and scientists (see online Appendix 3 for a random sample illustrating more- and less-famous individuals included in each of these categories). Because of the reduced sample sizes, the top categories are now too sparse for most domains, so we collapsed the 1,000,000 and 100,000 categories into the 10,000 mentions category.

The results presented in Figure 2 demonstrate that the overall pattern shown in Figure 1 is also evident in most of the specific domains. For men in most domains, including politicians, businesspeople, and entertainers, the coverage tone remains quite stable regardless of the level of fame. Famous male athletes and scientists exhibit somewhat worse coverage sentiment than their less famous counterparts, while men who commit crimes receive substantially better coverage when they are more famous (keeping in mind that the type of crime may vary). Conversely, the coverage tone for famous women in nearly all domains is significantly worse than for non-famous ones, perhaps with the exception of female athletes.

Figure 2 about here

The robustness of these results is confirmed in regression analysis. In table 3 we present results from OLS regression models predicting coverage sentiment from gender, fame, and their interaction. Model 2 is the same as model 1, except that it includes dichotomous variables measuring membership of six major social and occupational domains, into which names on Wikipedia are categorized. We use heteroskedasticity-robust standard errors because at very low fame levels, numbers of positive and negative mentions are naturally also low, leading to higher variance in the dependent variable. Both models show a significant interaction effect: On our sentiment scale from -1 to 1, for each unit (=10-fold) increase in fame the gender difference in sentiment is a full .03 points smaller. At low and intermediate levels of fame, women receive better coverage. At very high levels of fame, men receive better coverage.

Table 3 about here

5.1. Media as a mirror?

The robust interaction effect we present above lends support for the paper cut thesis that media discourse about women vis-à-vis comparable men becomes more negative as they acquire greater fame. However, the evidence presented so far may alternatively be interpreted as reflecting negative real-world actions or occurrences in the life of famous women (H3). That is, the media may be accurately reporting on true variability in the actions taken by or events occurring to men and women of different fame levels. Perhaps famous women “deserve” more negative coverage, for example because they are put into tougher situations or being deliberately thwarted, hindered, or blocked from succeeding.

To differentiate between these two alternative accounts (articulated in sections 2.1 and 2.2), we evaluate parallel patterns in biographical sentiment, testing H3 and H4. The “media as a mirror” account predicts that biographical sentiment patterns will match media sentiment patterns. The “paper cut” account instead predicts that biographical sentiment will be more positive for women at all levels of fame. These predictions do not consider the spillover problem mentioned earlier, whereby media coverage may be partly reflected in encyclopedic content. The setup of the test is thus stacked against the paper cut thesis, as it increases the likelihood of finding evidence for the media as a mirror thesis.

Figure 3 shows the three-way relationship between gender, fame, and biographic sentiment. The figure demonstrates that, unlike the pattern we found for media coverage, women *of all fame levels* receive significantly better biographic coverage when compared with men: Regardless of how often women appear in the media, their biographies on Wikipedia contain more positive adjectives than those of equally famous men. Model 3 in table 2 reports the results of OLS regression models predicting biographic sentiment from gender, fame, and their interaction effect, controlling for occupational categories. There is no evidence of disproportionately negative sentiment for famous women in biographical data. The interaction effect is even slightly positive in

this model, driven by the lowest fame levels where the gender difference in sentiment is somewhat smaller than at intermediate and high fame, as Figure 3 also shows. These results are consistent with a biographical basis for the more favorable media coverage that women in general enjoy (H3). However, they do not reveal a biographical reason for covering famous women but not famous men much more negatively. That is, although famous women have more positive public lives than famous men (rejecting H4), their lives are portrayed disproportionately negatively in the media. These results thus clearly support the theory of a paper cut.

Figure 3 about here

6. Conclusion and discussion

We have theorized and empirically investigated the three-way relationship between gender, fame, and coverage tone. In contrast to previous anecdotal evidence, we found that women receive more *positive* coverage sentiment than men do, overall and in various specific domains, such as politics, business, sports and entertainment. Yet, this coverage sentiment advantage varies dramatically with coverage volume. Women who are not often mentioned in newspapers clearly receive more positive coverage than men who are not mentioned often. However, for better-known women and men, this disparity is smaller and even reverses among the most famous celebrities, so that the coverage tone of the most often-mentioned men is more positive than that of the most often-mentioned women.

This pattern also holds within most professional and social spheres, including politics business, science, and crime. The women who are best-known (those who clearly break the glass ceiling) are those who receive the worst media coverage. In all these fields, the coverage sentiment for those women who received the highest levels of media attention was much worse than the coverage sentiment of women who were slightly less famous. Finally, we found that sentiment in encyclopedic content is more favorable for women than for men, *regardless of their level of fame*. This contrast between media and encyclopedic sentiment suggests that media report more

negatively on women when they become famous not because of negative actions and events surrounding them, but rather because their authority violates patriarchal norms regarding gender hierarchies and appropriate feminine behaviors and aspirations.

These tendencies are perhaps most pronounced in the domain of crime. While the coverage tone for relatively obscure women offenders was quite neutral, it was much more negative for more famous women offenders. For men offenders, however, the tendency was reversed, as famous male criminals enjoyed more positive coverage than less famous ones. These findings are consistent with feminist criminology views and critiques of the chivalry hypothesis (Lloyd 1995). Qualitative case studies report that petty crimes committed by women tend to be both overlooked and treated leniently (Chesney-Lind 1999, Grabe et al. 2006, Weimann and Fishman 1988). Conversely, these studies also found that women who commit serious acts of violence and breach normative gender expectations draw significant media attention and scrutiny, quickly becoming notorious, and their coverage tone tends to be particularly negative (Barnett 2005, Cavaglian 2008, Eastal et al. 2015, Grabe et al. 2006, Hinds and Stacey 2010).

In politics, we also found that women of greater fame receive worse media sentiment while the sentiment for men is largely invariant across levels of fame. This pattern again confirms the anecdotal observations of qualitative studies on successful politicians, which reported that their performance is heavily scrutinized and criticized (Heldman, Carroll and Olson 2005, Meeks 2012, Ryan 2013, Trimble and Treiberg 2010).

Our findings also align with those of a recent study on Twitter users (Nilizadeh et al. 2016) that reported a somewhat similar interaction between gender and fame. Nilizadeh et al. found that users who were perceived as female experienced a type of a glass ceiling. While for users in lower quartiles of visibility, being perceived as female was associated with more visibility, this tendency flipped among the most visible users, for whom being perceived as female was strongly associated

with less visibility. While the study by Nilizadeh et al. examined visibility rather than sentiment, as we did, we believe that the mechanisms underlying these patterns may be similar.

While our sentiment measure successfully captures negative coverage following negative events in the lives of famous people (see online Appendix 1), it is nonetheless possible that it misses important nuances in the quality of the coverage. Former research has suggested that the depictions of women in politics may not be straightforwardly negative, but rather might often trivialize or overlook their achievements and professional success by focusing on unrelated issues, such as their family, their looks, their marital status, and their attire (Heldman, Carroll and Olson 2005, Kahn 1994). Bligh et al. (2012) have argued that even descriptions that might seem positive, such as those focusing on warmth, gentleness, or emotions may in fact convey a negative message, as they simultaneously portray women in politics as weak and indecisive.

This limitation of our sentiment measurement is not restricted to the field of politics. Successful women in entertainment, business, science, and sports may also receive coverage that might appear to be positive in tone, but in fact focuses on issues such as their appearance and thus diminishes from their achievements. For example, Cranmer et al. (2014) have argued that the devaluation of female athletes occurs through a subtle focus on sexualized frames (see also Shugart 2003). Similarly, studies of sports photography, which our textual tool is unable to analyze, have commented on the focus on female athletes' bodies and the commonality in which female athletes appear in sexualized poses rather than while competing (Crossman, Vincent and Speed 2007, Fink and Kensicki 2002, Lumpkin 2009). Hence, our findings about the significant worsening of coverage tone for well-known women may fail to capture the full picture about this phenomenon.

Notwithstanding these methodological limitations, our study clearly shows that in various professional and social domains women receive more positive media coverage than men, except when they are famous. Our evidence suggests that this interaction between gender and fame is not

due to famous women experiencing more negative events. Instead, acts and achievements are portrayed more negatively in the media when involving a famous woman.

References

- Aaldering, Loes, Tom van der Meer and Wouter van der Burg. 2018. "Mediated Leader Effects: The Impact of Newspapers' Portrayal of Party Leadership on Electoral Support." *The International Journal of Press/Politics* 23(1):70-94.
- Aaldering, Loes and Daphne Van der Pas. 2020. "Political Leadership in the Media: Gender Bias in Leader Stereotypes During Campaign and Routine Times." *British Journal of Political Science* 50:911–31.
- Adams, Julia , Hannah Brückner and Cambria Naslund. 2019. "Who Counts as a Notable Sociologist on Wikipedia? Gender, Race, and the "Professor Test". " *Socius* 5:1-14.
- Alessio, John C. and Julie Andrzejewski. 2000. "Unveiling the Hidden Glass Ceiling: An Analysis of the Cohort Effect Claim." *American Sociological Review* 65(2):311-15.
- Antin, Judd, Raymond Yee, Coye Cheshire and Oded Nov. 2011. "Gender Differences in Wikipedia Editing." *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*:11-14. doi: <https://doi.org/10.1145/2038558.2038561>.
- Baker, Ted, Haward E. Aldrich and Liou Nina. 1997. "Invisible Entrepreneurs: The Neglect of Women Business Owners by Mass Media and Scholarly Journals in the USA." *Entrepreneurship & Regional Development* 9(3):221-38.
- Barden, Maria. 1996. *Women Politicians and the Media*. Lexington: University of Kentucky Press.
- Barnett, Barbara. 2005. "Perfect Mother or Artist of Obscenity? Narrative and Myth in a Qualitative Analysis of Press Coverage of the Andrea Yates Murders." *Journal of Communication Inquiry* 29(1):9-29.
- Bauer, Jill. 2020. *The Qualifications Gap: Why Women Must Be Better Than Men to Win Political Office*. Cambridge, UK: Cambridge University Press.
- Bautin, Mikhail, Lohit Vijayarenu and Steven Skiena. 2008. "International Sentiment Analysis for News and Blogs." *Proceedings of the International AAAI Conference on Weblogs and social media*:19-26.
- Bautin, Mikhail, Charles Ward, Akshay Patil and Steven Skiena. 2010. "Access: News and Blog Analysis for the Social Sciences." in *19th International World Wide Web Conference (WWW 2010)*. Raleigh, NC.
- Baxter, Janeen and Eric Wright. 2000. "The Glass Ceiling Hypothesis: A Comparative Study of the United States, Sweden, and Australia." *Gender & Society* 14:275-94.
- Berrington, Eileen and Paivi Honkatukia. 2002. "An Evil Monster and a Poor Thing: Female Violence in the Media." *Journal of Scandinavian Studies in Criminology and Crime Prevention* 3(1):50-72.
- Bishop Smith, Edward, Jillian Chown and Kevin Gaughan. 2021. "Better in the Shadows? Public Attention, Media Coverage, and Market Reactions to Female Ceo Announcements." *Sociological Science* 8:119-49.
- Bjursell, Cecilia and Lisa Backvall. 2011. "Family Business Women in Media Discourse: The Business Role and the Mother Role." *Journal of Family Business Management* 1(2):154-73.
- Bligh, Michelle, Michelle Schlehofer, Bettina Casad and Amber Gaffney. 2012. "Competent Enough, but Would You Vote for Her? Gender Stereotypes and Media Influences on Perceptions of Women Politicians." *Journal of Applied Social Psychology* 42(3):560-97.
- Byerly, Carolyn M. and Karen Ross. 2006. *Women & Media: A Critical Introduction*, Vol. Malden, MA: Blackwell Publishing.
- Castagnetti, Carolina and Luisa Rosti. 2013. "Unfair Tournaments: Gender Stereotyping and Wage Discrimination among Italian Graduates." *Gender & Society* 27(5):630-58.
- Cavaglia, Gabriel. 2008. "Bad, Mad or Sad? Mothers Who Kill and Press Coverage in Israel." *Crime Media Culture* 4(2):271-78.
- Chesney-Lind, Meda. 1999. "Media Mysogyny: Demonizing "Violent" Girls and Women." Pp. 115-41 in

- Making Trouble: Cultural Representation of Crime, Deviance, and Control*, edited by J. Ferrel and N. Websdale. New York: de Gruyter.
- Collier, B. and J. Bear. 2012. "Conflict, Criticism, or Confidence: An Empirical Examination of the Gender Gap in Wikipedia Contributions." *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, CSCW* 12:383-92.
- Collins, Rachel. 2016. "'Beauty and Bullets': A Content Analysis of Female Offenders and Victims in Four Canadian Newspapers." *Journal of Sociology* 52(2):296-310.
- Comrie, Margie. 2006. "The Media and Leadership." Pp. 175-95 in *Political Leadership in New Zealand*, edited by R. Miller and M. Mintrom. Auckland: Auckland University Press.
- Cook, Alison and Christy Glass. 2013. "Glass Cliffs and Organizational Saviors: Barriers to Minority Leadership in Work Organizations?". *Social Problems* 60(2):168-87.
- Cooky, Cheryl, Michael Messner and Michela Musto. 2015. "'It's Dude Time!': A Quarter Century of Excluding Women's Sports in Televised News and Highlight Shows." *Communication & Sport* 3(3):261-87.
- Cotter, David, Joan Hermsen, Seth Ovidia and Reeve Vanneman. 2001. "The Glass Ceiling Effect." *Social Forces* 80(2):655-81.
- Cranmer, Gregory, Maria Brann and Nicholas Bowman. 2014. "Male Athletes, Female Aesthetics: The Continued Ambivalence toward Female Athletes in Espn's the Body Issue." *International Journal of Sport Communication* 7(2):145-65.
- Creed, Barbara. 1996. "Bitch Queen or Backlash? Media Portrayals of Female Murderers." Pp. 20-30 in *The Thing She Loves: Why Women Kill*, edited by K. Greenwood. Canada: Allen & Unwin.
- Crossman, Jane, John Vincent and Harriet Speed. 2007. "'The Times They Are a-Changin': Gender Comparisons in Three National Newspapers of the 2004 Wimbledon Championships." *International Review for the Sociology of Sport* 42(1):27-41.
- Czarniawska, Barbara. 2004. "Women in Financial Services: Fiction and More Fiction." Pp. 121-37 in *The Sociology of Financial Markets*, edited by K. Knorr Cetina and A. Preda. Oxford: Oxford University Press.
- Czarniawska, Barbara. 2008. "Femmes Fatales in Finance, or Women and the City." *Organization* 15(2):165-86.
- De Swert, Knut and Mark Hooghe. 2010. "When Do Women Get a Voice? Explaining the Presence of Female News Sources in Belgian News Broadcasts (2003-5)." *European Journal of Communication* 25(1):69-84.
- Dezso, Cristian and David Gaddis Ross. 2012. "Does Female Representation in Top Management Improve Firm Performance? A Panel Data Investigation." *Strategic Management Journal* 33(9):102-1089.
- Eagly, Alice and Steven Karau. 2002. "Role Congruity Theory of Prejudice toward Female Leaders." *Psychological Review* 109(3):573-89.
- Easteal, Patricia, Lorana Bartels, Noni Nelson and Kate Holland. 2015. "How Are Women Who Kill Portrayed in Newspaper Media? Connections with Social Values and the Legal System." *Women's Studies International Forum* 51:31-41.
- Falk, Erika. 2010. *Women for President: Media Bias in Nine Campaigns*. Champaign, IL: University of Illinois Press.
- Farr, Kathryn. 2000. "Defeminizing and Dehumanizing Female Murderers: Depictions of Lesbians on Death Row." *Women & Criminal Justice* 11(1):49-66.
- Ferris, Kerry. 2007. "The Sociology of Celebrity." *Sociology Compass* 1(1):371-84.
- Fink, Janet and Linda Kensicki. 2002. "An Imperceptible Difference: Visual and Textual Constructions of Femininity in Sports Illustrated and Sports Illustrated for Women." *Mass Communication & Society* 5(3):317-39.
- Fowler, Linda L. and Gennifer L. Lawless. 2009. "Looking for Sex in All the Wrong Places: Press Coverage

- and the Electoral Fortunes of Gubernatorial Candidates." *Perspectives on Politics* 7:315-35.
- Fox, Richard and Jennifer Lawless. 2014. "Uncovering the Origins of the Gender Gap in Political Ambition." *American Political Science Review* 108(3):499-519.
- Fulton, Sarah. 2012. "Running Backwards and in High Heels: The Gendered Quality Gap and Incumbent Electoral Success." *Political Research Quarterly* 65(2):303-14.
- Gallagher, Margaret. 2010, "Who Makes the News? Global Media Monitoring Project (Gmmp) 2010": World Association of Christian Communication. Retrieved December 30, 2011, 2011 (http://www.whomakesthenews.org/images/stories/website/gmmp_reports/2010/global/gmmp_global_report_en.pdf).
- Gidengil, Elisabeth and Joanna Everitt. 2003. "Conventional Coverage/Unconventional Politicians: Gender and Media Coverage of Canadian Leaders' Debates, 1993, 1997, 2000." *Canadian Journal of Political Science* 36(3):559-77.
- Glick, Peter and Susan Fiske. 1996. "The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism." *Journal of Personality and Social Psychology* 70(3):491-512.
- Godbole, Namrata, Manja Srinivasaiah and Steven Skiena. 2007. "Large-Scale Sentiment Analysis for News and Blogs." *Proceedings of the Third International conference on Weblogs and Social Media*:219-22.
- Goddu, Jenn. 1999. "'Powerless, Public-Spirited Women,' 'Angry Feminists,' and 'the Muffin Lobby': Newspaper and Magazine Coverage of the Canadian Advisory Council on the Status of Women, the National Action Committee on the Status of Women, and Real Women of Canada." *Canadian Journal of Communication* 24:1.
- Grabe, Maria Elizabeth, K. Trager, Melissa Lear and Jennifer Rauch. 2006. "Gender in Crime News: A Case Study Test of the Chivalry Hypothesis." *Mass Communication & Society* 9(2):137-63.
- Grandy, Karen. 2013. "The Glossy Ceiling: Coverage of Women in Canadian and American Business Magazines". *Journal of Magazine & New Media Research* 14(1):1-20.
- Greenwald, Marilyn S. 1990. "Gender Representation in Newspaper Business Sections." *Newspaper Research Journal* 11:68-73.
- Halvorson, Heidi. 2015. *No One Understands You and What to Do About It*. Harvard: Harvard Business Review Press.
- Heldman, Caroline, Susan J. Carroll and Stephanie Olson. 2005. "'She Brought Only a Skirt': Print Media Coverage of Elizabeth Dole's Bid for the Republican Presidential Nomination." *Political Communication* 22(3):315-35.
- Henderson, Anne. 1999. *Getting Even: Women Mps on Life, Power, and Politics*. Sydney: Harper Collins.
- Hinds, Hilary and Jackie Stacey. 2010. "Imaging Feminism, Imaging Femininity: The Bra-Burner, Diana, and the Woman Who Kills." *Feminist Media Studies* 1(2):153-77.
- Huckerby, Jayne. 2003. "Women Who Kill Their Children: Case Study and Conclusions Concerning the Differences in the Fall from Material Grace by Khoua Her and Andrea Yates." *Duke Journal of Gender Law & Policy* 10:149-72.
- Huddy, Leonie and Nayda Terkildsen. 1993. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37(1):119-47.
- Jalalzai, Farida. 2008. "Women Rule: Shattering the Executive Glass Ceiling." *Politics & Gender* 4(2):205-31.
- Jia, Sem, Thomas Lansdall-Welfare, Saatviga Sudhahar, Cynthia Carter and Nello Cristianini. 2016. "Women Are Seen More Than Heard in Online Newspapers." *PLOS One* 11(2). doi: DOI:10.1371/journal.pone.0148434.
- Joo, J. H. 2002. "The Influence of News Frames on the Audience's Attitudes: Examining News Stories on Women Cabinet Members." *Unpublished Master's Thesis*. Korea University, Seoul, Korea.
- Kahn, Kim F. 1992. "Does Being Male Help? An Investigation of the Effects of Candidate Gender and

- Campaign Coverage of U.S. Senate Candidates." *The Journal of Politics* 54(2):497-517.
- Kahn, Kim F. 1994. "The Distorted Mirror: Press Coverage of Women Candidates for Statewide Office." *The Journal of Politics* 56(1):154-73.
- Khan, Walayet and Joao Paolo Vieito. 2013. "Ceo Gender and Firm Performance." *Journal of Economics and Business* 67:55-66.
- Koenig, Anne, Alice Eagly, Abigail Mitchell and Tiina Ristikari. 2011. "Are Leader Stereotypes Masculine? A Meta-Analysis of Three Research Paradigms. , (4), 616. ." *Psychological bulletin* 137(4):616-41.
- Krefting, Linda. 2002. "Re-Presenting Women Executives: Valorization and Devalorization in Us Business Press." *Women in Management Review* 173(4):104-19.
- Lazarus, Jeffrey and Amy Steigerwalt. 2018. *Gendered Vulnerability: How Women Work Harder to Stay in Office*. Ann Arbor: University of Michigan Press.
- Lloyd, Ann. 1995. *Doubly Deviant, Doubly Damned: Society's Treatment of Violent Women*. New York: Penguin.
- Lumpkin, Angela. 2009. "Female Representation in Feature Articles Published by Sports Illustrated in the 1990s." *Women in Sport and Physical Activity Journal* 18(2):38-51.
- Manning, Alan and Farzad Saidi. 2010. "Understanding the Gender Pay Gap: What's Competition Got to Do with It?". *Industrial and Labor Relations Review* 63(681-698).
- McGregor, J. 2000. "Stereotypes and Symbolic Annihilation: Press Constructions of Women at the Top." *Women in Management Review* 155(6):290-95.
- McGregor, Judy. 1996. "Gender Politics and the News: The Search for a Beehive Bimbo-Boadicea." Pp. 181-96 in *Dangerous Democracy? News Media Politics in New Zealand*, edited by J. McGregor. Palmerston North: Dunmore.
- Meeks, Lindsey. 2012. "Is She 'Man Enough'? Women Candidates Lead to Atypical Coverage?". *Howard Journal of Communication* 23:272-87.
- Mills, Kay. 1997. "What Difference Do Women Journalist Make?" Pp. 41-55 in *Women, Media and Politics*, edited by P. Norris. Oxford: Oxford University Press.
- Morgan, Laurie. 1998. "Glass-Ceiling Effect or Cohort Effect? A Longitudinal Study of the Gender Earnings Gap for Engineers, 1982 to 1989." *American Sociological Review* 63(4):479-83.
- Morrissey, Belinda. 2003. *When Women Kill: Questions of Agency and Subjectivity*. New York: Routledge.
- Naylor, Bronwyn. 1990. "Media Images of Women Who Kill." *Legal Services Bulletin* 5(1):4-8.
- Naylor, Bronwyn. 2001. "The 'Bad Mother' in Media and Legal Texts." *Social Semiotics* 11(2):155-76.
- Nilizadeh, Shirin, Anne Groggel, Peter Lista, Srijita Das, Yong-Yeol Ahn, Apu Kapadia and Fabio Rojas. 2016. "Twitter's Glass Ceiling: The Effect of Perceived Gender on Online Visibility." *Proceedings of the Tenth International AAAI Conference on Web and Social Media*:289-98.
- O'Brien, Diana. 2015. "Rising to the Top: Gender, Political Performance, and Party Leadership in Parliamentary Democracies." *American Journal of Political Science* 59(4):1022-39.
- Palmer, Barbara and Dennis Simon. 2008. *Breaking the Political Glass Ceiling*. New York: Routledge.
- Prokos, Anastasia and Irene Padavic. 2005. "An Examination of Competing Explanations for the Pay Gap among Scientists and Engineers." *Gender & Society* 19(4):523-43.
- Ridgeway, Cecilia. 1997. "Interaction and the Conservation of Gender Inequality: Considering Employment." *American Sociological Review* 62(2):218-35.
- Ridgeway, Cecilia. 2009. "Framed before We Know It: How Gender Shapes Social Relations." *Gender & Society* 23(145-160).
- Ridgeway, Cecilia. 2011. *Framed by Gender: How Gender Inequality Persists in the Modern World*. New York Oxford University Press.
- Ridgeway, Cecilia. 2013. "Why Status Matters for Inequality." *American Sociological Review* 79(1):1-16.
- Rodgers, Shelly and Esther Thorson. 2003. "A Socialization Perspective on Male and Female Reporting." *Journal of Communication* 53(4):658-75.

- Ross, Karen. 2007. "The Journalist, the Housewife, the Citizen and the Press." *Journalism* 8(4):440-73.
- Ross, Karen. 2009. *Gendered Media: Women, Men, and Identity Politics*. Lanham, MD: Rowman and Littlefield.
- Ross, Karen and Cynthia Carter. 2011. "Women and News: A Long Winding Road." *Media, Culture & Society* 33(8):1148-65.
- Rudman, Laurie, Corinne Moss-Racusin, Julie Phelan and Sanne Nauts. 2012. "Status Incongruity and Backlash Effects: Defending the Gender Hierarchy Motivates Prejudice against Female Leaders." *Journal of Experimental Social Psychology* 48(1):165-79.
- Ryan, Kaitlyn. 2013. "The Media's War on Women: Gendered Coverage of Female Candidates." *Xavier Journal of Politics* 4(1):13-25.
- Ryan, Michelle and Alexander Haslam. 2005. "The Glass Cliff: Evidence That Women Are over-Represented in Precarious Leadership Positions." *British Journal of Management* 16:81-90.
- Saner, Emine. 2014, "Feisty, Flounce and Bossy: The Words Used to Put Women Down": The Guardian. Retrieved April 29, 2016, 2016 (<http://www.theguardian.com/lifeandstyle/2014/sep/01/feisty-flounce-bossy-words-put-women-down>).
- Sanghani, Radhika. 2014, "Feisty, Frigid, and Frumpy: 14 Words We Only Use to Describe Women": The Telegraph. Retrieved April 29,, 2016 (<http://www.telegraph.co.uk/women/womens-life/11067727/Downton-Abbey-row-14-words-we-only-use-to-describe-women.html>).
- Scharrer, Erica. 2002. "An "Improbable Leap": A Content Analysis of Newspaper Coverage of Hillary Clinton's Transition from First Lady to Senate Candidate." *Journalism Studies* 3(3):393-406.
- Schlehofer, Michelle, Bettina Casad, Michelle Bligh and Angela Grotto. 2011. "Navigating Public Prejudices: The Impact of Media and Attitudes on High-Profile Female Political Leaders." *Sex Roles* 65:69-82.
- Shor, Eran, Arnout van de Rijt, Charles Ward, Saoussan Asakar and Steven Skiena. 2014a. "Is There a Political Bias? A Computational Analysis of Female Subjects' Coverage in Liberal and Conservative Newspapers." *Social Science Quarterly* 95:1213-29.
- Shor, Eran, Arnout van de Rijt, Charles Ward, Aaron Blank and Steven Skiena. 2014b. "Time Trends in Printed News Coverage of Female Subjects, 1880-2008." *Journalism Studies* 15:740-56. doi: 10.1080/1461670X.2013.834149.
- Shor, Eran, Arnout van de Rijt, Alex Miltsov, Vivek Kulkarni and Steven Skiena. 2015. "A Paper Ceiling: What Explains the Sex-Ratio Inequality in Printed News Coverage?". *American Sociological Review* 80(5):960-84.
- Shor, Eran. 2018. "Media Partisanship and Coverage Sentiment: Are Conservative Newspapers More Negative toward Women?". *Social Science Quarterly* 100:307-19.
- Shor, Eran, Arnout van de Rijt and Alex Miltsov. 2019. "Do Women in the Newsroom Make a Difference? Sentiment toward Women and Men as a Function of Newsroom Composition." *Sex Roles* 81:44-58.
- Shor, Eran and Alex Miltsov. 2020. "The Price of Greater Representation: A Cross-National Analysis of Women's Parliamentary Representation and Media Coverage Sentiment." *Newspaper Research Journal* 41:455-68.
- Shor, Eran, Arnout Van de Rijt and Babak Fotouhi. 2020. "A Large-Scale Test of Gender Bias in the Media." *Sociological Science* 6:526-50.
- Shugart, Helene. 2003. "She Shoots, She Scores: Mediated Constructions of Contemporary Female Athletes in Coverage of the 1999 Us Women's Soccer Team." *Western Journal of Communication* 67(1):1-31.
- Smith, Kevin B. 1997. "When All's Fair: Signs of Parity in Media Coverage of Female Candidates." *Political Communication* 14(1):71-82.
- Trimble, Linda and Nastasja Treiberg. 2010. "Either Way, There's Going to Be a Man in Charge: Media

- Representations of New Zealand Prime Minister Helen Clark." Pp. 115-36 in *Cracking the Highest Glass Ceiling: A Global Comparison of Women's Campaigns for Executive Office*, edited by R. Murray. Santa Barbara, CA: Praeger.
- Tuchman, Gaye. 2000. "The Symbolic Annihilation of Women by the Mass Media." Pp. 150-74 in *Culture and Politics: A Reader*, edited by L. Crothers and C. Lockhart. New York: St. Martin's Press.
- van Acker, Elizabeth. 2003. "Media Representations of Women Politicians in Australia and New Zealand: High Expectations, Hostility or Stardom." *Policy and Society* 22:116-36.
- van de Rijt, Arnout, Eran Shor, Charles Ward and Steven Skiena. 2013. "Only 15 Minutes? The Social Stratification of Fame in Printed Media." *American Sociological Review* 78(2):266-89.
- Van der Pas, Daphne and Loes Aldering. 2020. "Gender Differences in Political Media Coverage: A Meta-Analysis." *Journal of Communication* 70:114-43.
- Van Zoonen, Liesbet. 1988. "Rethinking Women and the News." *European Journal of Communication* 3(1):35-53.
- van Zoonen, Liesbet. 1994. *Feminist Media Studies*. London: Sage.
- Vincent, John, Paul Pedersen, Warren Whisenant and Dwayne Massey. 2007. "Analysing the Print Media Coverage of Professional Tennis Players: British Newspaper Narratives About Female Competitors in the Wimbledon Championships." *International Journal of Sport Management and Marketing* 2(3):281-300.
- Wagner, Martin. 2015. "It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia." *Proceedings of the Ninth International AAAI Conference on Web and Social Media* 9(1).
- Waldon, Theresa. 2005. "Pinned by Gender Construction?: Media Representations of Girls' Wrestling." *Women in Sport and Physical Activity Journal* 14(2):52-68.
- Wang, Alice, Aasish Pappu and Henriette Cramer. 2021. "Representation of Music Creators on Wikipedia, Differences in Gender and Genre." *Proceedings of the Fifteenth International AAAI Conference on Web and Social Media (ICWSM 2021)*.
- Weber, Andrea and Christine Zulehner. 2010. "Female Hires and the Success of Start-up Firms." *The American Economic Review* 100(2):358-61.
- Weimann, Gabriel and Gideon Fishman. 1988. "Attribution of Responsibility: Sex-Based Bias in Press Reports on Crime." *European Journal of Communication* 3:415-30.
- White, Alice. 2018. "The History of Women in Engineering on Wikipedia." *Science Museum Group Journal* Autumn 2018(10). doi: <http://dx.doi.org/10.15180/181008>.
- Wilczynski, Ania. 1991. "Images of Women Who Kill Their Infants." *Women & Criminal Justice* 2(2):71-88.
- Zoch, Lynn and Judy V. Turk. 1998. "Women Making News: Gender as a Variable in Source Selection and Use." *Journalism & Mass Communication Quarterly* 75:762-75.

Table 1: Descriptive Statistics (N = 42,862)

Variable	Mean	Std.
Fame		
Number of media mentions per person	4,400	50,152
Number of media sentences per person	3,297	36,088
Number of media articles per person	1,468	11,405
Media sentiment	.214	.480
Encyclopedic sentiment	.512	.354
Percentage		
Woman	29.9%	
Politicians	9.2%	
House representatives	1.5%	
Senators	0.4%	
Businesspeople	7.0%	
CEOs	1.8%	
Billionaires	1.1%	
Entertainers	33.8%	
Oscar nominees	1.1%	
Emmy nominees	1.2%	
Criminals	5.2%	
Athletes	9.2%	
Tennis Grand Slam winners	0.3%	
Olympic medalists	0.7%	
Scientists	13.5%	
Nobel Prize winners	0.5%	

Table 2: Media sentiment by gender and domain

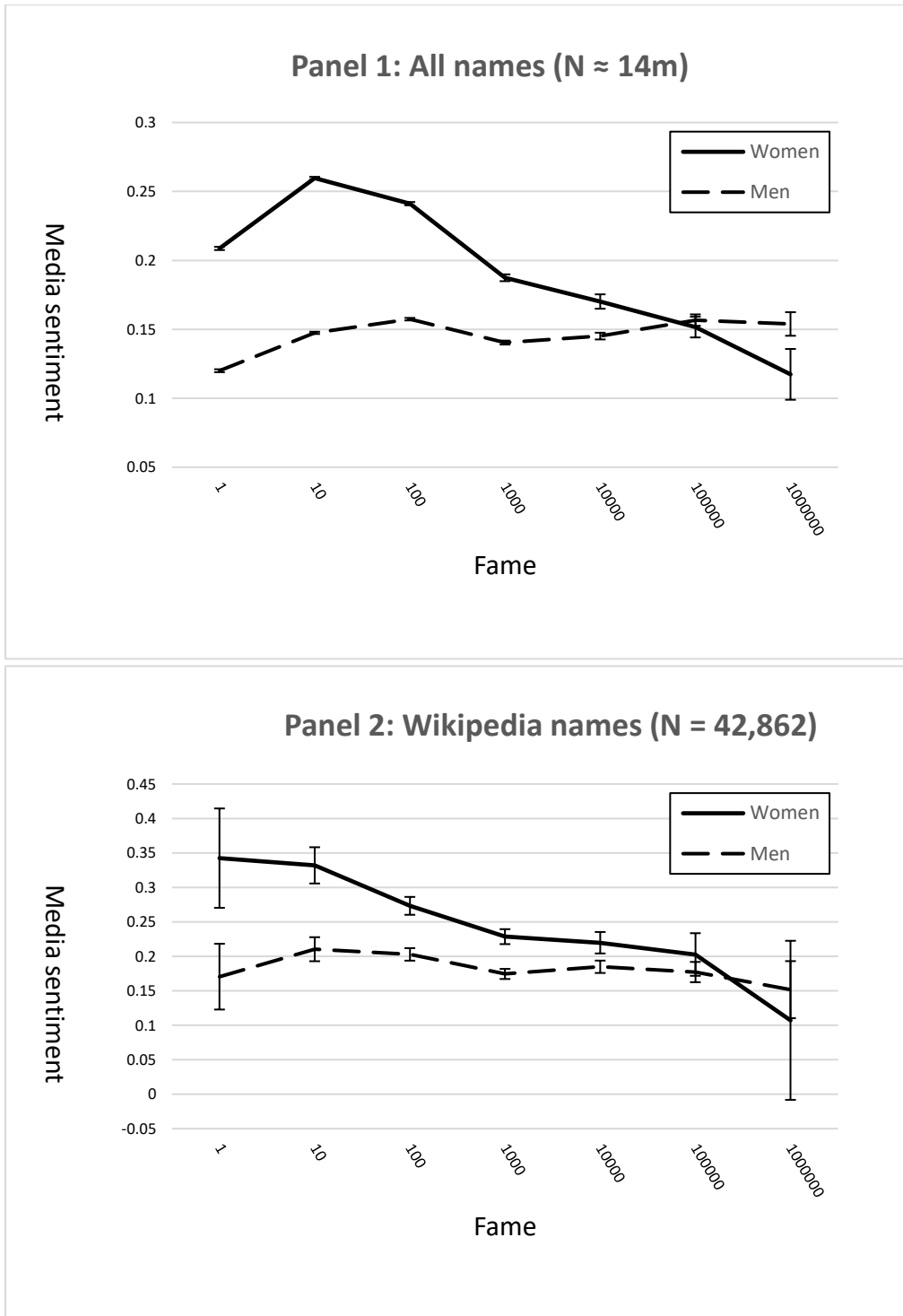
Field	Women			Men			Difference Women vs. men: sentiment difference
	N in sample	Median # mentions per year	Mean sentiment	N in sample	Median # mentions per year	Mean sentiment	
All	12,822	194	.27	30,040	217	.19	.08***
Politicians	1,061	339	.19	2,897	324	.16	.04*
House representatives	141	4,784	.24	296	7,615	.18	.05*
Senators	30	32,400	.23	146	33,451	.22	.01
Businesspeople	484	208	.39	2,543	161	.31	.08***
CEOs	149	378	.38	653	199	.37	.01
Billionaires	53	139	.45	421	652	.39	.06
Entertainers	5,950	202	.28	8,532	341	.21	.07***
Oscar nominees	234	6,920	.24	231	15,771	.18	.06**
Emmy nominees	235	3,000	.27	259	5,268	.20	.07**
Criminals	356	214	-.25	1,880	147	-.39	.14***
Athletes	953	814	.41	2,994	824	.32	.10***
Tennis Grand Slam winners	55	6,403	.28	54	6,223	.31	-.04
Olympic medalists	160	1,562	.51	128	2,927	.42	.09**
Scientists	1,220	63	.27	4,555	76	.22	.05**
Nobel Prize winners	20	442	.35	214	146	.41	-.06

* p < .05 ** p < .01 *** p < .001 (two-tailed independent sample t-test)

Table 3: OLS Regression of media and biographical sentiment

	Model 1: media sentiment		Model 2: media sentiment		Model 3: biographical sentiment		Model 4: biographical sentiment	
	β	SE(β) ¹	β	SE(β) ¹	β	SE(β) ¹	β	SE(β) ¹
Fame (log10)	-.01**	.00	-.02***	.00	-.02***	.00	-.02***	.00
Female	.15***	.02	.12***	.02	.05***	.01	.03**	.01
Female * Fame	-.03***	.01	-.03***	.01	.00	.00	.01	.00
Politician			-.06***	.01			-.06***	.01
Criminal			-.59***	.01			-.29***	.01
Businessperson			.10***	.01			-.03***	.01
Entertainer			.00	.01			-.02***	.00
Athlete			.13***	.01			-.07***	.01
Scientist			-.01	.01			-.02***	.01
Constant	.19***	.01	.25***	.01	.54***	.01	.58***	.01
N	42,862		42,862		42,862		42,862	
R ²	.01		.09		.01		.04	
¹ Heteroskedasticity-consistent standard errors								
* p < .05 ** p < .01 *** p < .001 (two-tailed)								

Figure 1: Media sentiment by gender and fame



Note: Error bars represent 95% confidence intervals around the mean values.

Figure 2: Media sentiment by gender and fame in 6 prominent domains

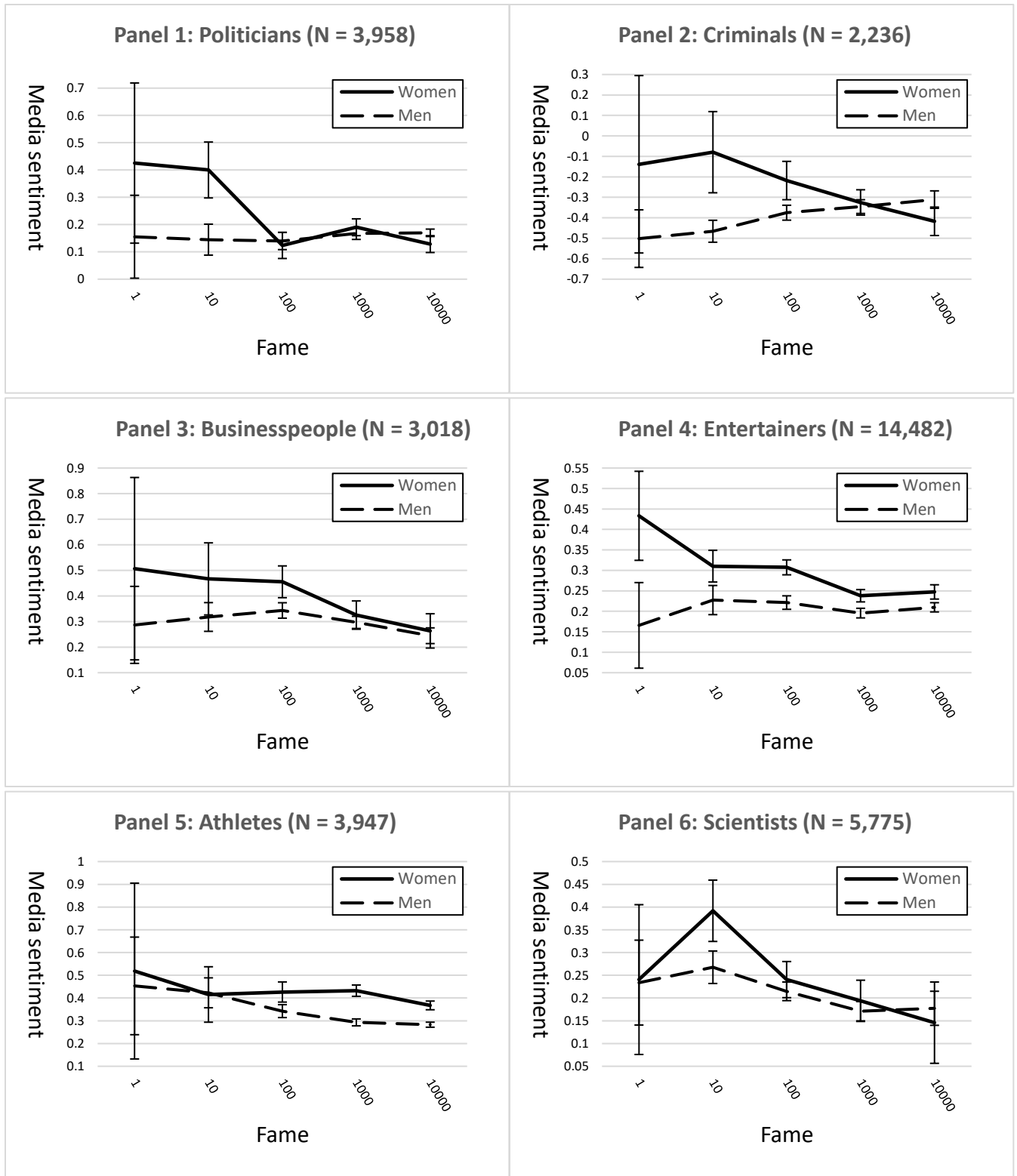
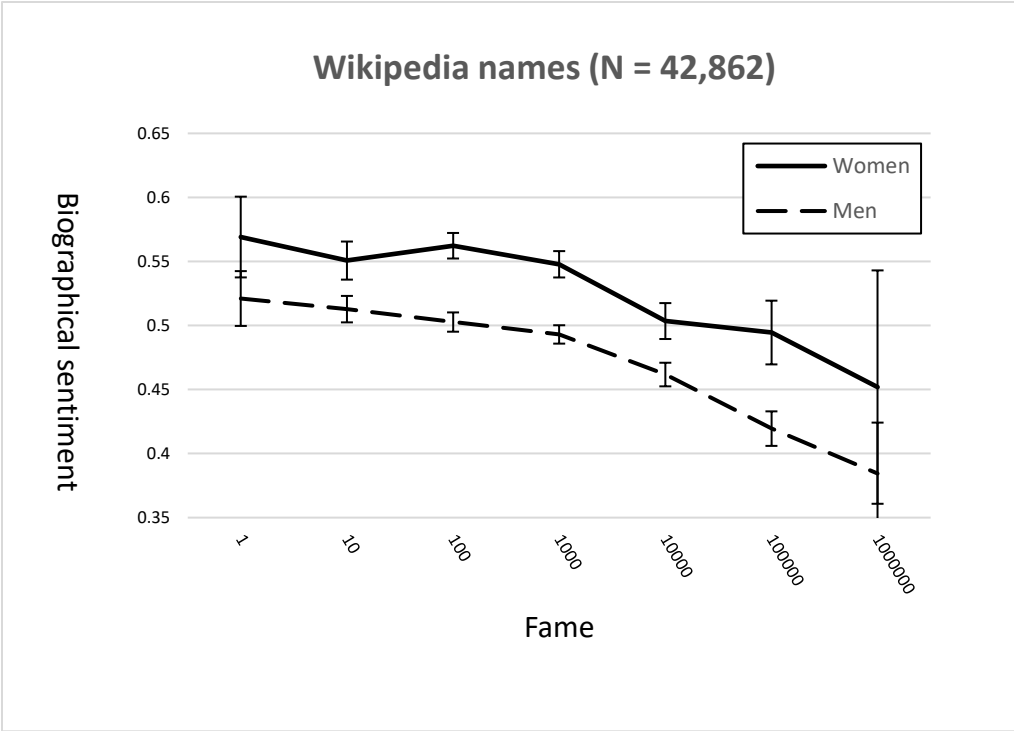
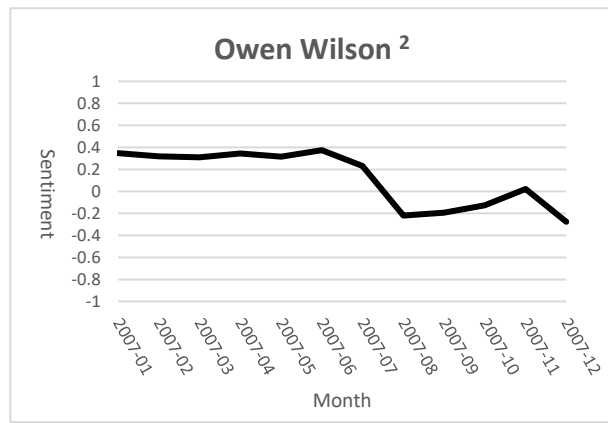
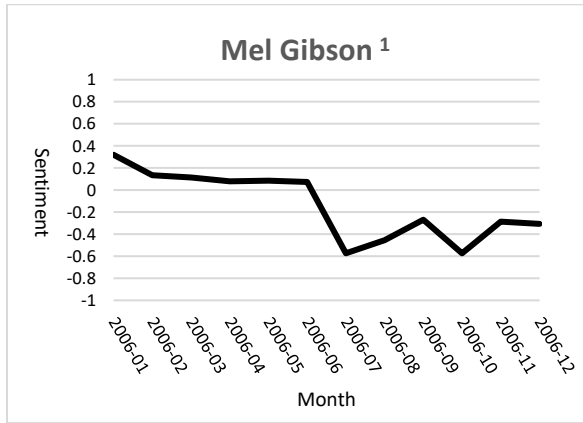


Figure 3: Biographical sentiment by gender and fame

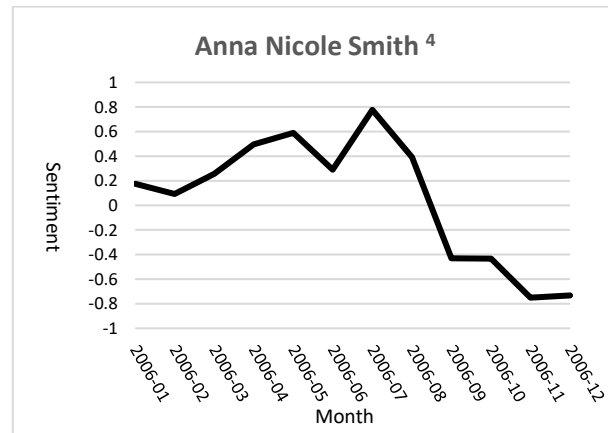
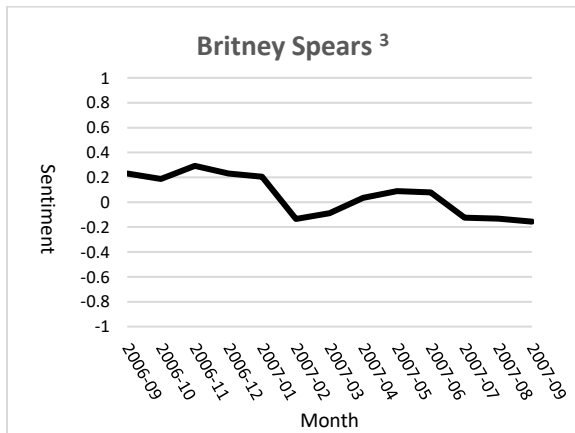


Appendix 1: Seven American celebrities and the sentiment they garnered over selected periods. We selected these individuals because they experienced at least one negative life event (to themselves or to a close one) between 2004 and 2009. The graphs demonstrate the ability of our sentiment analysis tool to detect changes in coverage sentiment following such events.



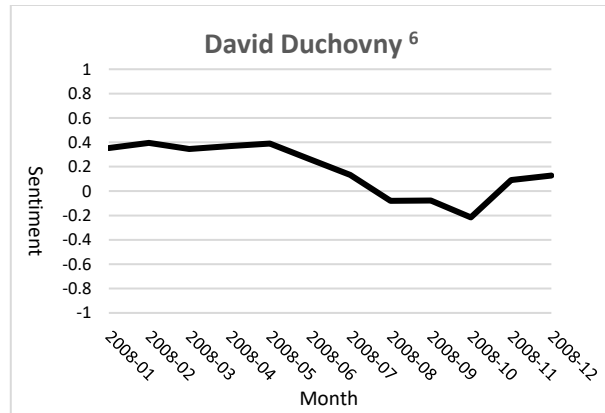
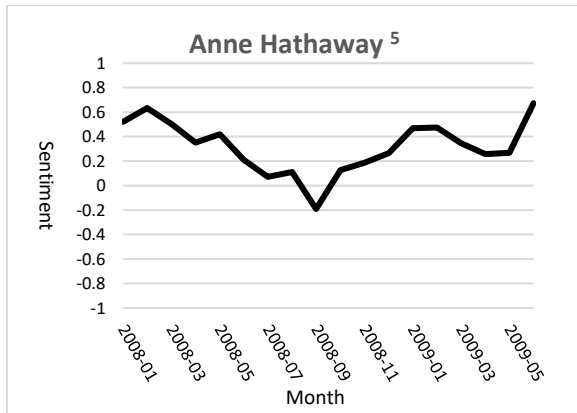
¹ In **July 2006**, **Mel Gibson** was arrested for driving under the influence and also separated from his wife.

² In **August 2007**, **Owen Wilson** tried to commit suicide.



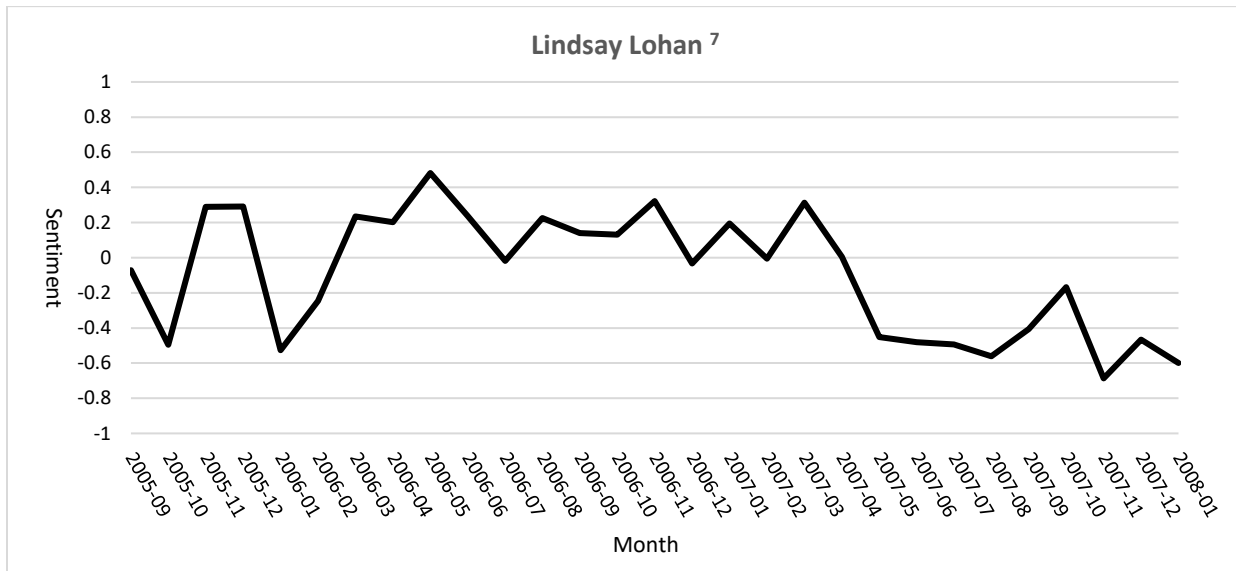
³ In **February 2007**, **Britney Spears** was admitted to a few drug rehabilitation facilities. She also shaved her hair.

⁴ In **September 2006**, **Anna Nicole Smith's** son died; in **November** of that year she got into trouble with the law.



⁵ In **June 2008**, **Ann Hathaway**'s partner, Farraello Follieri, was arrested on charges of defrauding investor scheme; in **September**, Follieri pled guilty and received a prison sentence; in **November**, Hathaway began dating a different partner.

⁶ In **August 2008**, David Duchovny announced that he checked himself into a sex addiction rehabilitation clinic; two months later, in **October** he separated from wife, Tea Leoni.



⁷ **Lindsay Lohan – Prominent life events from September 2005 to January 2008:**

October 2005: Accident with paparazzi photographer and minor injury.

January 2006: Another accident with paparazzi photographer and minor injury.

May 2006: Release of the romantic comedy *Just My Luck*, starring Lohan, for which she reportedly received a salary of seven million dollars.

May 2006: Lohan was arrested for driving under the influence and entered a drug rehabilitation center.

July 2006: Lohan was arrested again for driving under the influence.

August 2006: Lohan was arrested for Cocaine use, spent one day in jail, and subsequently entered a drug rehabilitation center again.

November 2007: Lohan was arrested again for drug use.

Appendix 2. Distinguishing Male and Female Names in Lydia

Names are first marked up in the Lydia natural language processing (NLP) pipeline based on a technique that involves lists of first names and surnames, grammatical information (part of speech tagging), and machine learning applied to the context of the entity. Anaphora Resolution (Lappin and Leass 1994; Mitkov 2002) then attempts to resolve multiple ways of referencing the same entity. Anaphora can be defined as a linguistic relation between two textual entities, which is determined when a textual entity (the anaphor) refers to another entity of the text which usually occurs before it (the antecedent). The process of determining the antecedent of an anaphor is referred to as anaphora resolution. In the following example, (1) and (2) are utterances; and together, they form a discourse: (1) Jack helped Connie; (2) He was generous. Human readers and listeners can quickly and unconsciously understand that the pronoun "he" in utterance (2) refers to "Jack" in (1). However, the underlying process of how this is done is not completely clear, presenting a challenge for natural language processing by computer scientists. Anaphora resolution is an algorithm designed to deal with this challenge. For example, an article that refers to John Smith and later to a Mr. Smith will resolve the latter to the former. To differentiate between male and female names in our news corpus, we used the 1990 U.S. Census data (U.S. Census Bureau 1990) for which common male and female first names were available. This list includes 1,219 male first names and 4,275 female first names, and it covers an approximately equivalent fraction of men and women in the U.S. population (about 90 percent for each).

In cases where the less common sex for a given name occurred over 10 percent of the time (e.g., when a typically male name was used in over 10 percent of the cases as a female name), the name was categorized as sex-ambiguous and excluded from all reported analyses. It is unclear whether these names significantly differ from the overall sex ratio, but their paucity minimizes their potential bias. Although sex-ambiguous names have become significantly more popular in recent years, they still make up less than 4 percent of all names, and their inclusion would likely not change our results substantially. Similarly, names that did not appear in any census source were removed from our analysis. The majority of these unknown names were, in fact, NLP artifacts, along with a small number of rare foreign names. Table S1 presents a random sample of 100 names from our analysis and demonstrates how these were classified as "male," "female," "sex-ambiguous," or "unknown." Of the names in this random sample, 52 were identified as male and 32 were identified as female (seven were identified as sex-ambiguous and nine as unknown). Examining these names, it appears there were no cases of major miscategorization (i.e., a female identified as a male or vice versa).

Table S1. A Random Sample of 100 Names and the Way They Were gender-Classified by the Lydia System Using Anaphora Resolution

Man	Man (continued)	Woman (continued)	Gender-Ambiguous
Anthony Arnold Holds	Michael Chan	Carolyn Robinowitz	Chris Dodds
Austin Feistl	Michael Sookiayak	Charmaine Harvey	Chris Yon
Brad Wiesley	Michael Tabor	Chelsie McGorry	Devin Logan
Brandon Henson	Miles Fairchild	Claire Sutton	Lee Thomas
Brandon McEndaffer	Miles Kane	Denise Doherty	Robin Samuelsen
Brett Backwell	Mitch Farrington	Diane Best	Tracy Corbett
Curtis Staley	Nathan Colbert	Eileen Wong	Tracy L.
Daniel A. Lashof	Nicholas Utphall	Elizabeth Teeter	
Don Moseman	Patrick Kilduff	Esther D. Halvorson	Unknown
Donald H. Catlin	Paul Mancino	Ethel Darline Naus	Ah Yin Eng
Doug Stenger	Peter Saraf	Gail Conti	Force Peter Teets
Edward N. Heath	Preston Falls	Gina Binkley	High Leg Kick
Frank Doucette	Raymond Isherwood	Joyce Cashman	Jylmarie Kintz
Harry Towns	Raymond Miller	Judith Wartels	Marino Salas
Jacques Cesaire	Robert A. Mulligan	Lacey Andresen	Ranjit Walia
James Ciccolini	Robert Abelson	Linda Leis	Ritch Price
Jeff Garlin	Robert W. Geyer	Melissa McCoy-Garziona	Shoehorn Michael Conley
John Michael Wallace	Scott Holliday	Michelle Dohm	Wease Day
Jonathan Lephass	Thomas DeMartino	Pamela Mondo	
Jonathan Marvel	Todd Stottlemire	Phyllis Wade	
Joseph Faretra	Tony Iniguez	Rachel Belanger	
Julio Aleman	Vincenzo Romeo	Rachel Rush	
Kevin Undershute	Xavier R. Donaldson	Ryann Richardson	
Margarito Brito		Sarah Walter	
Mark Colwell	Female	Sarah Wetherill Okumura	
Martin Devaney	Abby Wagner	Susan Relland	
Matt Gorski	Barbara O'Regan	Valerie Barnes	
Maurice Higgins	Beverly Stripling	Victoria Lloyd	
Michael Bloomquist	Bonnie Featherstone		

Appendix 3: Examples of names in various categories

Table S4 below shows examples of 5 not so famous names (fewer than 100 mentions in our database) and 5 famous names (more than 100,000 mentions in our database) for each of the major professional categories in our analysis. These examples were randomly sampled from each pool of names.

Table S4. Randomly Sampled Names by Category and Fame		
	Not so famous (fewer than 100 mentions)	Very famous (more than 100,000 mentions)
Politicians	Peter J. Biondi	Nancy Pelosi
	Myron T. Herrick	Donald Rumsfeld
	Robert Fiddler	Kathleen Sebelius
	Mable Thomas	Arnold Schwarzenegger
	Albert Gorton Greene	George Pataki
Criminals	Ronnie Ferrell	Phil Spector
	Benjamin Atkins	Kenneth Lay
	John Eric Armstrong	Andrea Yates
	Floyd Allen Medlock	Bernard Madoff
	Norman Bernard	Brian Nichols
Businesspeople	Rick Dalzell	Steven Spielberg
	David H. Long	Donald Trump
	Selina Tobaccowala	Warren Buffett
	Chester McNulty	Jerry Jones
	Don Hankey	Kenneth Lay
Entertainers	Julie Mitchum	Meryl Streep
	Clint Ford	Michael Moore
	Rock Florian	Paula Abdul
	Robert Cicchini	Jennifer Hudson
	Suzanne Crough	Stephen Colbert
Athletes	Alison Nicholas	Jim Furyk
	Andre Wakefield	Michael Jordan
	Dennis DuVal	Maria Sharapova
	Ray Owes	Isiah Thomas
	Jessica Rostedt	Carmelo Anthony
Scientists*	Howard R. Driggs	Paul Krugman
	Thomas Shenk	Lisa Nowak
	Vern Oliver Knudsen	Ken Thompson
	Michael Quinn Patton	Doris Kearns Goodwin
	Richard L. Schmalensee	Ralph Nader

* In the scientists category, there were not enough names with 100,000+ mentions, so we sampled among names with 10,000+ mentions.

Appendix 4: Common adjectives cooccurring with female and male names

Table S3 below shows from among thousands of adjectives those 40 that were most commonly cooccurring with either female or male names in our sample. Of note, the table shows only minor differences between the most common adjectives associated with women and men. In fact, the first 13 adjectives most commonly cooccurring with male names are the exact same 13 adjectives that co-occurred with female names, and in very similar order. Out of the 40 most common adjectives, only five are unique to men or women (loving, wonderful, beautiful, beloved, and safe for female names; guilty, offensive, responsible, fair, and critical for male names). The table further demonstrates that both female and male names are mostly associated with positive adjectives. Only 15 percent of the adjectives most commonly associated with female names, and only 22 percent of those most commonly associated with male names carry a mostly negative connotation. Finally, adjectives that focus on physical appearance (such as “beautiful”, “pretty”, or “attractive”) were not very common. While these findings do not preclude the possibility that women are more often associated with bodies and the private sphere and men with mind and the public sphere, they do suggest that gender differences in coverage may not be as large as previously suggested.

Table S3. The 40 Adjectives Most Commonly cooccurring with Female and Male Names

Female		Male	
1. good (n=1724)	21. perfect (360)	1. good (n=3377)	21. serious (779)
2. right (1476)	22. easy (350)	2. right (2852)	22. nice (766)
3. senior (1371)	23. successful (333)	3. best (2656)	23. original (757)
4. best (1300)	24. proud (322)	4. senior (2395)	24. fine (746)
5. special (1246)	25. fine (321)	5. able (1767)	25. true (734)
6. able (889)	26. popular (310)	6. better (1752)	26. offensive (728)
7. free (835)	27. legal (308)	7. special (1670)	27. easy (715)
8. better (801)	28. hot (307)	8. hard (1514)	28. perfect (715)
9. hard (753)	29. dead (298)	9. free (1483)	29. dead (708)
10. sure (628)	30. original (294)	10. sure (1395)	30. wrong (682)
11. important (614)	31. serious (281)	11. old (1374)	31. proud (671)
12. strong (588)	32. true (272)	12. strong (1332)	32. popular (650)
13. old (564)	33. outstanding (270)	13. important (1303)	33. responsible (635)
14. loving (458)	34. beautiful (270)	14. bad (1063)	34. outstanding (619)
15. happy (454)	35. grand (266)	15. guilty (932)	35. grand (608)
16. bad (395)	36. active (260)	16. difficult (914)	36. hot (602)
17. wonderful (379)	37. beloved (258)	17. happy (850)	37. active (601)
18. nice (379)	38. wrong (248)	18. successful (843)	38. fair (574)
19. difficult (368)	39. criminal (245)	19. legal (815)	39. favorite (551)
20. favorite (367)	40. safe (242)	20. criminal (799)	40. critical (550)