

# Exploring Systematic Bias through Article Deletions on Wikipedia from a Behavioral Perspective

Zena Worku, Taryn Bipat, David W. McDonald, Mark Zachry

Human Centered Design & Engineering

University of Washington

Seattle, WA

zenagworku@gmail.com, tbipat@, dwmc@, zachry@uw.edu

## ABSTRACT

The gender gap in participation on Wikipedia is well established. The impacts of the participation bias on content may be expressed through a number of different behaviors. This research considers potential content biases that may result from efforts to delete content. We collect deletion data on a sample of article pages selected based on a method designed to identify content of likely interest to a specific group of people; men and women in this case. The analysis illustrates that there does not appear to be a systematic bias resulting in more deletions or nominations for deletions against content of likely interest to women. We consider how these results improve our understanding of bias and the ways peer production systems can mitigate the potential for biases.

## CCS CONCEPTS

•Human-centered computing~ Collaborative and social computing~ Empirical studies in collaborative and social computing

## KEYWORDS

Article Deletion, Gender, Peer Production, Systemic Bias, Wikipedia

## INTRODUCTION

Sara loves to cook. She's been hearing about the trendy "mug cakes" and wants to learn more about them. She knows that Wikipedia is a good source for information. She navigates to Wikipedia and enters "mug cake" in the search bar. Her search yields a page of search results, but no page for mug cake. This causes Sara to wonder why there is no page for mug cakes?

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

*OpenSym 2020*, August 25–27, 2020, Virtual conference, Spain

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-8779-8/20/08...\$15.00

<https://doi.org/10.1145/3412569.3412573>

One of Wikipedia's stated goals is to be a repository of all human knowledge. That there is no article specifically on Sara's desired topic is a concern. Understanding how and why some content is omitted from the encyclopedia is an important question. It could reflect

- Gaps in the experiences and knowledge of those who create and edit content,
- Aspects of collective decision-making processes, or
- Specific biases of the collective or broader more pervasive societal biases from which individuals would have trouble breaking free.

Wikipedia is a peer production community which means that it relies on the voluntary contributions of many people. It seems natural that those volunteers will create and edit content that specifically interests them. Further, when deciding whether some content merits inclusion into the encyclopedia it would seem obvious that those volunteers would feel more confident about the merits of some content over other content.

In the case of Wikipedia, a common result of contributor surveys indicates that 80% or more of the contributors are male. This "gender gap" in participation raises the specter of broad potential biases in the encyclopedic content. Prior studies have validated the existence of biases. One study considered biographies of women scientists and found that the women covered were more globally notable than the men covered in similar articles [16]. Another study based on articles about movies of supposed interest to men and women found that men and women Wikipedia editors have slightly disparate category foci with men editing more in Geography and Science and women editing more in People and Arts [11].

While these studies importantly point to challenges for the encyclopedia's content, they also illustrate some challenges in assessing systemic biases. One challenge is in the scope of the existing bias. Wikipedia covers a wide range of topics and claiming systemic bias based on samples of narrow topics might be overgeneralizing. A second challenge is how to operationalize the bias. Most prior studies have often focused on "acts of omission." That is, they have sought to understand what content is missing. An act of omission may be somewhat innocent because individuals just may not know what should be included. "Acts of commission," the behavioral activities within Wikipedia that may be resulting in systemic biases are less frequently studied. Another study of biographies of women scientists found

that articles appeared to mention familial ties, specifically marital status more frequently for women when compared to men [15]. These differences are created by the editors as a function of their work activity. An act of commission in this case is reflective of the behaviors of the editors and makes visible biases that become incorporated into the encyclopedia.

We note that in raising this behavioral view of biased production we are specifically trying to avoid ascribing intent by the individual editors. Systemic biases may be a function of an individual's environment and, as such, it may be very difficult for individuals to see in themselves. The potential here is that by taking a clear analytical stance, maybe research can uncover ways that systemic bias happens during peer production and then work to develop techniques to mitigate and, potentially, correct systemic biases over time. Thus, the focus of this research is to consider one particular behavior that may be leading to systemic biases in Wikipedia content. Specifically, we attempt to answer:

### **Is content supposedly of more interest to women being actively deleted from Wikipedia?**

In the following sections we first cover some prior work that has considered deletion activity, aspects of systemic bias, and the gender gap on the English Wikipedia. We then describe the methods of this study. Our approach builds off a previous method to measure topical representativeness [12]. Our research uses that method to identify content that is likely to be of interest to women or men. Our elaboration then explores if efforts to delete content demonstrate behavioral content biases which may reflect systemic biases on Wikipedia. We close with some discussion that considers other places where systemic biases might be observed through actions taken by Wikipedia editors.

## **Considering Deletion and Bias**

With over 5 million articles on the English-language version alone, Wikipedia editors create hundreds of pages a day and also delete hundreds of pages a day [10]. As such, pages not found on Wikipedia may have existed at one time, but have been removed as a result of a deletion. Previous work has detailed the processes involved in assessing whether an article fits the project's encyclopedic standards [5]. Debates around what is considered encyclopedic and to whom the content is important surround the complex norms in the types of articles that are nominated for deletion [5]. When an article is deleted from Wikipedia, it is most commonly the result of one of three different processes: Proposed deletion (WP:PROD), speedy deletion (WP:CSD) or article for deletion (WP:AfD)<sup>1</sup>. The procedures for all three are well documented and generally reflect a deliberative process that focuses on the individual merits of any given topic. For the

purposes of this work, we focus on speedy deletion (CSD) and article for deletion (AfD) activities as those processes are currently reflected in logs of Wikipedia activity. PROD was not included in the deletion logs when this study was conducted.

Wikipedia exists in a difficult tension. On the one hand the goal of the encyclopedia is to contain the sum of all human knowledge. On the other hand, some topics have very minor value in the big scheme of things. This tension is reflected in a long running debate on Wikipedia between "inclusionism" and "deletionism." The debate hinges on where one stands with regard to the idea of "notability." In Wikipedia, a topic merits inclusion in the encyclopedia if it satisfies a complex set of criteria described in the notability policy (WP:NOTE)<sup>2</sup>. Inclusionists support a more liberal take on the concept of notability and advocate for "lower entry barrier rates" for including topics. By contrast, deletionists support a more selective understanding on the concept of notability and are "motivated by a desire that Wikipedia be focused on and cover significant topics." Prior work that considered how the decisions to delete articles were resolved showed that individual biases were at play in the decision making. Preference statements to keep or delete articles were subject to prior biases to keep or delete and herding, wherein groups of like-minded individuals all visit the same discussion; a study reveals that both had affected the final outcome of keeping or deleting a page [14].

Prior work provides some insight on other potential impacts on how deletion decisions are being made. The formalization of norms on Wikipedia through strict guidelines and procedures has discouraged newer editors from contributing to some of the key decision-making processes, including deletion decisions [7]. The consequence of excluding newcomers is that perspectives that are not part of the existing views of editors or topics with which current editors are not familiar have a harder time being included. Another study that considered the "long tail" of less popular content noted that the number of articles being deleted was increasing due to the "lack of notability" [10]. This further suggests a resistance to certain topics by the editors. This resistance to certain topics points to potential biases being systematic in the way the community and its processes function.

## **Systemic Bias, Systematic Bias, and The Gender Gap**

A bias can be systematic without yet being systemic. When a bias becomes rationalized as part of a decision or embodied in some piece of code, it clearly becomes systematic. Further, when that biased decision-making process is broadly adopted, say in a peer production community, or that little piece of code is widely applied to a platform, then the bias is more deeply ingrained in the social or technical system and it is better to describe that bias as systemic.

Two projects start to illustrate the subtle differences. One piece of work considered the development of articles over time

<sup>1</sup>. WP:PROD: en.wikipedia.org/wiki/Wikipedia:Proposed\_deletion  
WP:AfD: en.wikipedia.org/wiki/Wikipedia:Articles\_for\_deletion

WP:CSD: en.wikipedia.org/wiki/Wikipedia:Criteria\_for\_speedy\_deletion  
<sup>2</sup>. WP:NOTE: en.wikipedia.org/wiki/Wikipedia:Notability

and how they would become more elaborate as a function of developing the content. This 'knowledge maturation' model highly correlated with the time that the article spends in existence, resulting in a type of bias toward keeping 'mature' articles over potentially less mature and shorter-lived topics [2]. A slightly different view considered how different cultural and historic perspectives were present in the Polish edition of Wikipedia and related that to existing systemic biases in culture [3].

When noting the participation gender gap, popular media has often attributed gaps in content as resulting from targeted harassment of editors. One specific example pertains to the Gamergate controversy that began in 2014, where women in the video game industry were targeted leading to the decline of their games and input into the industry. These specific harassment efforts led to the deletion and removal of content presumed to be of interest to women [8,9]. And the popular press is not alone in suggesting that deletion is a potential way to identify bias. In unpacking the ways in which different systematic biases may be present on Wikipedia, previous work has cited deletion of content as a possible result or indicator of systemic bias [13].

The Wikipedia community has been concerned that systemic biases are undermining the encyclopedic content. Members of the community formed a WikiProject, a type of virtual team, focused on identifying and mitigating systemic biases. WikiProject Countering Systemic Bias<sup>3</sup> considers all systemic biases but has a strong justification for their work based on the existing gender gap and well publicized anecdotes that illustrate enactments of one or more biases—often biases focused at women or on topics that are of supposed interest specifically to women.

If one key way to counter a systemic bias is to rally supporters and effect positive change, then that implies one needs to have the participation of individuals who can see and clearly articulate the nature of the systemic bias. Prior work that has studied the participation gender gap has illustrated the different, systematic and systemic ways that some individuals are discouraged from full participation in Wikipedia [7]. A study by Collier et al. [4], considered whether differing participation resulted from either (1) the high levels of conflict in discussions, (2) dislike of critical environments, (3) lack of confidence in editing other contributors' work, or (4) that women have less discretionary time. In sum, Collier et al. found that participation was significantly discouraged by the contentious and conflict-driven nature of Wikipedia.

In the popular media, one well known debate centered around whether the deletion of a Wikipedia article about Kate Middleton's wedding gown was an indicator of a content bias resulting from the well recognized gender gap in participation [1]. Wagner et al. [15] analyzed article content bias on multiple dimensions and noted that biography articles about men and

women were not handled in an equivalent manner. The study showed that biographies of women on Wikipedia tend to be more linked to men while the reciprocal links from men's biographies were not as prevalent. Additionally, relationships and family-related issues are more discussed in women's biographies than in men's [15]. Analysis of metadata, language and network structure of women's biographies suggest that there is a gender bias that can be due not only to the external behaviors of society but also that this bias stems from the internal generation of content on Wikipedia. [6].

With this prior work in mind, our research is designed to link specific deletions to uncover (or discover) if the encyclopedia is marked by systematic bias from a behavioral perspective. The gender gap in editor participation motivates us to think that if there are some specific systematic behavioral activities that establish bias, then we should be able to find them directed at content that is supposedly of interest to women<sup>4</sup> when compared to similar content of supposed interest to men.

## Observing Biases: how and where?

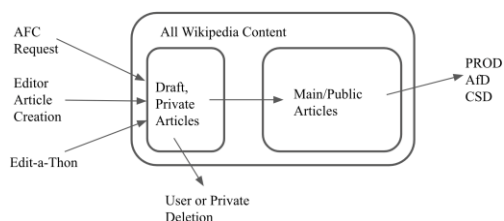
Observing bias can be somewhat challenging. In a technical sense a bias is simply when something is unbalanced, or unfair, in statistically demonstrable way. However, in a social setting people talk about bias in a more relativistic way. That is, we observe bias as relative unbalance or relative unfairness. This leads to the question – relative to what? This question matters because of endogeneity. In the context of one specific system, say Wikipedia, it is notoriously hard to simply observe the one system and identify the relative bias. A common approach, one seen in the far majority of research described above, relies on a baseline sample from some other system as the comparator. Those results are therefore making a claim that the bias observed in Wikipedia is relative to what is observed in the baseline. These claims of bias are naturally fraught with cultural or yet other systemic biases existing in the baseline. We all need to recognize that some biases may always be present in our baseline and that observing a bias may not speak to all biases present.

Another subtlety is whether the bias being observed is some outcome or a process. To try and make this point a little clearer we highlight the distinction through Figure 1 below.

<sup>3</sup>. en.wikipedia.org/wiki/Wikipedia:WikiProject\_Countering\_systemic\_bias

<sup>4</sup>. A careful reader might notice our somewhat pedantic use of "... content that is supposedly of interest ..." to a given gender. At this point we know of no specific method or prior work that has established a preference by one gender or another for

any specific Wikipedia content. We recognize this as a slight hedge on the strength of any claims we might make later in the paper, and this is one of the limitations of the work we have conducted. This content preference limitation is faced by all prior work as well.



**Figure 1. Artifacts and processes where we might be able to observe biases in Wikipedia.**

Contributor systems like Wikipedia typically have defined processes. These processes are mostly the actions and behaviors of people, which sometimes include bots. The traces of those processes can be inspected as a possible way to understand how bias might be enacted. In Wikipedia, AFC (Article for Creation), Edit-a-thons, and individual decisions by Wikipedia editors are example processes that result in content being potentially created and included in the encyclopedia. Processes like PROD, (Proposed Deletion), AfD (Article for Deletion), and CSD (Criteria for speedy delete) are examples of processes that result in content removal. Figure 1 illustrates these different processes as arrows.

Figure 1 suggests an alternative to studying bias in the processes, and that is studying or comparing content. This alternative sits partially in and partially out of the main, public article space because understanding biases by studying the content also requires an external baseline set of content for comparison.

In the work that follows, we focus on the arrow, the processes, on the right side of Figure 1. These are processes that result in the deletion or removal of content and we seek to understand how bias might be enacted there. Clearly, there are other potential sources for behavioral bias, such as bias on input processes, but this research leaves those for another study.

## Collecting Data on Deletions

The focus of this work is to understand whether there are specific differences in the number of deletions as a function of whether content is of supposed interest to women or men. This requires identifying a broad set of Wikipedia article pages that may have interest to a given gender, and then for that broad set of articles, assess the rates of deletion. There is a quirk in pursuit of this answer as we foreshadowed above. That is, there is possible content that was previously in Wikipedia, but which does not currently have an article. The method that we have adopted has curiously helped resolve that subtle issue.

We adopted the method described in Menking et al. [12]. Their method was designed to assess “topical representativeness” broadly across many potential topics in Wikipedia. One example case in the paper focused on potential bias in content of supposed interest to men or women. Following their method, we identified a set of popular magazines whose declared readership is primarily men or women. We used

EBSCO, a commercial indexing service, to collect meta-data for all of the content published in the selected magazines during 2014 and 2004. We bifurcated the meta-data depending upon whether it originated in magazines with primarily men or women readership, keeping the years separate, and then extracted the list of keywords (terms), removing any duplicates in each list. For each year, the men's and women's targeted term lists were compared, removing any identical terms that were present in both lists. In general, term lists have more terms than could be manually matched (as described in the Menking et al. paper). We took a random sample of terms from men's and women's term lists for both 2004 and 2014. We followed the procedure outlined in Menking et al. [12] for matching terms to likely articles.

We included terms from 2014 to see how a potential systematic bias might influence topics and articles that are more contemporary. Articles associated with terms from 2014 have the advantage of being more recent and thus potentially included in the encyclopedia at a time where norms were well understood, and reference materials are readily available. Articles associated with 2014 terms also face a disadvantage due to their novelty which might detrimentally impact their notability. In contrast, we included terms from 2004 in order to see how topics and their associated articles may have persisted across a longer stretch of time. A systematic bias would potentially impact these articles more severely as repeated efforts to remove or delete the content despite potential notability. Further, articles associated with terms from 2004 may (or may not) have survived the documented shift in notability standards for inclusion that the encyclopedia made in the early 2000s.

We recognized that the longevity of a given topic might also have a significant impact on notability and inclusion (or deletion) of a given article. We used our terms from 2004 and 2014 (the most recent year for which data was available when we conducted the study) to create a third set of terms that we call the decadal set (2004-2014). These are terms that occur in 2004 and also occur in 2014 for the respective men's and women's bifurcated lists of terms. Interestingly, there were many more men's terms that persist for the decade. However, both sets of decadal terms were small enough that we were able to include all of them. We repeated the article matching procedure for our decadal terms.

The data collection so far had generated three sets of terms, 2014, 2004, decadal, bifurcated by assumed interest to a target gender, with each term either associated with an article in Wikipedia or noted that there was no associated article. The next step was to collect deletions associated with the given articles/terms. The method nicely resolves the quirky problem of how to consider topics that are not in the encyclopedia. Given that our set of terms were all generated and sampled in the same way, we can use the unmatched terms—terms where we were not able to identify an associated article—to consider potential content that was present at one time but which is currently deleted.

We relied on Wikipedia's public deletion logs for Articles for Deletion (AfDs) and Speedy Deletes (CSDs) to collect the

		Terms for Analysis	Terms matched to articles	Terms with matched articles (Matched)	Terms without Matched Articles (DNE)	Articles with AfD (by Terms)	Articles with AfD (by Matched)	Articles with CSD (by Terms)	Articles with CSD (by Matched)	AfD+CSD of Unmatched Terms (by Terms)	AfD+CSD of Unmatched Terms (by unmatched)
2014	Women's	370	252	68.11%	31.89%	3.24%	4.80%	7.57%	11.20%	0.00%	0.00%
	Men's	365	310	84.93%	15.07%	3.01%	3.58%	6.85%	8.14%	0.82%	2.50%
2004	Women's	260	207	79.62%	20.38%	3.46%	4.35%	3.08%	3.86%	0.00%	0.00%
	Men's	244	210	86.07%	13.93%	2.87%	3.47%	9.02%	10.89%	1.64%	3.33%
Decadal	Women's	250	230	92.00%	8.00%	2.71%	2.79%	8.60%	9.41%	0.90%	1.67%
	Men's	402	383	95.27%	4.73%	3.73%	4.26%	9.05%	9.30%	0.75%	2.50%

Table 1. Table of ratios for article deletions and nominations for deletions of articles of supposed interest to men or women for 2004, 2014, and a set of terms that were present a decade apart. The total Terms for Analysis is different across samples because the method described in Menking et al. can result in term disqualifications that will vary across samples.

deletion information for all of the matched articles and for the unmatched terms. We resolved potential AfDs using Wikipedia's native search focused on the on-wiki AfD discussions. We manually conducted a search for each term, and recorded if a result revealed that an article had ever been nominated for deletion and how many times it had been nominated. If the search resulted in one or more nominations, we would record the separate logs for all nominations. We resolved potential CSDs using the Wikimedia Foundation Quarry<sup>5</sup> tool to query the deletion logs. Quarry queries "a set of live replica SQL databases of public Wikimedia Wikis." Our Quarry queries were focused on the English edition and were structured to consider all possible deletions for matched articles and unmatched terms. As in the case of the AfD logs, if an article or term<sup>6</sup> had one or more CSD-related records we logged the record.

## Comparing Systematic Deletions

Bluntly, there does not appear to be significant qualitative differences in the rates of AfD or CSD for articles of supposed interest to women compared to articles of supposed interest to men. We compare our data as ratios (percentages) because the numbers of terms/articles across our sets varies. Table 1 summarizes the number of terms, matched articles, AfDs and CSDs across terms for 2004, 2014 and decadal. The table represents the ratios both as a function of the total terms and as a function of terms matched to articles. Further, logged arguments for proposed deletions were also largely consistent across the articles presumed to be of interest to men and those presumed to be of interest to women. We consider each group of terms in more detail.

### 2014 Articles and Terms

For the 2014 terms, we found that articles presumed to be of interest to women had a slightly higher deletion rate for both the AfD and CSD processes. Under the AfD process, articles of

presumed interest to women had between 3.2% to 4.8% rate of being nominated, while articles presumed to be of interest to men ranged from 3.0% to 3.6% rate of nomination for AfD.<sup>7</sup> The CSD process is more aggressive, roughly nominating twice as many articles across both sets. Articles of supposed interest to women nominated for CSD ranged from 7.6% to 11.2%, while articles of supposed interest to men ranged from 6.9% to 8.1%.

From our initial list of topics of supposed interest to women, about 31.9% of topics could not be matched to an article using the matching method, and 15.1% of terms of supposed interest to men were in the same condition. These represent potential content that is not currently in Wikipedia. Our search for these generated only a few terms that came from the men's list only. There were no (zero) unmatched terms from the list of possible interest to women that resulted in a match with deleted content in the AfD and CSD logs. However, it is important to note that the numbers of terms, which may have had a corresponding article, and which are now deleted, is exceedingly small. Because the numbers are small we have combined the counts for AfD and CSD in Table 1.

### 2004 Articles and Terms

For the 2004 terms, the relative ranges and difference between AfD nominations for articles of supposed interest to women and those of supposed interest to men were very nearly the same as those for the 2014 articles. The real difference showed up in articles of supposed interest to men that were nominated for CSD. Articles of possible interest to men were nominated for CSD at a rate between 9.0% and 11.0% while for articles possibly of interest to women were only nominated from between 3.1% to 3.9%. This is a rather large difference even considering the relatively small numbers in our total sample. That content of supposed interest to men is more likely to be nominated for CSD runs contrary to common ideas regarding biases in content.

<sup>5</sup>. <https://meta.wikimedia.org/wiki/Research:Quarry>

<sup>6</sup>. There is a subtle issue to consider here. When "terms" are searched for AfD or CSD records, as in the method of Menking et al., there might not be an exact match

between term and deleted (or absent) article. In those cases, we attempted to apply the matching method (Menking et al., 2017) as best possible as judged by the authors.

<sup>7</sup>. The ranges reflect whether the denominator of the ratio is either the total number of terms (low end) or just the matched (or unmatched) set of terms (high end).

When we consider the terms that were not matched with articles from the 2004 lists, again, there were a few terms reflecting deleted content that came from the terms reflecting content of possible interest to men, and no (zero) terms that matched deleted content of supposed interest to women. Like the 2014 terms, we combined both AfD and CSD counts because the total numbers were so small.

## Decadal Articles and Terms

Recall, our decadal terms reflect terms from the respective men's or women's interest term lists that were present in both 2004 and also appearing in 2014, and which were not part of the random selection of terms for either 2004 or 2014 term lists. As such, these terms reflect topics that have demonstrated some longevity in the indexed sources.

The decadal terms showed very little difference in rates of articles nominated for deletion between articles that had been matched for possible interest to men and women. Interestingly, unlike the individual rates of 2004 and 2014 matched articles, where articles of supposed interest to women were nominated at a slightly higher rate, the decadal terms revealed that it was articles of supposed interest to men that were nominated at a higher rate.

Our results show a very small difference in CSD nomination rates for the terms that were not matched to Wikipedia articles in the first part of our process. Unlike the unmatched terms for 2004 and 2014, the decadal terms of possible interest to women yielded a very low number of deleted pages, while the rate of deleted pages matching unmatched terms for men was consistent with the rates for 2004 and 2014 terms.

## Considering AfD, CSD and Multiple Deletion Comments

We were somewhat surprised that our deletion data did not reveal a systematic bias. While the data hints at some issues, there are some clear instances where topics of likely interest to men are under more deletion threat. We took a closer look at the comments for deletion nomination, and comments in the deletion log, to try and understand whether small numbers were hiding a systematic bias that might be uncovered through careful qualitative reading.

### AfD Nomination Comments

We did not observe any blatant differences in the language used to propose deletions for articles presumed to be of interest to men or women, but one thing we did note was the surprisingly high count of articles deleted on April 1st, or what is popularly known as April Fool's day in the US. It is hard to know whether to take these nominations seriously. However, nominating a page for deletion could be perceived as a type of content harassment. That the nomination was performed on a day where practical joking is to be expected may simply provide cover for potential harassment.

### CSD (Speedy Deletion) Deletion Comments

CSD or Speedy Delete is a process to quickly and permanently remove content that no one contests deleting. When articles are deleted through CSD, administrators enter a deletion comment in a log. We considered the comments for all of the articles that were removed through CSD. The far majority of the articles deleted were deleted using criterion G6, also known as general technical deletions. Wikipedia CSD policy defines this as "uncontroversial maintenance." About 70.4 percent of articles nominated at least once for speedy deletion across our three sets of terms were done so as a part of the technical deletions.

The second most common reason for a CSD deletion was A7, which is identified as articles having "no indication of importance (people, animals, organizations, web content, events)." These articles are removed essentially because they do not meet Wikipedia notability requirements.

### Multiple Deletions

Across all three datasets, there were a few instances of articles being deleted or nominated for deletion two or more times. Some articles were nominated and/or deleted through the same process, and some articles were nominated and/or deleted multiple times through both AfD and CSD processes. We did not observe any differences in these articles or terms being targeted for deletion based on whether the articles were of supposed interest to men or women.

## DISCUSSION

Our goal with this work was to look for possible deletion behaviors that suggest a systematic bias relative to content. Prior work has suggested that efforts to delete content could be considered a type of harassment or assault on that content. Further, prior work has suggested sets of strong biases against content that might be of more interest to women. Therefore, our working hypothesis was that by considering specific deletion activities we might be able to see a systematic bias against content that was possibly of more interest to women.

These deletion data do not show a strong, systematic, bias resulting from deletion activity. While the data for 2004 and 2014 show slightly higher AfD nomination rates for content of possible interest to women, while CSD rates are mixed and have content of interest to men more heavily nominated for 2004 content. Further, considering content that is not in the encyclopedia, across all samples, content of possible interest to men was more likely deleted. However, to be fair, there is more content of possible interest to women that was, likely, never included and therefore it is not possible for it to be deleted.

A more positive interpretation of our results would be that content that is supposedly of more interest to women is staying in the encyclopedia once it is included. Our result clearly follows from, and supports, findings by Wagner et al. who found that biographies of women scientists were more globally notable [16]. Our result covers a wider selection of topics, and may highlight that the potential difficulty of getting content inserted may pay

some dividends in that less work is required to keep the content in the encyclopedia.

Given that our modest quantitative study did not yield a clear distinction, we took a little time to dig into some of the more qualitative aspects of the deletions. First, we considered the biographies and whether some biographies are more likely to be nominated. Second, we considered in the cases of CSD what types of reasons were given for the deletion when administrators took a deletion action.

## Biographies

Prior work on gender differences in Wikipedia have found strong biases related to biographies [6, 15,16]. Over all three sets of terms, 11.4% of articles deleted, or nominated for deletion, were biographies. We extracted names from both men's and women's lists to see if there were differences in nominations or deletions between groups. In this case there were more biographies originating on the list of terms of possible interest to men either nominated for deletion or deleted. We speculate, that this is more reflective of an act of omission. That is, biographies that might be of more interest to women are perhaps not being deleted or nominated for deletion simply because they are not there in the first place.

## Article Redirects

Prior work shows that a higher proportion terms of likely interest to women redirect to a topic that may or not be of interest to women [12]. The nature of a redirect is that two separate terms are determined to mean essentially the same thing. However, this consolidation of terms might have some gendered consequences. For instance, a search for "Stewardess" currently redirects to "Flight Attendant." While Flight Attendant is the currently accepted term for this career, the reality is that this policy essentially erases cultural knowledge and historical perspectives related to women that is not captured once terms are homogenized. This suggests that future work for understanding systemic bias should include further investigation of redirects.

## Bias and Content Inclusion Processes

As we attempted to distinguish in Figure 1, when examining biases one can attempt to focus on either the existing content or the surrounding processes related to content inclusion and deletion. This study considered key deletion processes.

Critical processes related to the way that content is included in the encyclopedia remain to be studied. One possible way to study content inclusion and potential systematic bias is to study the Article for Creation (WP:AfC) process. Through AfC, users request that an article be created. This process might expose the types of articles that are not included in the encyclopedia, and simultaneously understand which potential new articles are being ignored—a different potential explanation for content bias if deletion processes can be ruled out.

Further, the Wikimedia Foundation has been backing a large number of community driven Edit-a-thons as a way to broaden participation. These Edit-a-thons are an interesting and explicit content inclusion process. It seems a clear follow-on study would be to consider whether content that takes this alternate route to inclusion is fully included or whether it suffers from some biases from either not being accepted into the main article content space or whether it is under deletion process assault from day one.

## Conclusion

Behaviors that stem from systemic bias are often pernicious at best, and malicious at their worst because gender bias is so embedded in culture. Malicious forms of bias towards women on Wikipedia has been well-documented in numerous accounts of online harassment. These attacks blatantly expose dangerous attitudes and threats against women. In the case of deletions, we were not able to find this kind of deliberate, mal-intended attack on topics presumed to be of interest to women. Through our qualitative look at deletion log records we found no indication of targeted harassment in the arguments for deletion. Malicious exclusionary attitudes and practices around content exist, but our findings suggest that deletion talk pages are not generally where they play out. A large-scale trend of women's topics being deleted more frequently would have given us an understanding of some more pernicious behaviors that can be tied back to systemic bias. In the context of Wikipedia, we urge that future work should be done around the more pernicious ways that system bias is reinforced.

## ACKNOWLEDGMENTS

We thank Cole Lundell and Amanda Menking for their early contributions to this work. We would also like to acknowledge support from National Science Foundation (NSF) grant, IIS-1162114.

## REFERENCES

- [1] Torie Bosch. (2012, July). Kate Middleton's wedding gown and Wikipedia's gender gap. Slate. Retrieved from [http://www.slate.com/blogs/future\\_tense/2012/07/13/kate\\_middleton\\_s\\_wedding\\_gown\\_and\\_wikipedia\\_s\\_gender\\_gap\\_.html](http://www.slate.com/blogs/future_tense/2012/07/13/kate_middleton_s_wedding_gown_and_wikipedia_s_gender_gap_.html)
- [2] Simone Braun and Andreas Schmidt. (2007). Wikis as a technology fostering knowledge maturing: What we can learn from Wikipedia. In the 7th International Conference on Knowledge Management (IKNOW'07), Special Track on Integrating Working and Learning in Business (IWL).
- [3] Ewa S. Callahan and Susan C. Herring. 2011. Cultural bias in Wikipedia content on famous persons. *J. Am. Soc. Inf. Sci. Technol.* 62, 10 (October 2011), 1899-1915.
- [4] Benjamin Collier and Julia Bear. (2012.) Conflict, criticism, or confidence: An empirical examination of the gender gap in Wikipedia contributions. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12). ACM, New York, NY, USA, 383-392.
- [5] R. Stuart Geiger and Heather Ford. 2011. Participation in Wikipedia's article deletion processes. In Proceedings of the 7th International Symposium on Wikis and Open Collaboration (WikiSym '11). ACM, New York, NY, USA, 201-202.

---

<sup>a</sup>WP:AfC: [en.wikipedia.org/wiki/Wikipedia:Articles\\_for\\_creation](https://en.wikipedia.org/wiki/Wikipedia:Articles_for_creation)

- [6] Eduardo Graells-Garrido, Mounia Lalmas, and Filippo Menczer. (2015). First women, second sex: Gender bias in Wikipedia. In Proceedings of the 26th ACM Conference on Hypertext & Social Media (HT '15). ACM, New York, NY, USA, 165-174.
- [7] Aaron Halfaker, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. (2013). The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline. *American Behavioral Scientist* 57, no. 5, 664-688.
- [8] Jay Hathaway. (2014, Oct.). What Is Gamergate, and Why? An explainer for Non-Geeks. Gawker. Retrieved from <http://gawker.com/what-is-gamergate-and-why-an-explainer-for-non-geeks-1642909080>
- [9] Erik Klein. (2014, Sept.). GamerGate: A Closer Look At The Controversy Sweeping Video Games. Forbes. Retrieved from <https://www.forbes.com/sites/erikkain/2014/09/04/gamergate-a-closer-look-at-the-controversy-sweeping-video-games/#7ffb61534f8d>
- [10] Shyong (Tony) K. Lam and John Riedl. 2009. Is Wikipedia growing a longer tail?. In Proceedings of the ACM 2009 international conference on Supporting group work (GROUP '09). ACM, New York, NY, USA, 105-114.
- [11] Shyong (Tony) K. Lam, Anuradha Uduwage, Zhenhua Dong, Shilad Sen, David R. Musicant, Loren Terveen, and John Riedl. (2011). WP:Clubhouse?: An exploration of Wikipedia's gender imbalance. In Proceedings of the 7th International Symposium on Wikis and Open Collaboration (WikiSym '11). ACM, New York, NY, USA, 1-10.
- [12] Amanda Menking, David W. McDonald, and Mark Zachry. (2017). Who Wants to Read This?: A Method for Measuring Topical Representativeness in User Generated Content Systems. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '17). ACM, Portland, OR, USA, 2068-2081.
- [13] Jodi Schneider, Alexandre Passant, and Stefan Decker. 2012. Deletion discussions in Wikipedia: decision factors and outcomes. In Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration (WikiSym '12). ACM, New York, NY, USA, , Article 17 , 10 pages.
- [14] Dario Taraborelli and Giovanni Luca Ciampaglia. 2010. Beyond Notability. Collective Deliberation on Content Inclusion in Wikipedia. In Proceedings of the 2010 Fourth IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshop (SASOW '10). IEEE Computer Society, Washington, DC, USA, 122-125.
- [15] Claudia Wagner, David Garcia, Mohsen Jadidi, and Markus Strohmaier. (2015). It's a man's Wikipedia?: Assessing gender inequality in an online encyclopedia. In Proceedings of the International AAAI Conference on Web and Social Media (ICWSM2015).
- [16] Claudia Wagner, Eduardo Graells-Garrido, David Garcia, and Filippo Menczer. (2016). Women through the glass-ceiling: Gender asymmetries in Wikipedia. *EPJ Data Sci.* 5: 5. <https://doi.org/10.1140/epjds/s13688-016-0066-4>

## APPENDIX: Example Terms from the 2004, 2014, and Decadal Datasets

2004	
Men's	Women's
Activism	4-Androstenedione
Actor	Activities of daily living
Actor	Actor
Acupuncture	Acura
Adolescence	Adidas
Adoption	Aesthetics

Advertising	Air filter
Airline	Airpower
Airport	Alexander Payne
Alanis Morissette	All of Us
Alcohol	Alonzo Mourning
Altruism	Altitude
Alzheimer's Disease	Alzheimer's disease
Amusement	Amsterdam
Anna Quindlen	Andrew Sullivan
Anthocyanidin	Angels in America (miniseries)
Antidiarrheal	Ant
Antifungal	Aquatic Science
Antonio Banderas	Aramaic language
Archaeology	Artery
Arizona	Arthur Phillips
Aromatherapy	Ashton Kutcher
Asthma	Athlete
AT&T Mobility	Autism spectrum
Atlantic City, New Jersey	Automotive battery
Atmospheric temperature	Automotive electronics
Attention	Bagel
Attention deficit hyperactivity disorder	Balloons
Author	Bandage
Autoimmune disease	Bank robbery
Baby Bottle	Barbell
Baja California	Baseball
Ballot	Baseball (ball)
Balsamic Vinegar	Baseball uniform
Bar	Bedding
Bath & Body Works	Bilingual education
Battery (electricity)	Bill Clement
Beard	Biomedical engineering



Beaver Valley, Delaware and Pennsylvania	Blood Done Sign My Name
Beiersdorf	Blood sport
Beverly Hills, California	Body water
Bic (company)	Bookcase
Birth control	Bossa nova
Black & Decker	Bowls
Blanket	Boxing
Blender	Bread machine
Body image	Bridge
Boot	Bridget Jones: The Edge of Reason (film)
Botulinum toxin	Bulldog
Boxer (dog)	Bullet and a Target
Bridge	Bung
Bristol-Myers Squibb	Burn
Brother	Businessperson
Building	Cabinet (government)
Burping	Cancún
Cannabis (drug)	Capsaicin
Car	Car
Cardiovascular disease	Cashmere wool
Carol (music)	Centre (ice hockey)
Ceiling	Chianti
Centrepiece	Child labour
Chad Michael Murray	Coaches (sport)
Chewing Gum	Coleman company
Chickenpox	Contamination
Child	Convenience food
Child abandonment	Convention (meeting)
Children's literature	Costa Careyes, Jalisco
Chocolate Brownie	Creative Technology Limited
Chutney	Criminal Investigation

Class reunion	Cruise Line
Coffee preparation	Curb
Cold	Cybercrime
Commercial Bank	Dean and Dean Caten
Computer reservations system	Defense (legal)
Constipation	Delicatessen
Container garden	Demographic profile
Cooking	Destination spa
Cookware and bakeware	Destiny's Child
Coronary artery disease	Disc jockey
Cosmetology	Donald Trump
Counseling Psychology	Dopamine
Crabtree & Evelyn	Drug
Crime Prevention	E-book
Customer Service	Electric guitar
Death	Elementary School
Deep-submergence vehicle	Engine
Dental restoration	Enzyme
Developmental psychology	Enzyme inhibitor
Diane Sawyer	Ermenegildo Zegna
Dishwashing	Etymology
Disposable camera	Fear
Document	Fibula
Dog	Fish
Dog collar	Fishing bait
Door furniture	Flare
Door handle	Frank Bender
Doughnut	French West Indies
Dress Code	Funeral
Education Finance and Policy	Gavin Newsom
Employment	Gonorrhea
Encyclopedia	Government budget balance

Erotic Literature	Grappa
Etiquette	Gratuity
Eveready Battery Company	Guilt (emotion)
Family life education	Hank Aaron
Fashion	Harrison Pope
Feminist Ethics	Health risk assessment
Financial planner	Helicopter
Flight Attendant	Helmut Lang (artist)
Food	Hip hop
Food Manufacturing	Historic Site
Foot	Historical sociology
Fortune-telling	Honda
Fructose	Informed Consent
Handmade jewelry	Invention
Help (command)	Jack White
Herpesviridae	James Franco
Holiday	Javier Bardem
Holy Week	Jim Carrey
Housekeeper (domestic worker)	Joan Riudavets
Houseplants	Job
Human sexual response cycle	John Edwards
Ian Kerner	Kasey Chambers
Individual Family Service Plan	Ken Kesey
Indoor games and sports	Laird Hamilton
Intellectual giftedness	Larry Bird
Interpersonal relationship	Laura Bush
Jamie Lee Curtis	Lens, Pas-de-Calais
Jane Bryant Quinn	Lexicology
Jean Chatzky	Lie
Jennifer Connelly	Lining (Sewing)
Jennifer Love Hewitt	Love
Lancôme	Malapropism

Landscape Design	Malibu, California
Laura Bush	Man in the Mirror
Laura Hillenbrand	Managing up and managing down
Light	Measurement
Lipstick	Memorial
Long-distance calling	Mormons
Magnet	Mosquito control
Major depressive disorder	My Life Among the Serial Killers
Male	National Institutes of Health
Mandy Moore	Navigational aid
Manufacturing	NEC
Marble	Nellie McKay
Maybelline	Nike, Inc.
Melatonin	Nonsteroidal anti-inflammatory drug
Memorex	Nose
Michelle Bernstein	Odor-Eaters
Mobile phone	Officer (armed forces)
Mop	Paint
Musical instrument	Party platform
Myocardial infarction	Patient
Newspaper	Patronage
Nike, Inc.	Paul Bremer
Oral hygiene	Peru
Oscar Blandi	Pet
Parent	Pet store
Parent-Teacher Association	Pianist
Paul J. Coleman	Pregnancy
Pelvic inflammatory disease	Presenter
Pet	Public opinion
Physical attractiveness	Puma SE
Physical fitness	Quadriceps femoris muscle

Portion control (dieting)	Ray (film)
Professional audio store	Rhinoplasty
Psychological resilience	Richard Branson
Radio	Ringo Starr
Red Meat	Road trip
Red Springs, North Carolina	Robbery
Refugee	Robert Altman
Risk	Roberto Cavalli
Role Models	Scooter (motorcycle)
Romance (love)	Secondary School
Romantic comedy film	Sexual attraction
Secrecy	Seymour Cassel
Self-employment	Shoe size
Sense	Shopping
Sewing	Sin
Silver	Small business
Sister	Sneakers
Solanaceae	Sonny Liston
Squat	South Fork American River
Stress relaxation	Space environment
Student	Sport psychology
Subconscious	Starbucks
Surgery	Sunscreen
Telecommunication	Superior-subordinate communication
Television	The Waste Land
Television advertisement	Thomas P.M. Barnett
Tertiary sector of the economy	Thrombosis
Testicle	Tree climbing
The Apprentice (TV series)	Troy Aikman
The Bachelorette	Truth
The Haunted Bridge	Uma Thurman

The Tonight Show	United States Congress
Thought	Urban ecology
Tim Quinn	Uterus
Tom Cruise	Value (ethics)
Towel	Van Toffler
Unitized insurance fund	Vending machine
Video	Vice president
Watkins Incorporated	Workweek and weekend
Weight loss	XM Satellite Radio
Wife	Zero 7
Wine	
Wine making	
Woman	

2014	
Men's	Women's
300: Rise of an Empire	A.W. Tozer
3M	Abdominal exercise
A Hard Day's Night (film)	Accident
Acrylic paint	Acne vulgaris
Adam Silver	Acrophobia
Alex Gibney	Adoption
Alfred Dunhill Ltd.	Adrienne Bailon
Algiers (band)	Air travel
Allergic rhinitis	Airline
Allied health professions	Alexa von Tobel
American Psychiatric Association	Allergen
Amusement	American Psycho (film)
Amy Sedaris	Amy Brenneman
Anchorman 2: The Legend Continues	Angina pectoris
Ancient Evenings	Anna Kournikova
Andrew Powell	Anne Lamott
Anorectic	Ariana Grande
Ansell	Armani
Antacid	ASOS.com
Antibiotics	Baby transport
Archie Manning	Baking
Argument	Bathroom
Australia	Bed Bath & Beyond
Automotive industry	Bed rest
Axe (brand)	Bonfire
Backyard	Bra
Bank	Bread
Barber	Breast cancer
Baseball	Brian Wansink

Bathroom	Buttocks
Bathtub	Calvin Klein
Beats Music	Camping
Benedict Fitzgerald	Cancer (astrology)
Bitcoin	Candy
Blood test	Caroline Ingalls
Bob Livingston	Carpet
Bobby pin	Carrie Underwood
Bottle	Cassey Ho
Boxer shorts	Cat
Broadcasting	Charming Charlie
Bruce Springsteen	Cheap Monday
Bundy standoff	Chewing gum
Button	Chicago Fire (TV series)
C.O. Bigelow	Chicago III
Cable television	Christie Brinkley
Calvin Klein	Christmas
Cameo appearance	Christmas lights
Cannabis (drug)	Clarksdale, Mississippi
Carbohydrate	Color
Cardiopulmonary resuscitation	Common cold
Chef (film)	Compact (cosmetics)
Chevrolet	Cost of living
Chikyngunya	Cost-effectiveness analysis
Children of Men	Council of Fashion Designers of America
Chili pepper	Cream
Chris Coons	Credit risk
Chris Paul	Dani Evans
Christina Aguilera	Dave Franco
Cleveland	Demeter Fragrance Library
Coffemaker	Dermatology

Cognitive disorder	Destination spa
Cola	Detoxification
College basketball	Diabetes mellitus type 2
College football	Dietitian
College town	Distracted driving
Comme des Garçons	Doily
Comorbidity	DVD
Compulsive hoarding	E.L. James
Conflict management	Economist
Constantine (TV series)	Ed Sheeran
Credit card fraud	Ellie Goulding
Cutlery	Emergency contraception
Cynicism (contemporary)	Emoji
Dan Patrick	Emoticon
Dane DeHaan	Energy medicine
David Fincher	Environmental protection
David Haskell	Equation
Daymond John	Exploration
Deraillleur gears	Eyes Wide Shut
Desert	Family therapy
Dhani Jones	Female
Diamond blade	Fibromyalgia
Diesel (brand)	Final good
Dignity	Flextime
Disordered eating	Folic acid
Dixon Ticonderoga Company	Food bank
Doctor Who	For Today I Am a Boy
Dog collar	Forest conservation in the United States
Domino's Pizza	Foster care
Don Draper	Freida Pinto
Donal Logue	Gail Caldwell

Donald Glover	Garnier
Drill	Geneticist
Drywall	Grape
Dye	Grimm (TV series)
Ear	Guinea pig
Ejaculation	Hair coloring
Electric power	Handbag
Electrical conduit	Hangover
Electrical network	Healing
Elephant	Home equity loan
Ellar Coltrane	Housekeeping
Elvis Costello	Hunger
Engine	Ice cream cone
Epidural administration	Idina Menzel
Esquire Network	IHOP
Ethanol fuel	Impression formation
Failure to Launch	Incandescent light bulb
FIFA World Cup	Infection
Flooring	Instinct
Food industry	Interior architecture
Food rheology	Internet relationship
Foot	Investment
Football player	Jennifer Nicole Lee
Foreclosure	Jessica Morgan
Fracture	Jewelry
Front yard	Jill Abramson
Genetic engineering techniques	Jon Bon Jovi
Girls (TV series)	Julep (company)
Glenn Greenwald	Junk food
Gluten-free diet	Kate Moss
Goose bumps	Kendall Jenner
Grammy Award	Keratin

Granola	Khloé Kardashian
Guardians of the Galaxy (film)	Labor supply
Half Acre Beer Company	Lauren Miller
Hand sanitizer	Law enforcement agency
Handgun	Leah Vincent
Handle	Lee Brice
Harman International Industries	Leggings
Health risk assessment	Lemon
Heat therapy	Liev Schreiber
Heterosexuality	Lime (fruit)
Higher education	Long-term care insurance
Hip hop music	Lottery
Hobbit	Maeve Binchy
How I Met Your Mother	Maize
Human back	Makeup brush
Ice cream parlor	Maria Menounos
Ilan Hall	Marriage settlement (England)
Illegal drug trade	Mastication
Imagination	Matzo
Ingestion	Melissa McCarthy
Injury	Memory disorder
Intellectual giftedness	Miranda Kerr
Intimidation	Mirror
Janet Napolitano	Miu Miu
Jay Glazer	Mobile phone
Jay Leno	Mold (cooking implement)
Jeff Goldblum	Moleskine
Jeremy Renner	Montgomery Ward
Jessie James Decker	Motivation
Jimi: All Is by My Side	Mozzarella
John Carroll Lynch	Muskmelon

John Oliver (comedian)	Nails Inc.
Johnny Manziel	Neil Patrick Harris
Jon Hamm	Nina Turner
Kiefer Sutherland	Oatmeal
Ladder	Office management
Lantern	Old Navy
Lawless (film)	Onion
Lawrence Berkeley National Laboratory	Orderliness
L-DOPA	Oregon State University
Leadership	Orgasm
Leaf	Outsourcing
Leather jacket	Paper clip
Lionel Messi	Parenting
Liquid Swords	Parts cleaning
Lubricant	Pencil
Lubrication	Periodontitis
Luxury vehicle	Pesto
Magnesium	Pickled cucumber
Maroon 5	Picnic
Masked and Anonymous	Pinterest
Masonry	Plate (dishware)
Medical journalism	Portrait
Medicinal plants	Public policy
Mental disorder	Pyramid scheme
Mercedes (car)	Quality of working life
Michael Ian Black	Real property
Miele	Relationship counseling
Miles Davis	Retinoid
Misdemeanor	Rimmel
Morning Phase	Safe sex
Music industry	Sanitation
National security	Sara Shepard

Netflix	Sasha DiGiulian
Neuroscientist	Saving
New Girl	Sealant
Newspaper	Sears Holdings
Nick Cave	Seize the Night (novel)
Nissan	Sewing
Nitroglycerin	Sex differences in humans
No Place to Hide (Greenwald book)	Sex toy
Norovirus	Sexual fantasy
Ocean Springs, Mississippi	Sharing
Opinion poll	Ship transport
Pail (container)	Shopping
Pallet	Shortcake
Patient education	Sleep disorder
Paula Lambert	Snack food
Peninsula	Social relation
Penn Jillette	Socialite
Peppermint	Solar energy
Peyronie's disease	Sommelier
Philosophy	Sophia Amoruso
Pizza	Stool (seat)
Pizza Hut	Store-value card
Plastic bag	Sun tanning
Platelet	Sur La Table
Plywood	Sweater
Premature ejaculation	Table (furniture)
Prostate	Taste buds
Psyllium	Tax
Public relations	Temperature
Rachel McAdams	Tennessee
Ranitidine	Text messaging
Resort	The Leftovers (TV series)

Respiratory disease	Thyroid-stimulating hormone
Revolution	Tim Quinn
Richard Linklater	Time clock
Road bicycle	Tom Binns Design
Rob Howard	Toxic shock syndrome
Robert Bosch GmbH	Toys "R" Us
Robert Gates	Troian Bellisario
Robin Wright	Tunic
Rock climbing	Tyler Jacob Moore
Russell Westbrook	Umbrella
Rza	Unintended pregnancy
Sadomasochism	Unsecured debt
Sandfly	Vagina
Saw	Vein
Scandinavia	Vera Bradley
School district	Veronica Roth
Scotch whisky	Video
Scott Raab	Video game
Shaving	Vietnam War
Shawn Levy	Vinaigrette
Sigourney Weaver	Vince Vaughn
Slacker	Vineyard Vines
Snowboard	Virgin America
Socket wrench	Vulva
Sowing	Widow
Speculative fiction	Wind & Willow Home
Sport	Wine bottle
Sport management	Wine tasting
Sports car	Woman
Sports injury	Women's rights
Sports medicine	Working Girl (TV series)
Spray painting	Wrinkle

Squat (exercise)	Zucchini
Stephen King	
Stimulant	
Stock exchange	
Strategic planning	
Sturgill Simpson	
Subaru Outback	
Suspicion (emotion)	
Synthetic oil	
Tai chi	
Taran Killam	
Tarpaulin	
Tavi Gevinson	
Tax exemption	
Taxi (TV series)	
Technology	
Telecommunication	
Testosterone	
The David Letterman Show	
The Huffington Post	
The Knick	
The Last Song (novel)	
The Mind of a Chef	
The Practice	
Thermostat	
Tijuana Cartel	
Tilda Swinton	
Tinnitus	
Tobey Maguire	
Todd Snyder	
Toilet	
Tom Hardy	

Tool	
Toro (company)	
Travis Scott (musician)	
True Detective (TV series)	
Truss	
United States Naval Academy	
United States Secret Service	
University of Nebraska-Lincoln	
Valeria Lukyanova	
Waiting staff	
Warfarin	
Water industry	
Water supply network	
Waterfowl hunting	
Wes Anderson	
Wheelbarrow	
White House Down	
Will Smith	
Wood flooring	
Woodchips	
Wooden box	
Woodworking	
Workplace bullying	
Wound	
Wu-Tang Forever	



Decadal	
Men's	Women's
3M	401(k)
Actor	Abdomen
Adhesive	Abortion
Aerospace Manufacturer	Acne vulgaris
Aesthetics	Actor
Air filter	Acupuncture
Air travel	Adolescence
Airline	Adoption
Airline Seat	Aerobic exercise
Airplane	Ageing
Airport	Air Conditioning
Alcohol	Alcohol
Alcoholism	Alcoholism
Alec Baldwin	Allergy
Alexander (2004 film)	American Express
Allied health professions	American Heart Association
Alligator	Anger
Alternative Medicine	Animal Shelter
Alzheimer's disease	Anna Quindlen
Americans	Anniversary
Amusement	Anxiety
Anal sex	Arizona
André 3000	Arthritis
Animal Product	Attention deficit hyperactivity disorder
Anorectic	Autonomy
Antibiotics	Baking
Antidepressant	Bar
Antique car	Balsamic vinegar
Appetite	Basket

Argentina	Bath & Body Works
Armed forces	Beard
Arrested Development (TV series)	Beiersdorf
Art	Beverage
Ashton Kutcher	Bikini
Assassination	Birth control
Association football	Blanket
Asthma	Blouse
Aston Martin	Book
Astronaut	Bookcase
Athlete	Boot
Athletics (physical culture)	Bowls
Atrial fibrillation	Brest self-examination
Attic	Brunch
Audi	Buyer (fashion)
Australia	Car
Author	Cardiovascular disease
Auto racing	Caregiver
Autobiography	Celebrity
Automatic control	Chair
Automotive Industry	Charge card
Bachelor party	Cheeseburger
Back pain	Cheesecake
Bacteria	Chicken
Bally shoe factories	Child
Baltimore	Chocolate brownie
Bandsaw	Coffee preparation
Barber	Colgate-Palmolive
Barbering	Collagen
Barcelona	Comedy film
Baseball	Competition
Baseball cap	Container garden

Basketball	Convenience food
Bathroom	Cooking
Beach	Coronary artery disease
Bear Attack	Cosmetics
Beauty salon	Crabtree & Evelyn
Bedroom	Crying
Beef	Cucumber
Belief	Daughter
Benign Prostatic Hyperplasia	Death
Beta Blocker	Debt
Bicycle	Denise Austin
Bikini	Diane Sawyer
Bill (law)	Discounts and allowances
Bill O'Reilly (political commentator)	Divorce
Bioactive compound	Dog
Biologist	Domestic relations
Black Francis	Door
Blender	Dough
Blessing	Dreams
Blood test	Dress code
Blouse	Driving
Boat	Economic security
Boating	Education
Bob Dylan	Estée Lauder Companies
Bodybuilding	Eucalyptus oil
Bomb	Eveready Battery Company
Book	Everyday life
Boredom	Exercise physiology
Bottega Veneta	Exhibition
Bottle	Eye
Bottled Water	Family life education
Botulinum Toxin	Fashion

Bow tie	Fashion accessory
Boxing	Financial Plan
Bracelet	Floristry
Bread	Flower
Brick	Food allergy
Briefcase	Footwear
Brighton Rock (novel)	Fraud
Broadcasting	Gift
Bromance	Glasses
Brother	Gynecology
Budget	Halle Berry
Bugatti Automobiles	Health Promotion
Building	Herpesviridae
Bullying	Holiday
Burberry	Home Economics
Bus	Horoscope
Business plan	House Keeping
Businessperson	Hugh Jackman
Caffeine	Human eye
Calcium	Hygiene
California	Hypnotic
Calvin Klein	Ian Kerner
Camera	Interior design
Canada	Interpersonal relationship
Car	Jacqueline Kennedy Onassis
Car rental	Jamie Lee Curtis
Car seat	Job satisfaction
Car Wash	Jogging
Carbohydrate	Kate White
Card game	Lancôme
Carpentry	Law & Order
Carrot	Life partner

Casino	Lighting
Cast iron	Linens
Cat	Lip
Cate Blanchett	Lipstick
Cedrus	L'Oreal
Ceiling	Los Angeles
Celebrity	Magnetic resonance imaging
Centers for Disease Control and Prevention	Major depressive disorder
Central Intelligence Agency	Man
Central nervous system disease	Married people
Chair	Maybelline
Change (philosophy)	Middle age
Cheese	Mixer (cooking)
Chicken	Mood (psychology)
Child development	Muscle
Childbirth	Myocardial infarction
Chile	Natural foods
Christmas	Necklace
Christopher Nolan	Negotiation
Chronic pain	Nickname
Chrysler (brand)	Nicole Kidman
Cigarette	Nike, Inc.
Circadian rhythm	Non-fiction
City	OPI Products
Classic car	Oprah Winfrey
Clay Aiken	Oral Hygiene
Cleveland	Pain
Clinical trial	Pain management
Clive Owen	Papillomaviridae
Clock	Parent
Closer (2004 film)	Personal lubricant

Clothing	Pet
Clothing industry	Phone sex
Coach (sport)	Physical attractiveness
Coffeehouse	Physical Exercise
Cole Haan	Physical fitness
Collar (clothing)	Quality of Life
Collecting	Rachel Bilson
College Athletics	Reading (process)
Colorado	Religion
Comedian	Ribbon
Comedy	Rob Lowe
Comic books	Romance (love)
Comic strips	Romantic comedy film
Comics	Romantic orientation
Commercial bank	Rose
Communication	Salma Hayek
Commuting	Sandwich
Compact disc	Scarf
Competition	Self-confidence
Concrete	Sensation (psychology)
Condom	Sense
Conflict management	Sexual orientation
Corn Syrup	Shampoo
Cuff	Sibling
Culture	Soap
Curse	Solanaceae
Cutlery	Strawberry
David Bowie	Stress (psychological)
Demonstration (teaching)	Sweater
Deodorant	Sweepstakes
Developmental psychology	The Tonight Show
Dick Cheney	Thought

Diesel (brand)	Tote bag
Disc jockey	Towel
Documentary Film	Trisha Yearwood
Doubt	Wall
Dress	Washington, D.C.
Drill bit	Watch
Drug	Weight loss
Eating	Williams-Sonoma
Eau de Cologne	Winter
Ejaculation	Woman
Elevator	Women in music
Engine	Workplace
Enzyme	Workplace romance
Eye	
False (logic)	
Fear	
Film	
Fly	
Food	
Friday Night Lights (film)	
Frustration	
Gene	
General Motors	
George Clooney	
Gratuity	
Green tea	
Hair loss	
Headphones	
Health risk assessment	
High-density lipoprotein	
Hip hop	
Hollywood	

Human back	
Human sexuality	
Hunger	
Imagination	
Imprisonment	
Industrial mixer	
Injury	
James Franco	
Japan	
Jealousy	
Jim Caviezel	
John Waters	
Joke	
Jon Favreau	
Khaki	
Kobe Bryant	
Laird Hamilton	
Lawyer	
Linebacker	
Louisiana	
Low-density lipoprotein	
Macro ethics	
Manager (baseball)	
Mark Wahlberg	
Marlon Brando	
Masking tape	
Mayor	
Meeting	
Melanoma	
Michael Chabon	
Microwave	
Miles Davis	

Military personnel	
Mobile phone	
Mortar (weapon)	
Motor vehicle	
Motorcycle	
MTV	
Music industry	
Musical ensemble	
Nail (fastener)	
Nail (fastener)	
NASA	
Near-death experience	
Nonsteroidal anti-inflammatory drug	
OXO (brand)	
Oxygen	
Pail (container)	
Pamela Anderson	
Parenting	
Parenting	
Patient	
Pay television	
Pedro Almodóvar	
Performing Arts	
Peyton Manning	
PHV (company)	
Platelet	
Political Corruption	
Popular Culture	
Popularity	
Portion control (dieting)	
Poverty	
Power tool	

Preservative	
Procedural Law	
Produce	
Public Morality	
Public space	
Quadriceps femoris muscle	
Ralph Lauren	
Razor	
Reading (process)	
Recording studio	
Retaining wall	
Road	
Rock music	
Safety	
Same-sex relationship	
San Diego	
Sanitation	
Scientist	
Screenwriter	
Screw	
Sexual attraction	
Shark Attack	
Shaving	
Sherwin-Williams	
Shure	
Sigourney Weaver	
Sister	
Sleep	
Small business	
Sneakers	
Spain	
Sperm bank	

Sport car	
Sports club	
Sprint (running)	
Stand-up comedy	
Steve Coogan	
Stew	
Stove	
Stress (biology)	
Suit (clothing)	
Superior-subordinate communication	
Surgery	
Swimsuit	
Tailor	
The Apprentice (TV series)	
The Bronx	
The North Face	
Therapy	
Thom Browne	
Thorax	
Tobacco smoking	
Tom Wolfe	
Town	
Truck	
Trust (emotion)	
Truth	
UHF connector	
United States Congress	
University of Pennsylvania	
Vintage clothing	
Vitamin	
Vitamin E	
Volkswagen Group	

Watch	
Weed control	
Weight training	
Wes Anderson	
Will Ferrell	
Winery	
Wisdom	
Wool	
Workplace bullying	
Wound	
Yohji Yamamoto	