Our Browser Extension Lets Readers Change the Headlines on News Articles, and You Won't Believe What They Did!

FARNAZ JAHANBAKHSH, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, USA

AMY X. ZHANG, Allen School of Computer Science & Engineering, University of Washington, USA

KARRIE KARAHALIOS, University of Illinois at Urbana-Champaign, USA

DAVID R. KARGER, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, USA

Headlines play a critical role in how users perceive articles. But many headline publishers craft headlines in ways that either attract clicks in an attempt to earn ad revenue, or misinform users or manipulate their opinions for malicious intents. Such headlines can do harm since many users simply skim and share headlines without reading the articles in full. We present an exploratory browser extension that empowers users to suggest headlines they deem better for news articles. Users can view headlines suggested by other users that they follow as they browse websites. We conducted a study of 27 users who used the extension for one week to read news and suggest headlines. We found that users saw value in the tool and used it to change headlines that they found in need of improvement. We characterize the changes that people make to headlines if enabled. We also report on a followup study we conducted with 312 participants to evaluate headlines suggested by the tool. The purpose of the study was to examine whether headlines suggested by untrained users could be preferred over original headlines by professional editors. We found that a substantial number of the suggested headlines were indeed preferred. Our work explores the designs for, and opportunities and consequences of, empowering news consumers by giving them control over the content curation process.

CCS Concepts: • **Human-centered computing** → **Empirical studies in collaborative and social computing**; **Empirical studies in HCI**.

Additional Key Words and Phrases: Clickbait, News Headlines, Misinformation, Crowdsourcing

1 INTRODUCTION

Headlines play a critical role in how news consumers perceive articles [40]. But the interests of headline publishers are not always aligned with those of consumers. While consumers may want to be informed or entertained, publishers may wish to attract clicks to stories to earn ad revenue [50, 71], and malicious actors may wish to *dis*inform users or manipulate their opinions. The manipulation techniques, employed by disinformation news websites and also to some extent by legitimate sources, include using language that is exaggerated, sensationalized, teasing, misleading,
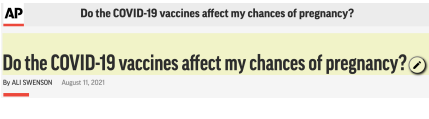
Fig. 1. Upon visiting an article page, the Re-headline extension recognizes the headline, highlights it, and adds an edit button next to it. Users can suggest a new headline by clicking on the edit button.
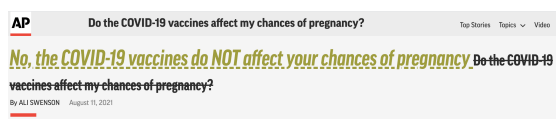

Fig. 2. If a headline for which an alternative headline has been suggested is encountered on any page, the original headline will be crossed out and the alternative headline will be placed next to it.

Fig. 3. Once a headline has been suggested for an article, the alternative headline will appear alongside the original headline on any page where the original headline (or a text similar enough to it) is encountered.
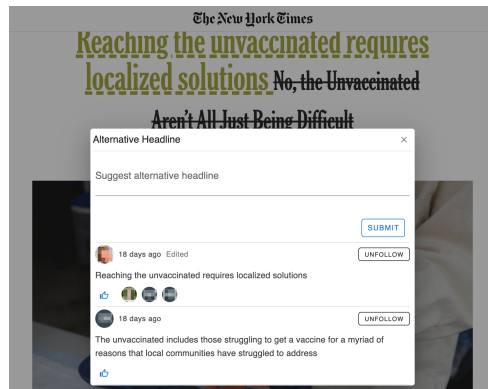

Fig. 4. By clicking on the displayed alternative headline, a user can see all the headlines that have been suggested for the article by the people the user follows. The user can additionally choose to suggest headlines of their own, endorse headlines, or unfollow users whose headlines they do not like.

or even inaccurate [13]. As described by a Macedonian teen who profited from exaggerated Facebook titles leading up to the 2016 election, *"The most-read news articles are usually the ones containing the click-bait [title] words. The click bait words, as you know, are, 'Oh my god, breaking news, wow,' and usually something that has never been aired before. Because if the title just says, 'Today this happened, today that happened,' no one will open that."* [71].

As it is headlines that news consumers first encounter and that can entice them to follow a link or that can leave an exaggerated or manipulated impression on them, the manipulation techniques are increasingly applied to news headlines [27]. Manipulative headlines, even if they redirect to articles that ultimately present accurate information, are harmful because they not only deceive people into clicking, but also misinform those who do not click. As a large portion of news consumption these days happens on social media streams where headlines are detached from an article's content and context [3, 13], misinforming headlines have the potential to be believed and further spread [13]. Even if users do navigate to the article, they are not likely to read it in full [47], or may only retain the message presented in the title. For instance, Kong et al. show that even if a visualization in an article does not align with the headline, it is the headline that most users will

remember [38, 39]. Misinforming or 'clickbaity' headlines therefore, can pose a similar threat as other types of misinformation and need to be addressed.

A body of work has leveraged machine learning (ML) approaches to find, label, or down-rank clickbaity headlines [11, 60, 66]. Prior work suggests that the effect of credibility labels on recognizing misinformation is limited [24, 57]. In addition, centralized labeling or down-ranking of problematic headlines is a type of truth governance by the platforms which can be at odds with freedom of speech and autonomy of individuals deciding what content to consume [78]. Users are also not in agreement on whether platform assigned labels are helpful, as some find the labels judgmental, paternalistic, and against the platform ethos [67].

In this work, we adopt a crowdsourcing approach; and instead of labeling problematic headlines for users, we enable them to modify headlines so that they inform rather than misinform. The motivation behind this design is to give users agency to change their environment as well as empower them to help others be informed, not misled, or otherwise not baited into clicking.

In pursuing this approach, we recognize that different users have different notions of what counts as a problematic headline and what changes to a headline would improve it. For instance, a user may consider a headline inadequate if it withholds certain information that the user is interested in knowing about at a glance or that they believe others should be aware of, while another user who is not as interested in the topic or on the contrary, knows about the particulars of the story, may be satisfied with the conciseness of the headline. Therefore, we do not seek after a single headline that appeals to all users, and instead give users the autonomy to alter headlines as they see fit for themselves and those who follow them. And it is worth noting, that while a tool may be designed to empower users to *inform*, other use-cases worth exploring will arise.

We present Reheadline[1], a browser extension that empowers users to collaboratively edit and share news headlines. The aim is for the collaborative editing to yield headlines that are more informative or better align with the article's content. Reheadline detects the headline on a news article page that a user is visiting, and allows the user to suggest alternative headlines for it. Other users running the extension who follow the headline suggester and who come across the article, or encounter the original headline on other pages such as their social media feed or the news website's homepage, will be shown the alternative headline next to the original. If there are multiple alternative headlines suggested for an article, one will be chosen for placement on the page based on a set of factors, including how many endorsements it has received from other users. Users can also choose to additionally view other alternative headlines that are not placed in-situ.

To evaluate the potential of such a design, we conducted a user study of the extension. The user study informed us about whether people perceive value in a tool that allows them to collaboratively curate alternate headlines for news articles, what kinds of headlines they find problematic, and what they believe are better replacements for these headlines. Our study involved 27 participants who used the extension over one week to read news articles of their choice and suggest alternative headlines for the articles whose headlines they found inadequate or in need of improvement. We left it to each user to decide which headlines need to be improved and how. The participants' daily tasks involved suggesting at least a couple of headlines and visiting a platform we built which presented a feed of articles for which participants had suggested new headlines. We followed up with a second user study that evaluated the headlines that our tool users had suggested. We found that many of the alternative headlines were preferred over their original counterparts and that different users had different headline preferences for various reasons, such as their interest in the article.

---

[1]https://chrome.google.com/webstore/detail/reheadline/iignpdlabbnnacdkchpnpljkhdlkblbh

We do not argue that the approach of empowering news consumers to rewrite headlines is categorically better—in terms of how much information scent, deception, sensationalism, etc. user headlines will contain—compared to the status quo of confining editorial power to news publishers. Rather, we begin to explore the designs for, and opportunities and consequences of, empowering news consumers to exert editorial control over headlines. We demonstrate that users are both capable of and interested in improving the headlines that they encounter. We also explore the types of changes, both positive and negative, that users choose to make. For instance, we found that some users removed the bias, deception, or sensationalism present in some headlines, while others added sensationalism or their opinions to headlines. We leave it to future work to investigate under what circumstances users make these types of changes. Our findings suggest that there is potential for including users in the content curation process and not treating them as passive consumers. We also offer insights on the potential and the challenges of user empowerment through the design of tools such as the one in our study.

2 RELATED WORK

We discuss prior work related to the impact of audience metrics on the production of online news and especially headlines, the characterization of problematic headlines, the approaches that attempt to classify such headlines or more broadly misleading or inaccurate content on the web, and how our approach relates to tools that enable users to modify webpages on which they do not have authoring permission.

2.1 The Impact of Audience Metrics on Online News Production

While journalists may have once considered audience metrics such as user clicks or engagement—imposed by the marketing departments of their publishers—at odds with their professional values, they can no longer dismiss them because metrics are increasingly used to allocate scarce resources and bonuses to staff writers [15, 52, 59]. Therefore, as journalists and editors have adopted these metrics to guide them in creating content, they also developed strategies to reconcile their journalistic values with maximizing audience metrics. One such class of strategies attempts to undermine the connection between audience metrics and profits. For instance, they include drawing a boundary between "bad" metrics such as the number of clicks and unique visitors and "good" metrics such as the amount of time the audience spent reading an article or the number of retweets [15]. Some of these metrics considered good however, can be misleading and not tied to true engagement [23]. Another example of such strategies is the careful cherry-picking of examples demonstrating that news stories can both be important and attract traffic. Yet another such strategy reformulates metrics as valuable, however imperfect, representations of the public's informational needs and desires. Nevertheless, other strategies acknowledge the role of profits in using audience metrics but justify the usage of metrics by asserting that it is the advertising revenue brought about by the advertising-friendly content that makes the publishing of important pieces possible [15, 72].

Critics of audience metrics however, warn of the negative consequences of metric-centric news production and delivery. One concern is that metrics may result in indulging users in what they find interesting rather than delivering the information that is in their best interest [6, 81]. Another is that metrics can contribute to segmenting audiences into narrow groups, for instance by market-related factors such as purchasing power, eliminating the common ground where civic discourse can take place [81]. Additionally, in some decisions such as which content gets placed on the homepage, audience metrics take precedence over editorial judgements [45].

Widely adopted audience metrics in newsrooms can play a substantial role in shaping news headlines [28, 42]. Dor argued that a headline carries a certain contextual effect—i.e., a change to the reader's prior assumptions by either adding new ones, or weakening or strengthens existing ones;

and that news headlines are written to optimize the ratio of contextual effect over processing effort (or in other terms, the relevance of the story) for the readers. To achieve a high contextual effect, headline writers craft headlines to contain new—but not overly new—information, include names and concepts with high "news value", have an interesting framing, or guide the reader to the desired context of interpretation as is the case of most tabloid headlines, among other approaches [19]. This "relevance optimization" results in headlines that sometimes summarize a story, other times highlight a detail extracted out of the story, or even contain material that does not appear in the news story itself [7, 19].

2.2 Clickbait and Other Types of Problematic Headlines

Studies on uninformative or sensationalist news headlines predate the age of online news media [68]. For instance, Lindemann argued that contrary to the popular belief that news headlines need to be telegram-like, informative, and non-redundant, tabloid headlines can disinform, puzzle, and be highly repetitive [46]. With the shift to online news consumption [53], uninformative or disinforming news headlines have the potential to do greater damage since the sheer volume of information that individuals encounter, by request or chance, has made it intractable to read and investigate every piece. In fact, research has reported that most people simply only read headlines and do not delve deeper into news stories [31]. Another study showed that the majority of the URLs mentioned on Twitter are not clicked at all [23]. Research reports that a large portion of news consumption happens on social media these days [3] and users share content without reading it. Therefore, misleading headlines, as a form of false information, have the potential to spread and reach a wide audience [87]. Indeed, much of the body of work on detecting misinformation focuses on articles, which would be inadequate when an article presents accurate information but it is the headline that misleads.

Even among those who do click on an article, most do not read it in full [47] or may only remember the message in the headline even if it disagrees with the content. Kong et al. studied the recall and comprehension of misleading visualization headlines and reported that people remembered a headline even if it was misaligned with visualizations in the article [38, 39]. Ecker et al. reported on a lab study using headlines that misrepresent the gist of an article by cherry-picking information pieces from the article. They found that misleading headlines affected people's inferential reasoning for opinion pieces. They argued that this effect arises because people struggle to update their memory and reasoning in order to correct initial misconceptions [21].

Cherry-picking misleading content from an article is one of the many manipulative ways to produce headlines that attract user clicks. In fact, a body of work attempts to identify the characteristics of clickbait, linguistically or content-wise. For example, one manipulation technique, reported in this body of work, employs the use of sensationalist and exaggerating language or teasing to draw the reader's attention to gaps in their knowledge [9, 27]. To study sensationalism in headlines, Molek-Kozakowska proposed a framework by analyzing a set of headlines from the Daily Mail. She reported that the illocutions performed by sensational headlines included exposing ostensibly hidden information, speculating future consequences of an event, generalizing, delivering a warning, and extolling [54]. In a large-scale content analysis of a set of headlines produced by four major global media corporations, Reis et al. reported that extremely negative and positive headlines attracted more clicks, while neutral headlines were less popular [20]. Chakraborty et al. found that clickbait tweets included more images, hashtags, and user mentions, which can draw the attention of consumers [12].

A body of research has focused on automatic detection techniques to combat clickbait. One such work by Chen et al. proposed potential methods for recognizing textual and non-textual clickbaiting cues, for instance, using SVMs or frequency analysis to identify common lexical patterns in clickbaits

such as overuse of numerals and unresolved pronouns [13]. Other studies have presented machine learning models based on linguistic and sentiment features, or other entities accompanying the headline such as images, or by simply using word embeddings [8, 36, 62, 64, 65]. Chesney et al. argued that ML techniques to label teasing or sensationalist headlines, which have been the focus of this body of work, are not appropriate for classifying incongruent headlines that do not accurately represent the information in the article and propose alternative approaches [14].

2.3   Tools that Help People Determine the Quality of Content on the Web

It is important to help users determine content credibility at reading time, rather than relying on post hoc corrections to previously seen content. While post hoc corrections are still valuable, they fall behind the misinforming or misleading content they aim to fix with respect to both the delay after which they are propagated and how much visibility they receive [76, 90]. Therefore, a body of work has studied and presented tools to help people determine the accuracy or quality of content that they encounter on the web, either by informing them what the high quality contents or sources are, or by nudging them to come to the realization on their own. Some of these tools are bound to certain social spaces and others can work across the web.

One such tool for instance is NewsGuard, a browser extension that shows detailed transparency and credibility scores next to links to different websites as users browse the web. These scores are based on assessments from journalists [44]. Recently, Twitter has introduced Birdwatch, where users can identify tweets they believe are misleading and write notes that provide context to the tweet as well as rate the helpfulness of other users' notes [16]. Epstein et al. argued for developing a toolkit that prompts users to assess news accuracy before sharing [22], building on research that reported shifting users' attention to accuracy at sharing time increases the quality of news that they share [32, 58]. Jahanbakhsh et al. design a set of user affordances aimed at empowering users to protect themselves and their social circle from misinformation. The affordances include accuracy assessment of posts by users as part of the data model, user-specified indications of which other users they trust to assess posts, and filters that users can configure to block from their feed posts assessed as inaccurate by other users they trust [33].

Such tools have also been proposed in the domain of news headlines. For instance, Chakraborty et al. built a classifier to label teasing headlines and integrated it into a browser extension to warn users of such headlines in the links on the page that they are viewing [11]. Rony et al. developed a clickbait classifier to detect teasing or sensationalist headlines and incorporated that into BaitBuster, a browser extension that flags clickbaits on a Facebook timeline or a Facebook page along with explanations of why an item has been flagged [66].

While tools that inform consumers about trustworthiness of news sources are valuable, not all headlines by a low credibility source are problematic and not all those by mainstream media are informative and nonsensationalist. In addition, the approaches that rely on ML mostly focus on the problem of teaser and sensationalist headlines that can be detected by investigating the headline's lexical features. These approaches therefore, cannot cover all accounts of problems in headlines such as incongruence between the message of a headline and that of the article it references or whether and how much the headline exaggerates. Automatically classifying such problems in a headline has been argued as extremely difficult because it requires analyzing the complex relationships between a headline and an entire article [14]. Instead of relying on third party algorithms or content raters that do not accept user input, in this work, we empower users to not only identify problematic headlines, logically or semantically, but also modify them for the better.

2.4 Enabling Users to Modify Web Pages

Our work in enabling users to modify news headlines on proprietary news websites is influenced by prior work on helping people add content to web pages on which they do not have authoring permission. Somewhat close to our scenario is Newsr, a mobile app that presents news stories that users can annotate through graffiti-like interactions. Through a study of 15 participants, Wood et al. examined the engagements that the tool led to and found that participants used the graffiti feature to question the importance or accuracy of content, among others [85].

A body of work in this area has focused on enabling users to annotate or add additional content to a page specified by its URL address. For instance, Hypothesis is a browser extension that allows users to select a piece of text to annotate and allows future visitors on the page to view and interact with the annotations. Users of Hypothesis have the ability to make their annotations visible to the public or a certain group of users [1, 35]. Similar is NB, a social annotation tool that focuses on classroom settings [91]. Other tools enable users to add comments that are not necessarily bound to a particular piece of content on a page. One such tool was Google Sidewiki, which presented a sidebar on any webpage where Google users could write comments and react to those written by others [5]. Another is Eyebrowse, a browser extension that allows users to follow others and see recent visitors and comments they have left behind on the webpage the user is viewing [89].

Annotation tools such as these allow for modifying a resource identified by its URL or present a modifiable instance of a resource. Therefore, the annotations placed on a particular instance of a webpage or a news story will only be displayed if users visit that specific page. Although these tools can potentially be repurposed for capturing alternative headlines or accounts of a news article, they fall short of protecting users from clickbait links, or from misinforming headlines that they do not click, because they display the annotations only *after* a user clicks through to an article. In addition, they cannot provide annotations for similar text encountered outside the specific resource.

Our extension on the other hand, shows headlines suggested by users even if an article's headline is encountered outside the article's page, e.g., on social media feeds. It is able to do so because it targets the specific problem of news headlines, rather than serving as a general-purpose annotation tool. Because a large portion of news consumption is through skimming headlines, it is important that other occurrences of an article's headline outside the article are detected and alternative headlines are suggested for them.

3 THE REHEADLINE SYSTEM

The re-headlining ecosystem that we built consists of an extension that allows users to interact with a webpage to suggest headlines and modifies pages to insert headlines suggested by users when the original headlines are encountered, as well as a platform that presents a feed of articles for which users have suggested headlines. We describe the workings of each component below.

3.1 The Reheadline Extension

When a user is on an article's page, the extension detects and highlights the headline, and places an edit button next to it, as shown in Figure 1. Users can click on the edit button to suggest an alternative headline for the article. Then, if the original headline is encountered on any page including the homepage of the news website, or on social media feeds, the alternative headline will appear first, and the crossed out original headline will be placed next to it with a smaller font, as shown in Figure 2. Figure 3 demonstrates a Twitter feed where the extension has found the original headline and has placed the alternative headline before it.

An article can have multiple alternative headlines suggested for it, but only one alternative headline will be placed in-situ on the page. The in-situ headline is determined based on a set

of factors, including whether the user viewing the page is the author of any of the alternative headlines, whether the user has endorsed any of the headlines, how many users (followed by the authenticated user) have endorsed each headline, and the recency of each alternative headline. Users can additionally click on the alternative headline to view and possibly endorse the rest of the alternative headlines, as shown in Figure 4.

Since as we will elaborate later, the notion of what makes a good headline is highly subjective, we allow users to *follow* other users. A user will only see alternate headlines authored by people they follow.

The extension uses text matching to find any headlines on a page for which alternative headlines have been suggested. Another approach would be to look for those links for which users have suggested headlines. We decided not to use links because it is common for articles to embed links to other articles for reference as part of a longer explanation, and substituting the few word references they use in these occasions with the longer alternative headlines would disrupt the text flow. See Figure 5 for an example of such embeddings. Another disadvantage of relying on links to fetch relevant alternative headlines is that sometimes the links to an article that are encountered outside the article's page, e.g., on social media or RSS feed readers, are not direct links to the article, but rather go through a series of redirects to reach the article. Retrieving alternative headlines based on links will result in missing rewrites for these articles. Furthermore, many news websites republish articles by other websites such as The Associated Press. Binding a headline to a link can result in failing to display the article's alternative headlines on all the other websites that have published the same article as well as on news aggregators and RSS readers.

A potential complication with finding alternative headlines based on text matching is that the headlines that appear on an article's page, the homepage of the news website featuring the same article, and on social media feeds are sometimes different from each other. This difference can be minimal, e.g., the addition or omission of a few words, or can be more drastic. To accommodate at least minimal changes, the extension supports finding headlines that are *similar enough* to the original, using fuzzy string matching. However, the headlines that are more drastically different may not be detected. Nevertheless, it is not clear whether an alternative headline should still replace a headline that is much different from one for which the alternative headline was suggested.

To find alternative headlines suggested for the headlines found on the page, the extension sends the hashes of the pieces of the text that can potentially be headlines (whether as part of the page's title, links, or inside the text body) to the server. The hashes are not computed on entire text nodes, but rather substrings. Hashes offer performance benefit—messages can be smaller than if whole or even partial headlines are sent. They also increase user privacy since the text of the user's page is never sent (note, however, that if a hash sent by the client is found by the server, the server will know that the corresponding substrings of the hashes can be found on the page that the user is viewing. Therefore, privacy is not perfect.). After having received the hashes from the client, the server checks to see if the hashes match the hashes of any (original) headlines for which alternative headlines have been suggested and returns a list of potential candidates along with their alternative headlines. When the extension performs fuzzy matching of the returned original headlines to see whether any of the candidates can be found on the page, it makes the fuzzy threshold more lax for those candidates whose domain matches the domain of the current page.

The extension additionally has an options page where the user can specify their blacklisted domains, i.e., those on which they do not want the extension to find and display alternative headlines. The source code for Reheadline is open source [2].

---

[2]The source code can be found at https://github.com/farnazj/Reheadline-Extension

3.2 The Reheadline Feed

Without a large user base, it is unlikely that users would serendipitously encounter headlines that other users have suggested. Therefore, for the purpose of investigating how users interact with headlines that other users have submitted, we developed a platform that they could visit to view a feed of the articles for which users had suggested alternative headlines. The feed is akin to social media feeds where interactions of other users that a user follows appear on the feed of the user. However, contrary to social media feeds where these interactions can increase the visibility of a misleading or false post, each post on the Reheadline feed is accompanied by a salient correction for its uninformative or disinforming headline.

On the homepage of the Reheadline platform, a user can view a feed of articles for which an alternative headline is suggested by those that the user follows. Each article snippet on the feed displays the selected alternative headline, the original title, the article's image, its description, along with a button that would redirect to the article's page, as shown in Figure 6. Similar to the extension, users can further view what other headlines have been suggested for an article, and choose to endorse them. Clicking on a user's avatar redirects to their profile, where one can see a list of all the headlines the user has suggested in the past. On the users page, a user can see the list of all the other users of the extension, and choose to follow or unfollow them. A user can additionally follow or unfollow another user by visiting their profile. The Reheadline platform as well as the extension show a user only those headlines which are submitted by the people the user follows.

4 USER STUDY OF THE TOOL

We conducted a user study of the tool to evaluate the potentials and the challenges of our design. We aimed to understand whether people perceive value in a tool that empowers them to collaboratively curate headlines they perceive to be better for news articles and what changes they would like to make to headlines.

We recruited participants to work with the extension to suggest and view alternative headlines for news articles over a period of one week. We advertised the study on a behavioral research platform managed by our institution with a pool of diverse participants as well as on Reddit (/r/samplesize). We recruited most participants from the behavioral research platform. The eligibility criteria for the study were being 18 years of age or older, being a US citizen or a permanent resident, reading news online at least occasionally, and using Google Chrome as browser. We asked those who were interested to fill out a survey which in addition to detailing the study and providing the consent form, asked about the news sources participants consumed as well as their demographics. The purpose of these questions was so that we could recruit those who frequently consumed news, who we envisioned would be the main contributors of alternative headlines if the tool were used in the wild. Such users could benefit others who may come across articles for which alternative headlines are suggested.

To set up the study, one member of the research team met with participants online to demonstrate the use of the extension and the Reheadline platform. Participants' daily tasks during the one week period of the study consisted of contributing at least *a couple* of headlines and visiting the Reheadline feed at least once. Participants were encouraged to endorse the alternative headlines which they liked and to suggest other headlines if they found the existing submissions inadequate. To bootstrap the study more efficiently, we configured every user to initially follow all other users upon signup, although they could later choose to unfollow one another.

"This is what it looks like when elected leaders take a step toward healing our country's divisions rather than feeding those very divisions," Senator Kyrsten Sinema, Democrat of Arizona and a key negotiator, said before the bill's passage.

Fig. 5. Articles often cite and refer to other articles inside their text using a few short descriptive words. The extension therefore, does not look for links to insert alternative headlines for an article, since the inserted headline could disrupt the flow of the text. The extension instead uses text matching. The excerpt above is from The New York Times.
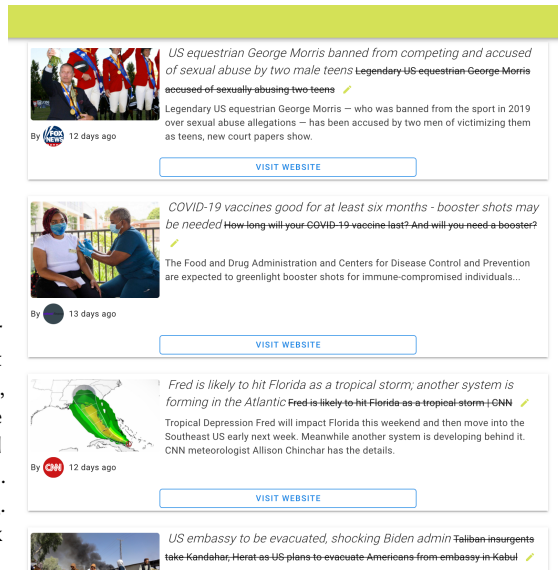
Fig. 6. The Reheadline platform where users could view a feed of the articles for which users they followed had suggested alternative headlines. Users could choose to visit the articles. Additionally, they could change the appearance of posts through a Settings UI to a compact form in which the image and the description were not displayed.

Not all participants started the study at the same time, with the latest start day being 3 days after the earliest.[3] At the study's conclusion, participants were asked to fill out a survey asking about the headlines that they changed during the study, why they changed the headlines the way they did, what kinds of alternative headlines they liked, and their perceptions about the tool and their ideas for improving it. Finally, participants answered demographics questions on their political and theistic ideologies, among others. We adopted the demographics section of the questionnaire from prior work studying misinformation [32]. The full questionnaire is included in the Supplementary Materials. The study was approved by our Institutional Review Board.

4.1 Logs

The extension and the Reheadline platform logged certain user interactions with the page. For instance, the extension measured and logged the time a user spent on a page before suggesting a headline. In addition, it logged whether the user clicked on a link to view the article for which an alternative headline existed.

4.2 Participants

A total of 31 users participated in and 27 completed the study, i.e., they kept performing the daily tasks until the study's conclusion. Of the 27, 13 were female and 14 were male. 19 identified as Democratic, 4 as Independent, and 4 as Republican. 11 were White, 7 Black, 5 Asian, and the rest either belonged to other ethnicities or identified as mixed. The median value for age was 30 (ranging

---

[3]The earliest set of participants started on August 5th, 2021, and the latest set of participants ended the study on August 14th, 2021.
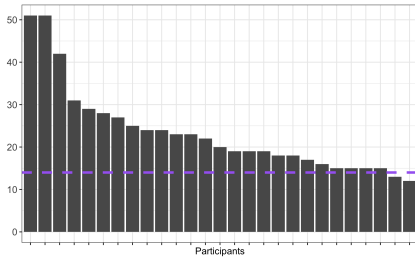
Fig. 7. Number of alternative headlines submitted by study participants over 12 days including the one week period of the study. Each bar represents one participant. The horizontal dashed line marks the expected number of headlines for each participant after their one week study period ended. Many participants submitted far more headlines than were required for the study.
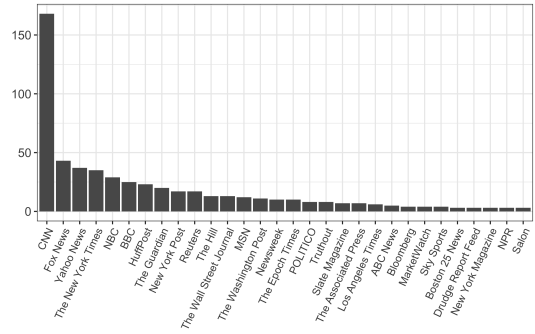

Fig. 8. the Distribution of alternative headlines submitted by participants for different news sources. The plot does not include those sources for which fewer than 3 headlines were suggested.

from 20 to 59); for highest degree achieved, Bachelor's Degree (ranging from some college but no degree to Doctoral degree or M.D./J.D.); and for income, $70,000 to $79,999 (ranging from less than $10,000 to $150,000 or more). We compensated each participant with a $30 electronic gift card at the study's conclusion. We determined the compensation amount a priori based on an estimate of how long performing the tasks as well as the onboarding and the final survey would take.

5 RESULTS

The data that we include in the analyses is from a 12-day period spanning from the time the earliest group of participants joined the study until 2 days after the latest joining participants finished the study. During this period, the 27 participants of our study submitted a total of 631 headlines. The rate of headline submission across the participants was different, with some submitting many more headlines than we requested for the study, as shown in Figure 7.

We informed the participants that they could remove the extension from their browser after the 7-day period of the study; however, a few decided to keep it installed and continued to suggest more alternative headlines [4].

Throughout the paper, where we present participants' responses, we identify them with a string of the form 'p-' + identifier for anonymity.

5.1 News Sources

Figure 8 shows the distribution of suggested alternative headlines across news sources. The figure does not include those sources for which fewer than 3 headlines were suggested. We found that CNN had by far the most alternative headline suggestions. 73% of the suggestions for CNN headlines came from participants who identified as Democrat (17% and 10% came from Independents and Republicans respectively). This difference partly reflects that there were more Democrats in our

---

[4]Participation eventually died off because as more participants left the extension, the value that the remaining participants received from interacting with the extension to see suggested headlines from or suggest headlines to others decreased. The last participant to leave the extension kept suggesting headlines until 18 days after their formal study participation had ended. Interestingly, we noticed that another participant had submitted a headline more than 5 months after submitting their last headline.

study. But even on a per-capita basis, the average number of headline suggestions for CNN by those Democrats who submitted at least one headline for CNN articles was 6.4 while for Republicans, it was 2.8. Of the headlines that were suggested for Fox News articles, 77% came from Democrats and the rest were suggested by Republicans. The average number of headline suggestions for Fox News by those Democrats who submitted at least one headline for Fox articles was 2.5 while for Republicans, it was 3.3. Therefore, it appears that users were more apt to change headlines by sources aligned with their party. Another outcome could have been for the members of each political party to attempt to modify headlines by sources of the opposing party as out-group members could have a higher level of disagreement with those headlines.

We caution the reader against concluding that CNN has more problematic headlines than other news websites, and note that CNN was one of the most frequented news websites by our participants according to the pre-study survey, with Yahoo News and the New York Times coming in 2nd and 3rd places respectively. In addition, with many users suggesting headlines for CNN articles and the articles appearing on the Reheadline feed, others who did not usually visit the website were exposed to its articles: *"I found that WP [the Washington Post] and CNN published the most problematic headlines, of the news sources I read. (I don't read CNN usually, but I did during this study for some reason. Maybe because other users would change CNN headlines and I found myself drawn to the site some.)" (p-3).* The over-representation of CNN or generally, left-leaning sources could also be due to a similar skew in the distribution of news sources returned by search engines such as Google [82].

In the post-study survey, participants were mixed about whether their opinion of various news sources had changed as a result of further scrutinizing their headlines. Some maintained similar views as before—*"Despite having retitled several of CNN's articles, I do generally trust it because much of their content aligns with over statistical/scientific research as well as substantiated anecdote" (p-8)*, others acknowledged problematic headlines regardless of their political alignment with the source—*"...my news preferences have held up overall, I'm just more aware of sensationalism, hyperbole, or other misleading titles, whether left leaning or right." (p-10).* Some however, became more negative towards sources of the opposing party—*"Probably my attitude toward CNN and WP [the Washington Post] is more negative, insofar as clickbait type headlines are concerned." (p-3).*

Interestingly, one participant stated that they had obtained a better view of a news source that did not align with their political ideology—*"I do not trust fox news because they are a very conservative news source and tend to provide conservative information. Some of the articles I would read were a lot less biased than the actual titles, and that made me trust the site a little more."(p-23)*

5.2 Changes to Headlines

To understand what kinds of headlines our users found problematic or in need of improvement in our study and how they would change such headlines, a member of the research team performed an inductive thematic analysis on the alternative headlines and assigned codes to them by comparing the alternative and the original headlines, as well as sometimes referring to the corresponding article. The reference to the article was needed to understand for instance, if the context that was added to a headline was taken from the article or if it was the participant's opinion. An alternative headline was often a result of multiple types of change applied to the original headline. Through subsequent passes over the data, the codes that had too much similarity were merged and others showing distinct changes were split. The research team then held multiple meetings to discuss and revise the codes by referring to representative participant suggested headlines for the categories. Subsequently, two members of the research team separately coded 7% of the data and resolved disagreements. We did not calculate inter-rater reliability because we developed the codes as part of a thematic analysis to yield concepts and themes [51].

Our Browser Extension Lets Readers Change the Headlines on News
Articles, and You Won't Believe What They Did!

530:13

Table 1. A summary of the types of changes that study participants made to news headlines. N denotes the number of changes we labeled as belonging to a particular category or subcategory.

Category	Subcategory	Example
Edit for Bias, Deception, or Accuracy	Neutralize sensationalism or bias (N=36)	✏️ Alternative headline: *"Arizona State Senator Tony Navarrete arrested on multiple counts of rape of minors" (p-12)* 🅾️ Original headline: The Democrats Have A Pedophile Problem: Democrat Senator Arrested for RAPING Multiple Children (enVolve)
	Correct deceptions (N=15)	✏️ Alternative headline: *"JetBlue Begins New York-to-London Direct Flights Starting at $202: Good Luck Getting a Seat at That Price" (p-3)* 🅾️ Original headline: JetBlue Begins New York-to-London Direct Flights Starting at $202 (Bloomberg)
	Correct headline and article mismatch (N=2)	✏️ Alternative headline: *"Top epidemiologist: Delta variant is 'maybe the most contagious virus that we've ever seen in living memory.'" (p-12)* 🅾️ Original headline: Top epidemiologist: Delta variant is 'maybe the most contagious virus' ever (The Hill)
Remove teasers	Add concealed information pieces (N=52)	✏️ Alternative Headline: *"Subway franchisees receive pushback after Megan Rapinoe, featured in one of their commercial ads, kneels during the National Anthem at the Olympics" (p-12)* 🅾️ Original headline: Subway franchisees are fed up with Megan Rapinoe's TV ads (New York Post)
	Make the topic known (N=24)	✏️ Alternative headline: *"Rep Marjorie Taylor Green alludes to Second Amendment rights in reference to door-to-door vaccination volunteers" (p-12)* 🅾️ Original headline: Marjorie Taylor Greene Is Going to Get Someone Killed (The Root)
Convey more nuanced information or Additional Context (N=244)		✏️ Alternative headline: *"Average Joes are not required to apply: NASA is looking for competent STEM candidates for a scenario of life on Mars." (p-19)* 🅾️ Original Headline: NASA Seeking People To Pretend To Live On Mars For A Year (HuffPost)
Add sensationalism, bias, or opinion (N=67)		✏️ Alternative headline: *"Democrats renounce one of their own for political gain" (p-11)* 🅾️ Original headline: New York Gov. Andrew Cuomo Announces Resignation Over Harassment Allegations (The Epoch Times)
Copy-Edit (N=210)		✏️ Alternative headline: *"Jennifer Aniston is no 'Friend' to the unvaccinated" (p-2)* 🅾️ Original headline: Jennifer Aniston defends staying away from the unvaccinated (CNN)

The changes that participants performed on headlines as well as the problems that the changes address are described below and summarized in Table 1. Note that these are the changes that our study participants made to the headlines that they modified, and they may not be a comprehensive list of all the changes that would be made to headlines if our tool were widely adopted in the wild. We labeled 32 changes as not belonging to any of the categories listed in Table 1. These changes came from a limited set of participants and often were nonsensical, unrelated to the article, or included very minor word changes. Their submission was likely incentivized by fulfilling the requirements for being granted the compensation.

Throughout the paper, we use the icon 🅾️ for differentiating **O**riginal headlines and the icon ✏️ for marking alternative (suggested) headlines by participants.

5.2.1 Many changes were made to rectify headlines that were on a spectrum of sensationalist to misleading to perceived as outright false. These changes involved the following:

Some changes neutralized the language of sensationalist or biased headlines. The headlines that underwent this change ranged from those that were simply provocative to those that were biased or exaggerated to the point of almost asserting falsehoods. Table 1 includes an example of such a change.

Some changes were made to correct misleading headlines. These were headlines that mislead through implication, e.g., by leaving out essential information that can drastically change the message of the headline or by juxtaposing terms that do not in effect have any connection between them but whose co-occurrence leads readers to draw a connection. Participants corrected these headlines by adding relevant facts from the article or elsewhere, by removing irrelevant terms, or by changing words that could deceive the reader to draw wrong conclusions.

✏️ Alternative headline: *"Texan man fatally shoots daughter's ex-boyfriend after ex breaks in and charges at him" (p-22)*; 🅾️ original headline: Texas father fatally shoots daughter's boyfriend for vandalizing her car (Boston 25 News) The alternative headline reveals that contrary to what the original headline implies, the reason why the father shot the boyfriend was in fact not that the latter engaged in vandalism, which would have arguably been a far more disproportionate response, perhaps shocking readers and enticing them to click on the article.

✏️ Alternative headline: *"Florida and CDC differ on Fri-Sun COVID case counts, by 8,643" (p-4)*; 🅾️ original headline: CDC Exaggerates Florida's Weekend COVID Case Counts (National Review via Yahoo News) In the post-study survey, the participant who suggested the above change explained *"'Exaggerate' implies deliberate deception, but the article presents no evidence of any. (I did my own arithmetic using the revised numbers from the CDC's web page.)"*

A few changes aimed to correct headline and article mismatches. With the article taken as ground truth (regardless of the actual veracity of the article), the headlines that underwent this change misreported what was in the article. Participants corrected these headlines by changing the words or terms in the headline that did not agree with the facts presented in the article. ✏️ Alternative headline: *"FL Gov. DeSantis doubts truth of hundreds of ventilators being sent to his state" (p-18)*; 🅾️ original headline: Federal government sends ventilators to Florida as DeSantis says he's unaware of shipment (CNN). In this example, while the article presents a quote from DeSantis saying that he "would honestly doubt that that's true", the headline inaccurately implies that he assumes no stance on the accuracy of the state's request for and receipt of medical supplies.

5.2.2 Many changes sought to fill the information gap that teasing headlines presented. These were headlines that withheld a piece of information and accentuated the reader's lack of knowledge about it to rouse the reader's curiosity and incite clicks. The distinction between the missing information in teasers and that of misleading headlines is that contrary to misleading headlines, in teasers, the reader is made aware that they do not know certain pieces of information. We further differentiated the changes made to these headlines into two groups, based on how much information was withheld.

Some changes added concealed information pieces. An example of such an alternative headline is as follows:
✏️ *"Martha Stewart's Grilled Cheese Recipe is Great, but Adding Butter Makes it Better" (p-23)*; 🅾️ original headline: I tried Martha Stewart's trick for the ultimate grilled cheese, and I found a way to make it even better" (Insider via Yahoo News)
The participant who suggested this headline explained the rationale behind their proposed change: *"I thought that this title did not provide enough details. Obviously, saying that there is a way to make*

the grilled cheese better is a hook to pull the audience in, but talking about butter in the title is not specific enough to give the entire article away. I thought adding the extra details were important, and if a reader enjoyed cooking they may be more likely to click on this article with the ingredient in the title."

Research has reported forward-referencing, or making a reference to the forthcoming parts of the discourse or the use of unresolved pronouns, as a characteristic of teasing headlines that present information gaps and incite clicks, because such a strategy makes a headline semantically dependent on the content of the article (e.g., This is an A-minus paper?) [9, 13]. Although we observed numerous examples of forward-referencing in teasing headlines, there were some that did not use forward-referencing but intentionally withheld information to increase anticipation:

✏️ Alternative headline: *"No, the COVID-19 vaccines do NOT affect your chances of pregnancy"(p-13)*; 🅾 original headline: Do the COVID-19 vaccines affect my chances of pregnancy? (The Associated Press)

Some changes aimed to better convey the topic of articles. These changes targeted articles whose headlines were too vague, too broad, or otherwise disconnected from the article. The information that was withheld in these cases was a clearer picture of the article's topic rather than the answer to a specific question posed or implied in the headline. These are headlines which Lindemann describes as posing a "fairly complex riddle" [46], and although these types of uninformative headlines have been reported as a characteristic of tabloid news papers [19, 46], we saw news sources traditionally considered broadsheet (e.g., NYT) also publishing such headlines.

✏️ Alternative headline: *"Governors Abbott and DeSantis's anti-mask mandate stances lead to higher rates of child hospitalization" (p-12)*; 🅾 original headline: The GOP's death cult comes for the children (Salon)

5.2.3 A frequently applied change was to convey more nuanced information or additional context. With this change, participants added information to the headline that they believed was important but not included. At times, this change was accompanied with removing information that participants considered of lesser importance or not the main topic of the article. The withheld information was not, at least overtly, left out with the intention to tease, but rather mostly to condense the headline. The difference between headlines in this category and the ones that left out information to mislead was that the omitted information in the former left readers with no information as opposed to false information. The information that participants added via this type of change included missing names, details, conclusions, etc.

Participants also applied this type of change to those headlines that they considered bait-and-switch, i.e., headline promises a content that is in fact not the main focus of the article. Because bait-and-switch headlines are uninformative rather than disinforming, we placed the changes targeting those headlines in this group rather than for instance, in those that edit for bias, deception, or accuracy.

✏️ Alternative headline: *"Pentagon to mandate Covid-19 vaccine for U.S service members in spite of requirement that vaccines mandated by the Pentagon have full FDA approval" (p-22)*; 🅾 original headline: Pentagon to Mandate Covid-19 Vaccine for U.S. Service Members (The Wall Street Journal)

The participant who suggested the above headline explained their reason for the change as *"...people don't usually read articles, just headlines, so I think it's important to note that the Pentagon requirement is illegal"*.

An example of a headline that was changed completely in favor of bringing in conclusions from the article was:

✏️ Alternative headline: *"Reducing greenhouse gas emissions presents last hope for stabilizing Atlantic Gulf Stream" (p-10)*; 🅞 original headline: A crucial ocean circulation is showing signs of instability. Its shutdown would have serious impacts on our weather (CNN)

Dor argues that a property of an appropriate headline is that it should be as short as possible, and while beginner copy-editors suggest longer headlines, senior editors condense headlines and decide which parts of the information can be left out [19]. The high number of changes that we observed in this category suggests that for many headlines, the style of beginner copy-editors seems to align more with how the crowd want to rewrite headlines.

5.2.4   Many changes added sensationalism, bias, or opinion to headlines. Somewhat contrary to the original motivation of the tool, we found that some users editorialized headlines in an attempt to make the headline grab attention or express the participant's perspective or sentiment on the subject. These self-expressions ranged from subtle to overt.

✏️ Alternative headline: *"Candace Owens Ignores Consensus of Medical Community and Common Sense as Covid-19 Crisis Continues" (p-26)*; 🅞 original headline: Candace Owens Says She's 'Proud' to Not Be Vaccinated Against COVID-19: 'I Trust My Gut' (Newsweek)

✏️ Alternative headline: *"Grifter greenlighted sweetheart deals for contractor buddies" (p-26)*; 🅞 original headline: How Trump stiff-armed Congress — and gaslighted the courts — to build his wall (POLITICO)

5.2.5   Participants often assumed the role of a copy-editor, and edited a headline with the intention to transform it into a syntactically or stylistically more appealing headline. Some of these changes were aimed at achieving a better arrangement of words or a more creative choice of vocabulary, some others attempted to shorten the headline into a more condensed form to reduce reading effort [19], etc.

Although some changes that aimed to improve the vocabulary or syntax of a headline simply enhanced those that did not have glaring problems, others were deemed as more essential:

✏️ Alternative headline: *"US equestrian George Morris banned from competing and accused of sexual abuse by two male teens" (p-23)*; 🅞 original headline: Legendary US equestrian George Morris accused of sexually abusing two teens (Fox News)

The participant who suggested the above headline explained: *"I changed this title because I thought using the term "legendary" glorified George Morris. Even if his equestrian status is legendary, I do not think a possible sexual assaulter should be glorified..."*

5.2.6   What makes a good headline is highly subjective. It is noteworthy that some users changed headlines in ways that were opposite to other users. Some users added information while others removed it. Some added bias while others removed it. Some made a headline more sensationalist while others made it less. Some added clever wordplay while others removed it. Our tool does not force all users towards a single consensus headline; rather, it permits a diversity of headlines to emerge that can satisfy a diversity of user preferences.

5.3   User Response to the Tool

In this section, we describe some of the user interactions with the tool and their experiences with re-headlining news articles and seeing headlines submitted by others.

5.3.1   Reading Articles Before Suggesting Headlines. To gain insight into how thoroughly participants tried to read an article before suggesting a headline for it, we examined the time participants spent on an article page before they submitted an alternative headline for it. The median time was 2 minutes and 9 seconds, and the minimum was under 8 seconds. We did not have the time spent

on the article for all the headlines because it was also possible to suggest a headline through the Reheadline feed (and not directly on the article page) and some headlines were suggested that way. While the reading times could suggest that some participants may have not read the articles for which they suggested headlines in detail, it is possible that they had opened and read the article before but did not suggest a headline for it at the time or were previously familiar with the story. Participants' responses in the post-study survey suggest that both explanations are accurate. In response to why they suggested headlines for articles they had not read in full, some said:

"I watch CNN often on tv, and so I had already heard their input on tv." (p-2)

"I felt that I could get the idea of it without reading it. Some articles were really not conveying serious info. They might have been about a celebrity or something minor." (p-14)

"I had read enough to see that some context was missing the headline. I think a few times I had run out of free articles, so I could only see the first paragraph or so" (p-26)

5.3.2 Perceptions of Headlines Suggested by Others. We have discussed how users changed headlines as they saw fit. Here we report on how they perceived alternative headlines written by other users. In the post study survey, we asked our participants whether they generally found headlines submitted by other users helpful. Some participants reported that they did:

"Yes, I thought most other users made the headlines more helpful. Other users tended to better highlight the point of the article with their headlines or to pick out the most relevant or important information." (p-20)

Other participants however, found faults with a fairly large number of alternative headlines. They reported that the flawed headlines were biased or editorialized, contained grammatical or spelling errors, or were simple a paraphrasing of the original headline or did not make meaningful improvements over it. Interestingly, although most participants who cited bias in alternative headlines mentioned that they tried to steer away from biased headlines in reading or creation, a few acknowledged that they were inclined toward those headlines that confirmed their biases: *"It was interesting to see how some went very political in their headline changes, and I found myself gravitating toward some users whose views I identified with, going so far as to check out what other headlines they had changed. It was also interesting to see what news sources people read, or what kinds of content they singled out eg sports, entertainment, political." (p-3).* This quote also signals the potential for echo chamber formation as people gravitate towards attitude consistent headlines. We discuss this further in section 7.6.

Our dataset contained 556 endorsements of alternative headlines by the 27 participants who completed the study. In the post-study survey, for each participant, we randomly displayed two headlines that they had endorsed and asked their reasons for doing so. Many of the reasons for endorsement align with those for rewriting headlines in the first place. participants' reasons for endorsement included because the alternative headline captured the crux of the article better than the original, brought out context that they would have liked to know, was more creative or amusing, removed the bias present in the original, conformed to the reader's own views or biases, had more preferable linguistic features (e.g., conciseness, choice of words or syntax, etc.) than the original, removed teasers, helped summarize the article for those who did not wish to read the article, or helped garner interest in the article for those who would otherwise have scrolled past the article: *"Very informative headline, way more details and way more useful than original headline. I instantly wanted to read the full story after reading this headline. Liked that the headline told me the article was about how private school vouchers funded by taxpayers are being used to teach racist Christian beliefs, pings multiple areas of interest." (p-20).* The headline in question was the following:

✏️ Alternative headline: *"Whitewashed Christian textbooks teach distorted history at private schools, with funding from U.S. taxpayers through 'school choice' voucher programs" (p-10)*; 🅞

The extent to which participants like to use a tool like Reheadline
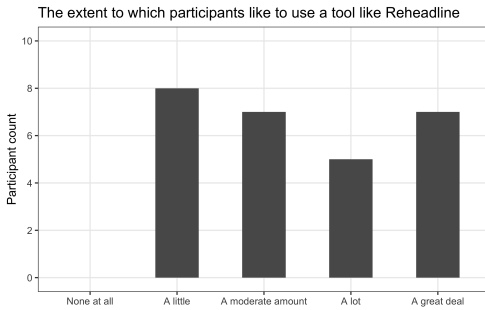
Fig. 9. The extent to which participants reported liking to use a tool similar to Reheadline that allows for suggesting alternative news headlines and seeing headlines suggested by others.

Benefit to Social Circle
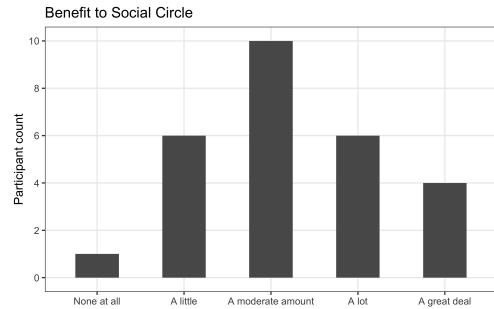
Fig. 10. The extent to which participants believed a tool like Reheadline can benefit their social circle.

Original headline: The rightwing US textbooks that teach slavery as 'black immigration' (The Guardian)

Interestingly, a participant noted that the perception of bias and even its direction in a headline can vary across users: *"The original article using the term "Biden administration" which can have either negative or positive connotation, depending on the reader. The focus of this article is about the urgency of climate issues, which is already a political issue, but should not be made into a bigger one. I liked the term "United State president" better." (p-23)*

In our design, users could not downvote headlines by other users. Instead, when they saw an alternative headline that they did not agree with, they could suggest another headline for the article. We observed that some users unfollowed a few others to not see their suggested headlines anymore. When asked about their reasons, they explained that the alternative headlines by those users were biased or did not meet their quality standards: *"[User 1 name]'s headlines were blatantly pushing their own political narrative. [User 2 name]'s were poor quality and irrelevant to the article." (p-13)*

5.3.3 Users' Perceptions of the Tool. Figure 9 summarizes the degree to which participants reported liking to use an extension like Reheadline that allows them to suggest alternative headlines for news articles and to see headlines suggested by others. The figure shows that participants overall liked to use a tool such as the one in our study, although as seen in earlier sections, the appeal to some was greater than to others. Participants also reported how much a tool like Reheadline can benefit their social circle, summarized in Figure 10. When asked about the potential benefits of such a tool to their social circle, participants cited empowerment of news consumers—*"It will help to save them the pain of been [sic] hopeless and helpless when you find a misleading headline and you can't do anything about it." (p-16)*, potential increased interest in or less hesitancy to read a larger range of news—*"They might read articles they'd skip, overlook, or otherwise avoid, depending on the quality of the headline, and that information might potentially benefit them...or not. But on the whole, it might widen their horizons or balance out their typical media diet." (p-10)*, potentially promoting critical thinking—*"...I think they would learn to think for themselves and get to the main idea of what they are reading" (p-14)*, learning about what others are thinking, and entertainment.

The value of any tool that enables collaborative editing of news headlines is dependent on the users who contribute alternative headlines. We asked participants from whom they would like to see headlines when using such a tool. 23 out of the 27 participants said they like to see headlines from journalists they respect, 21 wished to see headlines from the experts of the field the article is

about (e.g., a climate scientist on an article about climate science), 20 from the individuals they trust, 12 from anyone who uses the tool, and 11 from anyone as long as their submitted headline is endorsed by many users.

6 EVALUATING SUBMITTED HEADLINES

Our qualitative findings suggested that different participants preferred different headlines. We ran an additional user study to further investigate whether headlines suggested by untrained users could be preferred over the original headlines curated by professional editors. The core of this user study involved a new set of participants deciding whether they liked the original or the alternative headlines submitted by the tool participants for a series of news articles. This study was approved by our Institutional Review Board.

6.1 Procedure

Through an online survey developed and deployed by our research team, participants first reviewed a consent form in which we explained that the purpose of the study is to evaluate what news headlines news consumers prefer. Each participant was shown a series of 10 sets of questions. Each set contained a link to an article and a question asking whether the participant would have wanted to read the article at the time when it was published. We intended for this question to capture their interest in the topic. We logged whether the participant clicked on the article's link. In addition, each set presented a pair of headlines—one original and one alternative—for the article and asked which of the two headlines they would have preferred to see for the article and why. The participants could also specify that they liked neither headline. The 10 pairs of headlines that the survey presented to each participant were randomly drawn from a pool of 625 alternative headlines that the participants of the Reheadline study had submitted and their corresponding original headlines. We had excluded 6 alternative headlines for which the tool had not captured the original headline or the article link due to technical problems. Following the headline selections, participants then answered a set of demographics questions similar to the previous user study.

6.2 Participants

We recruited U.S. based participants from a behavioral research platform managed by our institution as well as from Amazon Mechanical Turk (MTurk). The participants from the Reheadline user study could not participate in this study. A total of 312 participants completed this study and provided demographics information of whom 50% were female. Of the participants, 43% identified as Democrat, 26% as Republican, 27% as Independent, and 4% as other. 80% identified as White, 11% as Asian, 5% as Black, and the rest either belonged to other ethnicites or identified as mixed. They were distributed across a wide age range (18 to 76) with a median of 35 years. The median for income was $50,000 to $59,999 (ranging from less than $10,000 to $150,000 or more); and for education, Bachelor's degree (ranging from high school diploma to Doctoral or Professional degree—M.D. or J.D.).

   We had set a payment of $1 for the study; however, at the study's conclusion we noticed that the median completion time for the task was approximately 15 minutes which was longer than what we had anticipated. Therefore, we increased the payment to $2 by granting a bonus of $1 to those users we had recruited through MTurk. We compensated those participants whom we had recruited from the behavioral research platform with a $2 Amazon gift card.

6.3 Results

Our dataset contained 3381 datapoints with each datapoint consisting of whether a participant preferred the presented original headline over the alternative or vice versa (or neither), along

Fig. 11. The empirical cumulative distribution function of headline preferences of participants for each article. The vertical portions (i.e., projection on the Y-axis) of the dash-dotted green, solid orange, and dotted purple lines on the right of the vertical line $x = 0.7$ show what portions of the articles had their original, alt, and neither headlines respectively preferred more than 0.7 of the time.

with other information that was captured about the participant's interaction with the article and their interest in it. Some of the datapoints in the dataset were provided by participants who had abandoned the task before its conclusion.

Overall, the majority (N=1756 or 52%) favored the original headlines. Nevertheless, a substantial number (N=1051 or 31%) preferred the alternative headlines, and 574 (17%) said they liked neither. In the majority of the datapoints that stated no preference, participants explained that they were not interested in the topic of the article for the choice of the headline to make any difference. However, the percentages for the alternative vs original headline preferences revealed that users did in fact prefer a large number of the headlines that untrained consumers had suggested over headlines by professional media editors. In addition, headline preferences across different participants were different, with some mostly preferring the original headlines, some the alternatives, and some often selecting that they liked neither headline. We examined the empirical (i.e., observed) cumulative distribution function (ECDF) of headline preferences for each article and by each participant (see Figures 11 and 12). The ECDFs show the diversity of interests. Figure 11 shows that there were articles for which the alternative headlines were often preferred over the original, and yet other articles for which the original headlines were often favored. For instance, the vertical and horizontal lines demonstrate that for 6% of articles the alternative headlines were selected more than 70% of the time, while for 23% of articles the original headlines were frequently selected. Similarly, Figure 12 shows that 28% of participants often (more than 60% of the time) preferred the original headlines, while 5% often preferred the alternatives.

It is conceivable that a factor influencing the perceived quality of an article's headline is the participant's interest in reading the article. For instance, an alternative headline that reveals all the necessary information in an article may be preferred by those who do not have enough interest in the article to read it in full. To understand whether participant's interest in an article was correlated with which headline they preferred, we performed a Chi-squared test of independence on the contingency table of participants' interest in reading the article and their headline preferences (see Table 2 in the Appendix). The test revealed that the distribution of the two were in fact not independent ($\chi^2(2) = 329.73$, $p < 0.001$). Table 2 suggests that when users are not inclined to read
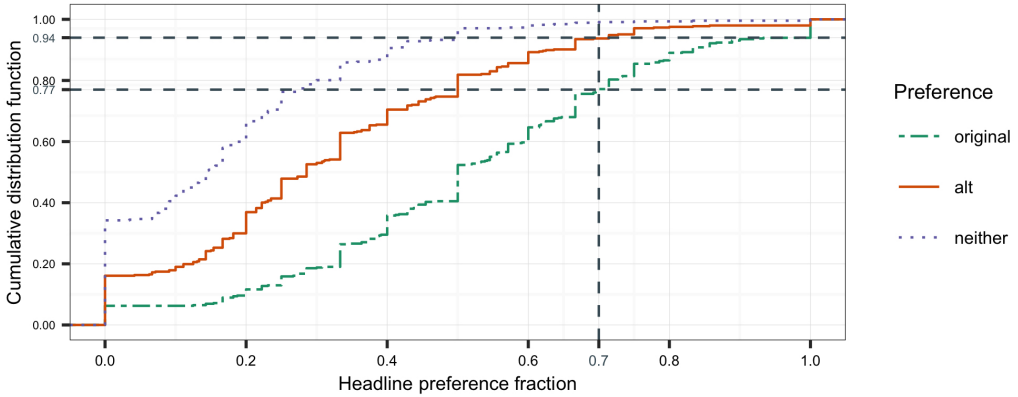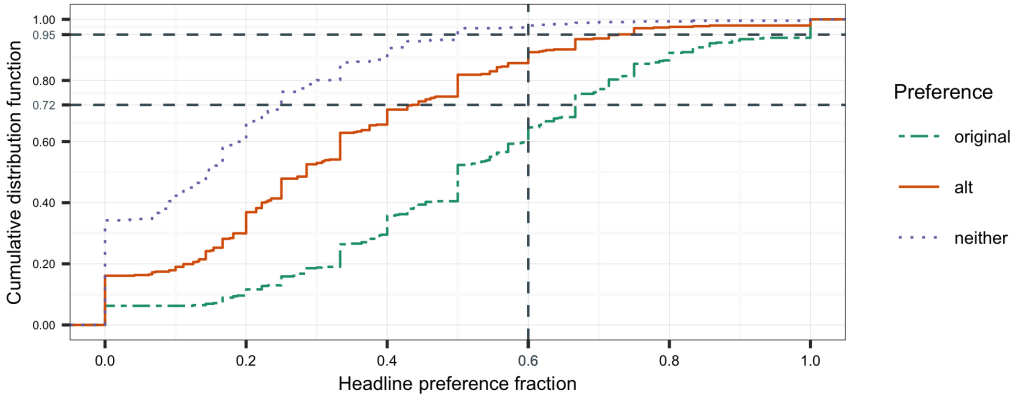
Fig. 12. The empirical cumulative distribution function of headline preferences for each participant. The vertical portions (i.e., projection on the Y-axis) of the dash-dotted green, solid orange, and dotted purple lines on the right of the vertical line $x = 0.6$ show what portions of the participants preferred articles' original, alt, and neither headlines respectively more than 0.6 of the time.

Table 2. The distribution of participants' headline preferences by their interest in reading the article they were presented.

Interested in reading the article	Alternative	Original	Neither
Yes	514	1060	96
No	537	696	478

an article, they still prefer the alternative headlines at a similar rate as when they are interested in an article. Their interest in the original headlines however, drops. The sustained interest in the alternative headlines could be because these headlines may disclose enough information to those users who have already decided they do not want to read an article in full to satisfy their information needs. Or rather, the alternative headlines release enough information scent to help users determine that they are not interested in an article. A participant who had indicated no interest in reading an article explained that they preferred the alternative headline because *"I like this one because you can just read the headline for information and skip reading the full article."*.

Participants' reasons for why they preferred a certain headline were for the most part consistent with those cited by the Reheadline users for why they endorsed certain headlines. They also cited a headline's catchiness, its being professional, absence of an accusatory tone, and even sometimes teasing and leaving questions unanswered as reasons why they preferred a headline. In a few cases, participants knowingly selected a sensationalist headline because they believed that the appeal to readers' emotion was toward a greater good: *"In the case of COVID stories, emotional appeal has been a big factor in getting people to get vaccinated. Data is important too, though. I would be more likely to read the other headline but I think the one I picked would move people more."*

6.3.1 How the Types of Headline Changes Influence Headline Preferences. We performed an exploratory analysis of the headline preferences dataset to understand how the changes characterized in Study 1 affected participants' choice of alternative vs original (or neither) headlines. We fit two types of models to our dataset, both with the headline choice (original, alt, or neither) as the response variable. One was a linear mixed effect model in which we assigned values of 0, 0.5, and 1

530:22 Farnaz Jahanbakhsh et al.

Table 3. Linear regression predicting headline preference with headline and participant as random effects.

Change Type	Estimate	Std Error	df	t value	p value
(Intercept)	0.31	0.05	496.13	6.90	**<0.001**
Remove bias or sensationalism	0.11	0.04	501.73	2.80	**0.005**
Correct deceptions	0.17	0.07	513.55	2.48	**0.013**
Add concealed information pieces	0.15	0.05	483.60	3.05	**0.002**
Make the topic known	0.18	0.06	566.44	2.76	**0.006**
Add nuanced info or context	0.11	0.04	483.28	2.58	**0.010**
Add bias or sensationalism	-0.08	0.04	480.46	-1.76	0.079
Copy-edit	0.07	0.05	490.34	1.39	0.165

to the preferences of original, neither, and alt respectively [5]. The other was a cumulative link mixed model which treated the preferences as ordinal. We present the results of the the linear mixed effect model in this section as it is more straightforward and describe the results of the cumulative link model in the Appendix Section A.

The independent variables for either model were each type of headline change. We included the headline identifier and participant id as random effects to account for the variation in choice attributed to (unobserved) characteristics of a headline or a participant, rather than the variation attributed to the change made to an original headline to yield the alternative. We removed those datapoints where the only change of the alternative relative to the original headline was that the former corrected headline and article mismatches because we had too few of those datapoints. In addition, we removed those datapoints where we had determined the change did not belong to any of the categories listed in 1, as explained in 5.2. These exclusions resulted in 3227 datapoints. The model we fit to the data was as follows:

$$\text{preference outcome}_{hp} = \beta_1 1(\text{neutralize bias})_h + \beta_2 1(\text{correct deceptions})_h + \beta_3 1(\text{add concealed info})_h$$
$$+ \beta_4 1(\text{make topic known})_h + \beta_5 1(\text{convey nuanced info})_h$$
$$+ \beta_6 1(\text{add bias})_h + \beta_7 1(\text{copy-edit})_h + b_h Z_h + b_p Z_p + \epsilon_{hp}$$

$$(1)$$

where the preference outcome$_{hp}$ is participant p's choice for a pair of headlines (one original and one alternative) denoted by h. The indicator variable $1(\text{change type})_h$ denotes whether for the pair h, the alternative incorporates the specified change type relative to the original. Z_p is the matrix for the random effects for observations for participant p, and Z_h for pair h. ϵ_{hp} is the error term. The coefficients of interest are β_1 through β_7.

Table 3 shows the result of the regression and the estimates for each of the change types. The results suggest that users' likelihood of preferring an alternative headline compared to its original counterpart increases when the alternative sought to remove bias or sensationalism, correct deceptions, add concealed information pieces, make the topic known, or convey more nuanced information or context. In contrast, the effect of changes of the type adding bias or sensationalism as well as copy editing were not statistically significant at the α = 0.05 level.

---

[5]In formulating the response variable as a numeric, we take "neither" to indicate that the value of the original or alt headline is the same to the user

7 DISCUSSION AND FUTURE WORK

With many people using social media to consume news [53] and not reading past the article headlines [23, 31, 47], misleading or otherwise problematic headlines have the potential to do great harm, especially as they can be easily shared and expose others to misinformation [13]. A body of work has tried to address this issue with machine learning models to detect and label clickbait headlines [8, 36, 62, 64, 65]. We take an experimental approach, where rather than labeling headlines, warning users, or downranking such headlines out of the users' view, we allow them to alter problematic headlines whenever they encounter them through the use of a browser extension. One study participant described that, *"...many people already comment on social media about problematic headlines and make suggestions about what the headline should have been - this is literally the evolution of that to a tool that allows them to directly "fix" problematic headlines and show it to all their social media followers." (p-20)*

We conducted a user study of the extension and saw that users perceive value in the tool and used it to make a variety of changes to news headlines, e.g., correcting misleading or teasing headlines by adding crucial missing information. While a potential concern around the findings or the participants' engagement is the existence of the Hawthorne effect, which leads participants to change their behavior due to an awareness of being studied [41, 49, 63], we observed the majority of our participants submitted far more headlines than what was required for the study (see Figure 7) and some continued to use the extension even after the formal study had ended and they had received their compensation, suggesting that they found the tool valuable. As one participant said: *"I think it is long overdo [sic] to give the public a voice in news now that all of them are ridiculously biased and not transparent nor honest about their bias." (p-9)*

In a followup study, we asked another set of participants to indicate whether they preferred the alternative headlines submitted by the users of Reheadline or their corresponding original headlines provided by news editors, while presenting them the article to which the headlines belonged. We found that a substantial number of the alternative headlines were preferred. In addition, our statistical analyses on the data from the followup study revealed that users were more likely to prefer an alternative headline compared to its original counterpart if the alternative sought to improve the headline, for instance, by removing bias or correcting deceptions. These results can be further evidence that the headlines provided by publishers do not fully align with the needs of news consumers and that although user empowerment in this domain can result in non-ideal changes, users are perceptive and appreciative of changes that are improvements over the original headlines.

In the sections that follow, we offer insights on the potentials and the challenges of the user empowerment that our tool brought about.

7.1 Problems and Improvements in Headlines Can Be Subjective

In giving users the agency to alter headlines, we recognize that not every user agrees on whether a headline is problematic and how to improve it. Indeed, our followup study of evaluating the headlines that the Reheadline users had submitted revealed that many of the alternative headlines were preferred over their original counterparts by some users, and not by others.

An interesting example of disagreement in the perceived problem that an alternative headline sought to fix arose when two members of the research team were resolving the differences in the codes they had assigned to participant responses. For the original headline: **O** Ditching Man Utd [Manchester United F.C.] to support my local team. Now they're in the Premier League (CNN), a participant had suggested the following alternative headline: ✏️ *"Beauty of football: Brentford fan switch [sic] back to his local team after many years of supporting Manchester United" (p-7)*. One of the

research team coders believed that the change was aimed at removing teasers, with the concealed information being the name of the local team, while the other believed that the alternative headline was simply adding more nuanced information the presence of which in the headline was not crucial. The opinion of the latter stemmed from their lack of interest in or knowledge about soccer. Indeed, to the latter coder, it was the presence of the original headline along the alternative that clarified that Brentford in the alternative headline was in fact the name of the local team in question.

Our approach leaves it to each user to decide how to improve headlines as it is the user who can best consider their and their social circle's information competency as well as knowledge and interest in the topic, among other factors.

7.2 Impact on the Livelihood of News Media

One potential implication of enabling news consumers to suggest headlines for news articles is that headlines that summarize an article well and provide enough details may deter users from opening the article. The reduction in clicks can subsequently impact the livelihood of the news media.

We wished to understand the effect of displaying or withholding alternative headlines on people's propensity to click on articles. Therefore, we set up the study such that the first time that a user encountered a headline in the wild (e.g., on a news website, social media feed, etc.) for which alternative headlines had been submitted, the server would randomly determine whether to display its alternative headlines or withhold them. The alternative headlines would be withheld with a probability of 0.2 unless the user had already suggested headlines of their own for the corresponding original headline. If the alternative headlines for an original headline were withheld on the first encounter, they would be withheld on subsequent encounters as well.

Unfortunately, we did not observe enough instances of participants coming across headlines, in the wild, for which alternative headlines were suggested. Therefore, we could not examine the effect through statistical tests. In the post-study responses however, participants reported that alternative headlines enticed them to click on articles included when they provided enough information for them to assess whether they were interested in the article, were much different from the original headline, or when users wanted to assess whether the alternative headline represented the article reasonably. In contrast, alternative headlines discouraged them from opening an article when the alternative headlines appeared biased or provided as much information as they wanted to know about the topic.

A naturalistic field study similar to ours examining the effect of displaying headlines would need to recruit a large user base to use the tool for a long period of time. This setup would increase the odds that there will be overlap between the news sources that the users frequent or follow on social media and the time that they visit or encounter the articles from those sources, despite homepages of news websites being updated frequently. Nevertheless, to make such a study more tractable, future work can opt for a more controlled lab experiment at the expense of some degree of naturalism in user click behavior related to news sources or topics of interest.

7.3 Rewriting Headlines of False Articles

A tool like Reheadline can enable people to signal that a headline does not accurately represent the content of an article. However, there are other types of false information where the fault is the article's and not the headline's. These can be either factual articles that are false or extremely biased, satire pieces, or op-eds with which the reader does not agree, and where the headline simply conveys the message of the article. In our study, we found instances of suggested rewrites for the headlines of op-eds or interviews, where users had tried to correct the points with which they disagreed in the headline by adding facts they knew from elsewhere or commentaries of their own:

✏️ Alternative headline: *"Philippa Rizopoulos, author of this article, claims they know better than everyone else what healthcare needs. But, what is their expertise and experience?" (p-9)*; 🅞 original headline: We Must Fight Privatization of Medicare and Medicaid to Win Single-Payer (Truthout)

Others however, refrained from suggesting alternative headlines for such articles:

"[I did not suggest headlines for articles that were false] because neither the article or the headline was accurate or should be promoted or viewed. I would not promote headlines from certain slanted sites, those type of sites don't deserve to get page views and it is disgusting that they get to spread their ignorant and harmful content." (p-20)

Different users taking different approaches to headlines of inaccurate articles can result in inconsistent usage of the tool, with some users re-purposing headline suggestion as a means to assess the accuracy of the article. This inconsistency can be problematic for those users who upon encountering the alternative headlines will not know whether to expect an informative summary of the article or commentaries by other users. However, the ability of users to choose which others they follow means that a user can over time follow people who rewrite headlines in ways that are predictable and appeal to the given user.

In addition, the fact that users were wary of promoting an article as a side effect of a headline suggestion or that users raised doubts about an article's veracity in its headline points to a need for tools that empower crowds to help each other discern false from true content by explicitly assessing individual articles. Jahanbakhsh et al. have proposed such a tool in [33] and discussed the potentials of enabling all users to assess content and for their assessments to be captured as structured metadata. Future work could add to the extension a checkbox to mark an article as false. Such an indicator could lead to the headline being struck out—without a rewrite—as a sign that it should not be read. A more aggressive design could make it possible to hide or delete such headlines entirely, to reduce the wasted consumption of other users' attention.

7.4 Towards User-Curated News Delivery?

A tool like Reheadline can act as a new community communication signal that allows the reader to counter a large establishment–a form of user empowerment that enables users to help their social circle or seek their help should they wish. This empowerment can help users customize headlines based on their needs as not all users agree on what the best headline for a story is. This disagreement was demonstrated by changes to headlines that were in opposite directions. For instance, some sought to condense headlines and some added yet additional information; similarly, some changes removed sensationalism and others made the neutral language of a headline sensationalist. News websites already perform A/B testing of headlines [28, 73, 80] and therefore, the headlines that they deliver are different across users. However, this difference in headlines is forced on users and optimized for profit. There exist other platforms such as Reddit or Hacker News that allow for user generated headlines of posted articles [30]. Tweets that link to articles are an example of posts on other social platforms that are also used for this purpose. One of the features that Reheadline additionally provides is that the user-generated content follows the original content regardless of on which platform the content appears. Furthermore, in Reheadline, the structure of the user generated content as well as the follower networks allow users to personalize the kind of headlines they want to see by modifying the list of the people they follow. Reheadline therefore, does not force users to all see a singular truth.

Taking this idea further, we can enrich user control over their information space by enabling them to endorse headlines on fine-grained categories such as informative, funny, etc., similar to platforms such as Slashdot [43], and providing them with the option to use these tags to filter the kind of alternative they wish to see. User preferences signalled by endorsements,

headline suggestions, and filter settings can also be taken as training data for machine learning models that display the headline most likely preferred by the user for the benefit of the user.

7.5 Should Users be Able to Choose Their Headlines?

Our system empowers users to choose what headlines are associated with stories they encounter. It is worth asking whether this freedom is a good thing. Much concern has been raised about filter bubbles—the risk that users may choose to filter out content that they do not like, through a phenomenon called selective exposure [69], and in doing so develop a picture of reality that diverges in important ways from others in their society, increasing polarization [37, 74, 79]. Allowing users to choose what headlines they read poses similar risks. If a user chose to see only those headlines rewritten in ways that always framed their political opponents' actions in negative terms, then that user would get a steady stream of headlines that reinforce their negative opinion of their political opponents.

However, despite the fact that filtering can lead to filter bubbles, we have chosen to provide consumers with numerous opportunities for filtering. Users can choose which other users to follow on social media or which news sites to read on the web or subscribe to in RSS readers; they can filter content based on tags, keywords, or upvotes [43].

Indeed, one can think of the many candidate headlines contributed by users as content in their own right—not too different from a tweet including a link to an article—and can therefore consider the selection of headlines as yet another example of the kind of content filtering we already offer to users.

Many online news sites, frequently reporting on the same news story, choose different headlines for it. Thus, there is already a diversity of headlines available. At present, there is no way for users to choose from the pool—they get whichever headline is chosen by the source they are reading. But one could certainly leverage NLP techniques, such as Topic Detection & Tracking [4, 88], to determine that several news sites are publishing the same story and that their different headlines are all alternative candidates for it. In this work, we have chosen instead to explore crowd-sourcing of those alternative headlines.

Additionally, Reheadline can render support to fact-checking organizations that may want to go beyond simply flagging headlines as false and instead replace them with accurate headlines. One could also imagine this role being adopted by news organizations; it would seem natural for journalists, who are already in the business of creating headlines, to also offer corrections to headlines they find problematic.

7.6 Echo Chambers

As discussed, a tool that gives users control over the content they consume can also give them the power to create echo chambers of like-minded people [55, 77]. To understand whether echo chambers could potentially be a problem with our tool, we asked our users whether they believed it was a bad idea to see headlines suggested by strangers. While a few preferred to see headlines from their smaller social circles only—*"...I had [sic] rather see those headlines changed by a smaller circle of friends versus the general public. Or perhaps some people I choose to follow." (p-3)*—surprisingly, the majority liked the idea of receiving headlines submitted by people they did not know:

"I think it was interesting. I definitely did find myself feeling angry at some of the headlines that were very conservatively written, but then I just unfollowed the users. I feel like if I were to only see headlines from people I follow, than [sic] my filter bubble would become extremely small and catered to myself, when the whole point of reheadlining is to see articles with less biased headlines." (p-23)

"It was really great, showing alternative headlines help other users to get insight on how people translate things differently." (p-7)

These observations are in alignment with prior work that reported on individual differences in people's preference of and receptiveness to collections of news articles that include only items confirming their views vs those that also include items that challenge their views [56]. Nevertheless, echo chambers could still be a possibility if our tool were widely adopted. And more work is required to fully understand the implications. Some initial suggestions to offer headlines outside of an echo chamber include displaying headlines tagged as "insightful" from people the user is not following, and recommending other people to follow who are different from the user with respect to the types of changes they suggest or the news sources for which they suggest, while being transparent about these processes. The hope that users will be receptive to such high quality suggestions is rooted in prior work which reports that people exclude opposing viewpoints not out of aversion to other opinions but because they perceive less benefit in them and that discusses factors that lead people to favor counter-attitudinal information in particular settings [25, 26]. In our experiment, we defaulted to having all users initially follow each other, while allowing them to *unfollow* users whose headlines were not satisfactory. While in a deployed system we would not want to follow *all* other users, we could periodically introduce a small number of random new users into a consumers' follow list, who would in the course of use be "vetted" by the consumer and unfollowed if their contributions were unsatisfactory. We might wish to switch from a framing of "following" to one of "blocking", but the results would the same.

7.7 Combating Misinformation

Harvard's Technology and Social Change Media Manipulation Casebook describes five phases in the media manipulation life cycle: (1) planning a manipulation campaign, (2) seeding a campaign across the web, (3) the response of activists, journalists, industry, and politicians, (4) mitigation, and (5) adjustments by manipulators responding to infrastructure changes [2]. Most mitigation strategies to date begin in the fourth phase. Leading disinformation researchers argue this is too late [17, 18, 70, 75, 76, 90]. Extensions such as the one we propose offer an opportunity to intervene in stages two to three. The approaches that tackle misinformation at this stage restrain the sharing of misinformation, e.g., as it is handed from user to user [32, 58], rather than provide correcting information after it has already spread. Our hope is that by presenting the more informative alternative headlines contributed by those users who assume the role of activists, others who come across the original headlines will receive better indications of their credibility and refrain from sharing misinforming headlines.

Moreover, opportunities exist to intervene in these earlier stages with community support. From our participant responses, there was some discussion of collaboration with local community serving and news organizations. With the support of regulations and local community expert groups, an opportunity exists to build distributed support networks that could cater to local news and broader topics.

8 ETHICAL CONSIDERATIONS

8.1 Impact on the Agency of News Author or Publisher

Our work may give rise to the question of whether enabling news consumers to contribute alternative headlines to news stories comes at the expense of the agency of news reporters in how they wish to present their work. While the practice of headline selection varies across newsrooms, traditionally it has been the case that although reporters can suggest headlines for their stories, the final say ultimately belongs to a copy-editor [19]. In fact, sometimes the decisions that should be made by the newsroom staff, such as the choice of headlines, are imposed on them from above [83]. Therefore, even in the current news ecosystem, the agency of headline selection is transferred

to the publisher, not the author of the work; and the tone, perspective, and information that is conveyed through the headline may not be in alignment with how the author wishes to present their piece.

A related question is whether by allowing the public to contribute alternative titles, we are taking the agency away from the publisher who retains the rights to the content piece. We argue that a succession of online platforms precede Reheadline that similarly follow the unchallenged practice of allowing for user-generated content to be appended to or else replace other content including news articles. Closest to our scenario are user written headlines on Hacker News and Reddit and the user commentaries in tweets or Facebook posts that link to articles [30].

8.2 Impact on the Revenue of News Organizations

It is possible that displaying alternative headlines on articles impacts users' likelihood to click on them or user retention. Although as explained in section 7.2, we did not succeed in measuring this impact in our study, other studies show that users' engagement with news posts are affected by a variety of sentiment and content features of headlines; notably that negative headlines receive more engagement [20, 30]. A tool like Reheadline therefore, can hamper the revenue that such headlines can bring into the journalism industry. This may jeopardize the publishing of quality content which can be underwritten by ad-friendly content [15, 72]. Therefore, a question that needs serious thought is whether increasing the quality of published content through crowdsouring tools like Reheadline can result in a reduction in the publishing of quality content overall, and if so, whether this entails that clickbait and other problematic headlines should be kept and fed to users despite users' interest in having the power to modify such headlines, for instance to be more informative or less biased, as we observed in our studies.

8.3 Weaponization

A reason why we designed Reheadline such that users can see alternative headlines only from the sources they follow was to protect against potential coordinated attacks. Such attacks usually leverage the power of bots or a multitude of fake accounts to mirror or engage with (e.g., retweet) disinformation [29, 84]. A disinforming piece of content that has received a high level of engagement can be further promoted into users' feeds by the platform's content curation algorithm. Reheadline does not have an algorithm that promotes headlines to users. For a user to see headlines created by another account on Reheadline, they need to explicitly follow that account. Therefore, even if a large number of coordinated accounts suggest inaccurate headlines, their efforts will be in vain if users do not follow those accounts. Nonetheless, we acknowledge that there is still the potential for abuse and weaponization of this tool by sleeper accounts—accounts that try to build a reputation by posing as legitimate for a period of time before taking active part in a disinformation campaign [86].

9 LIMITATIONS

Traditionally, copy-editors performed the labor of creating and editing headlines and titles [19]. An extension such as the one we propose could be seen as burdening users with additional labor for which they are not compensated [34, 48]. Today's internet ecosystem is full of participants performing labor of high value (which arguably should therefore be compensated) without compensation, because they view the work as a form of civic participation [10], or because it provides them with indirect value such a sense of accomplishment, community, or creativity [61]. One example is contributing content to Wikipedia; editors create a valuable knowledge artifact, without compensation, because they are passionate about the work. Indeed, Reheadline can be seen as "wikifying" the news ecosystem by allowing all users to contribute content they deem better. In

our exploratory study, participants were paid to explore and modify titles. Issues of time, effort, and affective labor did not surface in the participant responses. Moving forward, more research is necessary to understand the labor involved–visible and invisible–with such a tool.

Note also that our 27 participant study cannot capture or represent the complex dynamics that exist in current online follower network systems (e.g., we created follow networks of people who did not know each other in one study). While we viewed clusters of interaction behaviors in our participant pool, future work is needed to understand how the tool affects community dynamics, whether new norms emerge, if it could be weaponized, and how it would affect communities long-term.

We recruited participants from the U.S. and therefore, many of the news sources for whose articles our participants suggested alternative headlines were U.S. based sources. More work is needed to investigate whether our findings generalize to user populations and news sources that are culturally different from the ones in our study.

10 CONCLUSION

Headlines play a critical role in steering consumers to news. But many headline publishers—mainstream, tabloid, or misinformation sources—craft headlines in ways that either attract clicks in an attempt to earn ad revenue, or misinform users or manipulate their opinions for malicious intents. With many users simply skimming headlines and sharing articles on social media without reading them in full, such manipulated headlines have the potential to spread misinformation and do great harm. In this work, we explore this problem by equipping users with a browser extension that lets them suggest alternate headlines for articles. Using the extension, users can view alternative headlines alongside the original headlines wherever they encounter the original headlines, for instance, on a news website's homepage or on social media. We conducted a user study with 27 participants who used the extension for one week to read news articles of their choice and suggest better headlines for them. In this design, we found that users saw value in a tool that empowers them to have a part in curating the content that they consume. Moreover, we observed a desire to help others stay away from inaccuracies or hyperbole. Participants used the tool in a variety of ways, including to correct headlines that were misleading, had a mismatch with the article, were sensationalist, teasing, etc. And we noted the potential for misuse with such a system. We present the types of changes that users make to headlines if enabled. We also evaluated users' submitted headlines through another user study. Our findings revealed participant behaviors that modified headlines and suggest that mechanisms for rewriting headlines be further explored for collaborative content curation on the web.

REFERENCES

[1] [n.d.]. Annotate the web, with anyone, anywhere. https://web.hypothes.is/
[2] [n.d.]. *The Media Manipulation Casebook*. https://mediamanipulation.org/
[3] Ben Adler. 2014. Tabloids in the age of social media. *Columbia Journalism Review* (2014).
[4] James Allan. 2012. *Topic detection and tracking: event-based information organization*. Vol. 12. Springer Science & Business Media.
[5] Charles Arthur. 2009. *Google Sidewiki: the idea that won't die, but never lives*. Retrieved July 6, 2022 from https://www.theguardian.com/technology/blog/2009/sep/24/google-sidewiki-commenting
[6] Randal A Beam. 1995. How newspapers use readership research. *Newspaper Research Journal* 16, 2 (1995), 28–38.
[7] Allan Bell. 1991. *The language of news media*. Blackwell Oxford.
[8] Prakhar Biyani, Kostas Tsioutsiouliklis, and John Blackmer. 2016. " 8 amazing secrets for getting more clicks": detecting clickbaits in news streams using article informality. In *Thirtieth AAAI conf. on artificial intelligence*.
[9] Jonas Nygaard Blom and Kenneth Reinecke Hansen. 2015. Click bait: Forward-reference as lure in online news headlines. *Journal of Pragmatics* 76 (2015), 87–100.

[10] Brian Butler, Lee Sproull, Sara Kiesler, and Robert Kraut. 2002. Community effort in online groups: Who does the work and why. *Leadership at a distance: Research in technologically supported work* 1 (2002), 171–194.

[11] Abhijnan Chakraborty, Bhargavi Paranjape, Sourya Kakarla, and Niloy Ganguly. 2016. Stop clickbait: Detecting and preventing clickbaits in online news media. In *2016 IEEE/ACM int'l conf. on advances in social networks analysis and mining (ASONAM)*. IEEE, 9–16.

[12] Abhijnan Chakraborty, Rajdeep Sarkar, Ayushi Mrigen, and Niloy Ganguly. 2017. Tabloids in the era of social media? understanding the production and consumption of clickbaits in twitter. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–21.

[13] Yimin Chen, Niall J Conroy, and Victoria L Rubin. 2015. Misleading online content: recognizing clickbait as" false news". In *Proceedings of the 2015 ACM on workshop on multimodal deception detection*. 15–19.

[14] Sophie Chesney, Maria Liakata, Massimo Poesio, and Matthew Purver. 2017. Incongruent headlines: Yet another way to mislead your readers. In *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism*. 56–61.

[15] Angèle Christin and Caitlin Petre. 2020. Making peace with metrics: Relational work in online news production. *Sociologica* 14, 2 (2020), 133–156.

[16] Keith Coleman. 2021. Introducing Birdwatch, a Community-Based Approach to Misinformation.

[17] Joan Donovan. 2020. Concrete recommendations for cutting through misinformation during the COVID-19 pandemic.

[18] Joan Donovan. 2020. The Life Cycle of Media Manipulation". *The Verification Handbook 3 https://datajournal-ism.com/read/handbook/verification-3/investigating-disinformation-and-media-manipulation/the-lifecycle-of-media-manipulation* (2020).

[19] Daniel Dor. 2003. On newspaper headlines as relevance optimizers. *Journal of pragmatics* 35, 5 (2003), 695–721.

[20] Julio Cesar Soares Dos Rieis, Fabrício Benevenuto de Souza, Pedro Olmo S Vaz de Melo, Raquel Oliveira Prates, Haewoon Kwak, and Jisun An. 2015. Breaking the news: First impressions matter on online news. In *Ninth Int'l AAAI conf. on web and social media*.

[21] Ullrich KH Ecker, Stephan Lewandowsky, Ee Pin Chang, and Rekha Pillai. 2014. The effects of subtle misinformation in news headlines. *Journal of experimental psychology: applied* 20, 4 (2014), 323.

[22] Ziv Epstein, Adam J Berinsky, Rocky Cole, Andrew Gully, Gordon Pennycook, and David G Rand. 2021. Developing an accuracy-prompt toolkit to reduce COVID-19 misinformation online. *Harvard Kennedy School Misinformation Review* (2021).

[23] Maksym Gabielkov, Arthi Ramachandran, Augustin Chaintreau, and Arnaud Legout. 2016. Social clicks: What and who gets read on Twitter?. In *Proceedings of the 2016 ACM SIGMETRICS int'l conf. on measurement and modeling of computer science*. 179–192.

[24] Mingkun Gao, Ziang Xiao, Karrie Karahalios, and Wai-Tat Fu. 2018. To label or not to label: The effect of stance and credibility labels on readers' selection and perception of news articles. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–16.

[25] R Kelly Garrett. 2009. Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of computer-mediated communication* 14, 2 (2009), 265–285.

[26] R Kelly Garrett and Paul Resnick. 2011. Resisting political fragmentation on the Internet. *Daedalus* 140, 4 (2011), 108–120.

[27] Joshua Gillin. 2017. The more outrageous, the better: How clickbait ads make money for fake news sites. *PolitiFact, October* 4 (2017).

[28] Nick Hagar and Nicholas Diakopoulos. 2019. Optimizing content with A/B headline testing: Changing newsroom practices. *Media and Communication* 7, 1 (2019), 117–127.

[29] McKenzie Himelein-Wachowiak, Salvatore Giorgi, Amanda Devoto, Muhammad Rahman, Lyle Ungar, H Andrew Schwartz, David H Epstein, Lorenzo Leggio, Brenda Curtis, et al. 2021. Bots and misinformation spread on social media: Implications for COVID-19. *Journal of Medical Internet Research* 23, 5 (2021), e26933.

[30] Benjamin Horne and Sibel Adali. 2017. The impact of crowds on news engagement: A reddit case study. In *Proceedings of the Int'l AAAI Conf. on Web and Social Media*, Vol. 11. 751–758.

[31] American Press Institute. 2014. *How Americans Get Their News.* Retrieved July 6, 2022 from https://www.americanpressinstitute.org/publications/reports/survey-research/how-americans-get-news/

[32] Farnaz Jahanbakhsh, Amy X Zhang, Adam J Berinsky, Gordon Pennycook, David G Rand, and David R Karger. 2021. Exploring lightweight interventions at posting time to reduce the sharing of misinformation on social media. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–42.

[33] Farnaz Jahanbakhsh, Amy X Zhang, and David R Karger. 2022. Leveraging Structured Trusted-Peer Assessments to Combat Misinformation. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–40.

[34] Rae Jereza. 2021. Corporeal moderation: digital labour as affective good. *Social Anthropology* 29, 4 (2021), 928–943.

[35] Jeremiah H Kalir. 2019. Open web annotation as collaborative learning. *First Monday* (2019).

Our Browser Extension Lets Readers Change the Headlines on News
Articles, and You Won't Believe What They Did!
530:31

[36] Georgi Karadzhov, Pepa Gencheva, Preslav Nakov, and Ivan Koychev. 2018. We built a fake news & click-bait filter: what happened next will blow your mind! *arXiv preprint arXiv:1803.03786* (2018).

[37] Silvia Knobloch-Westerwick and Jingbo Meng. 2011. Reinforcement of the political self through selective exposure to political messages. *Journal of Communication* 61, 2 (2011), 349–368.

[38] Ha-Kyung Kong, Zhicheng Liu, and Karrie Karahalios. 2018. Frames and slants in titles of visualizations on controversial topics. In *Proceedings of the 2018 CHI Conf. on Human Factors in Computing Systems*. 1–12.

[39] Ha-Kyung Kong, Zhicheng Liu, and Karrie Karahalios. 2019. Trust and recall of information across varying degrees of title-visualization misalignment. In *Proceedings of the 2019 CHI Conf. on Human Factors in Computing Systems*. 1–13.

[40] Maria Konnikova. 2014. *How Headlines Change the Way We Think*. Retrieved July 14, 2022 from https://www.newyorker.com/science/maria-konnikova/headlines-change-way-think

[41] Frauke Kreuter, Stanley Presser, and Roger Tourangeau. 2008. Social desirability bias in cati, ivr, and web surveysthe effects of mode and question sensitivity. *Public opinion quarterly* 72, 5 (2008), 847–865.

[42] Kenza Lamot and Peter Van Aelst. 2020. Beaten by Chartbeat? An experimental study on the effect of real-time audience analytics on journalists' news judgment. *Journalism Studies* 21, 4 (2020), 477–493.

[43] Cliff AC Lampe, Erik Johnston, and Paul Resnick. 2007. Follow the reader: filtering comments on slashdot. In *Proceedings of the SIGCHI conf. on Human factors in computing systems*. 1253–1262.

[44] Issie Lapowski. 2018. Newsguard wants to fight fake news with humans, not algorithms. *Wired, August* 23 (2018).

[45] Angela M Lee, Seth C Lewis, and Matthew Powers. 2014. Audience clicks and news placement: A study of time-lagged influence in online journalism. *Communication Research* 41, 4 (2014), 505–530.

[46] Bernard Lindemann. 1990. Cheap thrills we live by-some notes on the poetics of tabloid headlinese. , 46–59 pages.

[47] Farhad Manjoo. 2013. You won't finish this article. *Why people online don't read to the end: Slate* (2013).

[48] J Nathan Matias. 2019. The civic labor of volunteer moderators online. *Social Media+ Society* 5, 2 (2019), 2056305119836778.

[49] Jim McCambridge, John Witton, and Diana R Elbourne. 2014. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *Journal of clinical epidemiology* 67, 3 (2014), 267–277.

[50] Terrance McCoy. 2016. Inside a Long Beach Web operation that makes up stories about Trump and Clinton: What they do for clicks and cash. *Los Angeles Times* (2016).

[51] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–23.

[52] Phillipa McGuinness. 2016. The people formerly known as the audience: Power shifts in the digital age. *Communication Research and Practice* 2, 4 (2016), 520–527.

[53] Amy Mitchell, Katie Simmons, Katerina Eva Matsa, and Laura Silver. 2018. Publics globally want unbiased news coverage, but are divided on whether their news media deliver. *Pew Research Center* 11 (2018).

[54] Katarzyna Molek-Kozakowska. 2013. Towards a pragma-linguistic framework for the study of sensationalism in news headlines. *Discourse & Communication* 7, 2 (2013), 173–197.

[55] Mohsen Mosleh, Gordon Pennycook, Antonio A Arechar, and David G Rand. 2021. Cognitive reflection correlates with behavior on Twitter. *Nature communications* 12, 1 (2021), 1–10.

[56] Sean A Munson and Paul Resnick. 2010. Presenting diverse political opinions: how and how much. In *Proceedings of the SIGCHI conf. on human factors in computing systems*. 1457–1466.

[57] Gordon Pennycook, Adam Bear, Evan T Collins, and David G Rand. 2020. The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science* 66, 11 (2020), 4944–4957.

[58] Gordon Pennycook, Ziv Epstein, Mohsen Mosleh, Antonio A Arechar, Dean Eckles, and David G Rand. 2021. Shifting attention to accuracy can reduce misinformation online. *Nature* 592, 7855 (2021), 590–595.

[59] JW Peters. 2010. Some newspapers shift coverage after tracking readers online. *The New York Times* (2010).

[60] Alex Peysakhovich and Kristin Hendrix. 2016. *Further Reducing Clickbait in Feed*. Retrieved July 6, 2022 from https://about.fb.com/news/2016/08/news-feed-fyi-further-reducing-clickbait-in-feed/

[61] Hector Postigo. 2009. America Online volunteers: Lessons from an early co-production community. *Int'l Journal of Cultural Studies* 12, 5 (2009), 451–469.

[62] Martin Potthast, Sebastian Köpsel, Benno Stein, and Matthias Hagen. 2016. Clickbait detection. In *European Conf. on Information Retrieval*. Springer, 810–817.

[63] Chris Preist, Elaine Massung, and David Coyle. 2014. Competing or aiming to be average? Normification as a means of engaging digital volunteers. In *Proceedings of the 17th ACM conf. on Computer supported cooperative work & social computing*. 1222–1233.

[64] Abinash Pujahari and Dilip Singh Sisodia. 2021. Clickbait detection using multiple categorisation techniques. *Journal of Information Science* 47, 1 (2021), 118–128.

[65] Md Main Uddin Rony, Naeemul Hassan, and Mohammad Yousuf. 2017. Diving deep into clickbaits: Who use them to what extents in which topics with what effects?. In *Proceedings of the 2017 IEEE/ACM int'l conf. on advances in social networks analysis and mining 2017*. 232–239.

[66] Md Main Uddin Rony, Naeemul Hassan, and Mohammad Yousuf. 2018. BaitBuster: a clickbait identification framework. In *Thirty-Second AAAI Conf. on Artificial Intelligence*.

[67] Emily Saltz, Claire R Leibowicz, and Claire Wardle. 2021. Encounters with Visual Misinformation and Labels Across Platforms: An Interview and Diary Study to Inform Ecosystem Approaches to Misinformation Interventions. In *Extended Abstracts of the 2021 CHI Conf. on Human Factors in Computing Systems*. 1–6.

[68] Deborah Schaffer. 1995. Shocking secrets revealed! The language of tabloid headlines. *ETC: a review of general semantics* 52, 1 (1995), 27–46.

[69] David O Sears and Jonathan L Freedman. 1967. Selective exposure to information: A critical review. *Public Opinion Quarterly* 31, 2 (1967), 194–213.

[70] Jieun Shin, Lian Jian, Kevin Driscoll, and François Bar. 2017. Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction. *new media & society* 19, 8 (2017), 1214–1235.

[71] Alexander Smith. 2016. Fake News: How a Partying Macedonian Teen Earns Thousands Publishing Lies. *NBC News* (2016).

[72] Gabriel Snyder. 2017. How The New York Times Is Clawing Its Way into the Future. *Nisan* 16 (2017), 2018.

[73] Jessica Soberman. 2013. *Designing from data — How news organizations use A/B testing to increase user engagement.* Retrieved July 6, 2022 from https://knightlab.northwestern.edu/2013/08/15/designing-from-data-how-news-organizations-use-ab-testing-to-increase-user-engagement/

[74] Dominic Spohr. 2017. Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review* 34, 3 (2017), 150–160.

[75] Kate Starbird. 2021. Online Rumors, Misinformation and Disinformation: The Perfect Storm of COVID-19 and Election2020. (2021).

[76] Kate Starbird, Jim Maddock, Mania Orand, Peg Achterman, and Robert M Mason. 2014. Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing. *IConf. 2014 Proceedings* (2014).

[77] Alexander J Stewart, Mohsen Mosleh, Marina Diakonova, Antonio A Arechar, David G Rand, and Joshua B Plotkin. 2019. Information gerrymandering and undemocratic decisions. *Nature* 573, 7772 (2019), 117–121.

[78] Frederik Stjernfelt and Anne Mette Lauritzen. 2020. *Your Post Has Been Removed: Tech Giants and Freedom of Speech.* Springer Nature.

[79] Natalie Jomini Stroud. 2010. Polarization and partisan selective exposure. *Journal of communication* 60, 3 (2010), 556–576.

[80] Alexandria Symonds. 2017. *When a Headline Makes Headlines of Its Own.* Retrieved July 6, 2022 from https://www.nytimes.com/2017/03/23/insider/headline-trump-time-interview.html

[81] Edson C Tandoc Jr and Ryan J Thomas. 2015. The ethics of web analytics: Implications of using audience metrics in news construction. *Digital journalism* 3, 2 (2015), 243–258.

[82] Daniel Trielli and Nicholas Diakopoulos. 2019. Search as news curator: The role of Google in shaping attention to news information. In *Proceedings of the 2019 CHI Conf. on human factors in computing systems*. 1–15.

[83] Suzanne Vranica. 2019. *Julia Angwin Ousted as Editor in Chief of Tech News Site She Helped Found.* Retrieved July 6, 2022 from https://www.wsj.com/articles/julia-angwin-out-as-editor-in-chief-of-news-website-she-helped-found-11556038543

[84] Patrick Wang, Rafael Angarita, and Ilaria Renna. 2018. Is this the era of misinformation yet: combining social bots and fake news to deceive the masses. In *Companion Proceedings of the The Web Conf. 2018*. 1557–1561.

[85] Gavin Wood, Kiel Long, Tom Feltwell, Scarlett Rowland, Phillip Brooker, Jamie Mahoney, John Vines, Julie Barnett, and Shaun Lawson. 2018. Rethinking engagement with online news through social and visual co-annotation. In *Proceedings of the 2018 CHI Conf. on human factors in computing systems*. 1–12.

[86] Leon Yin, Franziska Roscher, Richard Bonneau, Jonathan Nagler, and Joshua A Tucker. 2018. Your friendly neighborhood troll: The Internet Research Agency's use of local and fake news in the 2016 US presidential campaign. *SMaPP Data Report, Social Media and Political Participation Lab, New York University* (2018).

[87] Savvas Zannettou, Michael Sirivianos, Jeremy Blackburn, and Nicolas Kourtellis. 2019. The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality (JDIQ)* 11, 3 (2019), 1–37.

[88] Yun Zhai and Mubarak Shah. 2005. Tracking news stories across different sources. In *Proceedings of the 13th annual ACM int'l conf. on Multimedia*. 2–10.

[89] Amy X Zhang, Joshua Blum, and David R Karger. 2016. Opportunities and challenges around a tool for social and public web activity tracking. In *Proceedings of the 19th ACM Conf. on Computer-Supported Cooperative Work & Social Computing*. 913–925.

Our Browser Extension Lets Readers Change the Headlines on News
Articles, and You Won't Believe What They Did!

530:33

Table 4. Ordinal regression predicting headline preference with headline and participant as random effects. The results suggest that users were more likely to prefer an alternative headline if it changed the original headline for the better.

Change Type	Estimate	Std Error	z value	p value
Remove bias or sensationalism	0.50	0.18	2.81	**0.005**
Correct deceptions	0.83	0.32	2.61	**0.009**
Add concealed information pieces	0.76	0.23	3.33	**<0.001**
Make the topic known	0.85	0.29	2.98	**0.003**
Add nuanced info or context	0.57	0.20	2.87	**0.004**
Add bias or sensationalism	-0.36	0.19	-1.88	0.060
Copy-edit	0.37	0.22	1.70	0.089

[90] Arkaitz Zubiaga, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, and Peter Tolmie. 2016. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PloS one* 11, 3 (2016), e0150989.

[91] Sacha Zyto, David Karger, Mark Ackerman, and Sanjoy Mahajan. 2012. Successful classroom deployment of a social document annotation system. In *Proceedings of the sigchi conf. on human factors in computing systems*. 1883–1892.

A EFFECT OF CHANGE TYPE ON USERS' PREFERENCE OF HEADLINES USING CUMULATIVE LINK MODELING

We examined the effects of each change type on headline preferences of the participants in the second study using the same formula 1, but with a cumulative link mixed model instead of a linear model. Cumulative link models are appropriate for fitting ordinal values and find the cumulative probability of the ith rating (datapoint) falling in the jth category or below. The categories in our dataset are ordered preferences *"original"*, *"neither"*, and *"alternative"* headline. The cumulative link model assumes that there is a continuous but unobservable variable Y_i with a mean that depends on the predictors and that this underlying distribution has a set of cut-points $\theta_1, \theta_2, ..., \theta_j$ where if $\theta_k < Y_i < \theta_{k+1}$, the manifest response (headline preference decision) will take the value k.

To fit the model, we used the function "clmm" with a "logit" link from the package "ordinal" in R and set the threshold as symmetric. Table 4 shows the result of the ordinal regression. Similar to the result from the linear model counterpart, we observed that users' likelihood of preferring an alternative headline over its original counterpart increased when the alternative sought to modify the original by removing bias, correcting deceptions, adding concealed information pieces, making the topic known, and adding nuanced information or context. On the contrary, the effect of adding bias or sensationalism or copy editing were not statistically significant at the $\alpha = 0.05$ level.