NewsGuesser: Using Curiosity to Reduce Selective Exposure

LU SUN, University of California San Diego, USA
HENGYUAN ZHANG, University of California San Diego, USA
ENZE LIU, University of California San Diego, USA
MINGYANG LIU, University of California San Diego, USA
KRISTEN VACCARO, University of California San Diego, USA

Selective exposure has long been a concern of HCI researchers as it can lead to ideological polarization and distrust in society. Efforts have tried to reduce selective exposure online by serving diversified news content, but their effectiveness has been limited by users' lack of motivation to engage with the diverse content offered. To address this, we design the NewsGuesser system, which leverages the insight that curiosity can prompt motivation and engagement, by asking readers to guess the source of their news. In interviews with 40 participants, balanced for partisan affiliation, we use NewsGuesser as a probe tool to explore how guessing affects their perceptions of selective exposure. Participants struggled with the guessing game, which revealed a misalignment between users' expectations of different news sources and reality. Faced with the visualizations of the (often inaccurate) guessing results, participants were able to reflect on their own biases and selective exposure. In a number of cases, the guessing process changed participants' impressions of news organizations and some expressed an interest in engaging with more diverse news sources. While many also found the guessing game frustrating, the system and interview results suggest a number of new directions for designing social media and news media platforms.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**.

Additional Key Words and Phrases: filter bubble, selective exposure, gamification

ACM Reference Format:
Lu Sun, Hengyuan Zhang, Enze Liu, Mingyang Liu, and Kristen Vaccaro. 2024. NewsGuesser: Using Curiosity to Reduce Selective Exposure. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 99 (April 2024), 22 pages.
https://doi.org/10.1145/3637376

1 INTRODUCTION

Online news platforms have provided a rich and diverse source of information to society. However, they have also facilitated selective exposure to information — a tendency to favor information that supports pre-existing views [23, 32]. Individuals tend to click on ideologically consistent content more than cross-cutting content on their news feeds [4]. Such a phenomenon can potentially lead to ideological polarization, and consequently, damage the diversity of perspectives [4, 75]. There have been long standing concerns about filter bubbles and echo chambers, for example, that in the case of political news consumption, readers might never see contrasting viewpoints, strengthening their existing perspectives and biases [58].

Authors' addresses: Lu Sun, University of California San Diego, San Diego, CA, USA; Hengyuan Zhang, University of California San Diego, San Diego, CA, USA; Enze Liu, University of California San Diego, San Diego, CA, USA; Mingyang Liu, University of California San Diego, San Diego, CA, USA; Kristen Vaccaro, University of California San Diego, San Diego, CA, USA.

ACM 2573-0142/2024/4-ART99
https://doi.org/10.1145/3637376

Although a number of efforts have made diverse content more readily available, such as Ground News and AllSides[1], readers can still lack the motivation and willingness to engage with diverse content. Readers' choice and preference play a vital role in addressing this issue. Readers tend to select news from the sources that they are familiar with [4, 26] and from like-minded partisan sources [4, 21, 26, 75]. News consumers also engaged more deeply with the news content aligned with their own leaning[2] where they spent substantially longer time on these news than their visits to other news sources [26]. Recognizing these challenges, researchers have made a number of attempts to reduce selective exposure, using strategies like visualizing comparisons and providing additional metadata [17, 24, 28, 46, 59], pushing opposing content [18, 25, 55], and encouraging engagement with different points of view [38, 39, 79]. These approaches have had some success, but still face challenges around users' motivation to deeply engage with news they disagree with.

To address this, systems need to go beyond availability and increase readers' motivation to engage with cross-cutting content. In this work, we increase readers' intrinsic motivation, by encouraging their curiosity. Curiosity has been shown to prompt internal motivation and engagement [6, 7, 35, 43, 48] and has been used specifically in learning contexts to to motivate learners to engage with materials [6]. In this project, we attempt to improve readers' motivation and willingness to engage with cross-cutting content by triggering their curiosity using a guessing game. Having students make predictions (and engage in guessing) has been studied extensively in cognitive psychology [12, 50]. Making predictions and guessing has been found to enhance motivation and promote learning [19, 22, 53, 72], especially when learners find the answer surprising [11, 12, 33].

Drawing on this, we developed and evaluated a novel application, NewsGuesser, that gamifies the process of reading news. NewsGuesser asks readers to guess the leaning of the news source after reading a news article (Figure 1). Our aim is to trigger readers' curiosity while reading news and motivate them to engage with diverse content – and when they are surprised by the true leaning of news sources, to help them reflect and learn about media bias and selective exposure. NewsGuesser selects nine news articles each day. Following prior work [4], we focused on hard news. Articles were selected to be popular, distinct from each other, drawn from well known media organizations, and to span the political spectrum (equally sampled from left, center, and right leaning sources). NewsGuesser asks readers to guess the source leaning on a five-point scale, from far left to far right, and provides immediate feedback. The tool also presents a summary of readers' guessing accuracy when the game ends, drawing on visualization-based approaches to raise awareness of media bias [54]. NewsGuesser's visualizations showed three main types of information: accuracy, guess leaning, and time spent, to help users reflect on how they focused their attention, what their expectations were, as well as whether their assumptions may have been incorrect.

Using NewsGuesser as a probe tool, we conducted an interview study with 40 participants to explore readers' experiences with the guessing process, reflective visualizations, and selective exposure. The interview is designed to prompt Schön's reflection-in-action [71], a way of surfacing and criticizing tacit knowledge, including the influence of situational backtalk, i.e., any surprising results in the guessing game. Our qualitative analysis addressed the following research questions:
RQ1: How does guessing and reflecting with NewsGuesser influence readers' perceptions of their personal biases in news reading?
RQ2: How does guessing and reflecting with NewsGuesser influence readers' willingness to engage with diverse content?

[1]https://ground.news/ and https://www.allsides.com/
[2]"Leaning" is defined as the estimate of an individual's (or news outlet's) political ideological alignment with either a conservative or a liberal audience [66].

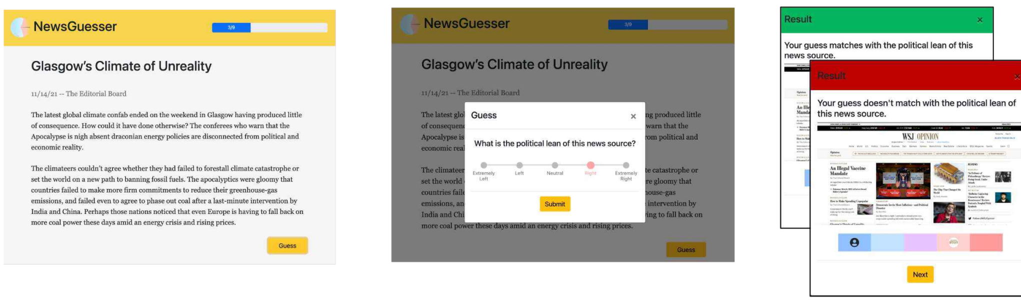

Fig. 1. NewsGuesser Guessing Game. NewsGuesser shows an article with the source omitted (left), the guessing game on a five point scale (center), and immediate feedback on the source's actual partisan lean (right).

Participants found the guessing game hard and guessed less than a third of source leanings correctly on average. Participants used their understanding of news biases to guess the sources, but their inaccurate results showed a misalignment between users' expectations of news bias versus reality. In many cases, participants were surprised by how neutral and fact-based the news articles from cross-cutting sources were. Participants commented that the guessing process helped them reflect on their impressions of some news sources and expressed an openness to engaging with more diverse – or even opposing – sources. While this is a one-shot study that does not measure the durability of this effect or changes in behavior, it suggests that surprise may be a useful mechanism for reducing selective exposure. We discuss further design implications of using curiosity and surprise to help readers engage with diverse news and reduce selective exposure.

2 BACKGROUND AND RELATED WORK

2.1 Selective exposure on news sources

While people have access to diverse information online, this does not guarantee they will hold diverse views [32]. Existing research has found that users have selective exposure to information, where individuals prefer to expose themselves to information that reinforces their existing attitudes and interests [23]. As personalized web recommendation algorithms try to engage users online, the information that users read only agrees with their views and is separate from the opposite [58]. This process can gradually narrow the sources from that people gather information and might lead to increasing ideological polarization [4, 32].

When users selectively pick news to read, news content organizations become one selection standard. The news sources that users selected to read inevitably reflect their leaning and selective exposure [4, 41, 42]. Lahoti et al. used user preferences and their news consumption to jointly estimate the users and news domain leaning on Twitter [42]. Prior studies also found evidence that online news consumers have partisan selectivity in news sources and pursue like-minded partisan sources for political information [26]. In addition, users engaged more deeply with news content from the same partisan leaning and their behavior shows polarized browsing patterns [26, 61].

One fundamental explanation of selective exposure is the cognitive dissonance theory. To reduce discomfort with holding conflicting beliefs, people are motivated to expose themselves to belief consistent information while rejecting inconsistent information, explaining away or avoiding new information. A big challenge here is how to raise users' motivation to engage with diverse content, especially from the opposite sources.

2.2 Systems that nudge readers to engage with diverse content

Many researchers have developed applications to raise readers' awareness of the filter bubble and nudge readers to engage with diverse content, using strategies like visualizing comparisons and providing additional metadata [17, 24, 28, 46, 59], pushing opposing content [18, 25, 55], and encouraging engagement with different points of view [38, 39, 79].

Researchers have developed applications that nudge people towards more diverse perspectives by presenting holistic views of information, often using visualization. For example, NewsCube highlighted different viewpoints [59] and Balancer tracked users' reading activities and showed their reading behavior and biases to increase awareness [54]. Others provided metadata to help users make sense of opposing viewpoints. Gao et al. developed an interface that annotates posts with others' reactions, which was able to motivate users to explore both sides of an issue [24]. Opinion Space automatically highlighted the comments found most insightful by users from a range of perspectives [17], Social Mirror showed users information about their Twitter network to mitigate political echo chambers [28], and Liao et al. found that that position indicators on controversial topics can encourage users to seek common ground [46]. Some have gone beyond nudging to push opposing viewpoints. For example, Spkr, a smart home device, unpredictably "pushed" socio-political discussion topics, with a purposefully assorted range of viewpoints, into users' homes [18]. It presented a novel means of engaging those who would not often be involved (i.e., other household members) in political discussions. Other researchers have developed recommender systems to expose opposing points of view [25] or to prioritize challenging news items [55].

A number of projects in the past sought to encourage activity by users to improve their openness to alternative points of view. For example, Kriplean et al. encouraged listening and perspective taking by enabling readers to restate key points made by others [40]. A similar approach encouraged users to articulate pros and cons of their policy positions, to guide reflections on tradeoffs and better understand opposing points of view [38]. However, these approaches focused on reducing skimming or misrepresenting opposing sides, rather than increasing motivation and engagement. More recently, Wang et al. used similar strategies of annotation and discussion, along with recommendations, but added moral framing to make cross-cutting content more appealing to users [79]. Finally, recent work has explored the possibility of inoculation as a strategy – that is, making users aware of the risk of echo chambers (and how they make be constructed) can help prevent them from falling victim in the future. This effort took a gamified approach, where users actually attempt to create an echo chamber in a game [34]. In our work, we also use gamification, but with an aim of increasing internal motivation to engage with opposing content.

Most existing systems attract readers to diverse information by presenting information in a holistic way, however, readers might still lack motivation to even engage with the information from an opposing view. To trigger this intrinsic motivation, we leverage curiosity. And while some of this prior work has prompted users' curiosity to encourage exploration [24], in this work we direct users' curiosity towards media bias and selective exposure itself.

2.3 Curiosity, Guessing, and Surprise

Curiosity arises due to a gap between what one knows and what one wants to know [48] and can promote user engagement, feelings of excitement and even self-awareness while interacting with new information [7, 35, 43, 48, 49]. Existing research has examined curiosity as an intrinsic motivational driver to incentivize crowd workers to finish tasks [43]. Our system creates an information gap by removing the news sources from articles and inviting readers to guess. We hypothesized that this guessing process can trigger readers' curiosity and motivate them to pay attention to diverse content.

Fig. 2. An overview of NewsGuesser's workflow. (1) For each news event, the system collects articles from each political leaning (left, center, and right), which are saved to a server. (2) The NewsGuesser interface displays these news articles to users as part of the guessing game. (3) Users are then are shown visualizations of their performance.

Guessing is beneficial because it can spark interest, evoke curiosity which enhances motivation, and boost learning [12, 33, 37]. When people make a prediction, it requires accessing prior knowledge and connecting to the new information. This retrieval and reflection process can not only promote learning but also stimulate curiosity when seeking the right answer.

When people guess the answer and nearly all answers are incorrect, the error can spark considerable interest in seeking information [12, 22, 53]. Previous research explained that when people found that their expectations were violated in the guessing or they cannot explain the results, people are willing to make greater efforts to figure out the explanations and recall prior knowledge, especially for the errors with high confidence [22, 50]. In the context of news reading, we hypothesized that asking people to guess the sources of the news can potentially trigger their willingness to explain their results and curiosity about the actual partisan leaning of each news source. When users make errors with high confidence, it may further motivate them to search for explanations and further reflect on their previous news selection behaviors. These cognitive processes may help readers be more open to learning and reduce selective exposure.

3 NEWSGUESSER: A SYSTEM TO ENGAGE WITH DIVERSE NEWS

The purpose of the NewsGuesser system is to trigger the reader's curiosity, by having them guess the partisan lean of the news source (Figure 1). The user reads a series of news articles, where information about the news organization that published the article has been removed. After reading each article, the user guesses the political leaning of the news source. After each guess, the system displays the correct answer with the real news organizations that published the news. When the user is finished reading, NewsGuesser provides a series of visualizations that show information like accuracy, partisan lean, and time spent reading each of the articles. Figure 2 gives an overview of NewsGuesser's overall system design, which includes an automated pipeline for selecting news articles (Sec. 3.2), a guessing interface for users (Sec. 3.3), and visualizations of the users' performance (Sec. 3.4).

3.1 Design Goals

Drawing on prior work, we identified three primary design goals for the NewsGuesser system: 1) providing news comparisons, 2) triggering curiosity through guessing, and 3) reducing cognitive biases through visualization and reflection.

3.1.1 Provide news comparisons. To reduce selective exposure, where users consume information aligned with their own predispositions, we must provide diverse or unaligned news. However, NewsGuesser not only shows unaligned news, it provides a comparison across news organization leaning. Prior work has shown that working through a comparison of different cases that share common underlying principles can be illuminating and promote learning [27]. In this case, comparing different news on the same events can prompt users to seek clues about different stances.

3.1.2 Trigger curiosity. Simply presenting different views side-by-side will not help users who have selective exposure, as they are still more likely to seek attitude-consistent information [45]. However, a reader's involvement in the topic is a critical moderator for the information-seeking process [62], and users reading about high-involvement topics extensively seek information, achieving relatively balanced exposure [45]. Based on its performance in prior work, we sought to increase engagement and trigger curiosity using a guessing game [22]. Initial designs explored the possibility of guessing other aspects of news articles: titles, authors, images, or captions. For example, prior work has shown that images accompanying news text can play an influential role in how people understand that content [64]. So initial designs considered the possibility of asking users to guess which image accompanied which news article. Our initial explorations suggested that news organizations were both mostly closely related to media bias and at a level of granularity that could be recognized and successfully guessed.

3.1.3 Reduce cognitive biases. An important factor in selective exposure is confirmation bias, where users prefer attitude-consistent information [36]. A similar effect of confirmation bias might also occur in users' recollection of their guessing performance. Thus we include a final stage to review and reflect on the guessing results, with the goal of focusing users' attention on all results, not only the ones that align with their existing beliefs. There are a number of strategies used in debiasing training that can reduce the effect of confirmation bias, including decision trees, practice identifying a selection of evidence, and careful testing of alternative hypotheses. These strategies are intensive, so we adopt a lighter-weight strategy: providing visualizations. Prior work has shown that visualization can reduce cognitive biases, such as anchoring and confirmation bias [69].

3.2 Selecting news articles

The news articles are selected to be recent (updated daily), popular (covered by a large number of organizations), spanning the political spectrum (equally sampled from left, center, and right leaning sources), and maximize the difference in article content. All news is collected from Ground News, an independent news aggregator designed to help readers compare news sources.

3.2.1 Selecting News Events. To maintain consistency, we choose three news *topics*. Every user sees stories for the same three topics: technology[3], politics, and COVID-19. While the news is selected daily, using the same topics offered a control for the kind of content read across users. These topics were selected to be frequently covered, with the largest number of news articles published during a three-month period prior to the study launch. For example, during this period, we found *Politics* had approximately 40,000 articles, while *Education* (not selected) had approximately 3,000. Following prior work [4], we focused on hard rather than soft content, although soft content topics (e.g., *Arts & Entertainment*, *Lifestyle*) were generally less common topics (with the exception of *Sports*, which was 4th). Selecting three categories allowed NewsGuesser to cover a breadth of topics while balancing reader fatigue. While many consumers are most interested in local news [13, 76], ensuring comparable content across readers meant we could not personalize with local news.

Each day, we collected all of the news events published within these three topics. For example, an event in the technology topic might be the European Union approving regulations restricting online

[3]This is officially "Science & Tech" but is largely technology content, with over two thirds of the articles technology-focused. For brevity, we refer to it as the technology category.

ad targeting. We selected the event that was: 1) covered by the most news sources and 2) covered by at least three left-leaning, three center-leaning, and three right-leaning news sources. News was selected daily, because most people report active news consumption: two-thirds of adults look at news several times a day or more [67]. And this is not just scanning headlines: 70% report closely reading the details of a story at least once a day [67]. Selecting news stories daily both matched typical news reading behaviors and minimized the likelihood that users would have previously encountered the selected articles.

3.2.2 Selecting News Articles. For each of the three news events of the day, NewsGuesser selects three news articles: one news story from a left-leaning, one from a center-leaning, and one from a right-leaning news organization. While NewsGuesser selects articles with distinct partisan lean, users are not informed of the process, to avoid influencing their guessing strategies.

To increase the likelihood of users recognizing the news organizations after their guess, we select news stories from more well-known news organizations. There are thousands of news organizations in the United States alone, one recent estimate places the number around 3000 [44]. News sites are ranked according to how frequently they are visited, measured using Alexa[4]. For each event, NewsGuesser considers the five most highly ranked (i.e., most visited as measured by Alexa) news sources for each partisan lean. This produces a set of 15 candidate articles for each news event.

To help users compare signals of leaning more directly, NewsGuesser shows three news stories about the same event, one from each partisan leaning. There is a risk, however, that users will become bored reading similar content three times. This is particularly true as many news organizations draw content from wire services (e.g., the Associated Press), which "*provide the basic news diet of most newspapers and broadcast stations*" [14]. To avoid this, NewsGuesser maximizes the difference in article content. To do so, it computes the cosine similarity between news articles. We calculate the similarity for all possible left/center/right triplets. The trio of news articles that maximize the difference are selected, with the goal of giving users the most diverse information possible on a given topic. To calculate cosine similarity between each pair, we convert the news content into a matrix of tf–idf features with the length as the total number of unique words in this pair of news and then calculate the similarity as the normalized dot product of them. After NewsGuesser selects the nine news stories, a researcher reviews each and manually removes any content that explicitly mentions the news organization.

3.2.3 Identifying News Organizations' Political Leaning. Political leaning is identified on a per-news organization basis, as in prior work [16, 30]. Prior work has shown that while the news industry as a whole does not appear to be biased, specific newspapers can show substantial bias, both to the left (e.g., Great Falls Tribune) and right (e.g., Indianapolis Star) [16]. For NewsGuesser, the political leaning is drawn from Ground News. We used Ground News ratings to label each news organization's partisan lean since their ratings are based on a collection of third-party independent organizations dedicated to monitoring and rating news publishers along the political spectrum. Those independent organizations include Media Bias Fact Check, AllSides, and Ad Fontes Media [56].

Each news source is assigned a single, fixed political leaning label. For example, Ground News labels CNN as left-leaning, the Associated Press as center, and Fox News as right leaning. While this may mask variations due to individual journalists or editors, prior work has found that these leaning labels are generally consistent with the organizations' published stories [57].

Finally, media bias is typically composed of three main issues: *selection bias* (which stories are covered), *coverage bias* (how much attention stories get), and *statement* or *presentation bias* (the

[4]https://www.alexa.com/ Alexa has since been discontinued, but was operational at the time of our study. Alexa was a popular metric for web traffic, reach, and popularity used for research and competitive analysis.

content of the story itself)[5]. NewsGuesser focuses only on presentation bias, which has been the focus of the vast majority of the media bias literature [30], asking users to reflect on the differences between how three stories are presented from different partisan leaning organizations. Future tools might explore ways to highlight or visualize differences in selection and coverage bias as well.

3.3 Guessing interface

Users read nine news articles (Figure 1a). After each news article, we invite readers to guess the political leaning of the news source (Figure 1b). We provide five selections: far left, left, center, right, far right. After users submit their guess, NewsGuesser shows how far their guess was from the correct answer, and the correct news organizations with their front page (Figure 1c).

Prior work has suggested that people decide which source to read based on political leaning [4]. One way to reduce this selective exposure is by removing the source. We go further by having users guess the source, to increase engagement and learning about media bias. Other guessing attributes are possible (e.g., guessing the headline, guessing the accompanying photo, etc.), but source leaning is the most closely related to our effort to reduce selective exposure.

The structure of the guessing game was partially inspired by GeoGuessr, a popular online game which places the player in a random location of Google Street View. The player then guesses where in the world they are. Similarly, the user of NewsGuesser tries to identify the partisan landscape of their reading. As researchers have noted, this process can be disorienting: "*The game quickly turns out to be much harder than it looks*" [20], particularly in trying to decide what to pay attention to or "*parse the relevant details from the irrelevant ones*" [20]. While Geoguessr also includes substantial competitive elements (e.g., points are exponential), NewsGuesser takes a more reflective approach. While users are given immediate feedback about their accuracy after each guess, the summary visualizations at the end are geared more towards reflection than optimizing performance.

3.4 Visualizations

The visualizations present a number of different views, which aim to help a user reflect on their expectations, biases, and potential selective exposure as they read news (Figure 3). As in prior work [54], our tool provides the visualization as a final stage. Studies on this prior work showed that visualizing feedback to news readers led to a modest move toward balanced exposure. Our tool used the same visualization stimuli to help people reflect on their guessing results. These visualizations aimed to show three main types of information: accuracy, guess leaning, and time spent. We prioritized these types of information because they would help a user reflect on how they focused their attention, what their expectations were, as well as whether their assumptions may have been incorrect. As discussed in Section 3.3, we aimed to prioritize reflection over competition, so we did not show comparative performance.

3.4.1 Overall Accuracy. The first visualization display's the overall accuracy of the guessing – the percentage of correct and incorrect guesses in a pie chart. This visualization was developed after initial pilot tests found that users wanted an overall summary of their performance at the guessing game.

3.4.2 Overall Leaning. The second visualization presents a seesaw visual, indicating the direction in which users' guesses were most often wrong. If a user's guesses were equally likely to be too far left and right, the seesaw is depicted as balanced. This visualization was designed to give readers a quick visual summary of which way their guesses leaned.

[5] There are also a number of more narrowly defined biases that contribute to these (e.g., concision bias, mainstream bias, incumbency bias, etc.)

3.4.3 Detailed Leaning. The third visualization provided a more detailed view of which way the participant leaned on their guesses. A histogram displayed how many times a guess was correct, off by one (lean left, lean right), or off by two or more (far left, far right).

3.4.4 Time Spent. The fourth visualization showed how long the participant spent reading the articles from the left, center, and right leaning organizations. Pilot tests also indicated that users typically spent the most time reading the first article of each topic, so articles were ordered so that readers saw a different partisan leaning (left, right, center) first for each of the three topics. If users were more engaged in articles from left or right leaning sources (i.e., they spent more time reading those articles), they could observe this in this visualization.

3.4.5 Accuracy of Individual Articles By Topic. The last visualization grouped the articles by topic, and again showed the each guess on the range of possible partisan leanings along with the ground truth (the partisan lean of the news source according to Ground News). This allowed users to directly compare their guesses to the ground truth for each article within a topic.

3.5 Implementation

The scripts to automate the collection of news were developed in Python, leveraging existing packages including news-please (an open-source Python library for news parsing [29]) and BeautifulSoup (a Python library for web scraping [73]). The NewsGuesser interface was developed using Node.js with Javascript. All data (including news articles and logs of user behavior) are stored in Google Firebase.[6]

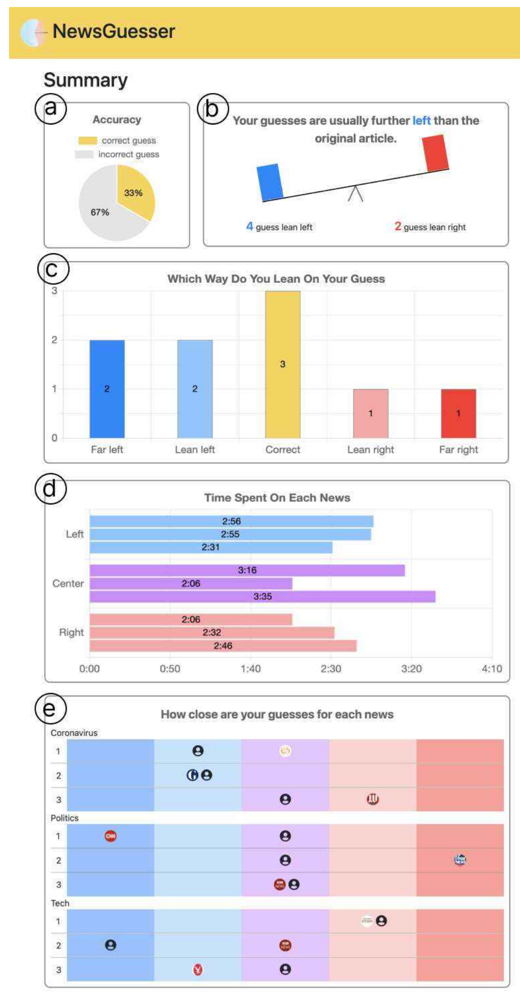

Fig. 3. Visualizations interface. Five visualizations capture various aspects of guessing performance: (a) guessing accuracy, (b) seesaw representation of overall leaning, (c) detailed leaning, (d) time spent on articles grouped by partisan lean, (e) accuracy of guesses for each article by topic

4 METHOD

A qualitative interview study, using NewsGuesser as a probe tool, was used to investigate our research questions. The interviews asked participants to reflect on their experiences while guessing and reviewing the visualizations. This approach leverages Schön's reflection-in-action, through which people often "*reveal a capacity for reflection on their intuitive knowing in the midst of action*" [71]. Each interview took place in a single, hour-long session (M = 58 min). Participants were paid $15/hour for their participation.

[6]Code and scripts: https://github.com/LusunHCI/filterbubble.git

All interviews were conducted online via Zoom. Interviews were conducted over two weeks in March–April 2022. In addition, a short demographics survey is hosted on Qualtrics but embedded within the NewsGuesser interface, accessed after the visualizations. The protocol was reviewed by our institutional IRB.

4.1 Participants

Participants were recruited through an online recruitment system, Prolific, and balanced for political leaning. Prolific provides five political spectrum options: *Conservative*, *Moderate*, *Liberal*, *Other*, *N/A*. We combine *Other* and *N/A* and treat this group as *Independent*. Ten participants are recruited from each group. We also report relevant background information on our participants.

4.1.1 Demographics. Participants roughly match the US demographics in most aspects: age (M=44, range = 22–76), gender (45% men, 55% women, 0% nonbinary), and household income (Median = $50,000 to $74,999). But participants are more White (82%) and educated (68% have a Bachelors degree or more) than would be representative. Full demographic details are included in the supplementary materials.

4.1.2 News reading habits. Among 40 participants, 34 reported that they read news daily, 4 weekly, and 2 monthly. Around 67% of Americans consume news daily [67], suggesting that our participants (85% consume news daily) may skew towards heavy news consumers. Participants reported reading news from diverse news organizations. Seven reported that they read news through news aggregators (e.g., Apple News, Google News, Twitter news feed, etc.), 12 reported that they read news from individual sources (e.g., Fox News, CNN, BBC), and 21 participants used both.

4.2 Study Procedure

The study, using NewsGuesser as a probe tool, consists of three primary phases: 1) the guessing game itself, 2) reflection on the performance visualizations, and 3) closing questions. The full protocol is included in the supplementary materials.

4.2.1 Guessing Game. Participants began by starting the NewsGuesser system. After reviewing the consent form, participants read nine news articles selected using the criteria described in Section 3.2.2. Participants are required to read each news article for at least two minutes. After each article, the participant guesses the partisan leaning of the news organization that published it. The participant is shown the accuracy of their guess (Figure 1).

Every three articles, the interviewer asks a series of questions about the guessing experience. This is designed to minimize interruptions, while still observing any learning effects or changes to the users' experience over the course of the guessing game. These questions included how the user planned to guess, what aspects of the article they paid attention to in making their guess, and how they reacted to the correct answer. For example, *What are you going to guess?*, *How did you make your guess?*, *How do you feel after seeing the result?* and *Does this change your impression of [the news source]?* The questions are designed to prompt reflection-in-action [71], a way of surfacing and criticizing tacit knowledge. A key insight is the way that a situation's back-talk may lead people beyond their initial perceptions: appreciating new aspects or reframing their perspective entirely. In this case, we use both the NewsGuesser system and interview questions to highlight back-talk of inaccurate guesses and prompt this kind of reflection.

4.2.2 Reflecting on Visualizations. Next participants view the visualizations of their guessing behavior and performance. These visualizations are detailed in Section 3.4. Questions aimed to help participants reflect on specific visualizations (e.g., *What do you notice comparing the time you spent on left, right, and center-leaning news?*) as well as thinking more generally about their beliefs and experiences, such as: *What are your takeaways or reflections on these visualizations?* and *Did the guessing process help you understand your impressions of any news sources?* In alignment with the goals of the visualization itself, which aimed to reduce the likelihood of confirmation bias, these questions similarly probed both accurate and inaccurate guesses and left- and right- leaning articles, as well as understanding participants' overall experiences.

4.2.3 Closing Questions. The closing questions asked participants to evaluate their experience with the guessing game and tool, including *Were there parts of guessing the source that you [enjoyed/disliked]?*, *How did the experience of reading news with this tool differ from how you consume news in your daily life?* and *If you had access to this tool, would you use it to read news in the future?*

This segment also asked more general questions about their political leaning, news consumption, and perceptions of media bias and selective exposure. These included questions like *How often do you read or watch or listen to the news?*, *How do you feel about news sources that don't align with your own political leaning?*, and *How do you think people should choose the news to read?* These questions, which were often most directly related to our research questions, were left to the end of the interview to avoid biasing participants' experiences with or expectations of the tool. After the closing questions, participants were invited to share any additional thoughts or suggestions. Finally, participants completed a brief demographic survey.

4.3 NewsGuesser Instrumentation and Measures

In addition to the answers to interview questions, we also recorded a limited set of data from the NewsGuesser interface itself. For each guess, we record: the accuracy, the leaning relative to the source leaning for incorrect guesses (far right, right, left, far left), and the time spent reading the article before guessing.

Over two weeks the studies were conducted, NewsGuesser selected 45 articles from 22 news organizations. On average, each article was read by 10 participants (range = 4—14). The average length of the articles shown to all participants is 658 words, and on average participants spend 3 minutes 37 seconds reading each article (range = 2:06−9:57).

4.4 Analysis

All interviews were recorded and transcribed. Three researchers conducted iterative coding on the transcripts using Dovetail[7]. Two researchers first segmented the interviews guided by the two research questions. This top-down approach resulted in sections focused on users' reactions to the guessing process, their reflections on the visualizations, and their closing thoughts on selective exposure [9]. From these sections of the interviews, three researchers shifted to an inductive, semantic, thematic analysis [10]. The segment of analysis was a sentence; multiple codes could be applied. Each researcher developed initial codes independently, and then through multiple iterations along with periodic discussions, these codes were consolidated. From these initial codes, discussions among all researchers identified three themes, discussed in Sections 5.1.3, 5.2.1, and 5.2.2. A final analysis sought to understand the user experience of NewsGuesser, with a particular focus on any negative experiences of or recommendations for the tool. This again used an inductive approach, albeit with a narrow focus, discussed in Section 5.3.

We also conducted a quantitative analysis of users' log data from the guessing game. For each guess, we calculated the distance between the partisan lean of the participants' guess with the ground truth, using a five point scale. If the lean is the same, we consider the guess correct. For an incorrect guess, if the absolute distance between the guessed and true lean is less than 2, we categorize the guessing result as lean left or lean right. For answers with an absolute distance larger

[7]https://dovetailapp.com/

than 2, we categorize the answer as far left or far right. We report the number of guesses with each leaning for participants in each of the four partisan affiliations (Conservative, Moderate, Liberal, Independent). We also report overall accuracy, in the mean number of correct guesses.

5 RESULTS

NewsGuesser removed the news source and let readers guess the political leaning of every article they read. The motivation was to trigger readers' curiosity while reading news and motivate them to deeply engage with cross-cutting content. We found that readers used their understanding of news biases to guess the sources in this game. However, the guessing results showed a misalignment between users' expectations of different news sources and reality. Faced with the visualizations of the (often inaccurate) guessing results, participants were able to reflect on their biases and selective exposure. Some participants noted the guessing process helped them reflect on their impressions of some news sources and expressed an openness to engaging with more diverse sources. While this is a one-shot study that does not measure the durability of this effect or changes in behavior, it suggests that surprise may be a useful mechanism for reducing selective exposure. At the same time, participants felt the guessing game was challenging, which led to frustration among some.

5.1 Guessing performance

5.1.1 Overall accuracy. The guessing game was very difficult. Only two (out of 40) participants guessed half of the source leanings correctly. And on average, participants had 2.4 correct guesses. As they discussed their guessing strategy, most people (n = 24) found guessing challenging, with experiences ranging from mild "*it was definitely difficult to try to guess*" [P27] to frustrating "*I didn't enjoy this because I got so much wrong*" [P30], which we discuss further in Section 5.3.

5.1.2 Guessing skews. We additionally separate the participants into the four partisan affiliations we use for balancing (*Conservative*, *Moderate*, *Liberal*, *Independent*) and investigate whether the guessing leans in similar directions for all groups. We find that every group makes guesses that are further left than the actual source leaning (Figure 4), though this effect was most extreme

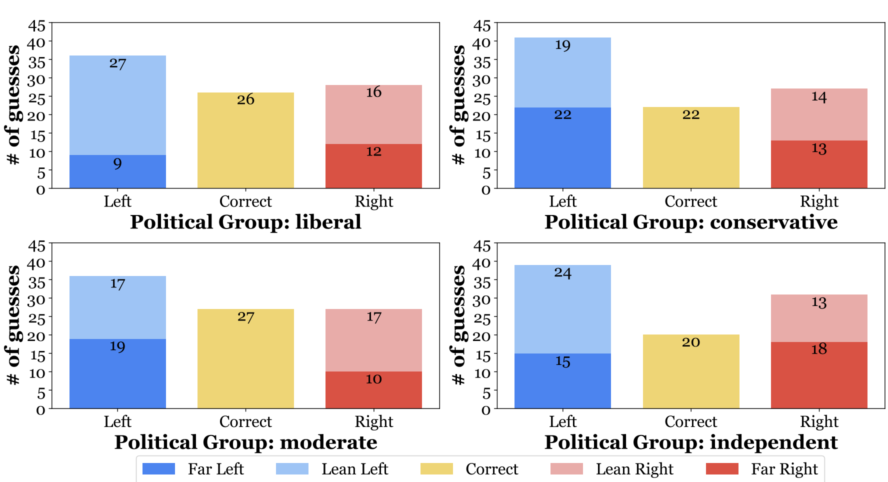

Fig. 4. Count of the number of guesses that are left or far left of the actual source leaning (blue), correct (green), or right or far right of the actual source leaning (red). Clockwise from top left: results for Liberal, Conservative, Independent, Moderate participants.

among Conservative participants. Two participants pointed out the culture of news leaning left. One participant said that, "*I guess left most of the time, because most of the positions on these people in the news are liberal.*" [P8]. It is possible that this was an unspoken assumption among other participants and related to the overall left shifting of the guessing results.

5.1.3 Surprise after seeing guessing results. Most participants explicitly mentioned that they felt surprised after seeing the guessing results (n = 31). Often this was after finding that news sources were more factual and less biased than they expected. Participants found differences between left, center, and right-leaning articles difficult to identify, "*all of these new sources, like the information was generally the same, like the, the actual hard facts, but the presentation was the part that was a little bit different for each of them. They were very similar.*" [P16]. This caused more surprise when the news was from far left and far right-leaning sources, such as Fox News or CNN (n = 23). For example, a left-leaning reader reflected that they expected Fox News to be "*really conservative and kind of nutty. And this article that I just read was actually calm and normal*" [P15].

Nevertheless, participants noted several ways that they expected bias to appear. First, many expected exaggeration, emotional appeals, or judgement in the text of the news itself (n = 11), "*I would think that there would be criticisms in there about either some of the Democrats that she worked for or for Madeline Albright herself. And I really didn't see anything. So I am completely surprised*" [P11]. Second, participants expected journalists to intentionally omit quotes in strategic ways (n = 5), "*I was surprised that, you know, the mention of the endorsement by President Clinton, for example, was included. I would think that that would just be something they would leave out*" [P13]. Finally, some participants expected left- or right-leaning news organizations to lie about factual information (n = 5), and expressed surprise when they did not, "*And you know, the facts were correct. And normally Fox News, well, in my experience, they misstate the truth*" [P15]. In all, participants had sophisticated understandings of how presentation bias could emerge in news (e.g., source selection), but were surprised when these biases were either absent or less noticeable than they expected.

In addition, a number of participants reflected on their own biases, expectations, and knowledge, after guessing incorrectly. Many participants remarked they had incorrect assumptions about news sources (n = 12). For example, one participant said, "*I assumed things that were objective were automatically left-wing because I assumed things that were not objective would be right-wing*" and noted "*I was wrong*" [P4]. Several others mentioned that their political affiliation influenced how they interpreted the news (n = 7). As one participant put it, "*I kept choosing right for more negative context in my mind*" [P20], simply because she associated right-leaning with negative things. Finally, some participants mentioned that their own bias in their typical reading behavior left them unprepared for what to expect from cross-cutting sources. One participant described this, "*I'm more familiar with a left leaning news sources than right ones. So select Breitbart, for example, I really don't read it all.* " [P13]. This meant that, despite our efforts to prioritize the most frequently visited news organizations, some participants were unfamiliar with them and weren't sure what it would mean to guess that organization.

5.2 Reflecting on selective exposure

5.2.1 Current selective exposure. Over the course of the interviews, participants reflected at length about selective exposure. Many participants noted that their current practices around news reading are influenced by selective exposure. Many participants said they currently read sources that are aligned with their political affiliation (n = 14) or avoid reading opposing sources (n = 14). For example, one participant noted that "*most of the news I read is all right-leaning just because that's what I agree with*" [P5]. Only four participants reported currently reading diverse sources, though 13 participants said they avoid extreme sources and stick with center-leaning sources, as with one participant who said, "*I do think of myself as moderate [and] I try to read things that are primarily fact-based*" [P6].

Participants describe themselves as avoiding cross-cutting content for a number of reasons. The three most common are that they believe the information included is inaccurate (n = 8), misrepresentations are strategic to further the goals of the news organization (n = 10), and the disagreements over facts and framing make the reader so angry they cannot continue to read it (n = 8). As one participant described this, "*I become very angry because a lot of it is incomprehensibly stupid. To the point that they're using obvious logical fallacies. And it causes me deep anger to know that there are people who actually believe lies and it, and it, it, it, it saddens me and angers me*" [P3].

Nevertheless, many participants saw the value of reading diverse news sources, Most notably, 19 participants mentioned that reading diverse news sources helps obtain a broader perspective on issues, and 10 of them specifically noted understanding the viewpoint of the other side. Other values include understanding the facts and getting more accurate information (n = 8) and even reducing their own biases (n = 3). For example, one participant said that this should inform your own ideas as they develop: "*I think it's important to get information and ideas from people who don't already completely agree with you at least a little bit. Otherwise I think that can lead to badly formed political ideas*" [P13]. And another noted it can help build empathy and community, "*if you can see where someone else is coming from and their point of view and understand [what] their concerns are or what they are scared of [...] having an understanding of someone else is never a bad thing.* " [P33]. So despite the challenges it creates, particularly around misinformation and personal frustration, participants saw the importance of exposing themselves to cross-cutting news sources.

5.2.2 Effects of NewsGuesser on selective exposure. As described in Section 5.1.2, we found that readers used their expectations to guess the sources, but the guessing results revealed a misalignment between users' expectations of different news sources and reality. Many readers, when they were surprised by the results, reflected on their expectations of and biases towards certain news sources (n = 19). For example, one participant reflected that "*some of these results definitely surprised me, which tells me that, you know, some of those biases that I have may not be correct. You know, maybe Fox's not that extreme*" [P10]. Another described their surprise after the source was revealed, "*I did not realize that it was Fox News because it didn't feel as blatantly in your face extremist as I expect them to be*" [P9]. Faced with their (often inaccurate) guessing results, many participants reflected on their own biases (n = 30). One participant mulled on this at length:

> *When I am reading the news and I see those different sites, there are times where I won't click on... If I was reading a headline and it had one of the other options as being from Fox News, I wouldn't read that one because I feel like it would lead to the right. And I personally would prefer to read neutral news articles. So I would skip that Fox News. But after doing this experiment, it kind of makes me feel like maybe I am unfairly biasing myself like that. Maybe I should read Fox News more [...] maybe I'm just biased* [P6].

Or as another put it, more succinctly, "*I like to learn new things. I may check out NPR and the Breitbart*" [P14]. Overall, participants shared that the guessing forced them to consider where they sought their news and why, "*Guessing has kind of opened my eyes a little bit, I guess about maybe not being quite so sure of what I'm reading*" [P11].

However, not all participants were as open to changing their perspectives. Several questioned the accuracy of NewsGuesser's labeling (n = 5), for example, "*BBC news is [listed as] neutral and it's kind of ridiculous*" [P25] or "*I don't agree that NPR is center, I don't trust any of these*" [P30]. This indicates that while curiosity and guessing may encourage some to reflect on and potentially reduce their selective exposure, it requires trust in the system and its labels.

5.3 Experience of NewsGuesser

At the end of the interview, most participants (n = 27) said they would be interested in at least occasional use of the tool in the future, while eight had a variety of reasons for preferring not to. In the interviews, participants also provided feedback about their likes and dislikes for our system. The most common reason for liking it was enjoying guessing and finding out the answers (n = 16), as one participant said, "*the fun aspect of like the fact that I did get a feedback at the end, so sort of made it fun, like spotting something*" [P7]. An equal number of participants said that they like getting useful information and evaluation from NewsGuesser (n = 7). Some participants liked the visualization that showed comparison (n = 5), "*I love the analysis at the end, because it shows you, do you have potential biases?* [P26]. Another popular reason was that they can get a range of news (n = 5), which is not only "*a variety of different sources to look at*" [P24] but also gave them a chance to compare. As a participant put it "*especially on the third one, you have a slightly better ideas as you've seen the other two and know where they are, then you can compare it to the last two*" [P1].

When reporting what they disliked, several participants mentioned that they got frustrated when they guessed incorrectly (n = 6). Guessing source leaning is very challenging. Participants felt especially frustrated if they failed many times or felt their beliefs were strongly challenged. One participant described this experience, "*when I was just so confident that it was one or the other and it wasn't, I was like, what is going on?*" [P20]. The other major reason participants disliked NewsGuesser arose from the design decision to present three articles on the same topic. Participants reported that there is duplicated information (n = 2) and no summary or comparison between articles (n = 2). And for some participants, reading closely related content led to boredom.

We also collected recommendations from the participants. Two participants were actually skeptical that the answer we showed was correct – or that the bias of a news source maps to all the articles it publishes. As one participant put it, "*I get the bias of the news source, but the individual story where in the individual story, does it show a bias?*" [P29]. This gap may also contribute to the difficulty of the game. Similarly related to the correct answer after each guess, another participant suggested that an explanation could be provided for each answer, and what about the article suggests that partisan leaning. Participants also mentioned that they would like to compare different news sources' viewpoints on the same topic to better understand issues. For example, one participant mentioned that "*It's awesome to have a left, center, and right view of the same topic. For instance, to know what's going on over in the Ukraine right now, I would like to know what the left says about it, what the right says about it and then what the neutral says*" [P17].

Participants suggested potential uses of this tool. One potential application is to support digital and media literacy education. Media literacy education intends to promote awareness of media influence, foster critical thinking skills and create an active stance on consuming media including online news [63]. This guessing game can potentially be used as an educational game for students to compare diverse news sources' stances. One participant mentioned that "*I would really like to use it with my students to help them see the power of words. And I think it would be good for lots of people to just see like perceptions of things and the power of phrasing of words and stuff*" [P35]. By reading news without sources on our app and guessing sources along the way, students might have a deeper understanding of news sources' stances and the functions of news media in society.

This guessing application can also be used for self-evaluation. In our interview, participants mentioned that they would like to reuse this tool periodically to evaluate their personal biases or use this tool when there are major news events. One mentioned that he would like to use this tool when he feels unsure about issues, "*the best thing about this is that I would use this to find out more about myself. Like let's say there was an election coming up and I really wanted to know how I felt about the issues...I probably would put articles through here*" [P36]. Similar tools could also be incorporated into news reading platforms, like Ground News or AllSides, which already help readers consider media bias. Readers could play this game to test their personal biases periodically.

5.4 Summary

Our study designed a new application to trigger readers' curiosity while reading news articles and raise their awareness of selective exposure by asking readers to guess the leaning of news sources. To answer the research question, we invited 40 participants from different political groups to use our application and conduct interviews as user evaluation. In the user interviews, we found that users were generally aware of the phenomenon of selective exposure and had reasons for avoiding reading from cross-cutting sources, such as triggering negative emotions. But we found that the guessing together with the visualizations helped readers to reflect on their personal biases and selective exposure. Some readers were surprised by how far their expectations of some news sources were from reality. Finding ways to surface the misalignment provides an efficient way to prompt readers to self-reflect on their news selection behaviors. Readers said that this process revealed and led them to rethink their biases towards some news sources and some stated that the experience opened their eyes to more diverse news sources.

6 DESIGN RECOMMENDATIONS

Research communities have made huge efforts to reduce news reader's selective exposure [17, 18, 24, 25, 28, 38, 39, 46, 55, 59, 79]. These efforts have used a variety of different approaches, including ones that attempt to engage users with opposing points of view [39]. In our study, we use curiosity and surprise to increase readers' awareness of selective exposure and willingness to engage with cross-cutting content, revealing their potential for the design of news reading applications.

6.1 Targeting The Right Difficulty Level

Our work revealed the potential for leveraging curiosity to reduce selective exposure, but participants found the guessing process hard. Particularly when they repeatedly guess incorrectly, frustration is common. Research in affective computing and learning science has found that this kind of frustration has the opposite effect of curiosity, where students have lower motivation, more negative emotional states and fewer learning gains [31, 70]. An ideal system should set the difficulty at the right level, where participants feel challenged enough but not too much, to trigger their intrinsic motivation [2, 3]. Prior work that has attempted to use gamification to train users on the risks of echo chambers faced similar challenges in establishing the right level of difficulty [34]. While our guessing mechanism is straightforward (simply presenting the news itself), we highlight several adaptations suggested by our findings that might make the guessing easier or more difficult.

6.1.1 Short, direct comparisons. NewsGuesser provided a seesaw visualization that showed the overall discrepancy between their guessing results and the actual news source leaning. Participants reported that visualizing that difference helped them better compare the news they read and prompted reflection. But they also expressed interest in direct comparisons between the news content. For example, instead of showing one news article at a time, tools could show similar snippets from left, center and right leaning at the same time. By reducing the overall cognitive demand (comparing only these shortened news snippets), readers might perceive the task as easier.

6.1.2 Annotations and hints. Alternatively, to make guessing easier, the application could provide iterative hints to readers that highlight some indicators or differences across news articles, a strategy that is often used in game design [60]. Researchers previously developed a tool, ConsiderIt, to surface the most salient pros and cons during policy discussions, to help readers compare and deliberate different points [38]. News reading applications could similarly show news from different sources with annotations of terms, sources, or facts used differentially. Whether they are generated through automation or through inviting collaboration, surfacing these differences can help readers engage more deeply with the news content and reasoning [5, 80].

6.1.3 Guessing about event focus, fact and opinion. In many cases, the surprise participants experienced was that sources were not as biased as they expected. However, as we discuss in Section 8, media bias can often take the form of the types of events that are covered. By matching news events, NewsGuesser cannot expose this aspect of organizational decision making. Future systems might invite guessing about which events are covered on the left and right instead. In addition, many journalists worry that readers are unable to distinguish factual reporting and opinion pieces [67]. Future designs might explore guessing around these differences.

6.2 Designing for Occasional Use

We observed several challenges in the use of NewsGuesser: participants became frustrated when their guessing accuracy was low and bored after reading multiple full articles on the same topic. Finally, the guessing itself is time-consuming. Researchers recognize a variety of news consumption practices, including *snacking*, *scanning*, and *checking*, that are shorter and less intensive than deeply reading [15]. Thus, NewsGuesser may not align with readers' news reading habits. There may be ways to design systems that integrate guessing and exposure to different leanings that do complement people's existing news consumption habits. Prior work has explored "pushy" systems that can increase users' exposure to cross-cutting content [18, 40, 55]. And our work suggests that it may be possible to add lightweight attempts to trigger curiosity (for example, guessing in direct snippet comparisons discussed in Section 6.1.1). Still, if new approaches are optional, they will need to be attractive to maintain reader interest over longer time spans. Prior work has found that adherence to even desired behavior changes can be low [51]. There are strategies for reducing lapses, including identifying strong social support networks [51], providing rewards [81], and tailoring the context or environment [8], but changes in behavior require considerable effort. While curiosity has been shown to increase task persistence in the past [43], future work would need to explore how to maintain long-term interest in a news consumption context.

However, these challenges suggest that using the tool to reflect on media bias and selective exposure may be more suitable to be used periodically instead of as people's daily news source. Participants suggested that they would like to occasionally use this tool to evaluate their personal selective exposure. This might work well when people are unsure about issues or when major news events occur (e.g. the presidential election). Users could then occasionally check-in, and use tools like NewsGuesser to test their exposure and leaning. Similar occasional use has been a successful approach in other fields, like privacy decision making online, where users find privacy settings useful, but only interact with them occasionally [47]. Designing interventions that do not need to disrupt users' everyday activities, but periodically encourage them to revisit their beliefs and behaviors may be the most productive approach.

6.3 Balancing Surprise and Reflection

Participants often experienced surprise after the news sources results were revealed. Surprise plays an important role in the guessing process as it can trigger participants' interest in comparing the differences between the given news and news they normally read. The feeling of surprise also forced them to consider and re-evaluate their reading habits. The function of surprise in our study is similar to its role in learning science – as it can prompt interest, curiosity, confusion and further promote knowledge exploration [65, 74, 77, 78]. Surprise, contrast, and intensity are often used to gain people's attention online [1] (often in dark patterns), but using this to increase people's awareness of selective exposure and media bias has the potential to help them be more critical media consumers. However, surprise is not always beneficial. Research found that happiness and surprise, in particular, are associated with believing and sharing false news headlines [68]. This suggests that gamified approaches should balance these feelings with reflection and consideration, to avoid users relying on their emotions for decision making rather than critical evaluation.

7 DISCUSSION

Our study observed that certain design features help readers reflect and prompt them to engage with more diverse content. We consider how these designs can be integrated into the many, diverse ways that people today already consume news, such as online news reading aggregators, news organizations and social media news feeds. For news reading aggregators, like Ground News or AllSides, which already help readers consider media bias, they may consider using this game to help readers test their personal biases once or periodically.

News organizations that aim to achieve a neutral or centered approach to their content could also use these gamification strategies to evaluate their own articles. News organizations can invite readers to guess the sources, headlines, or photos of their articles to observe whether significant amount of users guessed one leaning over the other. Although automated methods can also be used to test for potential biases [30], this approach has the potential to increase reader engagement with and loyalty to the news organizations.

Finally, this gamified approach can be leveraged on social media platforms. When readers obtain news from social media, they can check news shared by friends or produced by recommender systems. Previous research showed that people are more likely to click on news that is ideologically consistent with their own opinions and choose to consume more news shared by like-minded people in their social network [4]. To bring up more opportunities to expose to opposing viewpoints and prevent polarized news sharing, social news feeds can adopt our guessing approach, which could provide users an option to hide the news source while sharing news and ask others to guess [52]. This could help reduce the likelihood that readers pick news only based on news sources or based on their friends' ideology. Further, it could provide the opportunity for readers to share their results with friends on the social platform, motivating others to engage in the game or even read more diverse types of news. Social news feeds can also incorporate visualization interfaces like those we developed. For example, after collecting users' data over time, they could measure their selective exposure and use our seesaw visualization to reveal to users their personal biases, similar to the browser extension Balancer [54]. After understanding their personal selective exposure, readers might even use the results to customize their news feed. We anticipate a wide variety of ways to incorporate curiosity, guessing, and surprise into existing systems to encourage users to reflect on and reduce their selective exposure.

In doing so, it will be important to continue to study their effects, ideally with longitudinal studies that are able to investigate whether these systems result in any long-term behavioral or attitudinal changes. For example, systems that are instrumented to track engagement (reading time, guessing activity, annotation, or sharing), could look for changes as people either continue to use, or periodically check in about their selective exposure. However, as our work highlighted the benefits of reflection as part of this exposure, researchers might also consider diary studies or other recordings of self-reflections as participants use these tools over time.

8 LIMITATIONS

There is a long-standing interest in *selection bias* and *coverage bias* in the news media, which influence which stories are selected and how much attention is given to a story, respectively [16]. Since the system design of NewsGuesser required that news events be covered by left-, center-, and right-leaning news sources, this leads to an inability to reveal selection or coverage bias and is a limitation of this work. In addition, because this study involved only a single experience with NewsGuesser, we could measure neither actual changes in behavior nor the durability of any effects on changed perceptions. Future work could address both of these limitations.

9 CONCLUSION

We designed a news reading system called NewsGuesser, which invites readers to guess a news source's partisan lean as they read. We evaluate the system with an interview with 40 participants. The system was able to guide many participants to reflect on their own biases and selective exposure. And several participants even reported updating their impression of specific news organizations and expressed an interest in engaging with more diverse news. This effort showed the potential of curiosity in news reading systems to help combat selective exposure.

ACKNOWLEDGMENTS

We thank Kai-ning Keng for engaging in ideation and for his help in the system design.

REFERENCES

[1] 2010. Streams of content, limited attention: The flow of information through social media. *Educause Review* (2010).
[2] Hayward P Andres. 2019. Active teaching to manage course difficulty and learning motivation. *Journal of Further and Higher Education* 43, 2 (2019), 220–235.
[3] Maria-Virginia Aponte, Guillaume Levieux, and Stephane Natkin. 2011. Measuring the level of difficulty in single player video games. *Entertainment Computing* 2, 4 (2011), 205–213.
[4] Eytan Bakshy, Solomon Messing, and Lada A Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.
[5] Martina Balestra, Orit Shaer, Johanna Okerlund, Madeleine Ball, and Oded Nov. 2016. The effect of exposure to social annotation on online informed consent beliefs and behavior. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 900–912.
[6] Daniel Ellis Berlyne. 1954. A theory of human curiosity. *British journal of psychology* 45, 3 (1954), 180.
[7] Daniel E Berlyne. 1960. Conflict, arousal, and curiosity. (1960).
[8] Mark E Bouton. 2014. Why behavior change is difficult to sustain. *Preventive medicine* 68 (2014), 29–36.
[9] Richard E Boyatzis. 1998. *Transforming qualitative information: Thematic analysis and code development.* sage.
[10] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
[11] Garvin Brod. 2021. Predicting as a learning strategy. *Psychonomic Bulletin & Review* (2021).
[12] Garvin Brod, Marcus Hasselhorn, and Silvia A Bunge. 2018. When generating a prediction boosts learning: The element of surprise. *Learning and Instruction* 55 (2018), 22–31.
[13] Pew Research Center. 2019. For Local News, Americans Embrace Digital but Still Want Strong Community Connection. (2019).
[14] Philip Cook, Douglas Gomery, and Lawrence Lichty. 1992. *The Future of News: Television, Newspapers, Wire Services, Newsmagazines.* Woodrow Wilson Center Press.
[15] Irene Costera Meijer and Tim Groot Kormelink. 2015. Checking, sharing, clicking and linking: Changing patterns of news use between 2004 and 2014. *Digital journalism* 3, 5 (2015), 664–679.
[16] Dave D'Alessio and Mike Allen. 2000. Media bias in presidential elections: A meta-analysis. *Journal of communication* 50, 4 (2000), 133–156.
[17] Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. 2010. Opinion space: a scalable tool for browsing online comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1175–1184.
[18] Tom Feltwell, Gavin Wood, Phillip Brooker, Scarlett Rowland, Eric PS Baumer, Kiel Long, John Vines, Julie Barnett, and Shaun Lawson. 2020. Broadening Exposure to Socio-Political Opinions via a Pushy Smart Home Device. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
[19] Linda G Fielding et al. 1990. How Discussion Questions Influence Children's Story Understanding. Technical Report No. 490. (1990).
[20] Marta Figlerowicz. 2015. GeoGuessr's Digital Pilgrimages. *Room One Thousand* 3, 3 (2015).
[21] Seth Flaxman, Sharad Goel, and Justin M Rao. 2016. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly* 80, S1 (2016), 298–320.

[22] Meadhbh I Foster and Mark T Keane. 2019. The role of surprise in learning: different surprising outcomes affect memorability differentially. *Topics in cognitive science* 11, 1 (2019), 75–87.
[23] Dieter Frey. 1986. Recent research on selective exposure to information. *Advances in experimental social psychology* 19 (1986), 41–80.
[24] Mingkun Gao, Hyo Jin Do, and Wai-Tat Fu. 2017. An intelligent interface for organizing online opinions on controversial topics. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces*. 119–123.
[25] Kiran Garimella, Gianmarco De Francisc iMorales, Aristides Gionis, and Michael Mathioudakis. 2017. Mary, mary, quite contrary: Exposing twitter users to contrarian news. In *Proceedings of the 26th International Conference on World Wide Web Companion*. 201–205.
[26] Kiran Garimella, Tim Smith, Rebecca Weiss, and Robert West. 2021. Political Polarization in Online News Consumption. *arXiv preprint arXiv:2104.06481* (2021).
[27] Dedre Gentner, Jeffrey Loewenstein, and Leigh Thompson. 2003. Learning and transfer: A general role for analogical encoding. *Journal of educational psychology* 95, 2 (2003), 393.
[28] Nabeel Gillani, Ann Yuan, Martin Saveski, Soroush Vosoughi, and Deb Roy. 2018. Me, my echo chamber, and I: introspection on social media polarization. In *Proceedings of the 2018 World Wide Web Conference*. 823–831.
[29] Github. 2022. *news-please · PyPI*. https://pypi.org/project/news-please/
[30] Tim Groeling. 2013. Media Bias by the Numbers: Challenges and Opportunities in the Empirical Study of Partisan News. *Annual Review of Political Science* (2013).
[31] Neil Harrington. 2005. It's too difficult! Frustration intolerance beliefs and procrastination. *Personality and Individual Differences* 39, 5 (2005), 873–883.
[32] William Hart, Dolores Albarracín, Alice H Eagly, Inge Brechan, Matthew J Lindberg, and Lisa Merrill. 2009. Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin* 135, 4 (2009), 555.
[33] Kayoko Inagaki and Giyoo Hatano. 1977. Amplification of cognitive motivation and its effects on epistemic observation. *American Educational Research Journal* 14, 4 (1977), 485–491.
[34] Youngseung Jeon, Bogoan Kim, Aiping Xiong, Dongwon Lee, and Kyungsik Han. 2021. ChamberBreaker: Mitigating the Echo Chamber Effect and Supporting Information Hygiene through a Gamified Inoculation System. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–26.
[35] Celeste Kidd and Benjamin Y Hayden. 2015. The psychology and neuroscience of curiosity. *Neuron* 88, 3 (2015), 449–460.
[36] Silvia Knobloch-Westerwick, Axel Westerwick, and Benjamin K. Johnson. 2015. Selective Exposure in the Communication Technology Context. In *The Handbook of the Psychology of Communication Technology*, S. Shyam Sundar (Ed.).
[37] Nate Kornell, Matthew Jensen Hays, and Robert A Bjork. 2009. Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35, 4 (2009), 989.
[38] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting reflective public thought with considerit. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. 265–274.
[39] Travis Kriplean, Michael Toomim, JT Morgan, Alan Borning, and AJ Ko. 2011. REFLECT: Supporting active listening and grounding on the Web through restatement. In *Proceedings of the Conference on Computer Supported Cooperative Work, Hangzhou, China*.
[40] Travis Kriplean, Michael Toomim, Jonathan Morgan, Alan Borning, and Amy J Ko. 2012. Is this what you meant? Promoting listening on the web with reflect. In *proceedings of the SIGCHI conference on human factors in computing systems*. 1559–1568.
[41] Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P Gummadi, and Karrie Karahalios. 2017. Quantifying search bias: Investigating sources of bias for political searches in social media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 417–432.
[42] Preethi Lahoti, Kiran Garimella, and Aristides Gionis. 2018. Joint non-negative matrix factorization for learning ideological leaning on twitter. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. 351–359.
[43] Edith Law, Ming Yin, Joslin Goh, Kevin Chen, Michael A Terry, and Krzysztof Z Gajos. 2016. Curiosity killed the cat, but makes crowdwork better. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 4098–4110.
[44] Heidi Legg. 2021. Have you heard the news? But who owns what you're hearing and reading? We need to know. *USA Today* (2021).
[45] Q Vera Liao and Wai-Tat Fu. 2013. Beyond the filter bubble: interactive effects of perceived threat and topic involvement on selective exposure to information. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2359–2368.

[46] Q Vera Liao and Wai-Tat Fu. 2014. Can you hear me now? Mitigating the echo chamber effect by source position indicators. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 184–196.

[47] Yabing Liu, Krishna P Gummadi, Balachander Krishnamurthy, and Alan Mislove. 2011. Analyzing facebook privacy settings: user expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. 61–70.

[48] George Loewenstein. 1994. The psychology of curiosity: A review and reinterpretation. *Psychological bulletin* 116, 1 (1994), 75.

[49] Elisa D Mekler, Florian Brühlmann, Alexandre N Tuch, and Klaus Opwis. 2017. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior* 71 (2017), 525–534.

[50] Janet Metcalfe and Bridgid Finn. 2012. Hypercorrection of high confidence errors in children. *Learning and Instruction* 22, 4 (2012), 253–261.

[51] Kathryn R Middleton, Stephen D Anton, and Michal G Perri. 2013. Long-term adherence to health behavior change. *American journal of lifestyle medicine* 7, 6 (2013), 395–404.

[52] Lillio Mok, Michael Inzlicht, and Ashton Anderson. 2023. Echo Tunnels: Polarized News Sharing Online Runs Narrow but Deep. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 662–673.

[53] Edward Munnich and Michael A Ranney. 2019. Learning from surprise: Harnessing a metacognitive surprise signal to build and adapt belief networks. *Topics in cognitive science* 11, 1 (2019), 164–177.

[54] Sean Munson, Stephanie Lee, and Paul Resnick. 2013. Encouraging reading of diverse political viewpoints with a browser widget. In *Proceedings of The International AAAI Conference on Web and Social Media*, Vol. 7.

[55] Sean Munson and Paul Resnick. 2010. Presenting Diverse Political Opinions: How and How Much. In *Proceedings of CHI*.

[56] Ground News. 2022. Our Approach to Media Bias. https://ground.news/media-bias.

[57] Christine Ogan, Rosemary Pennington, Olesya Venger, and Daniel Metz. 2018. Who drove the discourse? News coverage and policy framing of immigrants and refugees in the 2016 U.S. presidential election. *Communications* (2018).

[58] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you.* Penguin UK.

[59] Souneil Park, Seungwoo Kang, Sangyoung Chung, and Junehwa Song. 2009. NewsCube: delivering multiple aspects of news to mitigate media bias. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 443–452.

[60] Neil Peirce, Owen Conlan, and Vincent Wade. 2008. Adaptive educational games: Providing non-invasive personalised learning experiences. In *2008 second IEEE international conference on digital game and intelligent toy enhanced learning*. IEEE, 28–35.

[61] Erik Peterson, Sharad Goel, and Shanto Iyengar. 2018. Echo chambers and partisan polarization: Evidence from the 2016 presidential campaign. *Unpublished manuscript. https://5harad. com/papers/selecfive-exposure. pdf* (2018).

[62] Richard E Petty and John T Cacioppo. 2012. *Communication and persuasion: Central and peripheral routes to attitude change.* Springer Science & Business Media.

[63] W James Potter. 2010. The state of media literacy. *Journal of broadcasting & electronic media* 54, 4 (2010), 675–696.

[64] Stacy Rebich-Hespanha, Ronald E Rice, Daniel R Montello, Sean Retzloff, Sandrine Tien, and João P Hespanha. 2015. Image themes and frames in US print news stories about climate change. *Environmental Communication* 9, 4 (2015), 491–519.

[65] K Ann Renninger and Suzanne E Hidi. 2015. *The power of interest for motivation and engagement.* Routledge.

[66] Ronald E Robertson, Shan Jiang, Kenneth Joseph, Lisa Friedland, David Lazer, and Christo Wilson. 2018. Auditing partisan audience bias within google search. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–22.

[67] Tom Rosenstiel, Nicole Willcoxon, Emily Swanson, Kevin Loker, Jeff Sonderman, Katie Kutso, Jane Elizabeth, Katherine Ellis, David Sterrett, Dan Malato, Liz Kantor, Jennifer Benz, Trevor Tompson, and Xian Tao. 2018. Americans and the News Media: What they do — and don't — understand about each other. *The Media Insight Project* (2018).

[68] Leah R Rosenzweig, Bence Bago, Adam J Berinsky, and David G Rand. 2021. Happiness and surprise are associated with worse truth discernment of COVID-19 headlines among social media users in Nigeria. *Harvard Kennedy School Misinformation Review* (2021).

[69] Daniel M Russell, Gregorio Convertino, Aniket Kittur, Peter Pirolli, and Elizabeth Anne Watkins. 2018. Sensemaking in a senseless world: 2018 workshop abstract. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–7.

[70] Jennifer L Sabourin and James C Lester. 2013. Affect and engagement in Game-BasedLearning environments. *IEEE transactions on affective computing* 5, 1 (2013), 45–56.

[71] Donald A Schon. 1984. *The reflective practitioner: How professionals think in action.* Vol. 5126. Basic books.

[72] Norman J Slamecka and Peter Graf. 1978. The generation effect: Delineation of a phenomenon. *Journal of experimental Psychology: Human learning and Memory* 4, 6 (1978), 592.
[73] Beautiful Soup. 15. *Beautiful Soup Documentation — Beautiful Soup 4.4.0 documentation*. https://beautiful-soup-4.readthedocs.io/en/latest/
[74] Joachim Stiensmeier-Pelster, Alice Martini, and Rainer Reisenzein. 1995. The role of surprise in the attribution process. *Cognition & Emotion* 9, 1 (1995), 5–31.
[75] Natalie Jomini Stroud. 2008. Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior* 30, 3 (2008), 341–366.
[76] Gallup & the Knight Foundation. 2019. State of Public Trust in Local News. (2019).
[77] Elisabeth Vogl, Reinhard Pekrun, Kou Murayama, and Kristina Loderer. 2020. Surprised–curious–confused: Epistemic emotions and knowledge exploration. *Emotion* 20, 4 (2020), 625.
[78] Elisabeth Vogl, Reinhard Pekrun, Kou Murayama, Kristina Loderer, and Sandra Schubert. 2019. Surprise, curiosity, and confusion promote knowledge exploration: Evidence for robust effects of epistemic emotions. *Frontiers in psychology* 10 (2019), 2474.
[79] Jessica Z Wang, Amy X Zhang, and David R Karger. 2022. Designing for Engaging with News using Moral Framing towards Bridging Ideological Divides. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–23.
[80] 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 conference on human factors in computing systems*. 1–12.
[81] Wendy Wood and David T Neal. 2016. Healthy through habit: Interventions for initiating & maintaining health behavior change. *Behavioral Science & Policy* 2, 1 (2016), 71–83.

Received January 2023; revised July 2023; accepted November 2023